Journal of Healthcare Engineering

Augmented Reality in Healthcare

Lead Guest Editor: Vincenzo Ferrari
Guest Editors: Gudrun Klinker and Fabrizio Cutolo

Augmented Reality in Healthcare

Journal of Healthcare Engineering

Augmented Reality in Healthcare

Lead Guest Editor: Vincenzo Ferrari
Guest Editors: Gudrun Klinker and Fabrizio Cutolo

Editorial Board

Contents

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 9321535, 2 pages
https://doi.org/10.1155/2019/9321535

Editorial
Augmented Reality in Healthcare

Vincenzo Ferrari,[1] **Gudrun Klinker**,[2] **and Fabrizio Cutolo**[1]

[1]*Dipartimento di Ingegneria dell'Informazione, Università di Pisa, Largo Lucio Lazzarino, 2, 56122 Pisa, Italy*
[2]*Institut für Informatik, München Technische Universität München, Boltzmannstraße 3 D-85748 Garching b. München, Munich, Germany*

Correspondence should be addressed to Vincenzo Ferrari; vincenzo.ferrari@unipi.it

Received 21 May 2019; Accepted 21 May 2019; Published 1 July 2019

Augmented reality (AR) technologies for the consumer market are nowadays mature for many potential fields of applications. In the healthcare sector, as demonstrated by the increasing number of publications on AR for surgery, medicine, and rehabilitation, there is a great demand for solutions that are able to improve current clinical practice. The aim of this special issue is to offer to engineers, computer scientists, and final users an overview of the potentials of AR technologies in fostering the development of useful applications in the early future and to steer the academic research towards overcoming the technological and human-factor issues still present among the current devices and among the most popular modalities for enriching the visual sensation with computer-generated elements.

Sixteen papers were submitted for this special issue. Our distinguished reviewers from respective research fields narrowed the field down to six papers which were finally accepted.

In this special issue, the reader can find useful examples of applications in the healthcare domain from doctor-patient communication up to surgery, rehabilitation, and phobia treatments.

Even if AR devices and applications are to date mostly devoted to augmenting the sense of sight, and the augmentation of different senses has not yet reached the same widespread diffusion, Z. Qin et al. show us in their work the potential of haptic feedback towards increasing the user's accessibility and allowing an intuitive and natural interaction with computer-generated elements.

From a technological standpoint, it is important to outline that, as confirmed in R. Touati et al., video-based tracking can be done through feature detection on the patient with a marker-less tracking approach.

Overall, it is often difficult to decide where exactly within the reality-virtuality continuum a specific AR application should be located. This is especially true for medical AR, where a lot of patient-specific data and images are available and sometimes it is almost impossible to clearly define to what extent a digital content shown on a display is real or virtual. In some research works in the healthcare sector, this debate becomes a pure comparison between VR and AR while the final goal of the application is lost. In this special issue, the reader can see that there are many ways by which the real and virtual information can be acquired and merged in a useful way for the user. M. Melero et al. show that, for some applications, the visualization of both VR and AR modalities can be an added value for the patient, while C.-F. Tsai et al. prove that the VR and AR visualization modalities stimulate different physiological reactions.

In almost all proposed applications, the AR view is shown on a traditional stand-up monitor. In the case of endoscopic procedures, as in V. Mamone et al., this choice is the best in terms of ergonomics since the users usually see the endoscope images directly in front of them. In principle, in the case of manual procedures, the optimal choice should be the use of a wearable display but, as confirmed by S. Condino et al., there are still perceptual limits to take into account, such as a parallax error and/or

a focus rivalry between real and virtual, that can compromise their efficacy in a real scenario.

Conflicts of Interest

The editors declare that there are no conflicts of interest regarding the publication of this special issue.

Vincenzo Ferrari
Gudrun Klinker
Fabrizio Cutolo

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 5613931, 12 pages
https://doi.org/10.1155/2019/5613931

Research Article

Low-Computational Cost Stitching Method in a Three-Eyed Endoscope

Virginia Mamone,[1,2] **Sara Condino**,[1,2] **Fabrizio Cutolo**,[1,2] **Izadyar Tamadon**,[3] **Arianna Menciassi**,[3] **Michele Murzi**,[4] **Mauro Ferrari**,[2,5] **and Vincenzo Ferrari**[1,2]

[1]*Information Engineering Department, University of Pisa, Pisa 56122, Italy*
[2]*EndoCAS Center for Computer-Assisted Surgery, Pisa 56124, Italy*
[3]*Sant'Anna School of Advanced Studies, Pisa 56127, Italy*
[4]*Department of Adult Cardiac Surgery, G. Pasquinucci Heart Hospital, Gabriele Monasterio Foundation, Massa 54100, Italy*
[5]*Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy*

Correspondence should be addressed to Virginia Mamone; virginia.mamone@endocas.unipi.it

Received 30 November 2018; Revised 1 March 2019; Accepted 24 April 2019; Published 17 June 2019

Academic Editor: Rafael Morales

Aortic valve replacement is the only definitive treatment for aortic stenosis, a highly prevalent condition in elderly population. Minimally invasive surgery brought numerous benefits to this intervention, and robotics recently provided additional improvements in terms of telemanipulation, motion scaling, and smaller incisions. Difficulties in obtaining a clear and wide field of vision is a major challenge in minimally invasive aortic valve surgery: surgeon orientates with difficulty because of lack of direct view and limited spaces. This work focuses on the development of a computer vision methodology, for a three-eyed endoscopic vision system, to ease minimally invasive instrument guidance during aortic valve surgery. Specifically, it presents an efficient image stitching method to improve spatial awareness and overcome the orientation problems which arise when cameras are decentralized with respect to the main axis of the aorta and are nonparallel oriented. The proposed approach was tested for the navigation of an innovative robotic system for minimally invasive valve surgery. Based on the specific geometry of the setup and the intrinsic parameters of the three cameras, we estimate the proper plane-induced homographic transformation that merges the views of the operatory site plane into a single stitched image. To evaluate the deviation from the image correct alignment, we performed quantitative tests by stitching a chessboard pattern. The tests showed a minimum error with respect to the image size of $0.46 \pm 0.15\%$ measured at the homography distance of 40 mm and a maximum error of $6.09 \pm 0.23\%$ at the maximum offset of 10 mm. Three experienced surgeons in aortic valve replacement by mini-sternotomy and mini-thoracotomy performed experimental tests based on the comparison of navigation and orientation capabilities in a silicone aorta with and without stitched image. The tests showed that the stitched image allows for good orientation and navigation within the aorta, and furthermore, it provides more safety while releasing the valve than driving from the three separate views. The average processing time for the stitching of three views into one image is 12.6 ms, proving that the method is not computationally expensive, thus leaving space for further real-time processing.

1. Introduction

In recent years, the main risk factors for heart disease, such as smoking, high cholesterol, and high blood pressure, have increased. Aortic stenosis (AS) is the third most prevalent form of cardiovascular disease in the Western world, with a 50% risk of death in the three years following the onset of symptoms [1]. AS develops from progressive calcification of leaflets, reducing the leaflet opening over time, until the valve is no longer able to control the blood flow. It is a highly prevalent condition, being present in 21–26% of elderly above 65 [2] and no pharmacologic treatment showed to be effective, nor attenuating the progressive valve calcification, nor improving survival [3]. The only definitive treatment for

AS in adults is aortic valve replacement (AVR). For decades, surgical AVR has been the standard treatment for severe AS. The traditional open-heart method gives the surgeon direct access to the heart through median sternotomy. But despite allowing excellent access to all cardiac structures, the open-heart method requires complete division of the sternum and sternal spreading, thus disrupting the integrity of the chest wall in the early recovery phase. Since the first surgical AVR intervention in 1960, less invasive methods have been investigated to complete the operation. In 1997 and 1998, surgeons performed the first intervention, respectively, in right mini-thoracotomy and mini-sternotomy [4, 5]. A few years later, a percutaneous transcatheter approach was tested, which to date is the standard intervention for high-risk patients. However, for low-risk patients, this approach is not advisable due to the increase in the chance of paravalvular regurgitation, implantation of pacemaker, and a worst 3-year survival [6].

For many patients, the best solution remains surgery via mini-thoracotomy or mini-sternotomy. These techniques offer the typical benefits of minimally invasive surgery, such as decrease in blood transfusion, hospital stay, and improved cosmesis, demonstrated not only in the cardiothoracic district but also in the vascular one [7]. Through robotics, further benefits can be reached in terms of telemanipulation, motion scaling, and even smaller incisions [8, 9]. Researchers proposed various robotic systems to assist heart surgery [10], allowing in some cases the preoperative planning to test the surgical case before the actual intervention [11].

A major challenge of minimally invasive techniques, including those based on a robotic solution, remain visualization: the surgeon lacks the direct view of the operative field and has a poor spatial awareness in the rather limited available space. Augmented reality is a promising asset in the context of image-guided surgery [12–14]. In endoscopic techniques, surgeon usually focuses on a stand-up monitor to perform the surgical task. Indeed, in such procedures, the surgeon operates watching endoscopic video images reproduced on the spatial display unit. Therefore, in endoscopic surgery, the augmented information is obtained by merging endoscopic video frames with virtual content useful for increasing spatial awareness and for aiding the surgical tasks [15]. Recent in vitro and cadaver studies have proven the efficacy of augmented reality in assisting endoscopic surgery [16–20].

To increase the field of view and offer a wider vision of the operative field, some solutions suggest the use of two cameras and the application of stitching techniques to merge the pairs of images into one [21]. However, traditional methods for stitching images such as the SIFT [22], SURF [23], or ORB [24] algorithms are computationally expensive, because they require the identification of features in each image and the search for correspondences between each image pair. This makes these powerful and effective methods restraining in real-time application. Other solutions consist of expanding the surgeon's field of view through dynamic view expansion: in a recent work, images from a single camera are merged using simultaneous localization and mapping (SLAM) to generate a sparse probabilistic 3D map of the surgical site [25]. The problem of visualization in minimally invasive systems is substantial, and several companies already provide systems such as the Third Eye Retroscope and Third Eye Panoramic, which allow framing larger areas through auxiliary systems, or the Fuse (full-spectrum endoscopy), a complete colonoscopy platform including a video colonoscope and a processor. The systems currently on the market, however, are mostly developed for colonoscopy or gastroscopy and cannot be integrated for use in heart operations due to the different morphology and surgical task.

This work focuses on the development of a computer vision methodology to increase the field of view and offer a wider vision of the operative field. The proposed methodology can be used for the navigation of any minimally invasive instrumentations, including robotic systems.

In this paper, the proposed approach was tested for the guidance of the robotic system for minimally invasive valve surgery developed in [26]. The robot is a flexible manipulator, having omnidirectional bending capabilities. Endoscopic vision through three small cameras on the robot tip is aimed to aid the surgeons in accurately positioning the aortic valve. The objective of this work was to evaluate the best way to merge the information coming from the cameras in real time, making it easier for the surgeon to orientate through the operative field. At the same time, it was also necessary to contain the computational cost to allow real-time operation while including the required computations to other device features.

2. Materials and Methods

The following sections describe computer vision issues to implement an image stitching method for a generic three-eyed endoscopic system. Then, the specific image-guided robotic platform for minimally invasive aortic valve replacement is presented together with the setup used to perform preliminary test.

2.1. Image Stitching for Three-Eyed Cameras Endoscope. The proposed approach was developed for three-eyed endoscopic systems, which compared to classical 2D mono-view endoscopic instruments offer improved navigation functionalities since it can allow for triangulation and stereo reconstruction and can offer a wider vision of the operative field. In these systems, however, the different off-axis viewpoints provide visual information not as usable if compared to the usual endoscopic view, where a single camera is centered on the workspace. This can be aggravated by the fact that, to facilitate the other functionalities, the cameras can be nonparallel oriented.

The proposed image stitching method can be performed by applying an appropriate image warping based on the estimation of the three plane-induced homographies between each camera and a virtual camera placed at their barycenter and oriented as one of them, chosen as a reference.

Figure 1 shows a possible camera configuration in an endoscopic instrument with a central operative lumen. We

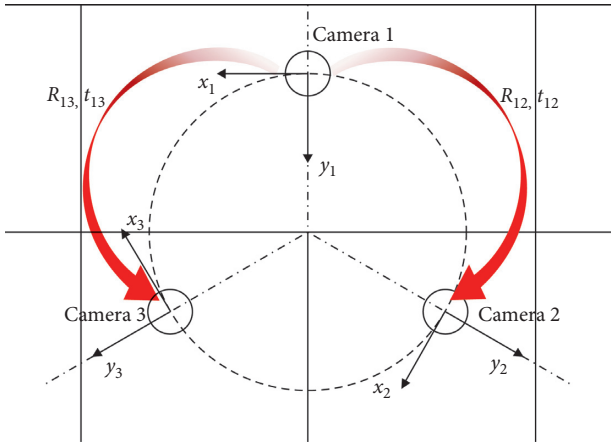


Figure 1: Possible camera configuration in an endoscopic instrument. The reference systems of the three cameras are oriented radially with respect to the manipulator axis. Camera 1 is the reference camera.

will refer to the three cameras with numbers from 1 to 3, number 1 being associated with the reference camera.

The following paragraphs describe the steps employed to achieve an accurate and reliable image stitching: we start from the description of the employed methods for camera calibration, and then, we introduce the basic concept of homographic transformations before describing the employed image stitching procedure.

2.1.1. Camera Calibration. Camera calibration, which involves the estimation of the camera intrinsic and extrinsic parameters, is the first essential procedure.

Plane-based camera calibration methods, as the well-known Zhang's method [27], which requires the camera to observe a planar calibration pattern at a few unknown orientations can be applied. At first, the matches between 3D world points (corners of a given chessboard) and their corresponding 2D image points are found. For this purpose, in this work, we used a 4×5 chessboard calibration pattern with 5 mm square side.

To estimate the matrix of the intrinsic linear parameters, K_i, and the radial distortion coefficients, the intrinsic calibration is carried out for each camera, i. In our method, we proposed the use of a radial distortion model with only two coefficients, neglecting the tangential distortion. The MATLAB calibration toolbox, which allows the selection of the most appropriate images and the elimination of any outliers based on the retroprojection error, was used. We ensured that all image areas were covered by the grid, to get the most accurate estimate of the distortion parameters.

After the estimation of the three cameras' intrinsic parameters, the extrinsic calibration is performed to get the relative poses of the three cameras: this phase allows estimating the reciprocal poses of the cameras, minimizing the total reprojection error for all the corner points in all the available views from the camera pairs. This phase can be performed in C++ language using the OpenCV libraries. The substantial advantage of using this environment is the possibility to perform the calibration whilst keeping the intrinsic parameters obtained in the previous calibration as fixed. By doing so, the estimation of the extrinsic parameters can be more accurate than estimating intrinsic parameters and relative poses together; the high dimensionality of the parameter space and the low signal-to-noise ratio in the input data can cause the function to diverge from the correct solution. Also, working with three pairs of cameras and dealing with the estimation of three reciprocal poses, it is essential to use univocal intrinsic parameters, so that the resulting poses are consistent.

The output of this step is two rigid transformations, in form of rototranslation matrices, relating the reference systems of camera 2 and camera 3 with respect to camera 1 reference system. As highlighted in red in Figure 1, in the following paragraph, we will refer to the following:

(i) R_{12} and t_{12} as the rotation and translation component of the rigid transformation from camera 1 to camera 2

(ii) R_{13} and t_{13} as the rotation and translation component of the rigid transformation from camera 1 to camera 3

2.1.2. Homographic Transform. A plane-induced homography is a projective transformation that relates the images of a reference plane in the world, grabbed by two generic cameras placed at different positions and/or orientations. Such homography describes the pixel-to-pixel relation between two camera images, x_i and x_j, as follows:

$$\lambda x_j = H_{ij}\left(R_{ij}, t_{ij}, K_i, K_j, \pi\right)x_i. \tag{1}$$

The image points from two cameras, x_i and x_j, are expressed in homogeneous coordinates, and λ is the generic scale factor due to the equivalence of homogeneous coordinate rule. The homography is a function of the relative pose between the two cameras (R_{ij}, t_{ij}), the intrinsic parameters of the two cameras (K_i, K_j), and the position and orientation of the reference plane in the scene with respect to the camera i. H_{ij} can be broken down as follows:

$$H_{ij} = K_j\left(R_{ij} + \frac{t_{ij} \cdot n_i'}{d_i}\right)K_i^{-1}, \tag{2}$$

where n is the normal unit vector of the reference homography plane with respect to the camera i and d is the distance between the origin of the camera i reference system and the plane.

2.1.3. Image Stitching. Image stitching can be performed by applying an appropriate warping of the camera images based on the estimation of the three plane-induced homographies between each camera image and a virtual camera, ideally placed at their barycenter [28]. This allows us to remap each camera view on an ideal and central viewpoint of the operatory site plane.

The three homographic transformations that relate the views of each camera to the virtual camera are

$$H_{1v} = K_v\left(R_{1v} + \frac{t_{1v} \cdot n_1'}{d_1}\right)K_1^{-1},$$
$$H_{2v} = K_v\left(R_{2v} + \frac{t_{2v} \cdot n_2'}{d_2}\right)K_2^{-1}, \quad (3)$$
$$H_{3v} = K_v\left(R_{3v} + \frac{t_{3v} \cdot n_3'}{d_3}\right)K_3^{-1},$$

where K_v represents the intrinsic matrix of the virtual camera, which is associated with a wider field of view than the single cameras to encompass all three individual views. The parameters n_i and d_i differ, although they refer to the same plane, as both are relative to the associated real camera reference system.

The rototranslations relating the poses of each camera to the virtual camera are provided by the calibration as follows:

$$\begin{aligned} R_{1v} &= I, \\ t_{1v} &= -b, \\ R_{2v} &= R_{12}^{-1}, \\ t_{2v} &= -R_{12}^{-1}t_{12} - b, \\ R_{3v} &= R_{13}^{-1}, \\ t_{3v} &= -R_{13}^{-1}t_{13} - b, \end{aligned} \quad (4)$$

where b is the center of gravity position in the camera 1 reference system.

2.2. Three-Eyed Cameras Endoscopic Vision for Robotic Aortic Valve Replacement. In minimally invasive heart valve surgery, surgeons can replace the aortic valve through a small incision (less than 40 mm) between the ribs. Controllable flexible manipulators, with appropriate endoscopic vision system, are ideal for accessing such areas of the patient's chest through a small entry point.

This application requires a great deal of accuracy in reaching and targeting the proper site of valve implantation, the annulus plane (the cross section with smallest diameter in the blood path between the left ventricle and the aorta). This system can greatly take advantage from the use of multicameras imaging system, with triangulation functionalities to extract in real-time the 3D position of the anatomical target and image stitching functionalities to offer the surgeon with a wider vision of the operative field.

The robotic platform is described in [26]. Figure 2 shows an overall view of the image-guided robotic system: the surgeon controls the robot through joysticks integrated in the control unit, and the navigation is guided by three camera views.

The robotic system is a 5-DoF cable-driven flexible manipulator with internal introducer and a visualization aid, named navigator in Figure 2.

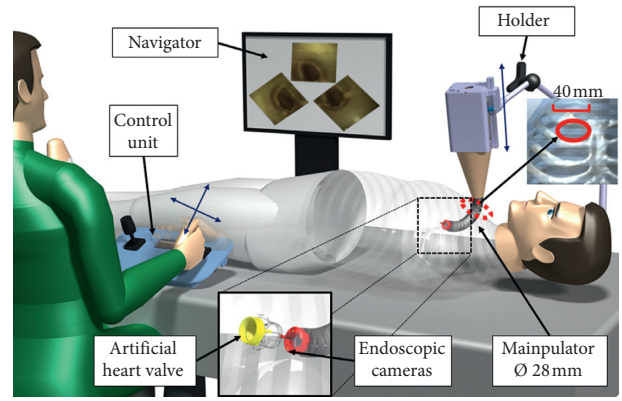


Figure 2: Proposed platform for aortic heart valve surgery. The camera images are processed and presented to the surgeon for navigation. The surgeon uses the joysticks in the control unit to operate the robot, which is held in place by the holder.

In the proposed surgical scenario, the flexible manipulator is attached to a linear actuator and is held fixed by a holder, attached to the patient's bed. The flexible manipulator has omnidirectional bending capabilities which are controlled by 4 set of cables and servomotors. The flexible part is 130–150 mm long with an external diameter of 28 mm and up to 120 degrees of maximum bending.

The surgical procedure requires a cardiopulmonary bypass that provides a bloodless field for the intervention. Next the manipulator, shown in Figure 3, is inserted into the aorta and advanced to the heart. When the manipulator is close enough, three flaps are opened, stabilizing the external part of the manipulator and allowing the internal structure with the new valve, named introducer, to advance. The introducer reaches the annulus, and the valve is rotated around the main axis of the manipulator to match the nadirs of the aortic cusps, as shown in Figure 4.

The surgery is carried out under the guidance of three cameras, positioned on the manipulator 120° from each other along a circumference of 21 mm in diameter. Microcameras (FisCAM, FISBA, Switzerland), 1.95 mm in diameter including illumination, were selected to fit into the reduced dimensions of the system. Illumination is given by LEDs from a separate control box, and it is directed through glass fibers. The specifics of the cameras are shown in Table 1.

2.2.1. Image Stitching for Navigation to the Aortic Annulus. In this application case, the operatory site plane is the aortic annulus, which is the target of the surgical task. Since the exact position and orientation of the annulus plane cannot be estimated a priori, the homography is calculated considering a plane oriented parallel to the virtual camera.

Figure 5(b) shows the plane normal vector, n, defined by this constraint, and the distance from the cameras reference system, d_i. The plane distance d_1 is set to 40 mm, corresponding to the average distance for valve releasing. Figure 5(a) highlights the composition of the stitched images from the views from cameras 1, 2, and 3.

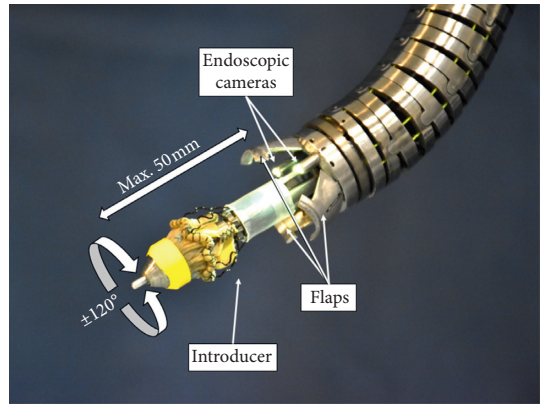


Figure 3: The manipulator. The robotic manipulator exposes the valve preloaded on the introducer. The maximum valve release distance is 50 mm, while the maximum rotation angle is 120 degrees.

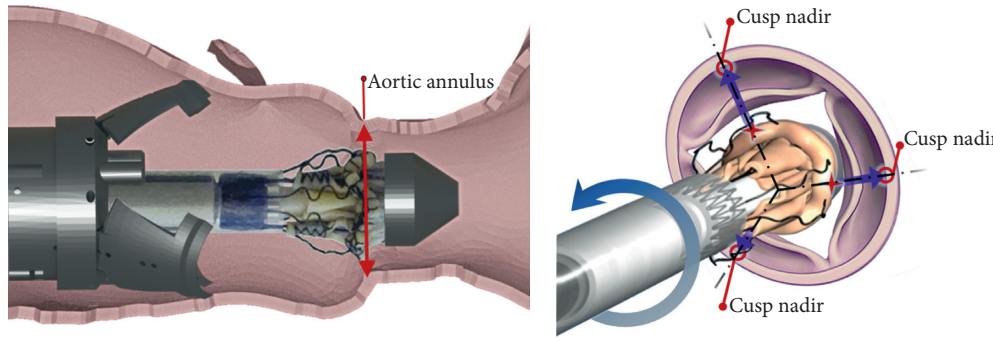


Figure 4: Correct positioning of the introducer. (a) The introducer is aligned with the plane corresponding to the aortic annulus, where the nadirs of the cusps reside. (b) Following a rotation on its own axis, the introducer is oriented to match the nadir of the replacement valve with the nadir of the old calcified valve.

Table 1: Technical specifications of the FisCAM cameras.

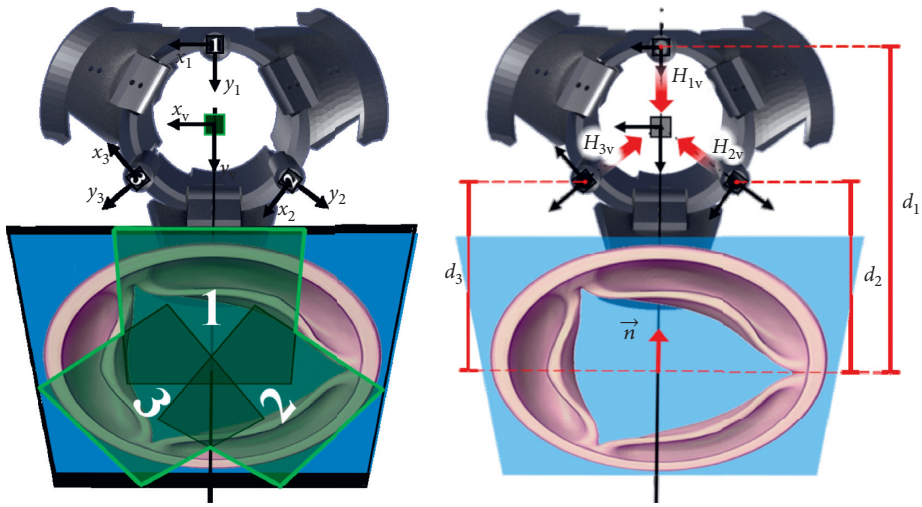
FisCAM, FISBA, Switzerland	
Resolution (px)	400×400 at 30 fps
Working distance (mm)	5–50
Diagonal field of view (degrees)	120°

Figure 6 shows on the left the single views of the three cameras. On the right, the result of the image stitching is shown; in this way, we obtain a single view that includes all the spatial information from the three cameras. The resulting stitched image can be enriched with virtual information content to provide the surgeon with an aid for the correct valve deployment. Figure 7 illustrates the basic concept of an augmented reality (AR) aid which can be implemented to simulate the final positioning of the valve by knowing the position/orientation of the manipulator at the deployment time (this functionality requires that the release of the valve from the manipulator is repeatable and predictable).

2.3. Test. Quantitative test and qualitative test were performed to respectively evaluate the precision of the stitching procedure and the usability of the stitched endoscopic image for the navigation to the nadir point.

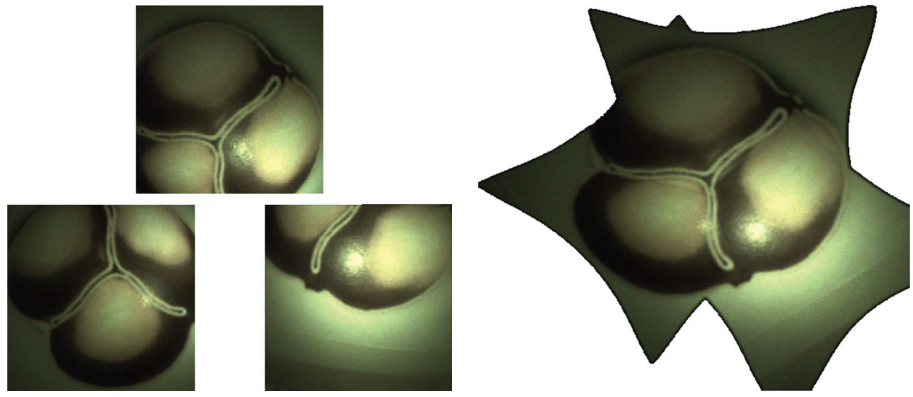
2.3.1. Quantitative Test. Quantitative measures aim at assessing the error in stitching the warped camera images. The error in terms of pixels was measured considering the misalignment of homologous features on pairs of warped images. Such error was evaluated respectively on a plane placed at 40 mm from the reference camera, i.e., the homography plane, and at incremental distances of 2.5 mm up to a depth of 20 mm. An 8×9 chessboard with a square side of 3 mm was used for the evaluation with the corners acting as reference features. In order to evaluate the mismatch introduced by the increasing distance from the homography plane, the chessboard was placed parallel to the image plane of the virtual camera by means of the support structure shown in Figure 8.

2.3.2. Qualitative Usability Tests. Experimental tests were conducted to evaluate the usability of the stitched view during navigation in the aorta to the nadirs and while releasing the valve. Three cardiac surgeons already



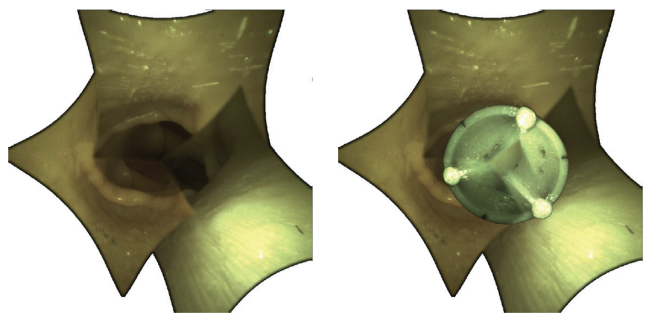
(a) (b)

Figure 5: Basic components of homography and image stitching. (a) Manipulator with the flaps open: the reference systems of cameras 1, 2, and 3 and of the virtual camera (in green) are shown. The homographic plane is highlighted in blue, and the contributions of the three views, merged and captured by the virtual camera, are distinguished. (b) Manipulator with the flaps open: H_1, H_2, and H_3 are the homographic transformations from the view of each camera to the virtual camera. Parameters d_1, d_2, and d_3 represent the distance of each camera from the homographic plane. The vector normal to the plane, n, is unique; however, it assumes different values in the reference systems of the three cameras.



(a) (b)

Figure 6: Single images and corresponding stitched image. The cameras capture a representation of a closed aortic valve at 40 mm. (a) Single views. (b) Stitching view: images are homographed and merged to reproduce the view from the virtual camera.



(a) (b)

Figure 7: Basic concept of the AR support planned to ease the correct valve deployment. The release of the valve is simulated by showing in AR the final positioning the open valve, deployed from the current manipulator position [15].

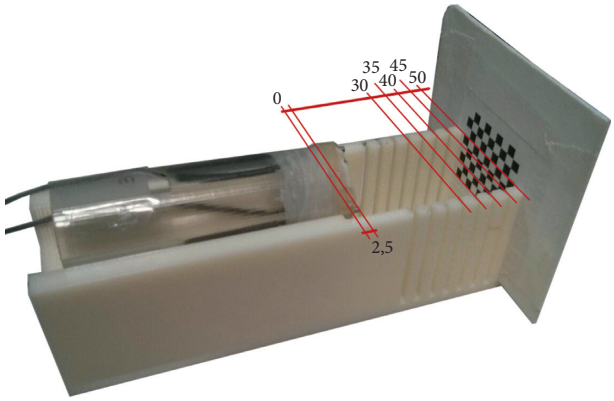


Figure 8: Support structure for quantitative error assessment. The tines are positioned so that the chessboard is parallel to the image plane of the virtual camera. Unit of measurement is in millimeters.

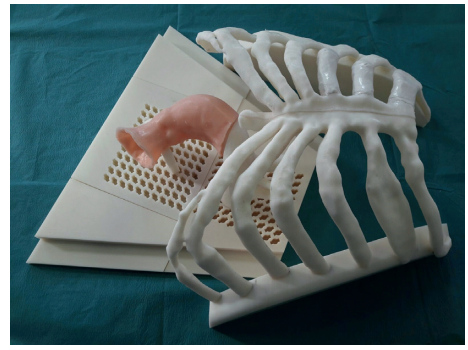


Figure 9: Patient-specific simulator used for qualitative usability tests. The tests were divided into two parts, I and II.

experienced in the AVR procedure via mini-thoracotomy and mini-sternotomy tested the view modalities.

Test was performed by using the simulation setup developed in [11], which includes a patient-specific replica of the rib cage, aortic arch, ascending aorta, and the aortic valve, as shown in Figure 9. The aortic arch is made of ABS, and it is provided with a pin to anchor it to a base, while the ascending aorta and the aortic valve are made of soft silicone for a realistic interaction with surgical instruments with casting technique, as described in [29–31].

In Test I, the stitched view mode is compared with the three views' mode in relation to three key points:

(i) Ability to use display mode to orient within the aorta

(ii) Ability to use display mode to navigate within the aorta

(iii) Safety in releasing the valve using the display mode under examination

Upon request, each mode offers an augmented reality view of the valve positioning once released. To isolate the contribution of merging individual views into a single image, images of the three views' mode are prerotated. The rotation angle is such that the horizons of cameras 2 and 3 coincide with the horizon of reference camera 1. This result is achieved by decomposing the R_i matrices into Euler angles and rotating the images from cameras 2 and 3 according to the Z-axis angle. Figure 10 shows the three views' mode with parallel horizon and the stitched view.

In Test II, the best way to manage transitions between images in the stitched view is investigated. Figure 11 compares the two transition modes. One mode clearly demarcates the transition from one image to another through black lines. The other consists in a gradual transition between images, so that it is not possible to identify any borderline between them. The two transition modes were evaluated on the basis of the eye strain and the disturbance to navigation.

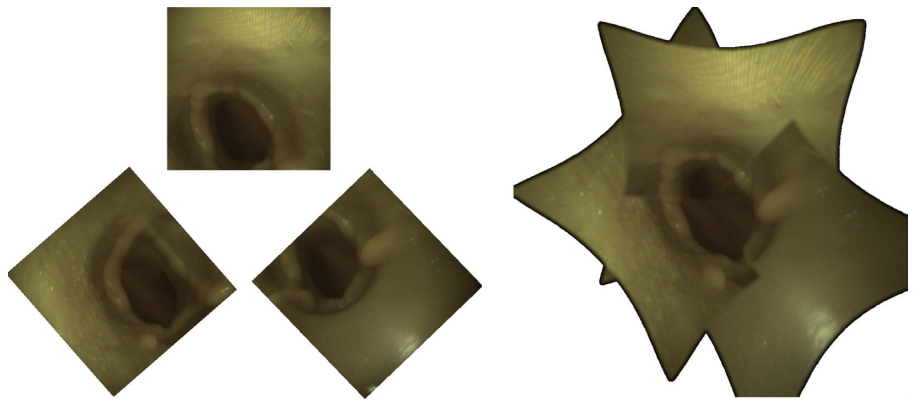
Surgeons were asked to navigate through the silicone aorta replica, reproducing the remains of the calcified aorta after removal. After completing the tests, surgeons filled out the questionnaire in Tables 2 and 3, organized in accordance with the 4-point Likert scale.

3. Results and Discussion

3.1. Quantitative Results. For each pair of warped images from cameras 1, 2, and 3, we computed the misalignment error in terms of mean and standard deviation between homologous corners of the chessboard pattern. Table 4 shows the statistical parameters as a percentage of the warped image side.

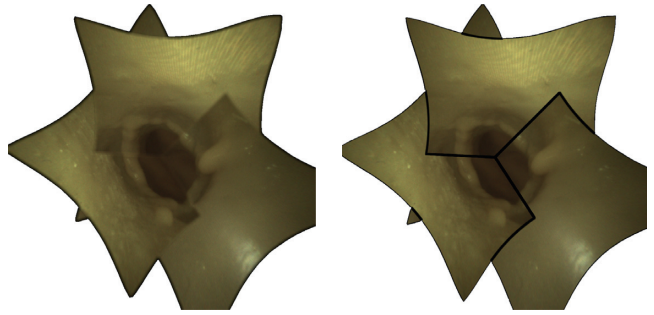
Plots in Figure 12 show the error trend in px. By way of an example, Figure 13 illustrates the misalignment between the camera images at distances of 37.5 mm, 40 mm, and 42.5 mm. The quantitative results show an increase in error when moving away from the homography plane. However, the constraints imposed by the specific surgical task and by the limited area surrounding the aorta significantly restrict the working area. Specifically, although the manipulator can adapt to tortuous and narrow paths through its link segments, the field of view of the cameras will always be obstructed by any curves. As a result, the length of the ascending aorta, which is about 5 cm [32], limits the maximum distance captured by the camera. Also, being the optimal position of the manipulator usually 3-4 cm away from the calcified valve [26], when positioning the valve, the translation along the aorta axis is limited to 1-2 cm. In this area, the maximum misalignment of $6.09 \pm 0.23\%$ is negligible compared to the size of the stitched image. This error is further decreased as the system approaches the optimal alignment, and it is reduced to $0.47 \pm 0.16\%$ at the homography plane. Therefore, in the most delicate phase of the operation, i.e., during the valve releasing, the error is kept to a minimum.

3.2. Qualitative Usability Results. Test results are expressed as median of the assessments. Table 5 shows that the three views' mode and the stitching view mode do not differ in terms of ability to orient and to navigate within the aorta. However, the stitching view provides surgeons with greater safety when releasing the valve. This may be due to the feasibility of orientation and navigation within the aorta even through the guidance of a single camera and eventually



(a) (b)

Figure 10: Comparison of view modes with and without stitching. (a) Single images view mode: the images from cameras 2 and 3 are rotated to align the horizon with that of the reference camera 1. (b) Stitching mode with semitransparent edges.



(a) (b)

Figure 11: Comparison of smooth and clear transition in the stitched image. (a) Smooth transition through semitransparent edges. (b) Clear transition through black lines.

Table 2: 4-point Likert scale questionnaire for comparison between single views and stitched view mode.

The view allows to easily orient inside the aorta.
Completely disagree Disagree Agree Completely Agree
It is possible to navigate the anatomy, visualizing the three nadirs.
Completely disagree Disagree Agree Completely Agree
This display mode would allow the valve to be released safely.
Completely disagree Disagree Agree Completely Agree

Table 3: 4-point Likert scale questionnaire for comparison between black line transitions and blurred transitions.

Black lines\blurring marking the transition between images, do not strain your eyes during navigation.
Completely disagree Disagree Agree Completely Agree
Black lines\blurring marking the transition between images, do not disturb navigation.
Completely disagree Disagree Agree Completely Agree

Table 4: Statistical parameters of the mismatch error as the distance from the reference camera varies.

Distances from virtual camera (mm)	Cameras 1-2		Cameras 1–3		Cameras 2-3		Cameras' mean	
	Mean value (%)	Standard deviation (%)	Mean value (%)	Standard deviation (%)	Mean value (%)	Standard deviation (%)	Mean value (%)	Standard deviation (%)
30	6.14	0.18	6.03	0.26	6.09	0.25	6.09	0.23
32.5	4.98	0.09	4.56	0.18	4.84	0.13	4.79	0.13
35	2.36	0.15	2.29	0.23	2.36	0.17	2.34	0.18
37.5	1.52	0.13	1.17	0.17	1.4	0.14	1.36	0.15
40	0.47	0.16	0.49	0.19	0.43	0.13	0.46	0.15
42.5	1.15	0.15	1.41	0.14	1.41	0.17	1.32	0.15
45	2.69	0.16	2.61	0.25	2.69	0.16	2.66	0.19
47.5	3.03	0.14	3.26	0.12	3.25	0.13	3.18	0.13
50	4.30	0.12	4.23	0.11	4.27	0.35	4.27	0.19

For each distance from the reference camera, the average value as a percentage of the homographed image side and the error standard deviation are reported. Values are given for each pair of cameras, 1-2, 1–3, and 2-3, and finally as an average of the three camera pairs.

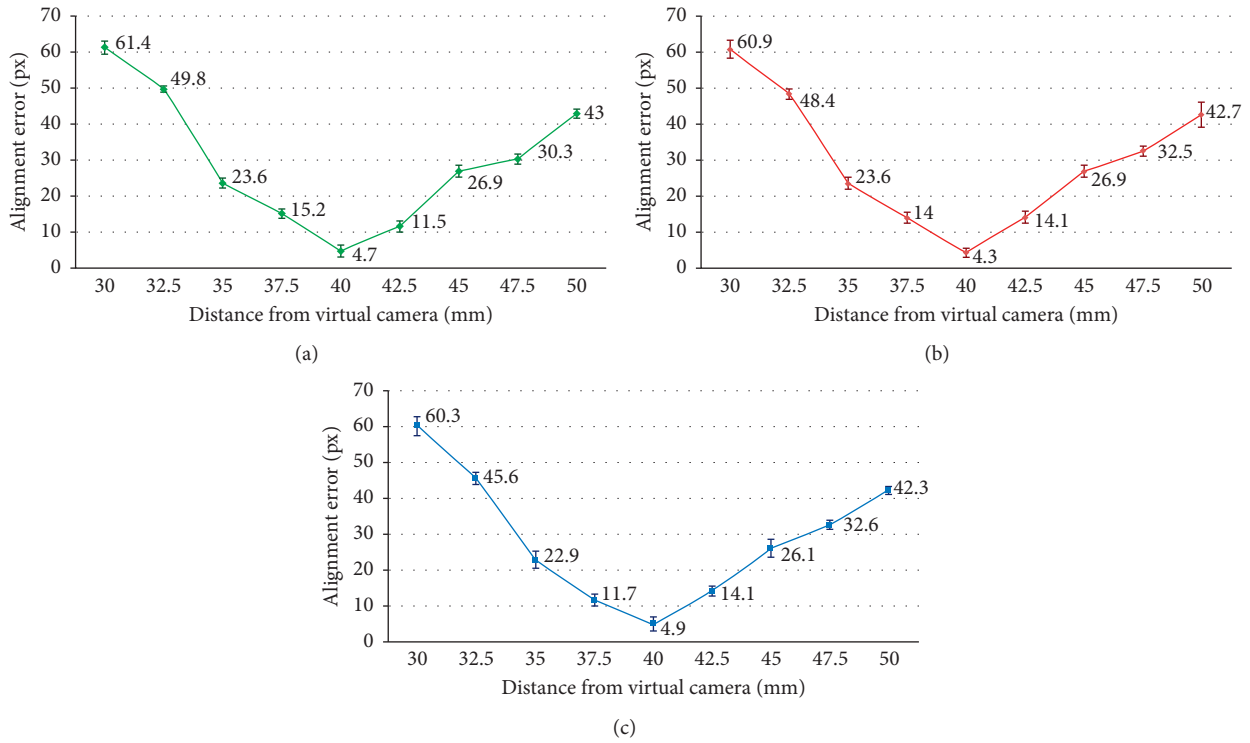


Figure 12: Error trend as the distance from the virtual camera varies. The error bars are expressed as twice the standard deviation. The number of samples varies in relation to the number of corners that can be identified in the images, and it is included in the range 32–56. Alignment error cameras: (a) 1-2, (b) 2-3, and (c) 1–3.

shifting the gaze to different views when necessary. But, when releasing the valve, an iterative alignment on the different views is more complex than simultaneously checking the match with all nadirs from a single image. Table 6 describes the results of the comparison between black borders and blurred borders transition mode, highlighting a preference for blurred border mode.

Tests were completed on a laptop with CPU 2.0 GHz processor, 8 GB RAM, and Windows 8.1 as an operating system. Experiments show that the proposed approach is effective in terms of computational complexity: time taken to stitch 3 images was of 12.6 ms averaged over an 8-minute video. Computational time is one of the most important parameters for measuring the stitching performance. The proposed method has low computational cost, as it does not require algorithms for the identification of common features in the images: experimental tests conducted with similar hardware shows that methods based on features detectors (Harris corner detector, SIFT, SURF, FAST, goodFeaturesToTrack, MSER, and ORB techniques) require from 60 ms up to 1.2 s for images with a lower resolution (320×225) only for detecting features [33], and the total computational time is further increased by computing and applying the image transformation.

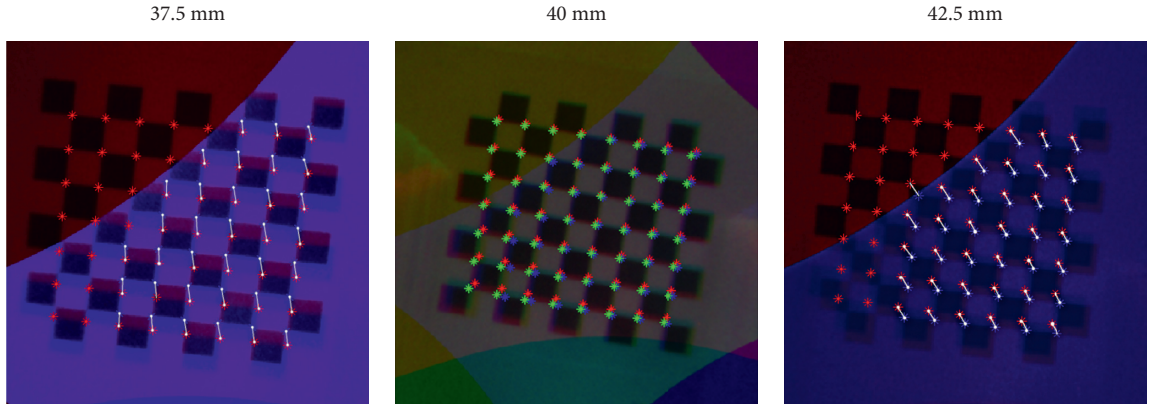


Figure 13: Error in alignment by moving away from the homographic plane placed at 40 mm. (a, c) The images of the reference camera are compared with the images of camera 2 at 37.5 mm and 42.5 mm, respectively. The distances between homologous pixels of the two cameras are highlighted in white. (b) The images acquired by the three cameras are compared, showing in red the chessboard corners from the reference camera, in blue the corners from camera 2, and in green the corners from camera 3.

Table 5: Results of the questionnaire comparing the three views' mode with the stitching view mode.

	Three views	Stitching
The view allows to easily orient inside the aorta	3 (agree)	3 (agree)
It is possible to navigate the anatomy, visualizing the three nadirs	3 (agree)	3 (agree)
This display mode would allow the valve to be released safely	3 (agree)	4 (completely agree)

Table 6: Results of the questionnaire comparing the black-edged and blurred-edged transition mode.

	Black lines	Blurring
Transitions between images do not strain your eyes during navigation	2 (disagree)	3 (agree)
Transitions between images do not disturb navigation	2 (disagree)	3 (agree)

4. Conclusions

The article proposes a method for merging the images acquired by three cameras into a single image that encompasses their single contributions. The cameras are placed off-center with respect to the axis of the manipulator, and the stitched image restores a central view to the user. The proposed method has low computational cost, as it does not require algorithms for the identification of common features in the images, but it is based on the knowledge of the reciprocal poses between the cameras and on the position and orientation of a reference plane in space. Taking advantage of the constraints imposed by the specific surgical procedure and by the aorta conformation, plane-induced homographies are used to merge the camera views. Quantitative tests showed that, although the misalignment grows moving from the homography plane, it remains negligible compared to the image size. Experimental tests with surgeons confirmed these results; they showed that the stitched view, allowing the visualization of the three nadir points in a single image, would allow surgeons to release the valve more safely, while not compromising orientation and navigation in the vessel.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research work was supported by VALVETECH project, FAS Fund-Tuscany Region (realization of a newly developed polymeric aortic valve, implantable through robotic platform with minimally invasive surgical techniques). The work was made possible by the FOMEMI project (sensors and instruments with photon technology for minimally invasive medicine), cofinanced by the Region of Tuscany, which aims to create a family of innovative devices and equipment for clinical applications, both therapeutic and diagnostic approaches.

References

[1] R. O. Bonow and P. Greenland, "Population-wide trends in aortic stenosis incidence and outcomes," *Circulation*, vol. 131, no. 11, pp. 969–971, 2015.

[2] M. Lindroos, M. Kupari, J. Heikkilä, and R. Tilvis, "Prevalence of aortic valve abnormalities in the elderly: an echocardiographic study of a random population sample," *Journal of the American College of Cardiology*, vol. 21, no. 5, pp. 1220–1225, 1993.

[3] G. W. Eveborn, H. Schirmer, G. Heggelund, P. Lunde, and K. Rasmussen, "The evolving epidemiology of valvular aortic stenosis. The Tromsø Study," *Heart*, vol. 99, no. 6, pp. 396–400, 2013.

[4] S. R. Gundry, O. H. Shattuck, A. J. Razzouk, M. J. del Rio, F. F. Sardari, and L. L. Bailey, "Facile minimally invasive cardiac surgery via ministernotomy," *Annals of Thoracic Surgery*, vol. 65, no. 4, pp. 1100–1104, 1998.

[5] F. J. Benetti, M. A. Mariani, J. L. Rizzardi, and I. Benetti, "Minimally invasive aortic valve replacement," *Journal of Thoracic and Cardiovascular Surgery*, vol. 113, no. 4, pp. 806-807, 1997.

[6] S. Rosato, F. Santini, M. Barbanti et al., "Transcatheter aortic valve implantation compared with surgical aortic valve replacement in low-risk patients," *Circulation: Cardiovascular Interventions*, vol. 9, article e003326, 2016.

[7] B. Gossetti, O. Martinelli, M. Ferri et al., "Preliminary results of endovascular aneurysm sealing from the multicenter Italian Research on Nellix Endoprosthesis (IRENE) study," *Journal of Vascular Surgery*, vol. 67, no. 5, pp. 1397–1403, 2018.

[8] R. C. Neely, M. Leacche, C. R. Byrne, A. V. Norman, and J. G. Byrne, "New approaches to cardiovascular surgery," *Current Problems in Cardiology*, vol. 39, no. 12, pp. 427–466, 2014.

[9] E. Rodriguez, L. W. Nifong, J. Bonatti et al., "Pathway for surgeons and programs to establish and maintain a successful robot-assisted adult cardiac surgery program," *Annals of Thoracic Surgery*, vol. 102, no. 1, pp. 340–344, 2016.

[10] K. Xu and X. D. Zheng, "Configuration comparison for surgical robotic systems using a single access port and continuum mechanisms," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3367–3374, Saint Paul, MN, USA, May 2012.

[11] G. Turini, S. Condino, S. Sinceri et al., "Patient specific virtual and physical simulation platform for surgical robot movability evaluation in single-access robot-assisted minimally-invasive cardiothoracic surgery," in *Proceedings of the 4th International Conference on Augmented Reality Virtual Reality and Computer Graphics*, Ugento, Italy, June 2017.

[12] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari, "Augmented reality in neurosurgery: a systematic review," *Neurosurgical Review*, vol. 40, no. 4, pp. 537–548, 2017.

[13] B. Fida, F. Cutolo, G. di Franco, M. Ferrari, and V. Ferrari, "Augmented reality in open surgery," *Updates in Surgery*, vol. 70, no. 3, pp. 389–400, 2018.

[14] F. Cutolo, "Augmented reality in image-guided surgery," in *Proceedings of the 4th International Conference on Augmented Reality Virtual Reality and Computer Graphics*, Ugento, Italy, June 2017.

[15] F. Cutolo, C. Freschi, S. Mascioli, P. D. Parchi, M. Ferrari, and V. Ferrari, "Robust and accurate algorithm for wearable stereoscopic augmented reality with three indistinguishable markers," *Electronics*, vol. 5, no. 4, 2016.

[16] M. J. Citardi, A. Agbetoba, J.-L. Bigcas, and A. Luong, "Augmented reality for endoscopic sinus surgery with surgical navigation: a cadaver study," *International Forum of Allergy & Rhinology*, vol. 6, no. 5, pp. 523–528, 2016.

[17] M. Carbone, S. Condino, F. Cutolo et al., "Proof of concept: wearable augmented reality video see-through display for neuro-endoscopy," in *Proceedings of the 5th International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, Otranto, Italy, June 2018.

[18] T. Finger, A. Schaumann, M. Schulz, and U.-W. Thomale, "Augmented reality in intraventricular neuroendoscopy," *Acta Neurochirurgica*, vol. 159, no. 6, pp. 1033–1041, 2017.

[19] R. Viglialoro, S. Condino, M. Gesi et al., "AR visualization of "synthetic calot's triangle" for training in cholecystectomy," in *presented at the 12th International Conference on Biomedical Engineering*, Innsbruck, Austria, February 2016.

[20] V. Mamone, R. M. Viglialoro, F. Cutolo, F. Cavallo, S. Guadagni, and V. Ferrari, "Robust laparoscopic instruments tracking using colored strips," in *Proceedings of the 4th International Conference on Augmented Reality, Virtual Reality, and Computer Graphics*, Ugento, Italy, June 2017.

[21] D. T. Kim and C. H. Cheng, "A panoramic stitching vision performance improvement technique for minimally invasive surgery," in *Proceedings of the 5th International Symposium on Next-Generation Electronics (ISNE)*, Hsinchu, Taiwan, August 2016.

[22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[23] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: speeded up robust features," in *Proceedings of the Computer Vision—ECCV 2006*, vol. 3951, pp. 404–417, Graz, Austria, May 2006.

[24] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2564–2571, Barcelona, Spain, November 2011.

[25] P. Mountney and G.-Z. Yang, "Dynamic view expansion for minimally invasive surgery using simultaneous localization and mapping," in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, MN, USA, September 2009.

[26] I. Tamadon, G. Soldani, P. Dario, and A. Menciassi, "Novel robotic approach for minimally invasive aortic heart valve surgery," in *Proceedings of the 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3656–3659, Honolulu, Hawaii, July 2018.

[27] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[28] F. Cutolo, U. Fontana, and V. Ferrari, "Perspective preserving solution for quasi-orthoscopic video see-through HMDs," *Technologies*, vol. 6, no. 1, p. 9, 2018.

[29] S. Condino, M. Carbone, V. Ferrari et al., "How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simulators," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 7, no. 2, pp. 202–213, 2011.

[30] S. Condino, K. Harada, N. N. Pak, M. Piccigallo, A. Menciassi, and P. Dario, "Stomach simulator for analysis and validation of surgical endoluminal robots," *Applied Bionics and Biomechanics*, vol. 8, no. 2, pp. 267–277, 2011.

[31] V. Ferrari, R. M. Viglialoro, P. Nicoli et al., "Augmented reality visualization of deformable tubular structures for surgical simulation," *International Journal of Medical Robotics*

and Computer Assisted Surgery, vol. 12, no. 2, pp. 231–240, 2016.

[32] S. A. Goldstein, A. Evangelista, S. Abbara et al., "Multimodality imaging of diseases of the thoracic aorta in adults: from the American society of echocardiography and the european association of cardiovascular imaging," *Journal of the American Society of Echocardiography*, vol. 28, no. 2, pp. 119–182, 2015.

[33] E. Adel, M. Elmogy, and H. Elbakry, "Image stitching system based on ORB feature-based technique and compensation blending," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 9, pp. 55–62, 2015.

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 7019046, 6 pages
https://doi.org/10.1155/2019/7019046

Research Article

Comparison of Two Innovative Strategies Using Augmented Reality for Communication in Aesthetic Dentistry: A Pilot Study

Romane Touati,[1,2,3] Raphaël Richert,[1,2] Catherine Millet,[1,2] Jean-Christophe Farges,[1,2,4] Irena Sailer,[3] and Maxime Ducret[1,2,4]

[1]*Université de Lyon, Université Lyon 1, Faculté d'Odontologie, Lyon, France*
[2]*Hospices Civils de Lyon, Service de Consultations et Traitements Dentaires, Lyon, France*
[3]*Division of Fixed Prosthodontics and Biomaterials, University Clinics of Dental Medicine, University of Geneva, Geneva, Switzerland*
[4]*Laboratoire de Biologie Tissulaire et Ingénierie thérapeutique, UMR 5305 CNRS/Université Lyon 1, UMS3444 BioSciences Gerland-Lyon Sud, Lyon, France*

Correspondence should be addressed to Maxime Ducret; ducret.maxime@gmail.com

Received 22 November 2018; Revised 28 February 2019; Accepted 14 March 2019; Published 3 April 2019

Guest Editor: Vincenzo Ferrari

During dental prosthetic rehabilitation, communication and conception are achieved using rigorous methodologies such as smile design protocols. The aim of the present pilot study was to compare two innovative strategies that used augmented reality for communication in dentistry. These strategies enable the user to instantly try a virtual smile proposition by taking a set of pictures from different points of view or by using the iPad as an enhanced mirror. Sixth-year dental students (n = 18, women = 13, men = 5, mean age = 23.8) were included in this pilot study and were asked to answer a 5-question questionnaire studying the user experience using a visual analog scale (VAS). Answers were converted into a numerical result ranging from 0 to 100 for statistical analysis. Participants were not able to report a difference between the two strategies in terms of handling of the device (p = 0.45), quality of the reconstruction (p = 0.73), and fluidity of the software (p = 0.67). Even if the participants' experience with the enhanced mirror was more often reported as immersive and more likely to be integrated in a daily dental office practice, no significant increase was reported (p = 0.15 and p = 0.07). Further investigations are required to evaluate time and cost savings in daily practice. Software accuracy is also a major point to investigate in order to go further in clinical applications.

1. Introduction

In dentistry, smile reconstruction is achieved using rigorous and detailed methodologies which are essential for communication between the practitioner, the laboratory, and the patient [1]. Several protocols were previously proposed, such as the "Digital Smile Design®" (DSD), developed by Christian Coachman [2]. Using only a set of photographs and presentation software, this picture-based strategy (PBS) offers a predictive view of the future patient's smile and makes treatment planning and communication with the patient easier. Until now, protocols have been limited by the following factors: they are handmade or only partly computer-assisted, are two-dimensional (2D), and are only partially immersive for patients. To improve the patient's experience and patient-practitioner communication, clinical protocols and technological evolutions were proposed, such as a mock-up, a video analysis, or a 3D facial conception [2, 3]. These tools provided a better immersivity for patients and additional details for practitioners, who were able to objectively evaluate facial movements in response to emotion and speech. However, all these features are complex to integrate for both the clinician and the laboratory, and they require a significant amount of time, energy, and cost [4].

Technological evolution of hardware and software aims to reduce the time and errors during information sharing between patients, practitioners, and laboratories. The aim of the technology presented in this pilot study is to improve the

communication with the patient using facial recognition (FR) and augmented reality (AR). FR is a technology capable of automatically identifying a person from a digital image, using reference lines of the face and mathematical algorithms [5]. AR is a type of technology in which an environment is enhanced through the process of superimposing computer-generated virtual content over a real structure [6, 7]. Even if AR tools are mainly used for video games and animations, the medical field is working to integrate these technologies for diagnosis, surgery, education, and communication with patients [8]. In dentistry, AR was firstly used for educational purposes as a tool to objectively evaluate students and give them direct feedback [8]. However, there is no study that evaluates AR as a tool to improve communication in aesthetic dentistry.

The present pilot study tested the user experiences using two innovative software of augmented reality for communication in aesthetic dentistry; one using a set of pictures and described as an automatized picture-based strategy (APBS), and the other using the front camera system of the touchpad called enhanced mirror strategy (EMS).

2. Materials and Methods

In this study, a recent application released for iOS 11 was evaluated, allowing for AR experiences to be created using a recent iPad or iPhone [9–11]. This application (IvoSmile®/Kapanu, Ivoclar-Vivadent) uses the captor camera integrated in a tablet to recognize the patient's face. After having determined virtual facial and oral landmarks [4, 12, 13], a second software proposes an artificial layer of smile propositions that is superimposed on the patient's smile (Figure 1).

Two strategies are possible: the first one (APBS) consists in taking a set of photographs in an automatized version of PBS. The user can instantly change the point of view by scrolling through the different photographs. In the second strategy (EMS), the patient can directly try and modify the proposition by looking at the iPad screen in motion, as an enhanced mirror (Figure 2).

Users can interact and change the shape, size, and color of the teeth using a large range of tools. The software gives the possibility to the user to modify the center of the arch according to the facial midlines (Figure 3(a)) and choose tooth form and proportion within different catalogues of the teeth (Figure 3(b)). The user can also modify the incisal edge position by raising or lowering length and width of the teeth (Figure 3(c)) or by changing the occlusal plane (Figure 3(d)) or the dental arch inclination and width (Figure 3(e)). Finally, the software allows the user to modify the shade and luminosity of the teeth (Figures 3(g)–3(i)).

In the present study, one operator (RT) presented the device to the sixth-year volunteer dental students (18 subjects, women $n = 13$, men $n = 5$; mean age: 23.8 years). After study subjects provided informed consent, they received some explanation and were requested to freely use the device and the different tools on their own smile (Figure 4).

After ten minutes of use, participants were asked to compare the two strategies (APBS and EMS). The experience of participants while using the application was rated using an anonymous questionnaire and a visual analog scale (VAS). The questionnaire included 5 questions and was adapted from a previous study [3] (Table 1). All supplementary declarative comments of participants were also collected and reported in the present report. VAS answers were converted into a numerical result ranging from 0 to 100 for statistical analysis. A statistical software (IBM SPSS Statistics v24) was used for analyzing data normality. Data were not normally distributed, and a Wilcoxon test was applied to evaluate the difference between the 2 camera systems ($\alpha = .05$).

3. Results

18 participants (13 women and 5 men; mean age: 23.8 years) were included in the study. Results of the questionnaire were reported in Table 2. In the present pilot study, participants' preference for one strategy over the other was not significant. Authors were not able to prove a difference between strategies in terms of handling of the device ($p = 0.45$), quality of the reconstruction ($p = 0.73$), and fluidity of the software ($p = 0.67$). According to the participants' experience, EMS was more often reported as immersive, but this study failed to report a significant advantage over the APBS ($p = 0.15$). Similarly, participants reported a preference regarding EMS, but the difference was not significant ($p = 0.07$). Participants reported that both AR strategies were complementary as they are not used for the same purpose. APBS was described by the participants as a pedagogic tool useful to explain the different smile possibilities to the patient, whereas EMS was used as the virtual try-in phase of the proposed smile project.

4. Discussion

The results of the study did not manage to report a significant difference between the two strategies in terms of handling of the device, quality of the reconstruction, fluidity of the software, and immersivity and interest for integration in a daily dental office practice. However, many questions still need to be discussed about this application.

The handling of this innovative software requires a learning curve, and many users reported that the overabundant offer of choices could make the decision process more difficult. Some suggested that the software could be simplified by creating, for example, a step-by-step version of the application, where the user is driven by the software through the different features in a logical and chronological way. Inversely, restricted freedom was reported for the determination of vertical facial and dental midlines, whereas these midlines play a significant role in the smile analysis and differences up to 2-3 mm between facial and dental midlines could be visually noticed [14]. The catalogue of teeth options was also limited, and a deep learning approach could be a valuable way of enhancing the catalogue of the teeth by collecting data from patients' and practitioners' projects.

It has been shown before that mobile devices could serve as an excellent way to communicate in dentistry [3]. Participants reported also a good immersivity for both

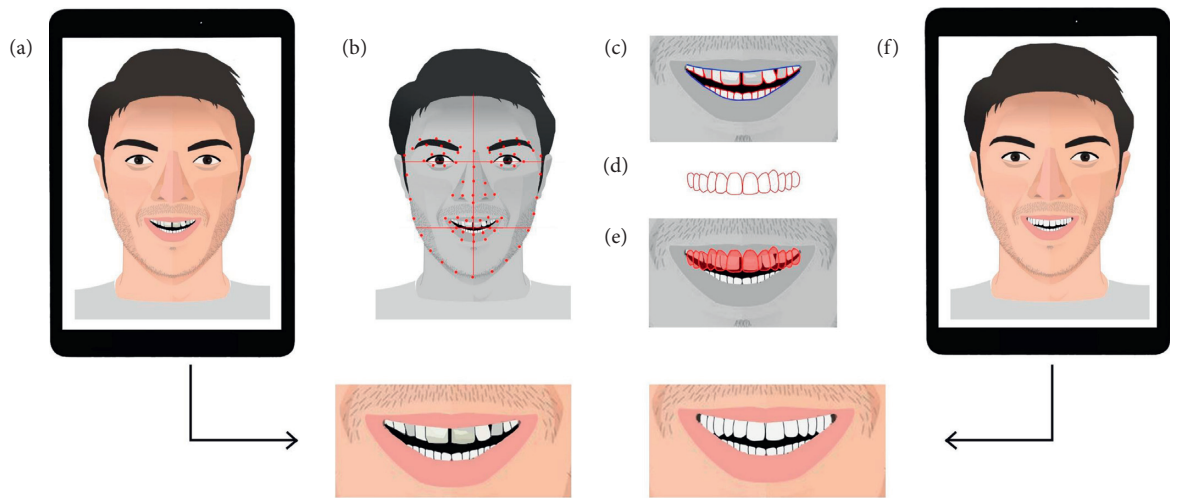


Figure 1: Schematic representation showing the basic principles of this technology. After having captured the patient's face with a picture or live with the touchpad camera (a), the FR software recognized virtual landmarks on the face (b), the lips and the smile of the patient (c). The software proposed a first mask on the patient's teeth (d). The overlay of the new mask enabled the visualization of the smile (e), and the patient was able to see the smile projected on the screen, with a set of pictures for APBS or in motion as a mirror in EMS (f).

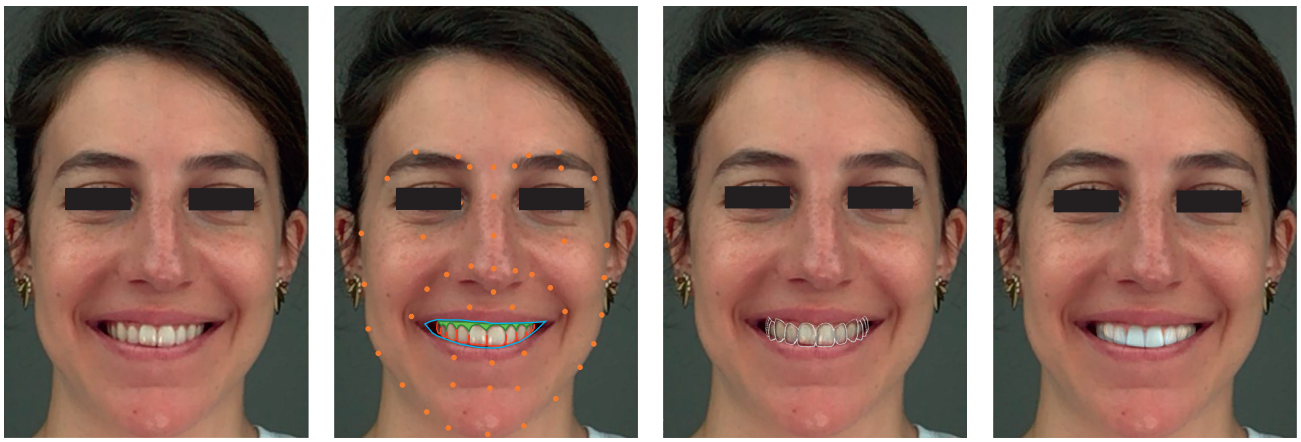


Figure 2: Illustration of the use of the software. (a) Using an iPad camera, the FR software is able to recognize nonfiducial markers (lips, smile, gum, and teeth) (b) and to propose a first mask overlaid on the initial face capture (c). A first smile design proposition is instantly obtained (d).

strategies. These results are close to those of Kim et al., which noted that AR technology was associated with excellent user experiences in education [14]. However, the present work failed to report that immersivity was significantly enhanced using EMS. These results could be explained by the fact that some participants reported a poorer picture quality using EMS, due to the video captor that leads to occasional mismatch or image pixelation. Even if a majority of participants reported their interest in using a similar tool in their daily practice for conception and communication with patients, further investigations are required to evaluate the cost and time savings brought by the device, compared to other PBS such as DSD [2]. It has to be noted that the present pilot study reported only experiences and analysis of the sixth-year students, and it could be interesting to propose this questionnaire to larger amounts of patients and clinicians in order to evaluate the impact of the device in daily professional practice.

Finally, another limitation reported by authors with present APBS or EMS was the impossibility to match the smile design with the digital cast of the patient. Indeed, in order to perform a realistic computer-assisted design (CAD) of the patient prosthesis, the software needs to be highly precise to prevent alignment mistakes during the matching process with the teeth [15]. Moreover, it was impossible to extract data from the software which prevented the analysis of software accuracy. Similar optical systems designed for AR software show a precision close to 5 mm [7]. This accuracy was considered sufficient for clinical applications in maxillofacial surgery, neurosurgery, or surgical endoscopy [16–18]. However, some limitations were reported for these optical systems, and the addition of infrared captors [19–21],

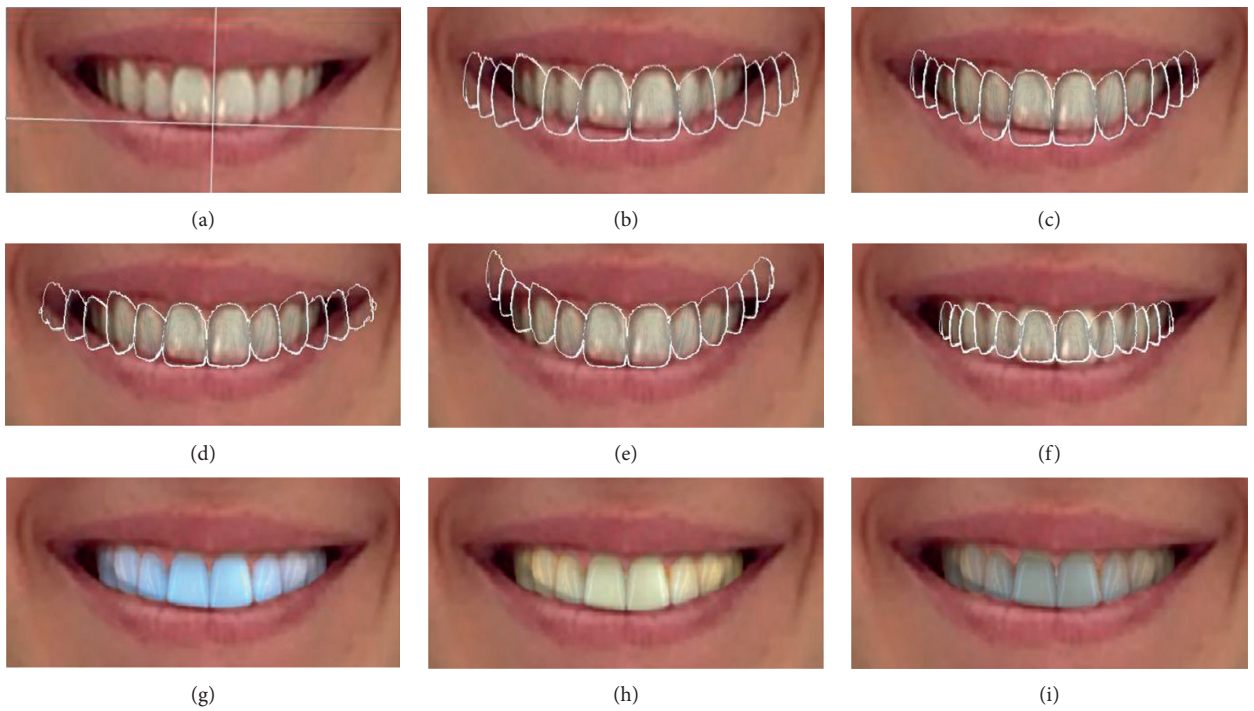


Figure 3: Illustration of some of the different features offered by the software and their impact on the smile rendering. (a) Software determination of the ideal dental midline according to the horizontal and vertical facial midlines, the interpapillary line, and the incisal edge position. (b) Proposition of form from the software catalogue. (c) Determination of the length and width of the teeth. (d)–(f) Determination of the occlusal plan height, inclination, width, and depth of the arch. (g)–(i) The final proposition can be chosen according to luminosity, shade, and color of the teeth.

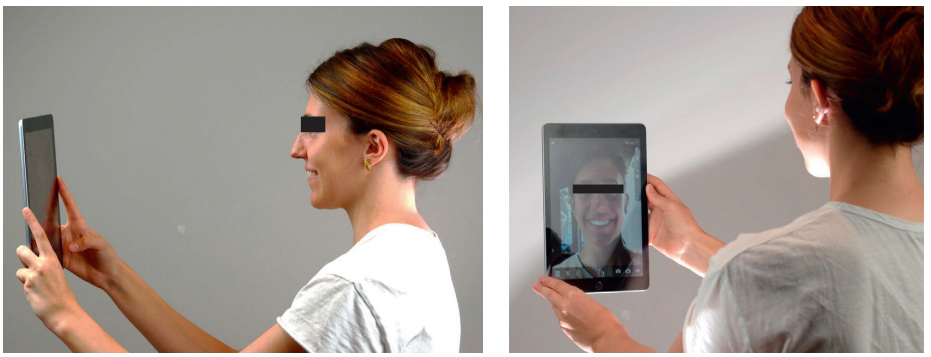


Figure 4: Use of the device. (a) Participant can use the technology by maintaining the tablet at a required minimal distance as a mirror. (b) User can see himself on the screen and interact with the software.

Table 1: Questionnaire for participants' perceptions after using APBS and EMS.

Questions for participants	Anchor terms
(1) How do you judge the handling of the device?	(0 = very difficult; 100 = very easy)
(2) How do you judge the quality of the smile reconstruction picture?	(0 = very low; 100 = very high)
(3) How do you judge the fluidity of the software?	(0 = very complicated; 100 = very easy)
(4) Do you find the experience immersive?	(0 = very low; 100 = very high)
(5) Would you be interested in using a similar device in your daily practice?	(0 = no, never; 100 = yes, with pleasure)

Table 2: Participants' perceptions of APBS and EMS.

Question	Automatized picture-based strategy			Enhanced mirror strategy			P
	Mean ± SD	95% CI	Range	Mean ± SD	95% CI	Range	
1	84 ± 2	80–88	73–100	82 ± 3	76–88	50–100	0.45
2	82 ± 2	77–86	65–100	80 ± 3	73–86	50–100	0.73
3	86 ± 2	81–90	64–100	85 ± 4	77–91	50–100	0.67
4	82 ± 3	75–88	60–100	89 ± 3	84–94	65–100	0.15
5	83 ± 3	76–89	50–100	88 ± 3	80–94	55–100	0.07

structured light, fiducial landmarks [20], or radiopaque markers attached to the patient's skin [22] has been proposed to help the accuracy for facial and dental recognition [5, 7, 22]. Further investigations are then required to evaluate the accuracy of this innovative device and to determine the precision needed in dentistry.

5. Conclusion

Although the size of the sample was limited, observations underline a good experience (handling of the device, quality of image, fluidity, and immersion) for users in both techniques. However, no statistically significant difference was observed between the two strategies. Further investigations are required for evaluating the efficacy of such a device in daily practice in particular regarding the economy of time and cost. The software accuracy is also a major point to investigate before going further in clinical practice.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

RT, RR, CM, J-CF, and MD declare that there are no conflicts of interest. IS is a scientific advisor of Kapanu, and she was implicated during the development of the tested software.

References

[1] C. Coachman and R. D. Paravina, "Digitally enhanced esthetic dentistry—from treatment planning to quality control," *Journal of Esthetic and Restorative Dentistry*, vol. 28, pp. S3–S4, 2016.

[2] C. Coachman, M. Calamita, and N. Sesma, "Dynamic documentation of the smile and the 2D/3D digital smile design process," *International Journal of Periodontics & Restorative Dentistry*, vol. 37, no. 2, pp. 183–193, 2017.

[3] I. Sailer, S. Liu, R. Mörzinger et al., "Comparison of user satisfaction and image quality of fixed and mobile camera systems for 3-dimensional image capture of edentulous patients: a pilot clinical study," *Journal of Prosthetic Dentistry*, vol. 120, no. 4, pp. 520–524, 2018.

[4] H. Popat, S. Richmond, R. Playle, D. Marshall, P. L. Rosin, and D. Cosker, "Three-dimensional motion analysis—an exploratory study. Part 1: assessment of facial movement," *Orthodontics & Craniofacial Research*, vol. 11, no. 4, pp. 216–223, 2008.

[5] M. Zollhöfer, J. Thies, P. Garrido et al., "State of the art on monocular 3D face reconstruction, tracking, and applications," *Computer Graphics Forum*, vol. 37, no. 2, pp. 523–550, 2018.

[6] H.-B. Kwon, Y.-S. Park, and J.-S. Han, "Augmented reality in dentistry: a current perspective," *Acta Odontologica Scandinavica*, vol. 76, no. 7, pp. 497–503, 2018.

[7] P. Vávra, J. Roman, P. Zonča et al., "Recent development of augmented reality in surgery: a review," *Journal of Healthcare Engineering*, vol. 2017, Article ID 4574172, 9 pages, 2017.

[8] T.-K. Huang, C.-H. Yang, Y.-H. Hsieh, J.-C. Wang, and C.-C. Hung, "Augmented reality (AR) and virtual reality (VR) applied in dentistry," *Kaohsiung Journal of Medical Sciences*, vol. 34, no. 4, pp. 243–248, 2018.

[9] M. Speicher, B. D. Hall, A. Yu et al., "XD-AR," *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, pp. 1–24, 2018.

[10] W. Zhang, B. Han, P. Hui, V. Gopalakrishnan, E. Zavesky, and F. Qian, "CARS: collaborative augmented reality for socialization," in *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications—HotMobile'18*, pp. 25–30, Tempe, AZ, USA, February 2018.

[11] C. Wu, D. Bradley, P. Garrido et al., "Model-based teeth reconstruction," *ACM Transactions on Graphics*, vol. 35, no. 6, pp. 1–13, 2016.

[12] T. Kilgus, E. Heim, S. Haase et al., "Mobile markerless augmented reality and its application in forensic medicine," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 5, pp. 573–586, 2015.

[13] V. O. Kokich, H. A. Kiyak, and P. A. Shapiro, "Comparing the perception of dentists and lay people to altered dental esthetics," *Journal of Esthetic and Restorative Dentistry*, vol. 11, no. 6, pp. 311–324, 1999.

[14] Y. Kim, H. Kim, and Y. O. Kim, "Virtual reality and augmented reality in plastic surgery: a review," *Archives of Plastic Surgery*, vol. 44, no. 3, pp. 179–187, 2017.

[15] R. Richert, A. Goujat, L. Venet et al., "Intraoral scanner technologies: a review to make a successful impression," *Journal of Healthcare Engineering*, vol. 2017, Article ID 8427595, 9 pages, 2017.

[16] G. Badiali, V. Ferrari, F. Cutolo et al., "Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning," *Journal of Cranio-Maxillofacial Surgery*, vol. 42, no. 8, pp. 1970–1976, 2014.

[17] D. Inoue, B. Cho, M. Mori et al., "Preliminary study on the clinical application of augmented reality neuronavigation," *Journal of Neurological Surgery Part A: Central European Neurosurgery*, vol. 74, no. 2, pp. 71–76, 2013.

[18] X. Kang, M. Azizian, E. Wilson et al., "Stereoscopic augmented reality for laparoscopic surgery," *Surgical Endoscopy*, vol. 28, no. 7, pp. 2227–2235, 2014.

[19] B. Fida, F. Cutolo, G. di Franco, M. Ferrari, and V. Ferrari, "Augmented reality in open surgery," *Updates in Surgery*, vol. 70, no. 3, pp. 389–400, 2018.

[20] K. A. Gavaghan, M. Peterhans, T. Oliveira-Santos, and S. Weber, "A portable image overlay projection device for computer-aided open liver surgery," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1855–1864, 2011.

[21] T. Okamoto, S. Onda, M. Matsumoto et al., "Utility of augmented reality system in hepatobiliary surgery," *Journal of Hepato-Biliary-Pancreatic Sciences*, vol. 20, no. 2, pp. 249–253, 2013.

[22] S. F. Rosenstiel, D. H. Ward, and R. G. Rashid, "Dentists' preferences of anterior tooth proportion-a web-based study," *Journal of Prosthodontics*, vol. 9, no. 3, pp. 123–136, 2000.

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 2163705, 9 pages
https://doi.org/10.1155/2019/2163705

Research Article

Upbeat: Augmented Reality-Guided Dancing for Prosthetic Rehabilitation of Upper Limb Amputees

Marina Melero,[1,2] **Annie Hou**,[2] **Emily Cheng**,[2] **Amogh Tayade**,[2] **Sing Chun Lee**,[2] **Mathias Unberath**,[2] **and Nassir Navab**[2]

[1]*Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, UK*
[2]*Department of Computer Science, Johns Hopkins University, Baltimore 21218-2683, USA*

Correspondence should be addressed to Marina Melero; marina.bernal.15@ucl.ac.uk

Received 29 November 2018; Accepted 26 February 2019; Published 19 March 2019

Guest Editor: Fabrizio Cutolo

Unsuccessful rehabilitation therapy is a widespread issue amongst modern day amputees. Of the estimated 10 million amputees worldwide, 3 million of whom are upper limb amputees, a large majority are discontent and experience rejection with their current prosthesis during activities of daily living (ADL). Here we introduce *Upbeat*, an augmented reality (AR) dance game designed to improve rehabilitation therapies in upper limb amputees. In *Upbeat*, the patient is instructed to follow a virtual dance instructor, performing choreographed dance movements containing hand gestures involved in upper limb rehabilitation therapy. The patient's position is then tracked using a Microsoft Kinect sensor while the hand gestures are analyzed using EMG data collected from a Myo Armband. Additionally, a gamified score is calculated based on how many gestures and movements were correctly performed. Upon completion of the game, a diagnostic summary of the results is shown in the form of a graph summarizing the collected EMG data, as well as with a video displaying an augmented visualization of the patient's upper arm muscle activity during gameplay. By gamifying the rehabilitation process, *Upbeat* has the potential to improve therapy on upper limb amputees by enabling the start of rehabilitation immediately after trauma, providing personalized feedback which professionals can utilize to accurately assess patient's progress, and increasing patient excitement, therefore increasing patient willingness to complete rehabilitation. This paper is concerned with the description and evaluation of our prototypic implementation of *Upbeat* that will serve as the basis for conducting clinical studies to evaluate its impact on rehabilitation.

1. Introduction

Limb loss is a recurrent problem all across the world. Every year, an estimated 185,000 people undergo upper limb amputations [1], and a significant portion of them add to the millions who live without the ability to comfortably perform activities of daily life (ADLs) [2, 3]. Efforts within the field have led to increased research in prosthetics over the years, enabling amputees to achieve higher degrees of motion and control aided with the development of myoelectric prosthetics. However, the functionality of these prosthetics remains limited, and coupled with the high rejection rate of these devices, development in the field has significant room for improvement.

A common cause of prosthetic rejection is unsuccessful rehabilitation therapy, in which the amputee is unable to develop the sufficient skills needed to successfully manage their prosthesis during ADLs [1]. Some of the major problems leading to rehabilitation failure include the late start of posttraumatic intervention due to wait time for a prosthetic fit, a lack of objective assessment of the patient's progress and performance [4], and poor patient motivation to commit to the repetitive practices involved in rehabilitation [5].

Augmented and virtual reality (AR/VR) has the potential to ameliorate the rehabilitation process. By using a virtual arm instead of waiting for a prosthetic, patients can start rehabilitation immediately after trauma, consequently reducing the acuteness of muscle atrophy [6]. An example is Anderson and Bischof's system, who developed an AR system involving a virtual arm overlaid on a patient's residual limb and controlled by residual limb muscle activity

[7]. This system, when compared to traditional, non-AR game-based systems, showed higher user experience and investment as well as comparable muscle isolation. Furthermore, the advantage of the virtual arm enables the patient to start the rehabilitation process earlier.

Another issue with current rehabilitation in upper limb amputees is the lack of a comprehensive objective assessment of the patient's progress. Commonly used methods to evaluate patient progress include the Box and Block Test (BBT) [8], the Southampton Hand Assessment Procedure (SHAP) [9], and the Clothespin Relocation Test (CRT) [10]. However, these are standardized tests only evaluating performance within a small range of movement tasks with limited degrees of freedom (DOF). Their assessment methods only account for completion rate (number of tasks successfully accomplished), lacking a more comprehensive quantitative and qualitative assessment of the patient's performance. Monitoring run-time dynamics to provide a more comprehensive assessment is important for practitioners to evaluate how well the patients are restoring mobility [11].

Among the solutions to tackle the lack of comprehensive assessment mentioned above stands the protocol developed by Chadwell et al. [1], which combines EMG signal monitoring, kinematic sensing using inertial measurement units (IMUs), and gaze tracking to determine the patient's proficiency of using an upper limb prosthesis. The results provide information on the quality of movement as well as the completion rate and are highly regarded for both its incorporation of gaze tracking as well as for accounting for the unpredictability introduced by the skin-electrode interface. However, despite these novel features, this protocol had a duration of approximately 4 hours, which was inconveniently long for efficient clinical use [1].

A year later, Hunt et al. developed the Prosthetic Hand Assessment Measure (PHAM), an alternative method to quantitatively assess performance in a range of manipulation tasks associated with object manipulation (e.g., pinch, key, and power, shown in Table 1) for upper limb amputees [12]. PHAM uses IMUs for motion tracking and presents a performance evaluation assessment metric that accounts for compensatory movements in the patient. Another method proposed by Yu et al. utilizes a Kinect-based system to introduce a personalized range of motion measurement with AR feedback [13]. The goals of the study were to establish the accuracy of the Kinect in measuring clinically relevant movements in patients with Parkinson's disease. The results of this system match experts' observations and show promising results for telerehabilitation scenarios [13], as well as, once again, the potential of rehabilitation within an AR system. As shown in the study by Yu et al., as well as later in *Upbeat*'s implementation, integration of motion tracking and electromyography (EMG) sensors within an AR system provides quantitative data physicians can use for objective assessment of patient progress. The flexibility of such a system also allows the therapy to be personalized to each patient's unique needs.

In addition to late posttraumatic intervention and lack of comprehensive assessment, one of the final main challenges of upper limb rehabilitation is maintaining patient motivation and commitment to practice, especially considering the prolonged and repetitive nature of this task. Previous work on gamified systems for AR-guided rehabilitation includes *mirrARbilitation* [14], a system based on gesture recognition and markerless motion tracking which recognizes and classifies biomechanical movements. The application provides exercise instructions, to prevent cheating via movement compensation, and has been proven capable of increasing patient success rate during rehabilitation, preventing wrong movements, and fostering an incentive to complete the process [14].

Table 1: Correspondence of object, hand gesture, and ADL used in the PHAM method [12].

Object	Hand gesture	Activity of daily living (ADL)
Cylinder	Power	Pouring a glass of water
Prism	Tripod	Picking up a pencil
Block	Pinch	Picking up coins
Card	Key	Grasping a credit card

As such, AR/VR-guided rehabilitation proves to be effective in increasing motivation and adding excitement to rehabilitation practices, consequently leading to increased investment by the patients themselves. Similar results have been extensively studied in rehabilitation for stroke patients [6], results which remain highly applicable towards upper limb rehabilitation programs [15]. Altogether, the advantages of integrating AR into rehabilitation therapy lead to a more effective restoration of mobility in amputees by providing more accurate performance evaluation methods, providing real-time guidance for improved performance, and increasing patient's excitement and motivation while performing therapy.

Upbeat takes the rehabilitation workflow presented in PHAM [12] and incorporates it into an AR-based dance game, simulating the idea of practicing a set of different hand gestures within a dynamic environment. PHAM focuses on monitoring gesture completion rate and accounts for compensatory movement. In *Upbeat*, we expand upon this idea through monitoring of patient's EMG activity as well as providing an AR feedback visualization system that allows the patient to see the muscles activated throughout the gameplay. The proposed system shall be understood as a proof of concept in order to quantify performance and validate design decisions such that, upon completion of this study, a refined version of the system can be used to evaluate clinical appropriateness for rehabilitation on a control group of amputees.

2. Materials and Methods

The proposed system for rehabilitation is based on AR guidance, gesture recognition, and markerless body tracking. A virtual dance instructor guides the patient through a set of dance movements containing specific hand gestures (Figure 1). A Myo armband, worn on the forearm, is used to detect the patient's muscle activity and classify the hand gestures using detected EMG data. The patient's position is

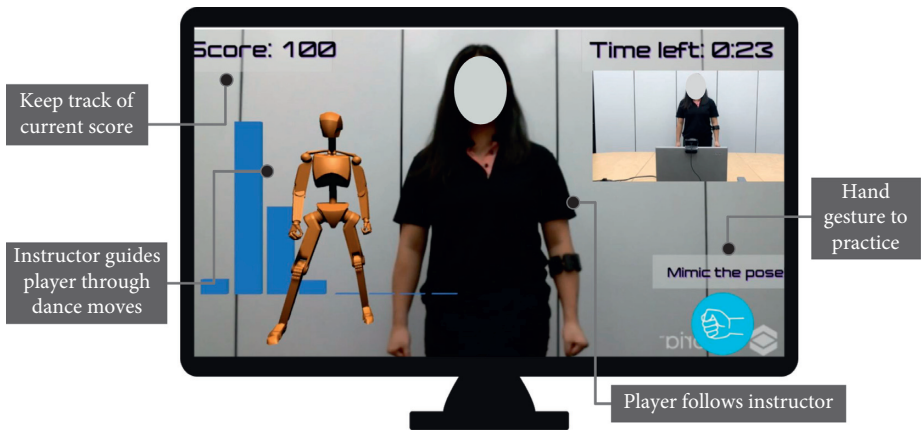


Figure 1: *Upbeat*'s gameplay screen.

tracked with a Microsoft Kinect sensor and used to display a visualization of the muscle activity upon completion of the session. The system's workflow is described in Section 2.1 and summarized in Figure 2. Moreover, a detailed explanation of the materials and methods to develop each component in *Upbeat* is provided in Sections 2.2–2.6.

2.1. Game Workflow. The game is composed of three major scenes for each stage of the game—Menu (song selection), Play, and Feedback. Navigation throughout the application can be done with either mouse inputs or with gestures detected from the Myo Armband (Figure 3). The Play scene (Figure 1) is where the majority of the gameplay occurs. This scene has two key components: a dance instructor and a hand gesture prompt. The virtual dance instructor that appears on the side of the screen shows the patient the dance movements to follow. The virtual dance instructor is to give the patient visual cues on the correctness of their dancing and encourage continuous engagement, similar to a real-life dance instructor.

The hand gesture prompt is an image icon on the bottom right corner of the screen that informs the patient of which hand gesture to perform at the given moment. Hand gestures are tracked with the Myo Armband, and if the correct gesture is performed in the expected time frame, the image is replaced with a green success symbol to tell the patient that correct gesture is performed. In the final Feedback scene (Figure 4), the patient can see a visual summary of the data collected during gameplay. These data can also be sent to a rehabilitation practitioner for further analysis. The Feedback scene is displayed immediately after the Play scene, a few seconds after the completion of the song and choreography. An overview of the game's workflow is shown in Figure 2.

2.2. Game Development. *Upbeat* was developed with the 3D game engine Unity, through which all the sensors and software used (Kinect, Myo, *Mirracle* [4]) were integrated. The Myo Armband SDK for Unity was used to feed data from the Myo Armband sensor wirelessly into the Unity application. Additionally, the Kinect SDK for Unity was used to capture the pose and position of the patient during gameplay and to utilize the data into the postgame feedback. Along with a standard RGB video recorded during gameplay, motion data from the Kinect sensors were fed into *Mirracle* [4], an AR application used to produce a color-coded visualization of the patient's muscle activity during gameplay.

2.3. Dance Choreography. The integrated dance used in *Upbeat* was choreographed and recorded with the Kinect. The motion-capture data were then used to animate the in-game virtual dance instructor which later guides the patient through the same movements. The choreography consists of a set of dance movements embedded with hand gestures inspired by the PHAM model. The dance choreography was designed to include full range of motion in the upper body, as well as regular, repetitive movements. These two factors (movements with various degrees of freedom and repetitions) have proven to be beneficial in state-of-the-art rehabilitation therapies [12].

Each dance move includes one of the four hand gestures (spread, wave right, wave left, and fist; see Figure 3) and is performed in time with selected music—in this case, the current demo song is What About Us by P!nk. The song was chosen because it had a tempo appropriate for novice users to effectively engage with the game. Through repeated sessions, the patients learned the dance by following the virtual dance instructor while also practicing the hand gestures for rehabilitation.

In the PHAM protocol [12], the patient is required to manipulate a set of objects within a physical frame by grabbing the object and changing its position in the frame (see Table 1). Each object requires the patient to perform a particular hand gesture, as shown in Table 1. While the hand gestures included in the PHAM protocol are useful for object manipulation, in *Upbeat* we selected gestures common to activities of daily living (ADLs) (see Table 2). The gestures selected were spread, wave right, wave left, and fist (see Figure 3). These gestures were selected due to two key advantages: First, they are highly suitable for the EMG classification. Secondly, they can easily be embedded in dance choreography.

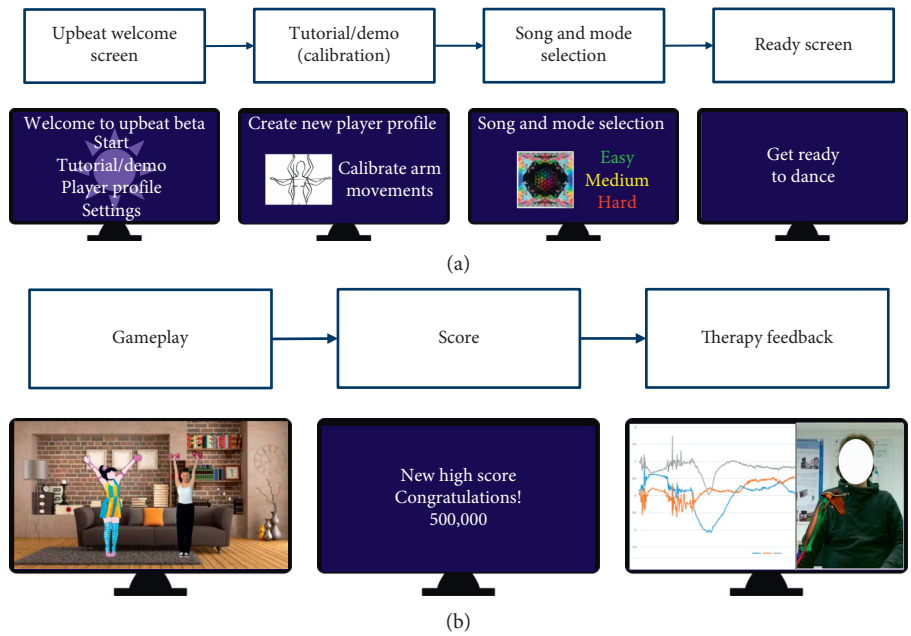


Figure 2: Overview of *Upbeat*'s game navigation.

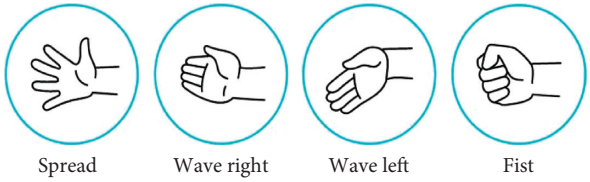


Figure 3: Hand gestures included in the choreography and tracked with Myo Armband ("Getting Starting With Myo On Windows". Welcome To Myo Support, 2018, https://support.getmyo.com/hc/en-us/articles/202657596-Getting-starting-with-Myo-on-Windows. Accessed 29 Nov 2018.).

2.4. Motion Capture and Model Animation. The Kinect sensor, along with NI mate software, was used to capture the movements of a dancer performing the choreography. The NI mate software allowed the motion-capture data detected with the Kinect to be fed into Blender, a free open-source 3D creation suite (Foundation, Blender. "Blender. Org–Home Of The Blender Project–Free And Open 3D Creation Software," Blender.Org, 2018, https://www.blender.org. Accessed 29 Nov 2018.), in the form of a rig, an animated 3D skeleton, which was later used to animate the dance instructor virtual model for the game ("Free Mannequin Male 3D Model." Turbosquid.Com, 2018, https://www.turbosquid.com/3d-models/free-mannequin-male-3d-model/1005602. Accessed 29 Nov 2018.). The motion data from the NI mate rigging were matched to the model's body (Figures 5(a) and 5(b)). Because the visualization of specific hand gestures was for a correct practice, each of the hand gestures was manually key-framed in Blender (Figure 5(c)) to increase the accuracy and clarity of the model hand gesture visualization.

2.5. Gameplay Setup. The Myo Armband was placed on the upper arm of the patient, right below where the forearm is the widest. In the case of an upper limb amputee, this would correspond to the phantom limb. The Kinect sensor was positioned at an approximate distance of 1.5 meters in front of the patient and at a height aligned with their upper chest. A 43″ TV monitor was used to display the game and located at the same distance as the Kinect (1.5 m), which was a clearly visible location for the patient. In addition, speakers were connected to the computer and used to play the music for the dance game.

Before the gameplay begins, each user undergoes a calibration protocol for the Myo Armband to ensure correct hand gesture detection. Calibration is done through the Myo Armband's proprietary interface, *Myo Connect*. In the calibration, the user is prompted to perform each of the five recognizable gestures (wave in, wave out, fist, fingers spread, and double tap) a sequence of times. The process takes approximately 3-4 minutes overall, and it has to be done every time a new user interacts with the system.

The recording of the patient's performance during the game is displayed in real time on the game's background, simulating the effect of a mirror. This setup is the most appropriate in order for the patients to clearly see the virtual dance instructor, as well as their own mirrored reflection (from the head to slightly above their knees), such that they could perform the dance movements with as much visual observance as possible.

2.6. Performance Monitoring and Feedback. Performance consisted in three key elements: gesture completion, muscle activity, and muscle activation. The Myo Armband was integrated into the application to track EMG activity and detect the hand gestures. The gesture completion is evaluated using a score based on the number of successfully completed gestures. The patient's muscle activity is then shown using an EMG

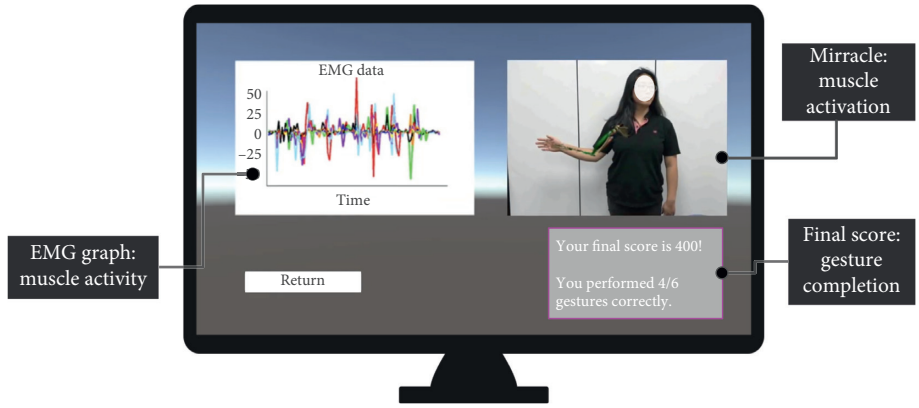


Figure 4: Feedback screen.

Table 2: Correspondence of hand gestures involved in *Upbeat* with ADL.

Hand gesture	Activity of daily living (ADL)
Spread	Greeting someone, offering help
Wave right	Indicating direction (right)
Wave left	Indicating direction (left)
Fist	Gripping a small object

graph, while a color-coded visualization of muscle activation is produced with the *Mirracle* AR mirror system [4].

The postgame feedback is designed to deliver a comprehensive evaluation of the user's performance, following methods that have proven to be effective for rehabilitation [6, 12]. The purpose of the feedback is to serve the user with an immediate quantitative evaluation of how well they performed the gestures during gameplay (given by the Game Score), as well as a qualitative understanding of their movements through the color-coded visualization of their muscle's EMG data. This EMG graph also provides the practitioner with a detailed understanding of the user's current progress in regaining control of their upper limb muscles. The feedback components are shown in Figure 4 and explained further below.

2.6.1. Game Score. The scoring system is based on the timing and accuracy of gesture completion. Each dance step shown by the virtual dance instructor contains one hand gesture that needs to be matched by the patient. Whether or not the patient successfully completes the gesture is tracked by the Myo Armband, which analyzes muscle activity in order to classify the movement into a recognized gesture. If the performed gesture both matches the one shown by the dance instructor and is performed within a certain time frame, it is deemed as the correct movement and adds 100 points to the player's total score. The time frame for each specific hand gesture and dance movement varies based on the choreography and music, but typically lasts between 6 and 12 seconds. The total game score is then calculated based on how many hand gestures are accurately performed by the player throughout the game, with points awarded for each successfully completed gesture.

2.6.2. EMG Graph. The Myo Armband contains an array of 8 bipolar surface electrodes that measure the EMG activity from the user. The raw data are then streamed wirelessly through the Myo Data Capture application at a frequency of 200 Mhz to populate a.csv file (stored locally) that is later used to produce a graph of the patient's EMG activity.

This graph is displayed upon completion of the game in the Feedback scene. Each color in the graph represents data collected by each individual sensor, and the overall analysis can be used by the practitioner to visually analyze the muscle activity patterns as an indication of the patient's progress through the rehabilitation process.

2.6.3. Color-Coded Visualization of Muscle Activity. In our system, the *Mirracle* application records a video of the patient performing the movements during the gameplay. It uses this video in combination with the Kinect depth sensor data to output the same video with an augmentation of the musculoskeletal system of the upper arm overlaid on top of the patient's right arm. The augmented muscles are color-coded (green for activated, red for relaxed) in real time to indicate the muscles being used.

3. Results and System Evaluation

The system was evaluated on three subjects, with 10 trials of the gameplay performed by each subject. In each trial, we measured the system's ability to correctly classify each specific gesture. Even though the classification is performed using the built-in Myo Armband software, measuring the classification accuracy within the *Upbeat* environment is important in order to evaluate whether or not the Myo Armband functions properly in a Unity environment.

For each trial, we also measured the system's operating time, reaction time, and detection time for each hand gesture involved in the gameplay. We define operating time (o_t) as the time taken for each hand gesture to be detected by the system from the moment it appears on the screen. Operating time can be broken down into detection and reaction times (Figure 6, equation (1)). Detection time (d_t) is defined as the time it takes for each specific gesture to be recognized by the

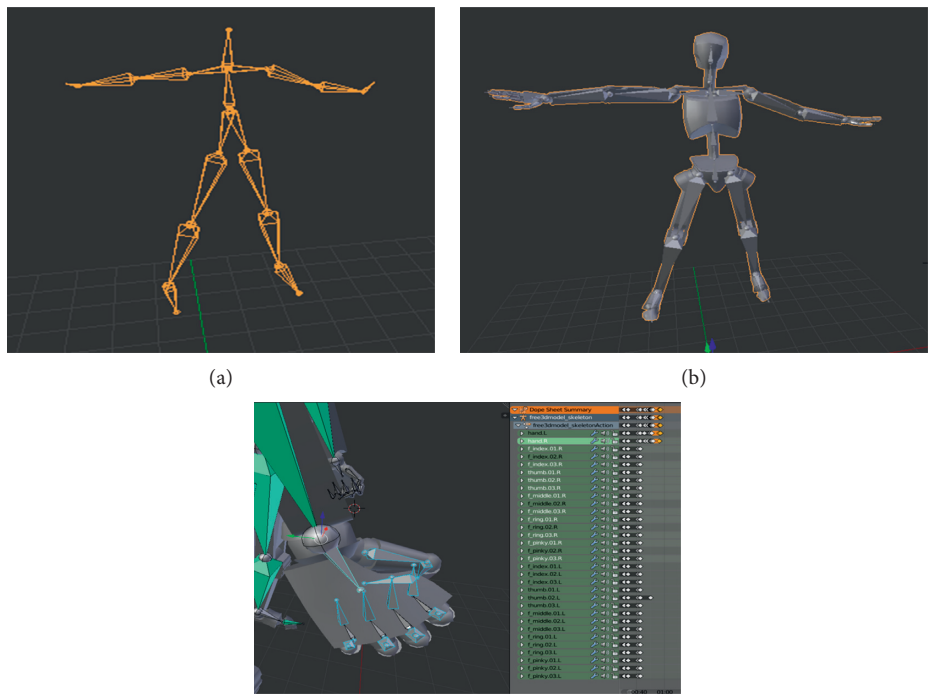


Figure 5: (a) NI Mate rig with motion-capture data using Kinect; (b) Blender model used for virtual dance instructor. (c) Manual keyframing of the model's hand bones with motion-capture data.

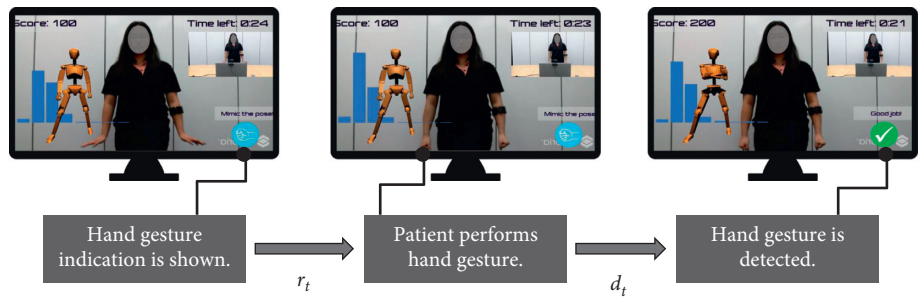


Figure 6: Detection time (d_t) and reaction time (r_t) measured during gameplay.

system, while reaction time (r_t) is defined as the time it takes the subject to perform a hand gesture from the moment it appears on the screen (Figure 6).

$$o_t = r_t + d_t. \tag{1}$$

The results for each hand gesture class are shown in Table 3. Summarizing these results, the system reported an average detection time across hand gestures of 0.24 ± 0.31 seconds, while the average reaction time was 0.92 ± 0.10 seconds. Overall, this gives us an average operating time of 1.15 ± 0.34 seconds. We also calculated the classification accuracy, expressed as the percentage of correctly classified gestures per class across the 10 trials for each patient (Figure 7).

Since more advanced versions of *Upbeat* would involve faster dance movements expected to be performed within a shorter time period, it is crucial for the system to be able to detect different gestures quickly and accurately, in order to accommodate for the different levels of proficiency for each patient as their rehabilitation therapy progresses. To assess the system's ability to accommodate faster dance movements, we used the detection time data to compute the percentage of hand gestures that could be efficiently detected within a time interval of no more than four seconds (Figure 8). The results show that it takes an average of 2.62 seconds for each gesture to be detected, meaning the system could effectively support faster-paced choreography. To set this in context, Figure 9 shows the dance movement time interval across the current gameplay, which currently ranges between 6 and 12 seconds.

4. Discussion and Further Directions

From the experiments we conducted during our system evaluation, 77% of the gestures performed by the subjects during gameplay were detected and accurately classified by

Table 3: System's average, maximum and minimum detection, reaction, and operating times.

	Wave left	Wave right	Fist	Spread
Average detection time (s)	0.06 ± 0.09	0.06 ± 0.05	0.34 ± 0.52	0.49 ± 0.68
Maximum detection time (s)	0.38	0.17	2.20	2.62
Minimum detection time (s)	0.02	0.02	0.02	0.02
Average reaction time (s)	0.88 ± 0.28	1.09 ± 0.47	0.87 ± 0.49	0.83 ± 0.36
Maximum reaction time (s)	1.53	2.02	2.55	1.33
Minimum reaction time (s)	0.22	0.05	0.02	0.03
Average operating time (s)	0.94 ± 0.37	1.14 ± 0.53	1.20 ± 1.00	1.33 ± 1.04
Maximum operating time (s)	1.92	2.18	4.75	3.95
Minimum operating time (s)	0.23	0.07	0.03	0.05

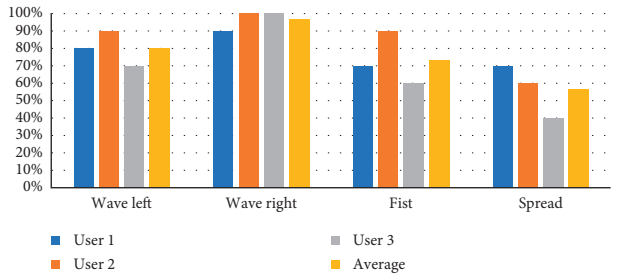


Figure 7: Hand gesture classification accuracy. This figure indicates the percentage of the gestures correctly classified across 10 trials, for each of the three users.

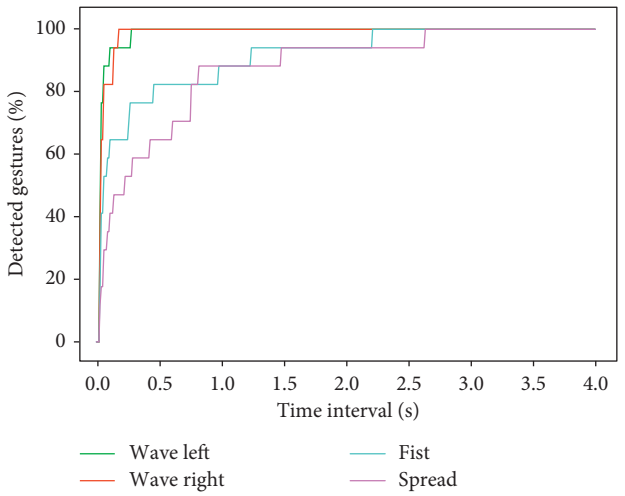


Figure 8: Optimal dance movement time interval. This figure reflects the percentage of hand gestures that would be detected, given different time interval thresholds in the range 0–4 seconds.

the system. Considering that the commercial Myo Armband has an average classification accuracy of 82.8% [16], *Upbeat's* results show that integrating the Myo Armband within *Upbeat's* local environment slightly compromises the classification accuracy.

Misclassification was more prevalent on the third subject. This was due to the difficulties this subject experienced during the calibration protocol, emphasizing the importance of a more robust calibration protocol compared to that of the commercial Myo Armband. Building a more robust calibration protocol and integrating it in the game's workflow

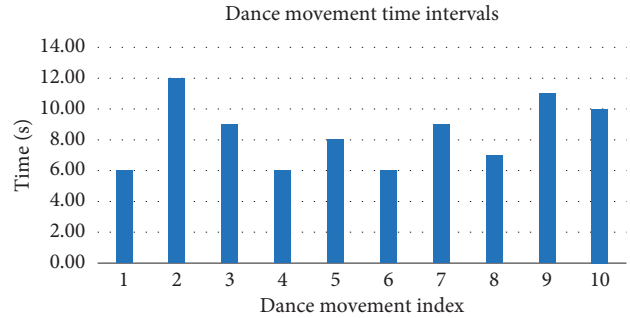


Figure 9: The time interval for each dance movement in the gameplay. Throughout the gameplay, there are sets of 10 movements, each containing one hand gesture.

with a tutorial/calibration feature are further steps for the system's improvement.

It is also important to consider that the hand gestures used in the current version of *Upbeat* have significantly differentiable EMG patterns. There is a trade-off between the complexity of the hand gestures and the accuracy of the EMG classification. If the sequence of hand gestures integrated in the game were to be expanded by introducing more complex hand gestures in a shorter time frame, the classification report of the system would be expected to achieve lower classification results. However, it is also expected that higher classification accuracy correlates with increased practice. The classification accuracy improves because, as the user becomes more experienced performing the rehabilitation exercises, the muscle signals become clearer and differentiable, which leads to better classification accuracy [12]. That is to say, as a patient becomes increasingly familiar with the choreography, it is expected that they can time and perform the gestures synchronously with the game with a higher degree of accuracy. As a result, this usage of gamified rhythm, time, and practice, likely contributes to higher success rates with the rehabilitative movements, making up for any initial complexity of the gestures. A further improvement to tackle this issue would be to add an initial learning session where the user learns the hand gestures and becomes familiar with the interface before the actual dance choreography begins.

Because the dance movements during gameplay take between 6 and 12 seconds (Figure 9), we expect that the system is able to detect the gestures faster than the minimum time of 6 seconds. With the purpose of assessing this, the

reaction, detection, and operating times are measured and presented in Table 3 of the results section. Amongst the correctly classified gestures, the maximum detection time across trials was 2.62 seconds, which was well below the minimum of 6 seconds. However, we also had to take into account the reaction time, namely, the time between the appearance of the gesture symbol on the screen and the patient actually performing the gesture. In our study, the average reaction time was 0.92 seconds. By adding the detection time to the average reaction time, we can conclude that the operating time of the system is 3.56 seconds, which is still well below our minimum requirement of 6 seconds to perform a given movement. As such, we are certain that the current system is suitable for accurately recognizing gestures in a choreographed sequence. Given this operating time, there would still be room for the system to introduce more dynamic movements as part of more advanced levels.

Upbeat is a proof of concept aimed at testing whether a gamified, AR version of upper limb rehabilitation therapy, based on the PHAM protocol, could successfully be utilized in a clinical environment. Currently, we assert that the designed system and workflow is successful at classifying hand gestures embedded in a dance routine taught by a virtual dance instructor with a success rate of 77%. Furthermore, the system was also successful in measuring EMG signals from the patient's upper arm muscle activity, as reported by the graphical summary of the data as postgame feedback. Finally, the system was able to display a recording of the gameplay with an accurate augmentation of the musculoskeletal system overlaid over the patient's body, allowing the visualization of the muscles being activated during each dance movement.

In order to make the system appropriate for rehabilitation, the next step is to implement a more complicated set of hand gestures better reflecting those used in PHAM [12]. This process includes developing the Myo Armband built-in classifier to detect a broader set of hand gestures given the raw EMG data. Another improvement of the proposed system is adding different levels to the current version of *Upbeat*, where the difficulty is based on the speed of the music, the complexity of the choreography, and the range of movements involved. Additionally, analyzing the motion between poses and accounting for the compensatory movements in the scoring system would give a further insight into the patient's performance, for example, calculating the actual accuracy of the performed gesture (as opposed to the currently binary system of whether or not the gesture was completed).

Most importantly, a clinical study with a group of upper limb amputees shall be conducted in order to evaluate their progress when using *Upbeat* in comparison to that of a control group following traditional rehabilitation therapy. As the system is intended to improve rehabilitation in upper limb amputees, it is important to understand how a system like *Upbeat* is received by its target group. Furthermore, while the qualitative and analytical aspects of the system already have strong support, both from this study as well as related studies, a more subjective assessment on how enjoyable an application like *Upbeat* will be is a future target of study.

5. Conclusions

Upbeat converts the proven success of PHAM rehabilitation therapy and transforms it into a fun and enjoyable gamified experience for rehabilitation. As such, the gamified aspect of *Upbeat* has the potential to improve the rehabilitation process by increasing user's excitement. Portability of the system allows for rehabilitation to begin immediately after trauma, rather than waiting for prosthetics to be made or for medical-guided therapies to be concretely established. *Upbeat* further facilitates the important element of personalized feedback which can prove to be essential for the amputee to understand their progress, as well as giving doctors the ability to simultaneously track the progress without being overbearing on the rehabilitative process. *Upbeat* is presented as a prototype for gamified AR rehabilitation therapy and, in future work, will be used to conduct a clinical trial to evaluate its efficacy in achieving the envisioned goals.

Data Availability

Release of source code and data will be considered on a per request basis.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Funding

This project was funded in part by Johns Hopkins University internal funding sources.

Acknowledgments

This project was realized in the framework of the Augmented Reality Course at Johns Hopkins University and we gratefully acknowledge the Department of Computer Science for making the required hardware and infrastructure available and Stryker for sponsoring the Best Project Award.

Supplementary Materials

Supplementary Material shows a demo video of the system. The user first selects the "start" option in the menu and browses through the available songs for the gameplay. This control is enabled by the Myo Armband, worn on the forearm, which detects the patient's muscle activity. Once the song has been selected, the gameplay starts and the virtual dance instructor (shown on the left of the screen) shows the user the dance movements to imitate. The hand gesture prompt is an image icon on the bottom right corner of the screen that informs the patient of which hand gesture to perform at the given moment and whether this is performed correctly. The labels on the top of the screen display the current score on the left and time left for the gameplay on the right. Finally, an additional view of the user during *Upbeat* gameplay has been added in the top right corner to

give a better understanding on the system's setup and the placing of the webcam. (*Supplementary Materials*)

References

[1] A. Chadwell, L. Kenney, S. Thies, A. Galpin, and J. Head, "The reality of myoelectric prostheses: understanding what makes these devices difficult for some users to control," *Frontiers in neurorobotics*, vol. 10, no. 7, 2016.

[2] K. Ziegler-Graham, E. J. Mackenzie, P. L. Ephraim, T. G. Travison, and R. Brookmeyer, "Estimating the prevalence of limb loss in the United States: 2005 to 2050," *Archives of Physical Medicine and Rehabilitation*, vol. 89, no. 3, pp. 422–429, 2008.

[3] M. Owings and L. Kozak, "Ambulatory and inpatient procedures in the United States, 1996," *Vital Health Stat*, vol. 139, no. 13, pp. 1–119, 1998.

[4] T. Blum, V. Kleeberge, C. Bichlmeier, and N. Navab, "Mirracle: an augmented reality magic mirror system for anatomy education," in *Proceedings of the 2012 IEEE Virtual Reality Short Papers and Posters (VRW)*, pp. 115-116, Amsterdam, The Netherlands, October 2012.

[5] D. J. Atkins and H. M. Robert III., *Comprehensive Management of the Upper-Limb Amputee*, Springer, Berlin, Germany, 1989.

[6] M. C. Howard, "A meta-analysis and systematic literature review of virtual reality rehabilitation programs," *Computers in Human Behavior*, vol. 70, pp. 317–327, 2017.

[7] F. Anderson and W. Bischof, "Augmented reality improves myoelectric prosthesis training," *International Journal on Disability and Human Development*, vol. 13, no. 3, pp. 349–354, 2014.

[8] J. S. Hebert and J. Lewicke, "Case report of modified box and blocks test with motion capture to measure prosthetic function," *Journal of Rehabilitation Research and Development*, vol. 49, no. 8, pp. 1163–1174, 2012.

[9] J. G. M. Burgerhof, E. Vasluian, P. U. Dijkstra, R. M. Bongers, and C. K. van der Sluis, "The southampton hand assessment procedure revisited: a transparent linear scoring system, applied to data of experienced prosthetic users," *Journal of Hand Therapy*, vol. 30, no. 1, pp. 49–57, 2017.

[10] A. Hussaini and P. Kyberd, "Refined clothespin relocation test and assessment of motion," *Prosthetics and Orthotics International*, vol. 41, no. 3, pp. 294–302, 2017.

[11] S. S. Johnson and E. Mansfield, "Prosthetic training," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 25, no. 1, pp. 133–151, 2014.

[12] C. Hunt, R. Yerrabelli, C. Clancy et al., *Prosthetic Hand Assessment Measure*, 2017.

[13] K. Yu, R. Barmaki, M. Unberath et al., "On the accuracy of low-cost motion capture systems for range of motion measurements," *Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications*, vol. 10579, 2018.

[14] A. E. F. Da Gama, T. M. Chaves, L. S. Figueiredo et al., "MirrARbilitation: a clinically-related gesture recognition interactive tool for an AR rehabilitation system," *Computer Methods and Programs in Biomedicine*, vol. 135, pp. 105–114, 2016.

[15] B. Hafner and J. E. Sanders, "Considerations for development of sensing and monitoring tools to facilitate treatment and care of persons with lower-limb loss: a review," *Journal of Rehabilitation Research and Development*, vol. 51, no. 1, pp. 1–14, 2014.

[16] A. Gogić, N. Miljkovic, and D. Đurđević, "Electromyography-based gesture recognition: fuzzy classification evaluation," in *Proceedings of the 3rd International Conference on Electrical, Electronic and Computing Engineering, IcETRAN*, At Zlatibor, Serbia, June 2016.

Hindawi
Journal of Healthcare Engineering
Volume 2019, Article ID 6813719, 10 pages
https://doi.org/10.1155/2019/6813719

Research Article

Towards Virtual VATS, Face, and Construct Evaluation for Peg Transfer Training of Box, VR, AR, and MR Trainer

Zhibao Qin,[1] Yonghang Tai,[1] Chengqi Xia,[1] Jun Peng,[2] Xiaoqiao Huang,[1] Zaiqing Chen,[1] Qiong Li,[1] and Junsheng Shi[1]

[1]*Yunnan Key Laboratory of Opto-Electronic Information Technology, Yunnan Normal University, Kunming 650000, China*
[2]*Department of Thoracic Surgery, Yunnan First People's Hospital, Kunming 650000, China*

Correspondence should be addressed to Yonghang Tai; yonghang@deakin.edu.au

Received 30 July 2018; Revised 31 October 2018; Accepted 29 November 2018; Published 6 January 2019

Guest Editor: Vincenzo Ferrari

The aim of this study is to develop and assess the peg transfer training module face, content and construct validation use of the box, virtual reality (VR), cognitive virtual reality (CVR), augmented reality (AR), and mixed reality (MR) trainer, thereby to compare advantages and disadvantages of these simulators. Training system (VatsSim-XR) design includes customized haptic-enabled thoracoscopic instruments, virtual reality helmet set, endoscope kit with navigation, and the patient-specific corresponding training environment. A cohort of 32 trainees comprising 24 novices and 8 experts underwent the real and virtual simulators that were conducted in the department of thoracic surgery of Yunnan First People's Hospital. Both subjective and objective evaluations have been developed to explore the visual and haptic potential promotions in peg transfer education. Experiments and evaluation results conducted by both professional and novice thoracic surgeons show that the surgery skills from experts are better than novices overall, AR trainer is able to provide a more balanced training environments on visuohaptic fidelity and accuracy, box trainer and MR trainer demonstrated the best realism 3D perception and surgical immersive performance, respectively, and CVR trainer shows a better clinic effect that the traditional VR trainer. Combining these in a systematic approach, tuned with specific fidelity requirements, medical simulation systems would be able to provide a more immersive and effective training environment.

1. Introduction

Video-assisted thoracoscopic surgery (VATS), which is the most common minimal invasive surgery (MIS) therapy for lung carcinomas [1–3], is the most widespread cancer in the world with only approximately 16% five-year survival rate. Furthermore, hospitalization outcomes show the patients' quality of life (QoL) and follow-up adjuvant chemotherapy endurance that are significantly promoted compared with traditional thoracotomy, without interfering with survival outcomes [4, 5]. Peg transfer training, as one of the essential modules of the fundamental thoracoscopic surgery curriculum, is the compulsory test requirement before the surgeries take the American Board of Surgery examination [6, 7]. With two Maryland clamps, trainees need to pick up six tiny blocks with nondominate hand, transfer to the other hand, and place them stably on the other side peg, respectively, whereby to achieve the bimanual dexterity and eye-hand coordinative skills training [7, 8]. The surgical simulators can be divided into box trainer, VR trainer, AR trainer, and MR trainer, and each class has advantages and disadvantages. The Box trainer is the traditional surgical training framework based on a real physical model that is portable and easy to operate; however, the disadvantage is this kind of trainer cannot be reused for multitimes. VR-based surgical simulations have attracted many researchers' attentions over the years and gradually turned into a real-life medical training simulation solution, providing repeatable training experience, without ethical or hygienic issues [9–11], and the disadvantage of this simulator is lack of high immersive visual and haptic rendering algorithms. Due to the recent advances in the field of AR and MR, cognitive sense has brought into the next level of surgical simulator with enhanced immersion and interactivity [12–14];

however, there are also several defects of these simulators, such as the high price, visual uncomfortable, and lack of the real surgical environment fidelity. Maciel et al. developed a virtual reality laparoscopic skill trainer named VBLaST, with real-time evaluation function for the peg transfer [15], and the Lap Mentor™ simulator [16, 17]. Loukas et al. compared the AR-based peg transfer with the box simulator and VR simulator [6], and Huber et al. added a highly immersive 360° real operating room environment to construct an MR training environments [18, 19]. Nevertheless, there is no comparative study that has focused box, AR, VR, CVR, and MR simulators on the peg transfer training, in another word, which one is the most effective simulator to shorten the surgical learning curve [20, 21]. The aim of this research is to determine which simulator is better by the advantages and disadvantages of the five simulators that were compared through the evaluation results of the face and content and construct. There are three main innovative contributions in this study, and the overarching one is we compared the experimental data and simulation results among five different kinds of trainers of peg transfer training and summarized the conclusion which may benefit for the advanced research of the virtual surgery. The second one is we addressed a full immersive and accurate scenario to peg transfer training with a detailed assessment tool, and the last one is we addressed the detailed design of the first commercial available VATS VR simulator in both hardware and the training implementation.

The structure and content of this article are organized as follows: we briefly make an introduction on peg transfer training and challenges of virtual medical training in the abovementioned part, after that, we review the previous related works on the virtual peg transfer simulator. Thirdly, the simulator design, with virtual training evaluation experiments, is designed in the methodology part. Fourthly, the experimental data and evaluation results are demonstrated and discussed. Finally, we summarize the results and the potential contributions this paper makes.

2. Materials and Methods

2.1. Simulator Design

2.1.1. Hardware. In the light of the development of commercial VR-based laparoscopic surgical simulators, LapSim® (Surgical Science, Sweden) and LAP Mentor (3D System, USA) [1, 15], and after observed considerable VATS lobectomy operations in the thoracic surgery department of Yunnan First People's Hospital, we chose the three-port standardized anterior VATS surgery as the simulated content. As the commonest surgical approach in lung carcinomas lobectomy, the three-port VATS are described as opened 3 single diameters of 1.5 cm incisions in the side of the patient's thorax: one is for endoscopic camera recording the operation field and the other two are the operative tunnels for surgical instruments stretching in and out to dissect and staple the lesion [2, 16]. The surgical simulator is called VatsSim-XR, as shown in Figure 1, based on the procedures need to be simulated in the abovementioned, and hardware of VatsSim-XR (VR, AR, and MR) is $60 \times 67 \times 160$ cm and principally composed of a 24-inch naked eye 3D display, a footswitch pedal, two VATS surgical devices connected with dual haptic devices with one held as the drag instrument and the another mimics the stapling device, an endoscope connected with a 3D mouse (3Dconnexion, Germany), an HMD, and a workstation. To improve the surgical immersion, we proposed a solution to incorporate multiple higher-fidelity factors towards a surgeon's sensations (vision, touch, and hearing) during practical surgery to achieve total immersion. We developed a versatile simulation platform VatsSim-XR that is able to implement 5 different simulation modes, specifically, the AR and MR modes. To ensure the same training block between the real and virtual peg transfer platform, the .STL file is exported from the blender firstly, and then, the 3D printer is employed to print all the training pegs for the box training simulator. For the AR peg transfer simulation, we utilized a web camera Logitech CC2900ep HD1080p as the detected sensor for the display, and for the CVR training and the MR training, HTC VIVE (HTC Corporation, Taiwan, China) is the first virtual reality helmet, and it is used to show the surgical environments [22]. Especially, for the real surgical environments rendering, we employed the Samsung Gear 360 (Samsung, South Korea) to record the 360° video in the operation room (OR) of an upper-right lobe VATS in Yunnan First People's Hospital [18, 19, 23, 24].

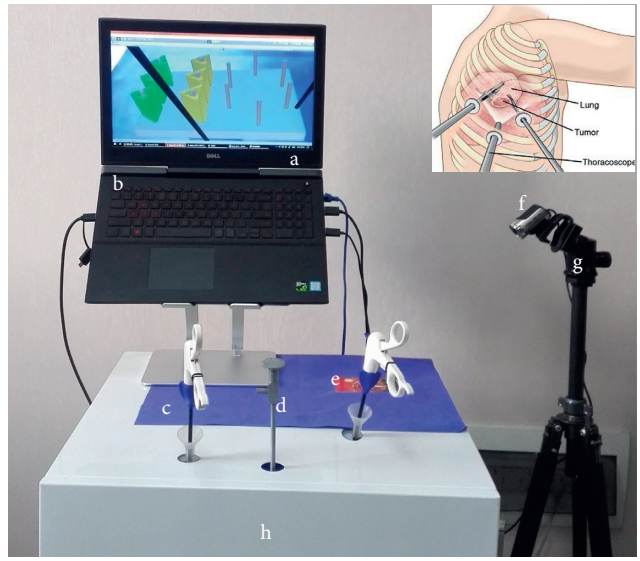


Figure 1: The immersive training platform for virtual peg transfer consists of (a) guide video, (b) a laptop with i7 6700 (3.4 GHz) CPU, 4 GB memory, and 1070 NVIDIA graphics GPU, (c) surgical clamp, (d) endoscope, (e) AR tracking target, (f) AR camera, (g) camera tripod, and (h) haptic devices (PHANTOM Omni).

2.1.2. Software. We designed a framework for the implementation of peg transfer training simulation, as shown in Figure 2. The virtual simulator mainly includes two parts: visual (physics and graphic) and haptic (PHANTOM Omni hardware and OpenHaptic software). The corresponding OpenHaptic function will be invoked when pressing the

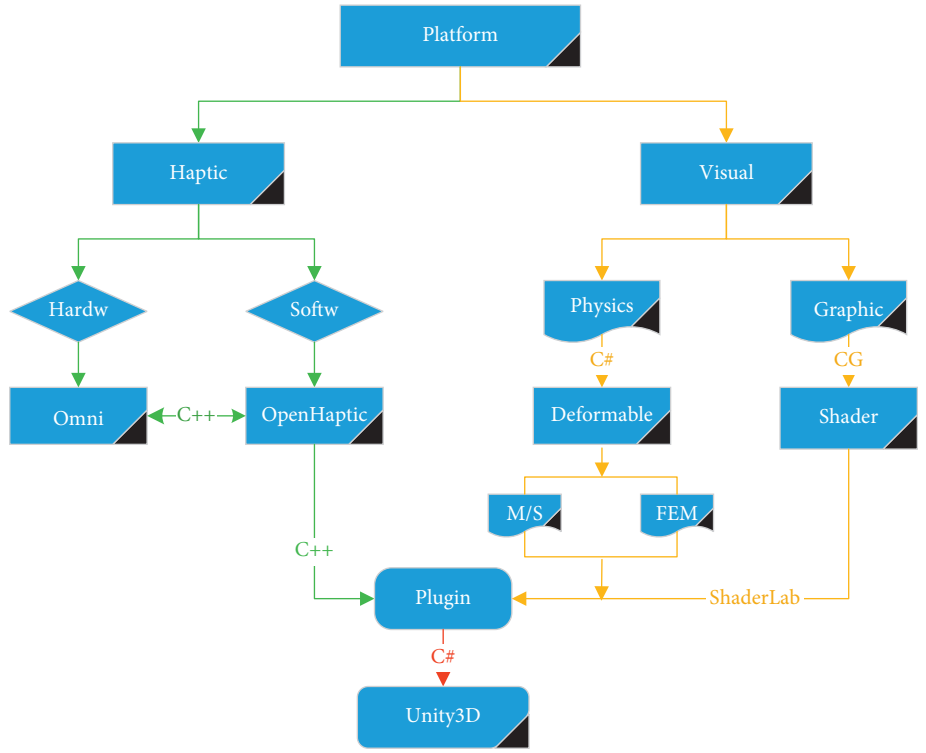


Figure 2: Visual and haptic rendering pipelines of the virtual peg transfer simulator.

physical button of PHANTOM Omni (Geomagic, USA) or there is contact between the virtual surgical instrument and the model with a rigid body, and then, there will have a corresponding force feedback. There are physics with rigid or deform and graphics with a shader in the visual aspect. We combine the visual plugin, such as Bullet, AR kit and SteamVR, and the haptic plugin OpenHaptic into the Unity3D (version 2017.3.1) for simulation training.

2.2. Participants. The evaluation cohort of 32 trainees comprising 24 novices and 8 experts underwent the peg transfer procedures on the five simulators conducted in the department of thoracic surgery of Yunnan First People's Hospital. Both expert and novice trainers firstly receive a didactic teaching from a developer of these simulators with the tasks and techniques of peg transfer. The medical experience of the novice group is 3 to 6 years and the expert group is 11 to 30 years. To evaluate if the virtual game experience may affect the training result, the experience of using both VR and HMD had been recored during the experiments. The demographic details of the trainees are shown in Table 1.

2.3. Simulator Tasks. Both novice and expert groups after didactic training session are towards to evaluate the simulator validity. The order of training is the box, VR, CVR, AR, and MR, and the 30-minute break should be interspersed between each peg transfer simulator test. Firstly, the trainee adjusted the virtual endoscope to reach the proper position and angle to obtain the best viewing perspective of the operating scene; after that, two clamps are used to grasp and transfer the block model on the peg base. The left-hand clamp grabs the block on the left side and passes it to the right clamp; then, the right clamp places the block on the right side, which is demonstrated in Figure 3.

Table 1: Demographic data of the thoracic surgery trainees.

	Group A (novices)	Group B (experts)
Number	24	8
Age (years)	25.5 (24–29)	42.2 (39–56)
Postgraduate year of training	4 (3–8)	12 (11–30)
Male (%)	92	83
Right-handed (%)	100	87
Box trainer experienced	<10	>50
VR game experienced	5/24	1/8
HMD experienced	3/24	1/8

After the aforementioned experiments, all 32 trainees' performance is recorded by the objective questionnaires to evaluate the detailed surgical skills and be compared with each other. Subjective questionnaires also need to be filled to evaluate the face and content validity of five simulators, and the detailed evaluation flow chart is demonstrated in Figure 4.

2.4. Evaluation. A questionnaire consisting of 13 questions about the visual and haptic aspects of the simulators was created for face and content assessment validity. Due to the lack of experience of the novice, subjective judgment may have a large error in the understanding of the simulator.

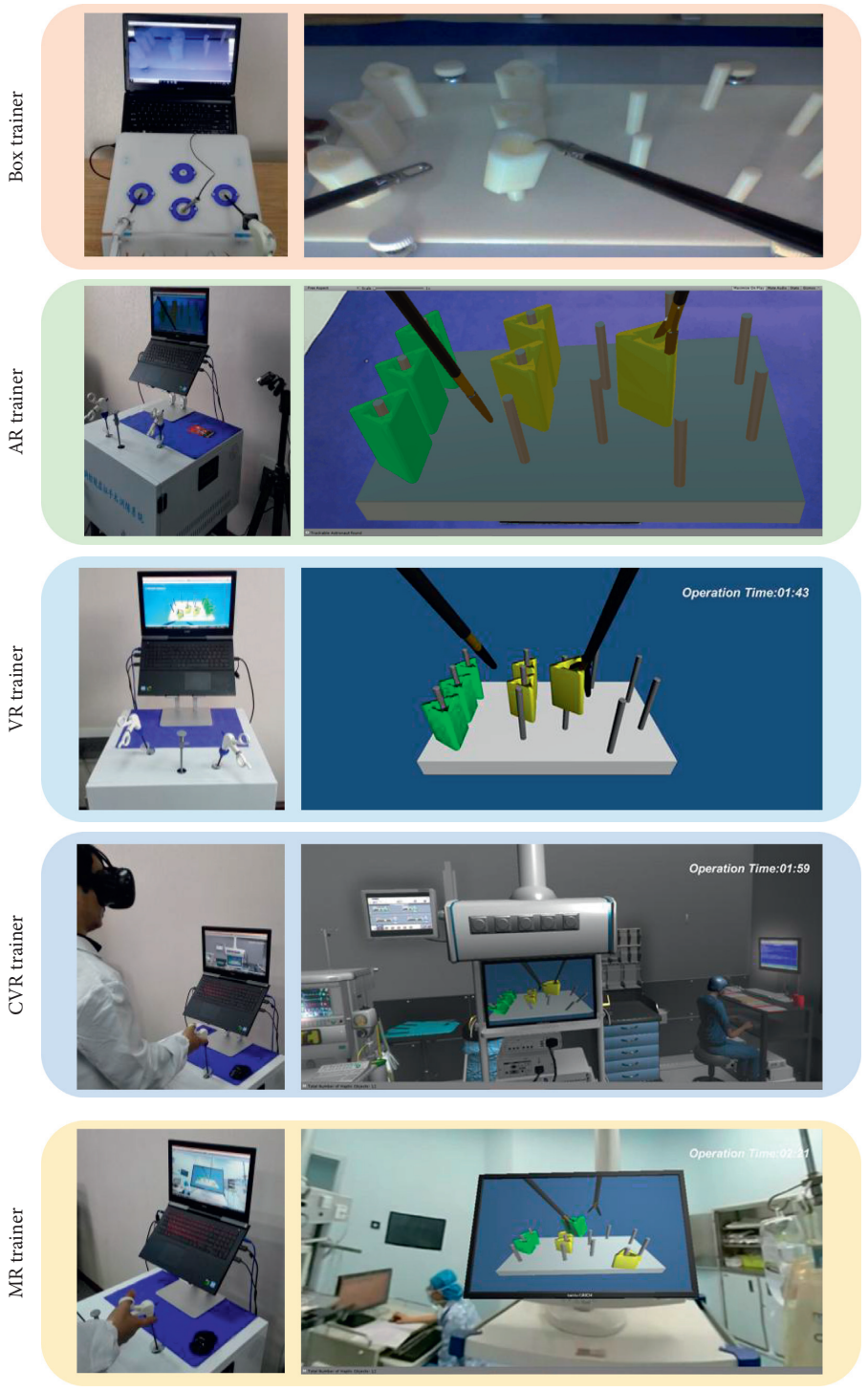


Figure 3: Peg transfer training platform. (a) The manipulative platform. (b) The operation interface. CVR training and MR training need the HMD device, and the AR training needs the camera module.

Face and content validities were established by 8 experts that standardized subjective questionnaires parameters. This questionnaire utilizes 5-point Likert-type scales to evaluate the visual and haptic of the simulators (1 = poor to 5 = excellent). The detail subjective questionnaire of the face and content validity is demonstrated in Table 2. The objective evaluation of the construct has six assessed items that are totally operation time (T), surgical clamps track length (CL), endoscope track length (EL), surgical clamps angle accumulation (CA), endoscope angle accumulation (EA), and the numbers of block drop (ND). First, verify whether each set of data obeys a normal distribution, and

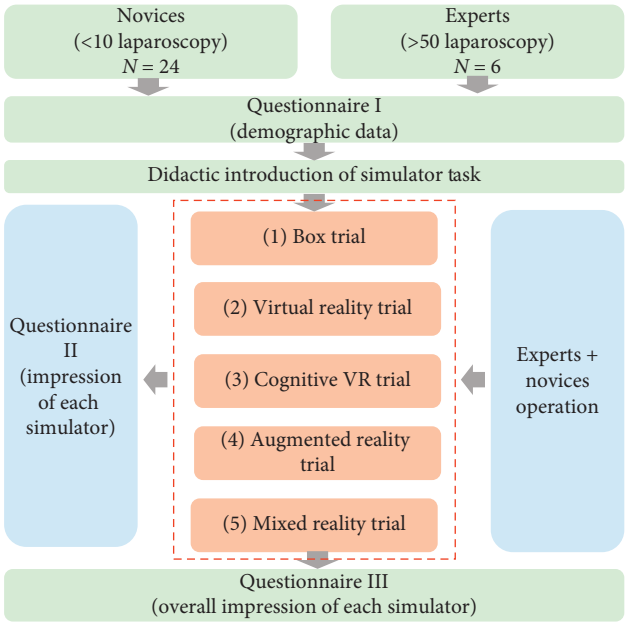


Figure 4: Evaluation system procedures design underwent our immersive virtual peg transfer simulator.

Table 2: Subjective questionnaire of the face and content validity.

Face and content validity questions (score: 1–5, 1 = poor to 5 = excellent)
Q1: Realism of peg model (visual)
Q2: Realism of endoscope model (visual)
Q3: Realism of surgical clamps (visual)
Q4: Realism of surgical environment (visual)
Q5: Comfortable of training content in HMD (visual)
Q6: Overall realism of visualization
Q7: Realism of peg manipulation (haptic)
Q8: Realism of endoscope manipulation (haptic)
Q9: Realism of surgical clamps manipulation (haptic)
Q10: Realism of interaction between the surgical clamps and peg (haptic)
Q11: Overall realism of manipulation
Q12: I would like to recommend this simulator to VATS surgical training for medical students
Q13: I would like to recommend this simulator as an assessment tool for VATS surgical skills

then compare the performance of multiple simulators under six parameters.

2.5. Data Analysis. In terms of face and content validity, descriptive statistics were used to analyze the questionnaire data by calculating the mean value of subjective question and their standard errors. For the construct validity, assessing the test results of novice and expert on six parameters of five simulators with Shapiro–Wilk test, $p > 0.05$, is considered to be subject to normal distribution. Box plots are used for analysis of objective parameters. Each group of T of novices and experts on the five simulators follows a normal distribution and was compared between groups using an independent-sample t-test, and each group of T of five simulators on novices and experts was compared between groups using the one-way ANOVA test. But, most of the CL, EL, CA, EA, and ND were not a normal distribution, each group of those data of novices and experts on the five simulators was compared between groups using the two-tailed Mann–Whitney U test, and each group of those data of five simulators on novices and experts was compared between groups using the Kruskal–Wallis test. A p value < 0.05 was considered significant. Analyses were performed using the software package, IBM SPSS Statistics, version 20.0.

3. Results

3.1. Face and Content Validity. Aiming to the perceptions of visual and tactile sensations during the use of the five simulators, we set up a subjective questionnaire with 13 questions using the scoring method of the 5-point Likert-type scale. After the experiment of the simulation operation, the questionnaire was filled by 8 experts for the experience during the operation. Through the Shapiro–Wilk test, each set of score data is an approximately normal distribution. The thirteen subjective questionnaires were collated, and the score data of each simulator were assessed by the Shapiro–Wilk test. The average and standard deviation of each set of data are shown in Table 3, and the distribution of the average score is shown in Figure 5. It can be found that the box and MR scores demonstrated a higher score than the AR and VR scores in a visual sense. In terms of haptic sense, the box scores are higher than the AR, VR, CVR, and MR scores, which demonstrated the best haptic immersion among these virtual simulators.

3.2. Construct Validity. During the operation of 32 trainees, the VR, CVR, AR, and MR simulators automatically recorded the trajectory length and rotation angle of the surgical instruments and endoscopes in real time, as well as the total time spent and the numbers of block drop on each operation. However, the box simulator cannot record the length and angle of the motion track and the numbers of the block drop, and the total time of the operation can only be roughly recorded by the timer. Six box plots for the experimental data of six parameters T, CL, EL, CA, EA, and ND were made by SPSS, and those box plots which include mediums and means are shown in Figure 6. Similarly, after calculating these p values of the parameters of the different simulators for the novice and expert groups, the results are shown in Table 4. Most of the evaluation parameters of the expert group are lower than that of the novice group in each simulator. $p < 0.05$, which is only for the EL of VR simulator. EL and EA of CVR simulator data of the expert group demonstrated a higher score than that of the novice. Comparing VR, CVR, and MR scores of the expert group, those scores of the VR group demonstrate the best performance. The clamp moving trajectory of AR, VR, CVR, and MR groups during the simulation is demonstrated in Figure 7, and it shows that the constructive difference between these four training groups and the AR group demonstrated a better-concentrated trajectory than other groups.

Table 3: Subjective questionnaire results of the face and content validity (experts group).

Questionnaires	Box		AR		VR		CVR		MR	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
	Score (1–5)									
Q1	3.75	0.71	2.63	0.92	3.00	0.53	3.38	0.92	4.25	0.71
Q2	4.13	0.64	2.63	0.92	2.63	0.91	3.38	0.92	3.63	0.92
Q3	4.25	0.71	3.50	0.93	3.50	0.53	3.88	0.64	4.00	0.76
Q4	1.50	0.76	1.63	0.74	1.63	0.74	2.75	0.71	4.25	0.71
Q5	—		—		—		2.75	1.04	2.50	0.93
Q6	3.13	0.64	2.50	0.93	3.13	0.64	3.13	0.64	4.13	0.64
Q7	4.25	0.89	2.50	0.93	2.00	0.75	1.88	0.64	2.00	0.76
Q8	3.75	0.71	2.00	0.76	3.13	0.83	2.00	0.76	2.50	0.93
Q9	4.75	0.46	3.88	0.64	4.00	0.76	3.00	0.76	2.13	0.64
Q10	4.63	0.52	3.03	0.64	2.00	0.76	1.63	0.74	1.25	0.46
Q11	3.50	0.93	2.88	0.64	2.00	0.76	2.13	0.64	2.00	0.76
Q12	3.13	0.64	2.00	0.76	2.88	0.64	2.00	0.76	3.25	1.04
Q13	3.75	0.71	1.75	0.71	2.75	1.04	1.75	0.71	3.25	0.89

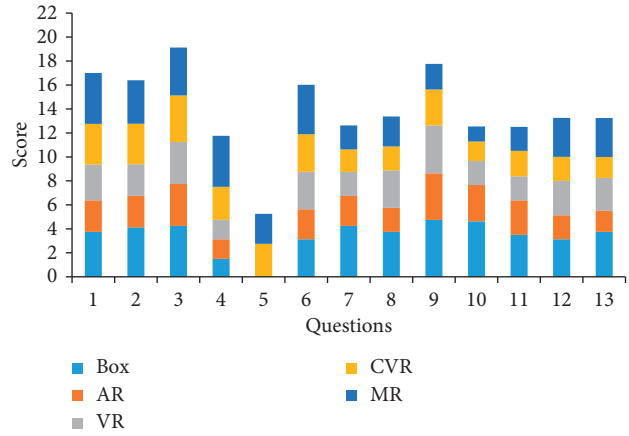


Figure 5: Face and content validity score of the expert group on these five simulators.

4. Discussion

There is an increasing use of simulators to learn, improve, and rehearse surgical skills in medical training. The VatsSim-XR is a versatile simulator that is able to perform multiple surgical training scenarios and implement 4 different simulation modes (VR, CVR, AR, and MR) on a single device. The Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) and the American College of Surgeons (ACS) have established a standard for fundamentals of laparoscopic surgery (FLS) training and surgical skills assessment [25–27]. The FLS consists of five basic foundations: peg transfer, pattern cutting, ligation loop, and suturing with either intracorporal or extracorporeal knot tying, and these tasks are designed to train the operational skills and assessment of medical interns and residents. Peg transfer is the basic training task to train the dual-hands coordination and the hand-eye coordination for the medical interns and residents [8]. FLS simulation training is mostly based on box and VR simulators in the market.

In terms of haptic, the box simulator is superior to the other four simulators, but an operation object is an object that is easily destroyed after repeated use. In terms of visual, the box and MR simulators are better. The AR simulator is the best in terms of interactive, and in the interactive environment part, where the simulator is more immersive: MR simulator > CVR simulator; however, the box, VR, and AR simulators have almost no surroundings and are less immersive.

Compared with novice groups of six parameters on five simulators, most of the expert groups show short operation time, short track length, small angle accumulation, and small drop number, but only these three groups, the EL of the VR simulator, EL, and EA of CVR the simulator, show the opposite result, and the main reason is that when adjusting the endoscope, the novice does not move and rotate the endoscope to an optimal viewing angle in strict accordance with the surgical standard, resulting in greater difference. For the VR, CVR, and MR simulators of six parameters, the data size of expert groups basically follows a rule VR < CVR < MR, but only the MR simulator data of the CA parameter are smaller than VR and MR simulators, because the experts have adapted to the 360° real operating room scene in the MR simulator in advance. Although the MR simulator is not perfect in terms of T, CL, EL, CA, EA, and ND, it provides a highly realistic operating room for hospital

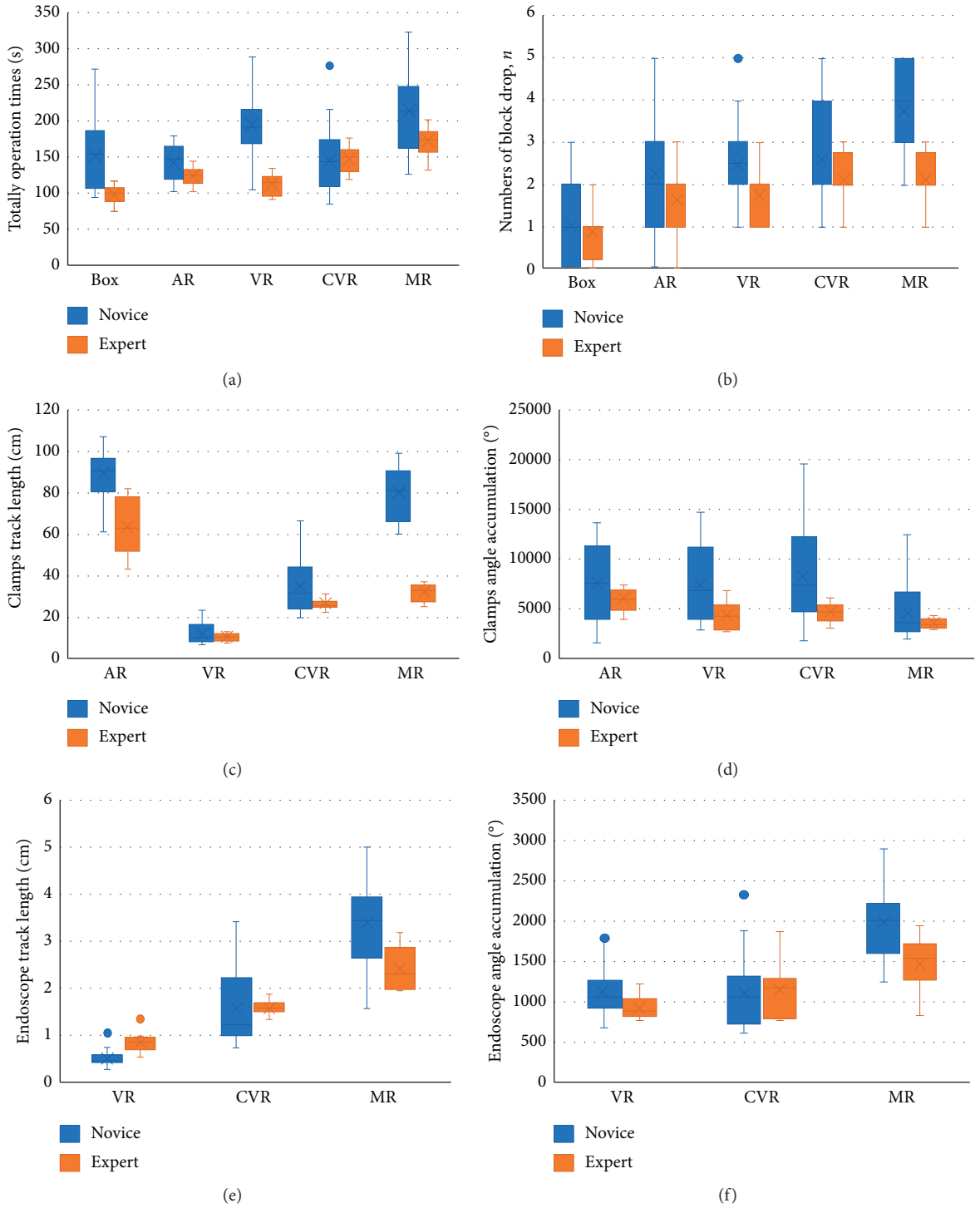


FIGURE 6: Performance comparison of six parameters in five simulators. (a) T: totally operation time. (b) ND: numbers of block drop. (c) CL: surgical clamps track length. (d) CA: surgical clamps angle accumulation. (e) EL: endoscope track length. (f) EA: endoscope angle accumulation.

interns and residents. The environment, closer to the real surgery site, has great potential for development. Experiments and evaluation results conducted by both professional and novice thoracic surgeons show that the AR trainer is able to provide a more balanced training environment on visuohaptic fidelity and accuracy, the box trainer and MR trainer demonstrated the best realism 3D perception and surgical immersive performance, respectively, and the CVR trainer shows a better clinic effect that the traditional VR trainer.

The advantages and disadvantages of the five simulators are shown in Table 5. The haptic, visual, and surrounding

Table 4: The comparison of objective parameters for construct validity (experts and novices).

		p value					
		T	CL	EL	CA	EA	ND
Box	Experts Novices	0.002	—	—	—	—	0.781
AR	Experts Novices	0.053	0.000	—	0.277	—	0.302
VR	Experts Novices	0.000	0.761	0.001	0.024	0.037	0.084
CVR	Experts Novices	0.744	0.041	0.177	0.030	0.486	0.405
MR	Experts Novices	0.045	0.000	0.003	0.965	0.009	0.001
p value	Experts Novices	0.000 0.000	0.000 0.000	0.000 0.000	0.005 0.013	0.012 0.000	0.014 0.000

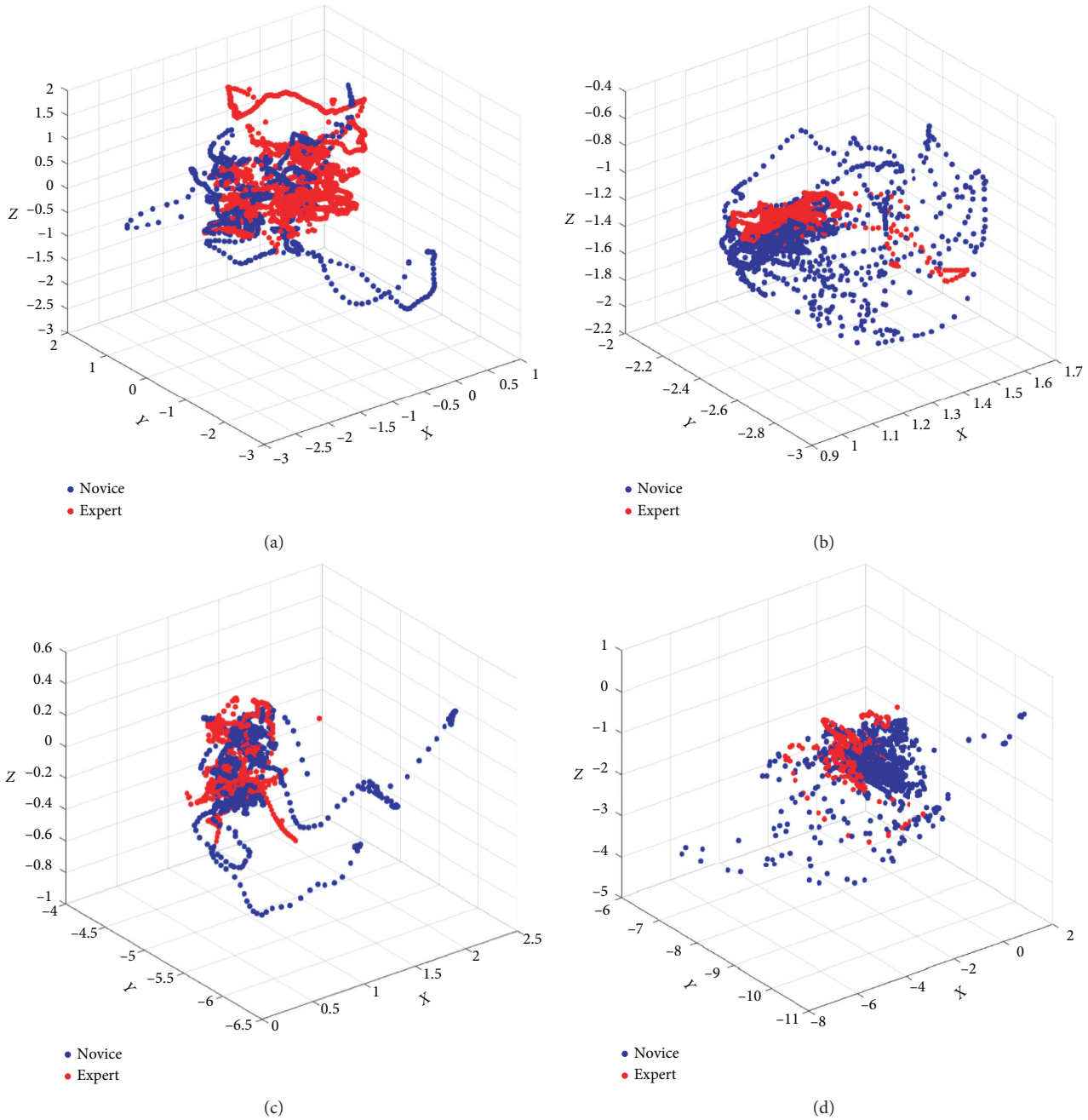


Figure 7: Clamp moving trajectory comparison during the simulation. (a) VR group. (b) AR group. (c) CVR group. (d) MR group.

Table 5: Comparisons of the five peg transfer simulators.

	Advantages	Disadvantages	Immersive		
			Haptic	Visual	Environments
Box	Realistic force feedback Realistic 3D perception Cost-effective Portable and easy to use	Low integration Lack of objective assessment Low immersion Without OR environment	III	I	I
VR	Objective assessment High interactivity Unlimited use	Low immersion Without OR environment Low 3D perception	I	I	I
CVR	Objective assessment High interactivity High immersion Motion track (HMD)	Virtual OR environment Narrow FOV Low 3D perception	I	II	II
AR	Objective assessment High interactivity High immersion Realistic 3D perception	Low immersion Without OR environment Environmental influence	I	I	I
MR	100% OR environment Objective assessment High interactivity High immersion	OR environment distortion Narrow FOV	I	III	III

environments are compared separately by levels (III means the highest and I means the lowest). The complex design of our integrated system will provide a more immense and effective training environment for the medical surgery simulation.

5. Conclusions

In this paper, we have detailed a virtual surgical educative simulator with realistic performance in both visual and haptic sensations for the peg transfer procedures, build the face, content and construct validation on the peg transfer training by use of the box, VR, AR, and MR trainer, thereby to compare advantages and disadvantages of these simulators. However, during the interaction, the sense of touch is not immersive enough; furthermore, there is a cross between the gripper and the interacting virtual object. What is more, in the comparison experiments design, the box simulator is not able to automatically record the trajectory and rotation angle parameters of the instrument and endoscope, which means the comparison of the six parameters is not comprehensive enough in this manuscript. In the future works, we will improve the visual and haptic experiment design parts to achieve a high immersive and realistic simulation environment both visual and haptic perception.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

We thank Professor Minghui, Xiao, Dr. Min Zou, Dr. Jie Shen, and Mr. Ran Cao of Yunnan First People's Hospital for the helpful surgical suggestions. This research was funded by the National Natural Science Foundation of China (61741516) and the National Science Foundation of Yunnan Province, China (ZD2014004), and Yunnan Key Laboratory of Opto-Electronic Information Technology, Kunming, China.

Supplementary Materials

The supplementary video demonstrated the peg transfer experimental procedures on the box, VR, CVR, AR, and MR trainer. (*Supplementary Materials*)

References

[1] B. Solomon, C. Bizekis, S. L. Dellis et al., "Simulating video-assisted thoracoscopic lobectomy: a virtual reality cognitive task simulation," *Journal of Thoracic and Cardiovascular Surgery*, vol. 141, no. 1, pp. 249–255, 2011.

[2] H. J. Hansen and R. H. Petersen, "Video-assisted thoracoscopic lobectomy using a standardized three-port anterior approach—the Copenhagen experience," *Annals of Cardiothoracic Surgery*, vol. 1, no. 1, pp. 70–76, 2012.

[3] K. Jensen, F. Bjerrum, H. J. Hansen, R. H. Petersen, J. H. Pedersen, and L. Konge, "Using virtual reality simulation to assess competence in video-assisted thoracoscopic surgery (VATS) lobectomy," *Surgical Endoscopy*, vol. 31, no. 6, pp. 2520–2528, 2016.

[4] D. Gossot, J. Lutz, M. Grigoroiu, E. Brian, and A. Seguin-Givelet, "Thoracoscopic anatomic segmentectomies for lung cancer: technical aspects," *Journal of Visualized Surgery*, vol. 2, p. 171, 2016.

[5] J. He and X. Xu, "Thoracoscopic anatomic pulmonary resection," *Journal of Thoracic Disease*, vol. 4, no. 5, pp. 520–547, 2012.

[6] C. Loukas, V. Lahanas, and E. Georgiou, "Surgical simulation training systems: box trainers , virtual reality and augmented reality simulators," *International Journal of Advanced Robotics and Automation*, vol. 1, no. 2, pp. 1–9, 2016.

[7] E. M. Ritter, T. W. Kindelan, C. Michael, E. A. Pimentel, and M. W. Bowyer, "Concurrent validity of augmented reality metrics applied to the fundamentals of laparoscopic surgery (FLS)," *Surgical Endoscopy*, vol. 21, no. 8, pp. 1441–1445, 2007.

[8] V. S. Arikatla, W. Ahn, G. Sankaranarayanan, and S. De, "Towards virtual FLS: development of a peg transfer simulator," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 10, no. 3, pp. 344–355, 2013.

[9] N. Suzuki, A. Hattori, S. Ieiri, M. Tomikawa, H. Kenmotsu, and M. Hashizume, "VR training system for endoscopic surgery robot," in *6th International Workshop on Augmented Environments for Computer-Assisted Interventions (AE-CAI 2011)*, pp. 117–129, Toronto, ON, Canada, September 2012.

[10] A. Boejen and C. Grau, "Virtual reality in radiation therapy training," *Surgical Oncology*, vol. 20, no. 3, pp. 185–188, 2011.

[11] M. P. Schijven and J. J. Jakimowicz, "Validation of virtual reality simulators: key to the successful integration of a novel teaching technology into minimal access surgery," *Minimally Invasive Therapy & Allied Technologies*, vol. 14, no. 4-5, pp. 244–246, 2009.

[12] Fida2018_Article_AugmentedRealityInOpenSurgery.pdf.

[13] A. Moglia, V. Ferrari, L. Morelli, M. Ferrari, F. Mosca, and A. Cuschieri, "A systematic review of virtual reality simulators for robot-assisted surgery," *European Urology*, vol. 69, no. 6, pp. 1065–1080, 2016.

[14] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari, "Augmented reality in neurosurgery: a systematic review," *Neurosurgical Review*, vol. 40, no. 4, pp. 537–548, 2017.

[15] A. Maciel, Y. Liu, W. Ahn, T. P. Singh, W. Dunnican, and S. De, "Development of the VBLaST™ : a virtual basic laparoscopic skill trainer," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 4, no. 2, pp. 131–138, 2008.

[16] N. Muralha, M. Oliveira, M. A. Ferreira, and J. Costa-Maia, "Virtual reality simulation as a tool to monitor surgical performance indicators: VIRESI observational study," *Acta Médica Portuguesa*, vol. 30, no. 5, p. 388, 2017.

[17] A. Zhang, M. Hünerbein, Y. Dai, P. M. Schlag, and S. Beller, "Construct validity testing of a laparoscopic surgery simulator (Lap Mentor®)," *Surgical Endoscopy*, vol. 22, no. 6, pp. 1440–1444, 2008.

[18] T. Huber, T. Wunderling, M. Paschold, H. Lang, W. Kneist, and C. Hansen, "Highly immersive virtual reality laparoscopy simulation: development and future aspects," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 2, pp. 281–290, 2018.

[19] T. Huber, M. Paschold, C. Hansen, H. Lang, and W. Kneist, "Artificial versus video-based immersive virtual surroundings: analysis of performance and user's preference," *Surgical Innovation*, vol. 25, no. 3, pp. 280–285, 2018.

[20] G. P. Moustris, S. C. Hiridis, K. M. Deliparaschos, and K. M. Konstantinidis, "Evolution of autnomous and semi-autnomous robotic surgical systems: a review of the literature," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 7, no. 4, pp. 375–392, 2011.

[21] A. D. L. Sihoe, "Uniportal video-assisted thoracoscopic lobectomy," *Annals of Cardiothoracic Surgery*, vol. 5, no. 2, pp. 133–144, 2016.

[22] J. Egger, M. Gall, J. Wallner et al., "HTC Vive MeVisLab integration via OpenVR for medical applications," *PLoS One*, vol. 12, no. 3, Article ID e0173972, 2017.

[23] S. Yoganathan, D. A. Finch, E. Parkin, and J. Pollard, "360° virtual reality video for the acquisition of knot tying skills: a randomised controlled trial," *International Journal of Surgery*, vol. 54, pp. 24–27, 2018.

[24] S. Mohiuddin, P. Voronov, H. Nishioka et al., "Applying 360-degree videos to interdisciplinary training and simulation," *Anesthesia & Analgesia*, vol. 124, no. 5, pp. 83-84, 2017.

[25] S. M. B. I. Botden, S. N. Buzink, M. P. Schijven, and J. J. Jakimowicz, "Augmented versus virtual reality laparoscopic simulation: what is the difference?," *World Journal of Surgery*, vol. 31, no. 4, pp. 764–772, 2007.

[26] J. H. Peters, G. M. Fried, L. L. Swanstrom et al., "Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery," *Surgery*, vol. 135, no. 1, pp. 21–27, 2004.

[27] G. M. Fried, "FLS assessment of competency using simulated laparoscopic tasks," *Journal of Gastrointestinal Surgery*, vol. 12, no. 2, pp. 210–212, 2008.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 6357351, 8 pages
https://doi.org/10.1155/2018/6357351

Research Article

The Effect of Augmented Reality and Virtual Reality on Inducing Anxiety for Exposure Therapy: A Comparison Using Heart Rate Variability

Chai-Fen Tsai,[1] Shih-Ching Yeh,[2] Yanyan Huang,[3,4] Zhengyu Wu,[3] Jianjun Cui,[2] and Lirong Zheng[2]

[1]*Taipei Veterans General Hospital, Taipei 11217, Taiwan*
[2]*Fudan University, Shanghai 200433, China*
[3]*Huashan Hospital, Fudan University, Shanghai 200040, China*
[4]*Tianqiao and Chrissy Chen Institute for Brain Diseases in Shanghai, Shanghai 200040, China*

Correspondence should be addressed to Shih-Ching Yeh; shihching.yeh@gmail.com

Received 16 July 2018; Revised 1 October 2018; Accepted 12 November 2018; Published 25 November 2018

Guest Editor: Vincenzo Ferrari

Claustrophobia is an anxiety disorder characterized by the fear of enclosed spaces. Although medication treatment can effectively control symptoms, the effects quickly disappear once medication is discontinued. Many studies have shown that combining psychotherapy and medication is more efficacious than solely using medication. However, the weaknesses of the traditional psychotherapy are that it is time-consuming and expensive. Alternatively, vivo exposure therapy is proposed in which anxiety is gradually triggered with stimuli. Targeting claustrophobia is diagnosed using the traditional method, and this study established virtual reality (VR) and augmented reality (AR) environments consistent with claustrophobic characteristics, comparing the two using an experimental process to examine whether VR and AR environments are equally capable of triggering anxiety in participants. This study further analysed the efficacies of VR and AR by measuring changes in participant's heart rates variability (HRV) and examining data from survey questionnaires. HRV results indicated that the proposed VR system and AR system were both able to trigger anxiety. Furthermore, the AR environment produced a stronger experience for the participants and caused physiological reactions more evident than those caused by the VR environment. Regarding the anxiety questionnaire, the participants suggested that their anxiety was significantly higher in the VR environment than in the AR environment.

1. Introduction

Claustrophobia is an anxiety disorder characterized by the fear of enclosed spaces [1]. Under certain circumstances, such as being in elevators, trains, or airplane cabins, sufferers can exhibit symptoms of panic or fear of panic [2]. Causes of claustrophobia are likely to be extremely small tonsils, genetic predisposition, or emotional responses induced by the classical condition. The two main symptoms are the fear of enclosed spaces and the fear of constriction [3–5]. Psychological literature suggests that people with claustrophobia do not necessarily fear the enclosed spaces themselves; instead, they fear that some dangerous event will occur in this type of environment, leading to insufficient air within the space and causing suffocation.

Cognitive therapy is a commonly accepted treatment for anxiety disorders [6]. The goal is to correct the person's misunderstanding toward the objects of their fear. A study by Rachman and Taylor [4] showed that cognitive therapy is effective in nearly 30% of individuals with claustrophobia, effectively reducing their fear and negative thoughts regarding specific environments [7, 8]. In vivo exposure therapy, which forces patients into the environments they fear, allows individuals to experience their fear. During treatment, therapists gradually increase the degree of live exposure. Booth and Rachman [7] found that live exposure

therapy is effective in reducing the fear of and negative thoughts regarding enclosed spaces in nearly 75% of patients [7]. In addition to these therapies, interoceptive exposure therapy, psychoeducational intervention, reverse conditioning, and breathing retraining are somewhat effective in treating claustrophobia. In addition, antidepressants or medications used for treating high blood pressure and heart disease can reduce the discomfort felt by people with claustrophobia during anxiety attacks [7, 9–12]. Although medication treatment can effectively control symptoms, the effects quickly disappear once medication is discontinued. Many studies have shown that combining psychotherapy and medication is more efficacious than solely using medication, and that the effects of treatment are longer lasting [13]. The weaknesses of the traditional psychotherapy are that it is time-consuming [14], generally requiring 1 or 2 years to conduct a full psychiatric analysis and complete regular treatments, and expensive. Exposure therapy is similarly time-consuming, and some patients refuse returning to treatment because of the fear they experience from the method [15, 16].

Virtual reality (VR) originated from Sutherland's (1965) concept of "The Ultimate Display" [17], which uses computer simulation to produce a 3D virtual world, provides users with visual, auditory, and tactile sensory simulations and allows them to view a simulated world using computers and related equipment. VR enables real-time, unrestricted observation of objects in a 3D space and allows for user interaction. The three basic elements of VR systems for the user are immersion, interaction, and imagination [18]. It emphasizes that the user could have a better control or dominance over the virtual environment.

VR applications enjoy the advantages of simulating environments that are difficult or impossible to find in everyday life at a low cost, and these environments can be customized based on requirements to meet specific demands. It was difficult for early VR technologies to enable immersion because their graphics were unrealistic. However, years of development have led to constant innovation and improvement in both software and hardware, providing stable and reliable environments for a variety of entertainment, simulation, and training purposes. Many studies of phobias now involve experiments performed in VR [19, 20]. Instead of making a role-playing activity in real environment, VR provides physicians and patients a way for exposure therapy that is safer, more comfortable, and less resource-intensive. In addition, it is possible to use VR to construct environments that are difficult to find in real life, a feature particularly useful for patients who encounter difficulties with imaginal exposure therapy [21, 22]. Numerous research experiments have demonstrated that VR is an effective tool for treating several phobias, such as acrophobia [23], arachnophobia [24], aviophobia, claustrophobia, and agoraphobia.

Augmented reality (AR) is an approach that integrates virtual objects with the scene of real world that enables the user to perceive an augmented world, as defined by Azuma [25] and Milgram and Kishino [26]. AR technologies can calculate the spatial positions of camera images in real time and provide corresponding information through display equipment, which allows VR images to be virtually embedded in the real world and relevant interactions therein. AR applications are considerably diverse. For example, AR can provide industrial support by aiding technicians in repairing automobiles or offer more distinctive and interesting gaming experiences. In recent years, AR has been suggested to be similar to VR in its efficacy for treating phobias [27, 28]. Juan et al. [29, 30] compared AR technologies with VR for effectiveness in treating acrophobia, finding that AR and VR systems are equally capable of triggering fear of heights. In a study of VR-triggered anxiety associated with acrophobia, Juan and Pérez [31] compared the degrees of presence induced by the cave automatic virtual environment (CAVE) display and head-mounted display (HMD) devices, finding that anxiety is highly correlated with presence. Furthermore, the study showed that the degree of presence provided by the CAVE display device was significantly higher than that of the HMD.

The purpose of this study was to compare established VR and AR environments to determine whether the environments are equally capable of triggering anxiety. The triggering characteristics in VR and AR environments considered claustrophobia and some other anxiety disorder-related diseases, such as agoraphobia. Outcome measurements included heart rate variability and state/trait survey results. Yeh et al. [32] had preliminary work on the comparison of VR and AR on induced anxiety using heart rate and skin conductance as indicators of anxiety. This study presented in this paper was more advanced, while heart rate variability was measured in which more indicators in regard to anxiety were analysed. Also, some preliminary data or raw data in early stage was published [33]. This paper presented compete results with a full-scale statistical analysis.

2. Materials and Methods

2.1. System Design. This experiment compared the degree of anxiety stimulated by VR and AR environments and the degree of realism experienced by participants. An enclosed elevator was used as the simulated environment because of its relevance to everyday life and relatively high likelihood of triggering claustrophobia. In addition, an interactive virtual skyscraper elevator system was constructed using VR and AR technologies, integrating visual and auditory stimuli to trigger anxiety in the participants. The 3D virtual environment created for the proposed system was developed using the Unity 3D game engine with Windows 7 Enterprise Edition. To increase the degree of interaction between the participants and the virtual environment, the system used a Z800 3D Visor Head-Mounted Display (HMD) for the output, and an embedded posture recognition device was used in conjunction with the official software development kit (SDK) to convert users' head movements into mouse motions. Furthermore, an MSI MyECG E3-80 First Professional-Grade Portable Electrocardiogram (ECG) was employed to measure and extract the changes in the heart rates of participants during the experiment for later analysis.

For the VR system, all visuals were composed of virtual scenery. The 3D modules used therein were developed using

3DMAX software in conjunction with texture mapping. The modules were incorporated into Unity 3D for use after development. The system framework is shown in Figure 1. For the AR system, full high-definition images, provided by a Logitech HD Pro C910 WebCamera, were used for the visuals. The Unity 3D engine imported the live video streaming and put virtual object over the real-scene image via texture mapping in order to come out with the AR scene. Next, the researchers added stimulating incidents, such as flashing lights, an electricity blackout, a fire, and thick smoke. It was not necessary to add tags to the surrounding environment during filming; the desired effects were achieved by intuitively adding virtual objects into the environment at the appropriate locations. The system framework is shown in Figure 2.

To enable adjustments in the stimulus level based on the viewer's circumstances, the following events were designed: (1) elevator door closing: when the participants enter the virtual environment, they are located in an open opaque elevator; the elevator doors suddenly close after a certain amount of time. Visually, the entire environment switches from an open to an enclosed environment; the participants are unable to see the environment outside of the elevator, generating a sense of constriction and the sound of the elevator door closing heightens the presence of the scenario. (2) Brightness level: after the elevator doors have been closed for some time, the lights inside the elevator begin to flash and then eventually turn off. Consequently, the elevator enters a state of blackout. Visually, the lights switch from light to dark, preventing the participants from seeing the objects around them and inducing a level of psychological stress. (3) Alarm sound: from an auditory perspective, fire alarms, evacuation alerts, and impact noises begin to sound, when the elevator lights begin to flash and when the elevator enters blackout, causing the participants to begin doubting the circumstances outside the elevator and eliciting fear. (4) Heartbeats sound: after the elevator enters blackout, participants cannot see any visuals for a certain period of time. A series of faint heartbeats start sounding in the darkness and, coupled with the quiet surroundings, make it difficult for participants to distinguish whether the heartbeats are sound effects or the sound of their own heartbeating. (5) Flames and heavy smoke: after participants gradually become acclimated to the surrounding darkness, flames suddenly burst out (Figure 3). Visually, the surrounding sparks and thick smoke cause the participants to believe they are in the midst of a fire. The crackling noises of the sparks add to the sense of realistic burning. (6) Screaming sounds: after the fire bursts out, participants hear screaming sounds from other virtual passengers and staff inside the building. There are four types of screams, exclamations, and cries intended to cause the participants to feel they are in an emergency situation. The frequency and type of screams alternated according to the change of the fire.

2.2. Experiment Design. This experiment focused on examining whether VR and AR environments elicit fear of enclosed spaces and comparing the effectiveness of these two

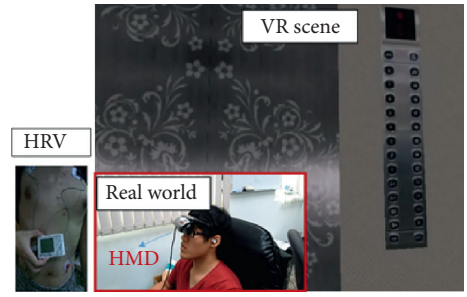


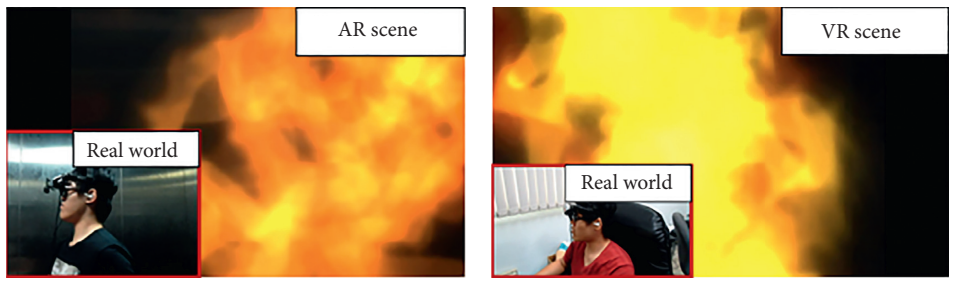
Figure 1: VR system.

Figure 2: AR system.

methods as a model of claustrophobia treatment. We used heart rate variability (HRV) as an objective measure of participants' physiological status and survey questionnaires to examine the participants' experiences.

We recruited 30 participants ranging in age from 18 to 35. These participants had no medical history in regard to claustrophobia or other types of fears. The order of environment conditions was counterbalanced to reduce order effect errors. The experiment took approximately two hours to complete both conditions. The participants were first fitted with the HRV physiological data measuring the instrument to collect their normal HRV for 1 h, after which the system timestamps were synchronized for the MyECG instrument and the computer was used for the experiment. The VR/AR environmental conditions lasted 5 min each. The physiological data regarding the HRV of the participants were collected continuously. The participants were given a 10 min rest between conditions. After both conditions were completed, they were asked to complete a survey in regard with technology acceptance [34]. The SD memory cards were then removed from the MyECG instruments worn by the participants and entered into the computer for statistical analysis.

In this experiment, both the VR and AR environments were comprised of an enclosed elevator. In the VR experiment, the participants were unable to move their body after entering the environment but were allowed to rotate their heads to change their viewing angles. In the AR environment condition, to be consistent with the circumstances of the VR environment, the participants were asked to stand in the center of the elevator and to avoid making movements besides rotating their heads to view their surroundings. Baseline time was recorded as the time from the start of the experiment to the start of the first anxiety-inducing event. An equivalent baseline time was in



(a) (b)

Figure 3: Scene with flames.

both conditions. Table 1 corresponds the experiment time elapsed with event occurrences.

2.3. Measurements. Heart rate variability (HRV) was analysed using ECGs. The analyses were divided into time domain and frequency domain. Time-domain analysis [35] uses ECG records over 24 h as the baseline data, detecting the gaps between each QRS complex wave in a continuous ECG. Adjacent R waves represent the cycles of heartbeats (i.e., the gaps or intervals between R waves (R-R)). The continuous gaps formed by consecutive R-R intervals represent HRV, defined as a normal-to-normal (NN) interval. Commonly used time-domain analyses include the following: standard deviation of the NN interval (SDNN), standard deviation of the averages of NN (SDANN), SDNN index, root mean square of successive NN interval differences (RMSSD), the percentage value of NN20 count (pNN20), and the percentage value of the NN50 count (pNN50). Because the lengths of the VR and AR environments were approximately 5 min, the researchers selected SDNN, RMSSD, pNN20, and pNN50 as the bases for comparison due to their correlation to short-term variability. The most commonly used calculation for frequency-domain analysis [35] is the fast Fourier transform (FFT), which analyzes the distribution of powers at different frequencies. Common frequency-domain analyses include total power (TP), high frequency (HF), low frequency (LF), very low frequency (VLF), ultralow frequency (ULF), normalized LF (nLF), and normalized HF (nHF). The biggest indicators of emotional influence are LF and HF. At a HF, TP of a heartbeat is subject to greater influence from the parasympathetic nervous system. Although the activity of the sympathetic nervous system increases at LF, the parasympathetic nervous system must synchronously adjust to suppress excessive excitement in the sympathetic nervous system and achieve a balanced state. Thus, LF is not necessarily directly correlated to the sympathetic nervous system. When the autonomic nervous system encounters stress from nervous emotions, activity in the sympathetic nervous system increases, whereas the opposite occurs for the parasympathetic nervous system. Therefore, this study utilized HF and log of nHF (LnHF) as the primary indicators for observation, with LnLF values serving as a supplement.

Table 1: Script of events.

Time elapsed for the game		Event occurred
Time 0	Baseline	Start of first stimulus
Time 1	After 30 s	Elevator door opens
Time 2	After 45 s	Elevator door closes
Time 3	After 55 s	Lights flash and alarms activate within the elevator
Time 4	After 1 min 40 s	Lights turn out, and evacuation sounds can be heard
Time 5	After 2 min 5 s	Collision sound and heartbeating sound can be heard
Time 6	After 3 min 40 s	Flames rise, and screaming sounds can be heard
Time 7	5 min	Simulation ends

To assess whether anxiety was triggered for the participants in the VR and AR environments, a questionnaire was used to measure their degree of anxiety. This questionnaire was based on the State-Trait Anxiety Inventory and modified to utilize a 5-point scale for measuring the degree of the anxiety experienced. This scale is typically used to measure anxiety in adults. The questionnaires were divided into the VR and AR environment sections, and the participants were asked about the degree of anxiety they experienced in the VR and AR experiments. More specifically, items in questionnaires were in regard to each stimulus event (Table 1) associated with the degree of anxiety, respectively.

In addition, the technology acceptance model (TAM) [34] was used to evaluate the behavioral intentions for executing the behavior while an individual engages in a specific behavior. The TAM shows that the perceived usefulness and perceived ease of use of information technology are the two primary deciding factors for behavioral intention to use. Additionally, the perceived ease of use has a direct influence on perceived usefulness, thereby indirectly influencing intention to use. After the participants completed the task, they were asked to complete the 5-point survey items regarding their acceptance of this technology in four dimensions: awareness and presence, usefulness, ease of use, and playfulness. These dimensions represented the degree of realism of the game environment, whether the game environment was able to induce feelings of anxiety in specific scenarios, the ease of use of the game controls and the physiological

feedback system, the entertainment value provided by the game, and the curiosity of the participants regarding this technology. The TAM was applied one time per participant disregarding the AR or VR system because we assumed AR and VR both laid on the same track of technology from the perspective of users.

We measured the correlation of the collected survey questionnaires using a *t*-test and performed a paired sample analysis and an independent analysis. The statistical analysis was conducted by the tool of SPSS™.

3. Results

3.1. HRV. For HRV, we divided the measured data into three main parts: data from 1 h before the experiment and data obtained during the VR and AR experiments. To compare the effectiveness of the VR and AR environments in triggering fear, we compared and analysed the normal HRV physiological data with those obtained from the VR and AR experiments. The results are shown in Tables 2 and 3. The VR and AR data were compared with the normal data, as shown in Tables 4 and 5. Finally, the VR and AR conditions' results were compared with each other, as shown in Table 6.

A comparison of the participants' HRV under normal circumstances versus HRV in the two conditions can be seen in Table 2. In the time-domain analysis, both short-term indicators (i.e., SDNN and pNN20) exhibited significant differences. In the frequency analysis, significant differences appeared in the LnHF indicator, and the other indicators approached significance. The lack of significance was likely because of an insufficient sample size; however, a descending trend was observed.

In a comparison of the means and standard deviations of the various values (Table 3), the AR conditions values were lower than normal. Because of the small sample size, only SDNN exhibited a significant decrease (Table 5). However, decreasing trends can be observed indicating that the participants were in a nervous emotional state. In the VR condition, the SDNN was again the only indicator that was significantly lower than normal (Table 4). The other indicator values were nearly equal to those in the normal state, sometimes even higher, implying that although the participants from the two conditions experienced nervousness during the VR condition, their reactions were not as pronounced as in the AR condition. As shown in Table 7, the study extracted the physiological data from the VR and AR conditions for comparison, finding that nearly all indicator values were significantly lower in the AR environment experiment than in the VR condition. This shows that anxiety was experienced more strongly by the participants in the AR condition than in the VR condition.

3.2. Anxiety Questionnaire. We measured the correlation of the collected survey questionnaires using a *t*-test and performed a paired sample analysis and an independent analysis (Table 8). The paired sample analysis compared the average degree of anxiety between the AR and VR environments for the 30 participants. The participants indicated that they experienced significantly greater anxiety in the VR environment than in the AR environment (Table 8).

3.3. TAM. In terms of the survey results for TAM (Table 6), the satisfaction was above neutral (3 points).

4. Discussion

For the HRV, as shown in the data in Tables 4–8, the AR condition generated better results than the VR condition. This finding may be due to that the participants in the AR condition were physically present in an actual environment, causing them to experience a more natural presence and become more engaged. The effects of the VR condition were inconspicuous, possibly because the participants did not suffer from claustrophobia and therefore showed less-pronounced reactions. Although the HMDs achieved immersive surround effects, previous studies [6, 7] have shown that HMDs are limited in their ability to achieve surround effects compared to other surround displays, such as CAVE displays, resulting in a poorer sense of presence and posing challenges for participants in immersing themselves in the virtual environment.

Regarding the anxiety questionnaire, the participants suggested that their anxiety was significantly higher in the VR condition compared to the AR condition. This finding substantially differed from the HRV physiological data measured. Practically speaking, however, these results are not improbable. Anxiety involves both psychological and physiological factors. Although participants may not have felt psychologically nervous, they may have experienced physiological reactions in response to the stimuli. Physiological signals are more objective data and were synchronously measured during the experiment. By contrast, the questionnaire responses were subjective, naturally creating a possible discrepancy. Nevertheless, the two sets of results are not necessarily contradictory. Because the participants recruited for this study did not suffer from claustrophobia, the simple act of entering an elevator scenario was not likely to cause substantial subjective emotional fluctuation. Thus, the participants may have overlooked their own anxious emotions.

Although the experiment utilized sufficient display equipment and used HMDs to create surround visuals and immersive effects, the overall average score of presence was only 3.50. This result was similar to that achieved by Juan and Pérez in a study comparing the presence and degree of anxiety induced by HMD and CAVE [29] devices. Their study showed that the presence of HMD was a mean of 3.59 (out of a maximum of seven points), slightly higher than that of normal circumstances. Usefulness was the second-highest scoring item in the TAM survey (mean = 3.85). The participants exhibited a positive attitude toward the use of the HMD, suggesting that the HMD helped them perceive their correct positions in the virtual environment, increasing the quality and effects of the VR task. In contrast, ease of use was the lowest-scoring item (mean 3.38). The researchers inferred that the HMD and HRV instruments were relatively

Table 2: One-way ANOVA of the HRV physiological data values.

		Sum of squares	Degree of freedom	Mean sum of squares	F	Significance
SDNN	Between-group	9123.245	2	4561.622	12.253	0.000
	Within-group	32389.664	87	372.295		
	Overall	41512.909	89			
RMSSD	Between-group	1298.241	2	649.120	2.559	0.083
	Within-group	22067.628	87	253.651		
	Overall	23365.869	89			
pNN20	Between-group	2518.745	2	1259.372	6.138	0.003
	Within-group	17851.371	87	205.188		
	Overall	20370.116	89			
pNN50	Between-group	897.829	2	448.914	2.563	0.083
	Within-group	15236.540	87	175.133		
	Overall	16134.369	89			
HF	Between-group	11054376.62	2	5527188.311	2.601	0.080
	Within-group	184856864.5	87	2124791.546		
	Overall	195911241.1	89			
LnHF	Between-group	3.891	2	1.945	3.478	0.035
	Within-group	48.662	87	0.559		
	Overall	52.553	89			
LnLF	Between-group	1.244	2	0.622	2.457	0.092
	Within-group	22.016	87	0.253		
	Overall	23.259	89			

Table 3: Descriptive statistics for the HRV physiological data between groups.

		Quantity	Mean	SD
SDNN	Normal	30	80.81	24.90
	AR	30	56.19	12.23
	VR	30	67.21	18.64
RMSSD	Normal	30	48.19	14.29
	AR	30	42.82	12.99
	VR	30	52.08	19.70
pNN20	Normal	30	47.90	12.95
	AR	30	42.67	15.30
	VR	30	55.55	14.62
pNN50	Normal	30	21.51	11.53
	AR	30	18.45	12.01
	VR	30	26.13	15.76
HF	Normal	30	2025.77	1330.49
	AR	30	1536.63	1285.64
	VR	30	2392.17	1717.94
LnHF	Normal	30	7.40	0.70
	AR	30	7.04	0.79
	VR	30	7.53	0.75
LnLF	Normal	30	7.77	0.44
	AR	30	7.49	0.47
	VR	30	7.65	0.59

Table 4: Comparison of HRV physiological data between VR and normal circumstances.

	Quantity	t-value	Significance (two-tailed)
SDNN	30	2.396	0.020
RMSSD	30	−0.875	0.385
pNN20	30	−2.144	0.036
pNN50	30	−1.298	0.199
HF	30	−0.924	0.360
LnHF	30	−0.706	0.483
LnLF	30	0.935	0.353

Table 5: Comparison of HRV physiological data between AR and normal circumstances.

	Quantity	t-value	Significance (two-tailed)
SDNN	30	4.861	0.000
RMSSD	30	1.525	0.133
pNN20	30	1.431	0.158
pNN50	30	1.006	0.319
HF	30	1.448	0.153
LnHF	30	1.866	0.067
LnLF	30	2.452	0.017

unfamiliar to the participants, requiring instructions regarding use. After instructions were provided, the participants quickly learned how to use the equipment, requiring only one round of operation. Finally, the average overall score for playfulness was 3.86, the highest scoring item for TAM in this experiment.

In this experiment, the researchers identified a number of challenges regarding the VR and AR conditions. Regarding the display equipment, the VR conditions required better audiovisual effects to create an immersive experience (as with the CAVE display device). In contrast, the AR conditions required integration with their surroundings, requiring consideration of equipment for sound and light effects, as well as an emphasis on maneuverability. In addition, regarding AR conditions, participants must be physically present in the environment, leading to differences in presence when compared to VR conditions. Although the two conditions utilized the same equipment and stimuli, other factors may influence the experience of users, such as changes in

Table 6: TAM results.

	Awareness + presence	Usefulness	Ease of use	Playfulness
Mean score	3.50	3.86	3.38	3.56

Table 7: Comparison of HRV physiological data between VR and AR.

	Quantity	t-value	Significance (two-tailed)
SDNN	30	−2.706	0.009
RMSSD	30	−2.150	0.036
pNN20	30	−3.335	0.001
pNN50	30	−2.124	0.038
HF	30	−2.184	0.033
LnHF	30	−2.467	0.017
LnLF	30	−1.176	0.244

Table 8: Paired sample analysis for anxiety.

Group	Sample size	Mean	SD	t-value	Significance
AR mean	30	3.16	0.55	−4.29	**
VR mean	30	3.58	0.48		

Significance level = 0.05; **indicates $P < 0.01$.

temperature, standing posture, and external noises beyond the control of the researchers. Assessing the presence of AR environments may require a different approach than that for VR environments to allow participants to adequately evaluate their experiences.

In summary, the researchers found that, in this experiment, the AR environment produced a statistically significantly stronger experience for the participants and caused statistically significant physiological reactions than those caused by the VR environment. However, in clinical therapy for claustrophobia, AR environment experiments are more difficult to construct than are VR environments. Furthermore, patients have a lower degree of acceptance during exposure therapy. Therefore, in conjunction with developing AR-based therapy, improving VR display equipment to provide greater presence may help induce the anxiety associated with enclosed spaces and therefore an opportunity to provide an intervention for claustrophobia.

5. Conclusions

This study successfully developed virtual reality (VR) and augmented reality (AR) environments for claustrophobia treatments. A test was conducted to validate these two systems and examine the effect between these two systems using HRV and anxiety questionnaires. HRV results indicated that the proposed VR system and AR system were both able to trigger anxiety. Furthermore, the AR environment produced a stronger experience for the participants and caused statistically significant physiological reactions than those caused by the VR environment. Regarding the anxiety questionnaire, the participants suggested that their anxiety was significantly higher in the VR environment than those in the AR environment. In the future, a large-scale clinical test is planned to further verify the therapeutic effect of the proposed systems.

Data Availability

The heart rate variability data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Tianqiao and Chrissy Chen Institute for Brain Diseases in Shanghai (grant number: 201802).

References

[1] L.-G. Öst, "The claustrophobia scale: a psychometric evaluation," *Behaviour Research and Therapy*, vol. 45, no. 5, pp. 1053–1064, 2007.

[2] S. F. Lourenco, M. R. Longo, and T. Pathman, "Near space and its relation to claustrophobic fear," *Cognition*, vol. 119, no. 3, pp. 448–453, 2011.

[3] L. M. Harris, J. Robinson, and R. G. Menzies, "Evidence for fear of restriction and fear of suffocation as components of claustrophobia," *Behaviour Research and Therapy*, vol. 37, no. 2, pp. 155–159, 1999.

[4] S. Rachman and S. Taylor, "Analyses of claustrophobia," *Journal of Anxiety Disorders*, vol. 7, no. 4, pp. 281–291, 1993.

[5] F. Mehdi, S. Farzeen, and S. S. Jain, "Healing phobias using augmented reality," *Journal on Science Engineering and Technology*, vol. 2, no. 4, pp. 292–296, 2015.

[6] Y. Choy, A. J. Fyer, and J. D. Lipsitz, "Treatment of specific phobia in adults," *Clinical Psychology Review*, vol. 27, no. 3, pp. 266–286, 2007.

[7] R. Booth and S. Rachman, "The reduction of claustrophobia—I," *Behaviour Research and Therapy*, vol. 30, no. 3, pp. 207–221, 1992.

[8] R. Shafran, R. Booth, and S. Rachman, "The reduction of claustrophobia—II: cognitive analyses," *Behaviour Research and Therapy*, vol. 31, no. 1, pp. 75–85, 1993.

[9] G. W. Alpers and R. Sell, "And yet they correlate: psychophysiological activation predicts self-report outcomes of exposure therapy in claustrophobia," *Journal of Anxiety Disorders*, vol. 22, no. 7, pp. 1101–1109, 2008.

[10] M. M. Antony, R. E. McCabe, I. Leeuw, N. Sano, and R. P. Swinson, "Effect of distraction and coping style on in vivo exposure for specific phobia of spiders," *Behaviour Research and Therapy*, vol. 39, no. 10, pp. 1137–1150, 2001.

[11] F. D. McGlynn, P. M. Moore, S. Lawyer, and R. Karg, "Relaxation training inhibits fear and arousal during in vivo exposure to phobia-cue stimuli," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 30, no. 3, pp. 155–168, 1999.

[12] I. Kirsch, H. Tennen, C. Wickless, A. J. Saccone, and S. Cody, "The role of expectancy in fear reduction," *Behavior Therapy*, vol. 14, no. 4, pp. 520–533, 1983.

[13] J. Cottraux, "Combining psychological and pharmacological treatment for specific phobias," *Psychiatry*, vol. 3, no. 6, pp. 87–89, 2004.

[14] L.-G. Öst, J. Johansson, and A. Jerremalm, "Individual response patterns and the effects of different behavioral methods in the treatment of claustrophobia," *Behaviour Research and Therapy*, vol. 20, no. 5, pp. 445–460, 1982.

[15] K. B. Wolitzky-Taylor, J. D. Horowitz, M. B. Powers, and M. J. Telch, "Psychological approaches in the treatment of specific phobias: a meta-analysis," *Clinical Psychology Review*, vol. 28, no. 6, pp. 1021–1037, 2008.

[16] G. C. L. Davey, "Psychopathology and treatment of specific phobias," *Psychiatry*, vol. 6, no. 6, pp. 247–253, 2007.

[17] I. E. Sutherland, "The ultimate display: information processing techniques," in *Proceedings of IFIP Congress 1965*, pp. 506–508, New York, NY, USA, May 1965.

[18] G. Burdea and P. Coiffet, *Virtual Reality Technology*, John Wiley & Sons, New York, NY, USA, 1994.

[19] C. Botella, R. M. Baño, H. Villa, C. Perpiñá, and A. García-Palacios, "Virtual reality in the treatment of claustrophobic fear: a controlled, multiple-baseline design," *Behavior Therapy*, vol. 31, no. 3, pp. 583–595, 2000.

[20] C. Botella, R. M. Baños, C. Perpiñá, H. Villa, M. Alcañiz, and A. Rey, "Virtual reality treatment of claustrophobia: a case report," *Behaviour Research and Therapy*, vol. 36, no. 2, pp. 239–246, 1998.

[21] M. Hunt and M. Fenton, "Imagery rescripting versus in vivo exposure in the treatment of snake fear," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 38, no. 4, pp. 329–344, 2007.

[22] M. Hunt, L. Bylsma, J. Brock et al., "The role of imagery in the maintenance and treatment of snake fear," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 37, no. 4, pp. 283–298, 2006.

[23] E. B. Foa and M. J. Kozak, "Emotional processing of fear: exposure to corrective information," *Psychological Bulletin*, vol. 99, no. 1, pp. 20–35, 1986.

[24] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. A. Furness, and C. Botella, "Virtual reality in the treatment of spider phobia: a controlled study," *Behaviour Research and Therapy*, vol. 40, no. 9, pp. 983–993, 2002.

[25] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[26] P. Milgram and F. Kishino, "A taxonomy of mixed reality visual displays," *IEICE Transactions on Information and Systems*, vol. 77, no. 12, pp. 1321–1329, 1994.

[27] K. B. Wolitzky and M. J. Telch, "Augmenting in vivo exposure with fear antagonistic actions: a preliminary test," *Behavior Therapy*, vol. 40, no. 1, pp. 57–71, 2009.

[28] C. Botella, J. Breton-López, S. Quero et al., "Treating cockroach phobia using a serious game on a mobile phone and augmented reality exposure: a single case study," *Computers in Human Behavior*, vol. 27, no. 1, pp. 217–227, 2011.

[29] M. C. Juan and D. Pérez, "Using augmented and virtual reality for the development of acrophobic scenarios. Comparison of the levels of presence and anxiety," *Computers and Graphics*, vol. 34, no. 6, pp. 756–766, 2010.

[30] C. Botella, M. Juan, R. M. Baños, M. Alcañiz, V. Guillén, and B. Rey, "Mixing realities? An application of augmented reality for the treatment of cockroach phobia," *Cyberpsychology and Behavior*, vol. 8, no. 2, pp. 162–171, 2005.

[31] M. C. Juan and D. Pérez, "Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE," *Presence: Teleoperators and Virtual Environment*, vol. 18, no. 3, pp. 232–248, 2009.

[32] S.-C. Yeh, Y.-Y. Li, C. Zhou, P.-H. Chiu, and J.-W. Chen, "Effects of virtual reality and augmented reality on induced anxiety," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, pp. 1345–1352, 2018.

[33] S.-C. Yeh and J.-F. Tsai, "A comparison between virtual reality and augmented reality for anxiety disorder treatment using heart rate variability," in *Proceedings of International Conference on Electrical Engineering and Information Technology*, 2014.

[34] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: four longitudinal field studies," *Management Science*, vol. 46, no. 2, pp. 186–204, 2000.

[35] T. F. ESC/NASPE, "Heart rate variability,stabdards of measurement,physiological interpretation,and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

Hindawi
Journal of Healthcare Engineering
Volume 2018, Article ID 5435097, 12 pages
https://doi.org/10.1155/2018/5435097

Research Article

How to Build a Patient-Specific Hybrid Simulator for Orthopaedic Open Surgery: Benefits and Limits of Mixed-Reality Using the Microsoft HoloLens

Sara Condino,[1,2] **Giuseppe Turini**,[3] **Paolo D. Parchi**,[1,4] **Rosanna M. Viglialoro**,[1] **Nicola Piolanti**,[4] **Marco Gesi**,[5,6] **Mauro Ferrari**,[1] **and Vincenzo Ferrari**[1,2]

[1]*EndoCAS Center, Department of Translational Research and of New Surgical and Medical Technologies, University of Pisa, Pisa, PI 56124, Italy*
[2]*Department of Information Engineering, University of Pisa, Pisa, PI 56124, Italy*
[3]*Department of Computer Science Department, Kettering University, Flint, MI 48504, USA*
[4]*1st Orthopaedic and Traumatology Division, Department of Translational Research and of New Surgical and Medical Technologies, University of Pisa, Pisa, PI 56124, Italy*
[5]*Department of Translational Research and of New Surgical and Medical Technologies, University of Pisa, Pisa, PI 56124, Italy*
[6]*Center for Rehabilitative Medicine "Sport and Anatomy", University of Pisa, Pisa, PI 56124, Italy*

Correspondence should be addressed to Sara Condino; sara.condino@endocas.org

Received 8 May 2018; Accepted 30 September 2018; Published 1 November 2018

Academic Editor: Olivier Beuf

Orthopaedic simulators are popular in innovative surgical training programs, where trainees gain procedural experience in a safe and controlled environment. Recent studies suggest that an ideal simulator should combine haptic, visual, and audio technology to create an immersive training environment. This article explores the potentialities of mixed-reality using the HoloLens to develop a hybrid training system for orthopaedic open surgery. Hip arthroplasty, one of the most common orthopaedic procedures, was chosen as a benchmark to evaluate the proposed system. Patient-specific anatomical 3D models were extracted from a patient computed tomography to implement the virtual content and to fabricate the physical components of the simulator. Rapid prototyping was used to create synthetic bones. The Vuforia SDK was utilized to register virtual and physical contents. The Unity3D game engine was employed to develop the software allowing interactions with the virtual content using head movements, gestures, and voice commands. Quantitative tests were performed to estimate the accuracy of the system by evaluating the perceived position of augmented reality targets. Mean and maximum errors matched the requirements of the target application. Qualitative tests were carried out to evaluate workload and usability of the HoloLens for our orthopaedic simulator, considering visual and audio perception and interaction and ergonomics issues. The perceived overall workload was low, and the self-assessed performance was considered satisfactory. Visual and audio perception and gesture and voice interactions obtained a positive feedback. Postural discomfort and visual fatigue obtained a nonnegative evaluation for a simulation session of 40 minutes. These results encourage using mixed-reality to implement a hybrid simulator for orthopaedic open surgery. An optimal design of the simulation tasks and equipment setup is required to minimize the user discomfort. Future works will include Face Validity, Content Validity, and Construct Validity to complete the assessment of the hip arthroplasty simulator.

1. Introduction

Surgical simulation, a key enabling technique to revolutionize patient care and patient safety, can provide a standardized method for surgical training without the risks that come with operating on real patients [1].

Orthopaedic simulation has generally lagged behind other specialties, with fewer validated simulators available; this trend is now changing and recent studies support the notion that orthopaedic simulators have the potential to translate useful technical skills into the operating theatre [2].

Several techniques of simulation are available today, including virtual reality (VR) simulation, physical simulation, and hybrid (virtual-physical) simulation.

Existing VR orthopaedic simulators are limited by a poor haptic feedback. One of the major issues to be addressed is the simplification of the computational models to speed up the interactive simulation without compromising the effective realism of the tissue response [3]. Moreover, conventional haptic interfaces are limited in the magnitude of the forces being rendered, so they do not enable a realistic simulation of the surgical instruments/bone interaction, particularly in open surgery where the interaction forces can be of considerable magnitudes. This could explain why, in a recent study [2], Morgan et al. found that commercially available VR simulators are mainly focused on arthroscopy, a minimally invasive procedure.

As for physical simulation, companies like Sawbones [4] offer orthopaedic training models for open surgery procedures such as joint replacement surgery. The strength of these simulators lies in the realism of the synthetic bone, which requires no special handling or preservation and exhibits mechanical properties similar to human bone [5–7]. This is very important for a good simulation experience to allow the surgeon to develop a force-feedback memory, which is crucial for the success of a surgical procedure including tasks such as bone drilling. However, standard commercial mannequins lack objective assessment of performance and cover a very limited range of individual differences and pathologies. Patient-specific simulation, a new frontier that promises great benefits for surgical training and rehearsal [8–10], can overcome this latter limitation.

As suggested by a literature review on orthopaedic surgery simulation [11], *"an ideal simulator should be multimodal, combining haptic, visual and audio technology to create an immersive training environment."* Hybrid simulation technologies, which combine VR with physical models of the anatomy, are the best candidate to meet these requirements. Hybrid systems indeed have the advantages of physical simulators, which can mimic the properties of human tissue [12–14] offering the trainee the possibility to use actual surgical instruments and experience a realistic haptic feedback; and, at the same time, they exploit the benefit of computer visualization and simulation, offering also objective tools for assessing the surgical performance. Moreover, augmented reality (AR) elements can be added to enrich the synthetic environment, to make hidden structures visible, and to present additional information for the surgical tasks guidance [10, 15–19]. Finally, spatial sound can be added in AR applications to improve the realism of the simulated scenario.

Available display technologies for AR include spatial displays (screen-based and projection-based); hand-held displays (such as phones and tablets); and head-mounted displays (HMDs). HMDs are deemed as the most ergonomic solution for applications including manual tasks performed by the user under direct vision, like what happens in open surgery. HMDs indeed intrinsically provide the user with an egocentric viewpoint and they allow the user to work handsfree [20].

This work explores the potentialities offered by mixed-reality (MR) using the HoloLens [21], an head-mounted display designed by Microsoft for MR applications, to develop an hybrid training system with immersive and interactive content.

Hip arthroplasty (HA), which involves replacing a damaged hip joint with a prosthetic implant, was chosen as a benchmark to evaluate the benefits/limits of the proposed system because it is one of the most widely performed procedures in orthopaedic practice [22], and there is a gap in the market for a high-fidelity hip replacement training simulators [11].

In a previous work [23], we have presented a lower torso phantom for HA including a patient-specific hemi-pelvis replica embedded in a soft synthetic foam. In this paper, we present the HipSim app: an evolution of our former simulator, focusing on the details for the implementation of wearable AR functionalities using the HoloLens. Quantitative and qualitative test were carried out to perform a preliminary evaluation of our multimodal surgical simulator and to explore advantages and limits of the new design and novel technologies being used.

2. Materials and Methods

The following paragraphs describe the peculiarities of the adopted HMD; the virtual content and the physical components of the simulator, with details on the implementation/fabrication strategy; the calibration and registration methods to align the VR content with the physical word; and the testing strategies to preliminary validate the simulator.

2.1. Selection of the Head-Mounted Display. HoloLens is an Optical See-Through (OST) HMD, which enables optical superposition of virtual content onto the user direct view of the physical world. Being an OST system, it offers an unhindered and instantaneous full-resolution view of the real environment which assures that visual and proprioception information is synchronized [24].

Differently, in Video See-Through (VST) HMDs, the virtual content is merged with the camera images captured by one or two external cameras rigidly fixed on the visor frame. This more obtrusive technology block out the real-world view in exchange for the ability to offer higher geometric coherence between virtual and real content, without requiring a user-specific calibration eye-to-display [25]. A complete comparison of OST and VST technologies is reported in [26].

Assuming that for simulation purposes the perceived positioning accuracy of the VR content is not as important as the possibility to give the user a naturalistic experience, we have opted for an OST system. More in particular, the HoloLens was chosen for our application since it provides significant benefits over other commercial HMD from human factors and ergonomics standpoints [27] and integrates important functionalities for an immersive and interactive simulation experience. In fact, the HoloLens offers head tracking, hand gesture controls, and voice commands and

enables binaural audio to simulate effects such as spatial sound within the user environment. Additionally, HoloLens has no physical tethering constrains that can limit the movements/gestures of the user during the simulation of the surgical tasks.

A recent literature study on the evaluation of OST-HMD suitability for mixed-reality surgical intervention [28] shows that Microsoft HoloLens outperforms other currently available OST HMDs (Epson Moverio BT-200, ODG R-7), in terms of contrast perception, task load, and frame rate. The same study shows that the integration of indoor localization and tracking functionalities, enabled by HoloLens environmental understanding sensors, provides significantly less system lag in a relatively motionless scenario.

For all these reasons, HoloLens can be considered a good candidate for the implementation of mixed-reality open surgery simulators. However, some well-known technical issues of HDMs have to be considered, such as a small overlay field of view (FOV); the vergence-accommodation conflict (VAC) [29]; the perceptual issues, intrinsic to standard optical see-through HMDs, due to mismatched accommodation between the virtual content and the real-world scene [30]; and the difficulties of OST systems in handling occlusion between the real and virtual contents [26].

The overlay FOV can be defined as "the region of the field of view where graphical information and real information are superimposed" [26] which, in the HoloLens, is about 35°.

As for the vergence-accommodation conflict, users wearing HoloLens are forced to accommodate their eyes to a fixed focal distance of approximately 2.0 m (Figure 1) to maintain a clear image of the virtual content, while the depth of the virtual objects (and hence the binocular disparity) varies depending on the application. This results in conflicting information within the vergence-accommodation feedback loops causing visual discomfort [30].

Moreover, the focal distance of each physical object in the real world depends on its relative distance from the user: if the distance gap between the display focal plane and real-world objects is beyond the human eye deep of field, the user cannot keep in focus both the virtual and real content at the same time [20].

The discomfort due to the vergence-accommodation conflict can be reduced by keeping the virtual content positioning stable over the time [31, 32]. However, the mismatch between the focal distances of real and virtual objects, together with the difficulties in handling the occlusions of overlapping objects, can affect the accuracy of the rendered depth [26].

For this reason, quantitative and qualitative tests were performed to evaluate if the perceived positioning accuracy matches the requirements of the target application. Moreover, qualitative tests were also performed to evaluate the visual discomfort and the usability of the proposed HDM for our specific scenario: orthopaedic open surgery simulation.

2.2. Design and Implementation of the Simulator Components: The Virtual Content. The development of the simulator starts from the segmentation and surface extraction of the anatomical organs of interest from a real CT dataset (Figure 2). The stack of medical images in DICOM format is processed using a semiautomatic tool, the EndoCAS Segmentation Pipeline [33] integrated in the open source software ITK-SNAP 1.5 [34]. Then, mesh reconstruction and optimization (artefacts removal, holes filling, simplification, and filtering) stages are performed to generate the 3D models of the patient anatomy necessary for the surgical simulation. Optimization stages are performed using the open source software MeshLab [35] and Blender [36]. The bone models included in the present version of the simulator are: hip bones, sacrum, coccyx, and femoral heads. Moreover, a model of pelvis and the principal muscles around the hip joint (such as *gluteal muscles*, *piriformis*, *inferior gemellus*, *superior gemellus*, *obturator internus*) are included to increase the anatomical knowledge of the user-trainee and form a solid basis for a complete surgical simulation system. Other key surgical structures to be added for further improving the simulation are fasciae, nerves, tendons, and blood vessels.

Finally, the virtual environment is enriched with information from a simulated planning phase with the 3D Hip Plugin [23]: a pair of viewfinders and a dotted line are added to the virtual anatomical model to show the surgeon the optimal trajectory for the reaming tool. This information, coupled with the real-time tracking of the surgical instrument, could also be used for a quantitative evaluation of the surgical performance on the basis of the deviation of the reaming tool from the optimal trajectory.

Moreover, a selection of radiological images (a hip radiograph, a CT slice, a CT volumetric rendering) (Figure 3) is added to the virtual content enriching the digital information available to the learner during the simulation.

2.3. Design and Implementation of the Simulator Components: The Physical Components. The development of the physical simulator starts from the CAD design (Figure 4). 3D virtual models are imported in the Creo Parametric 3D Modelling software, and each physical component is designed, including a support for the registration target (an Image Target as described in the following section). This support is rigidly anchored to the bone synthetic replica to guarantee a precise registration of the virtual content to the real scene.

A 3D printer (Dimension Elite 3D Printer) is used to turn the 3D CAD models into tangible 3D synthetic replicas made of acrylonitrile butadiene styrene (ABS). This plastic is commonly used for the manufacturing of bone replicas for orthopaedic surgery simulation since it adequately approximates the mechanical behaviour of the natural tissue [37]. Finally, silicone mixtures and polyurethane materials are used for the manufacturing of the soft parts.

The final mannequin includes a replica of the acetabulum embedded in a soft synthetic foam. Moreover, a skin-like covering is provided for an accurate simulation of palpation and surgical incision.

2.4. Calibration and Registration of the Virtual and Physical Content. Display-eye calibration and registration should be

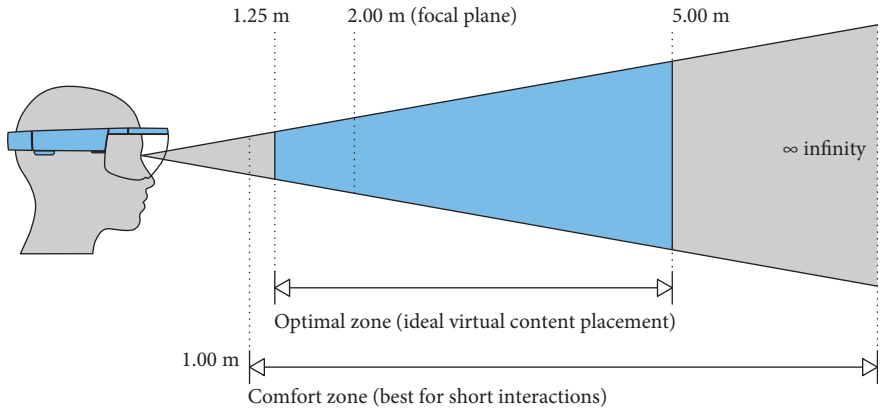


Figure 1: Optimal and comfort zones for placing virtual content as declared by Microsoft for HoloLens mixed-reality applications. Discomfort from the vergence-accommodation conflict can be avoided or minimized by keeping content that users converge to as close to 2.0 m as possible. When the content cannot be placed near 2.0 m, the discomfort can be reduced by keeping the virtual content positioning stable over the time.

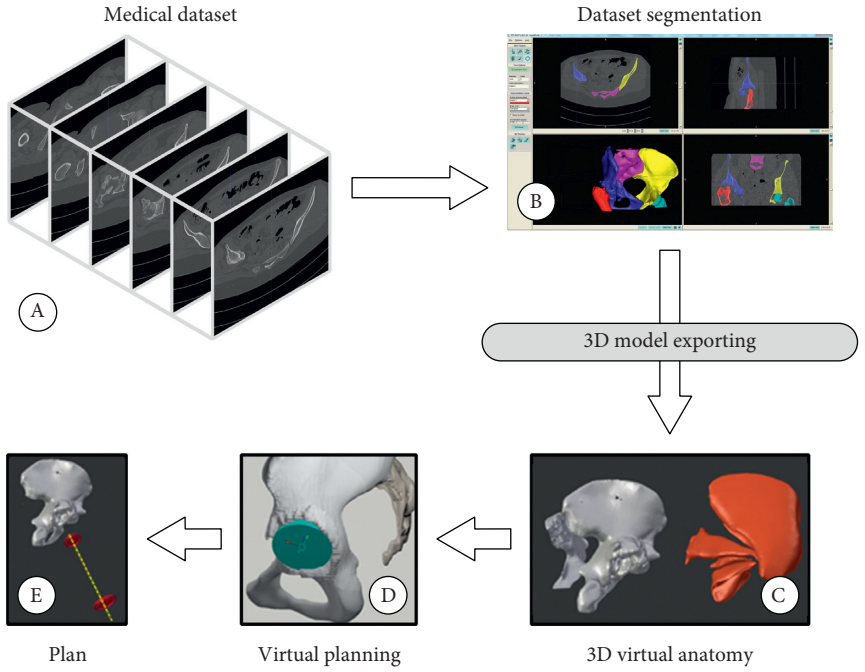
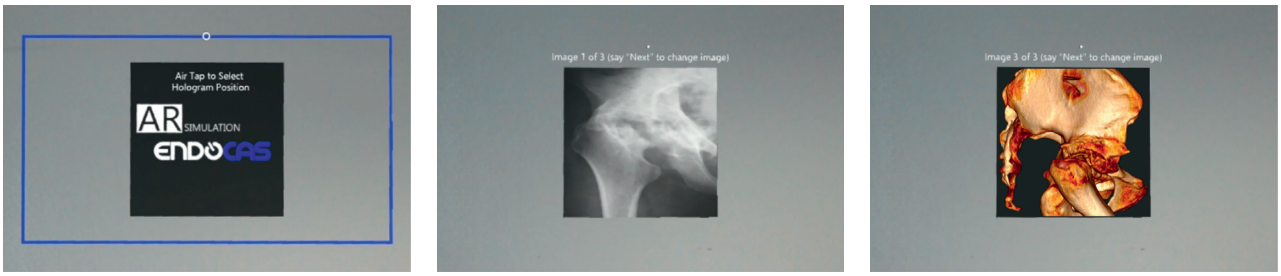


Figure 2: Schematic representation of the steps involved in the generation of the simulator virtual content: (A) the medical dataset of the patient; (B) the segmentation process using ITK-SNAP and the EndoCAS Segmentation Pipeline; (C) the 3D virtual anatomy generated by exporting the 3D models; (D) the virtual planning including the positioning and sizing of the acetabular component; and (E) the final preoperative plan.



(a) (b) (c)

Figure 3: Example of AR images illustrating the medical image navigation: (a) first image presented at the beginning of the application and (b, c) two of the medical images in the collection that the user can visualize. The Air Tap gesture is used to anchor the position of medical image navigator in the physical space, whereas the voice command "Next" is used to switch the radiological images.

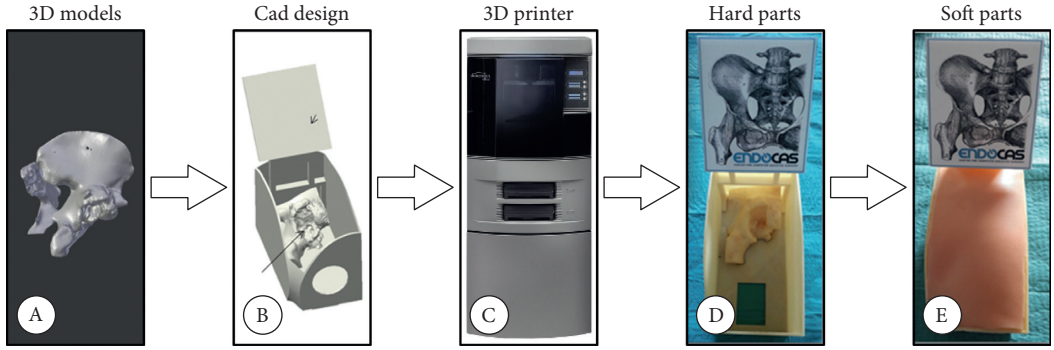


Figure 4: Schematic representation of the steps involved in the development and manufacturing of the physical components of the hybrid simulator: (A) the 3D model of the bones as generated from the CT dataset of the patient; (B) the CAD design for 3D printing, including the acetabulum and the support for the Image Target; (C) the 3D printer Dimension Elite; (D) and (E) the hard and soft components (respectively) of the hybrid simulator, including the Vuforia Image Target placed on top of an *ad hoc* support.

performed to properly align the virtual content with the real objects. The calibration procedure is necessary to model intrinsically and extrinsically the virtual viewing frustum to the user viewing volume. To perform this calibration, the Microsoft HoloLens includes an official "Calibration" app, which however does not offer a complete user-based calibration procedure, but it is designed to solely determine the interpupillary distance (IPD) [38].

The registration can be accomplished in real time by tracking the relative position and orientation of the real objects with respect to the rendering camera; this information is then used to update the corresponding transformations within the virtual world.

HoloLens includes a world-facing camera; thus, the optical detection and tracking of a target can be used for real-time registration purposes, with no need for an external tracking system. At this end, in our application, we use the detection and tracking functionalities offered by the Vuforia SDK [39].

More in particular, we employ an Image Target (Figure 5). Image Targets represent images that Vuforia Engine can detect and track at runtime. The Vuforia Engine detects and tracks the features that are naturally found in an image. These features, extracted from the original image, are stored in a preprocessed database, which can then be integrated in the software application. This database can then be used by Vuforia Engine for runtime comparisons. Once the Image Target is detected, Vuforia Engine will track it as long as it is at least partially visible by the camera. The fundamental attributes for an accurate tracking of an Image Target are good contrast, no repetitive patterns, and wealth of details. Moreover, for near-field applications, a physical printed Image Target should be at least 12 cm in width and of reasonable height [39]. For a more detailed definition of Vuforia Image Targets, please refer to the Vuforia SDK [39].

2.5. Implementation Details. From the software aspect, Unity3D (5.6.1f) was used to create the application (the HipSim app). The MixedRealityToolkit (2017.1.2), a collection of C# scripts and Unity components to develop mixed-reality applications, was utilized for the development of the surgical simulator. This toolkit allows the user to interact with the virtual content by means of head movements (Gaze), gestures (Air Tap, Bloom, etc.), and voice commands (via Cortana). A virtual cursor is added to the application to indicate the head/view direction: this interaction through head movements is called Gaze. The Gaze is estimated from the position and orientation of the user's head, without considering the user's eyes direction (since the current version of HoloLens does not include any eye-tracking sensor).

A Fitbox (a MixedRealityToolkit tool) is used in Unity to anchor in the physical space the virtual collection of radiological images according to the user preferences (Figure 3).

A virtual menu with multiple toggle buttons has been implemented to select the virtual components (pelvis, bones, and muscles; preoperative plan) to be visualized during each surgical task. Figure 6 shows examples of AR images captured by the HoloLens word-facing camera during a surgical simulation trial.

Operating room ambient sound, including voices of surgical staff and sounds of medical equipment, has been included in the HipSim app to improve the realism and immersion of the surgical simulation.

2.6. Quantitative Study. Quantitative tests were performed to estimate the accuracy of the system by evaluating the perceived position of AR targets.

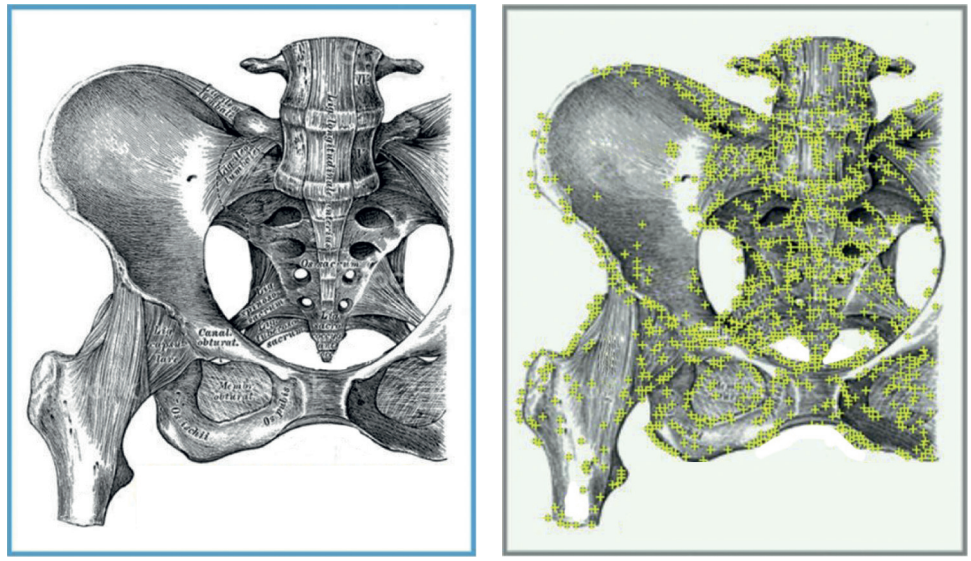
Five (5) subjects (gender: 2 males, 3 females, 0 nonbinary; years of age: 24 min, 32 mean, 39 max, 6 STD) with 10/10 vision were recruited to participate in this study. The HoloLens were used to present four (4) virtual targets consisting of red spheres (0.5 mm radius) virtually located on the *acetabulum* surface (Figure 7(a)). Targets were designed in the CAD environment and their 3D positions were acquired in the virtual environment reference frame.

Figure 7 shows the experimental setup consisting of:

(i) the Microsoft HoloLens HMD;

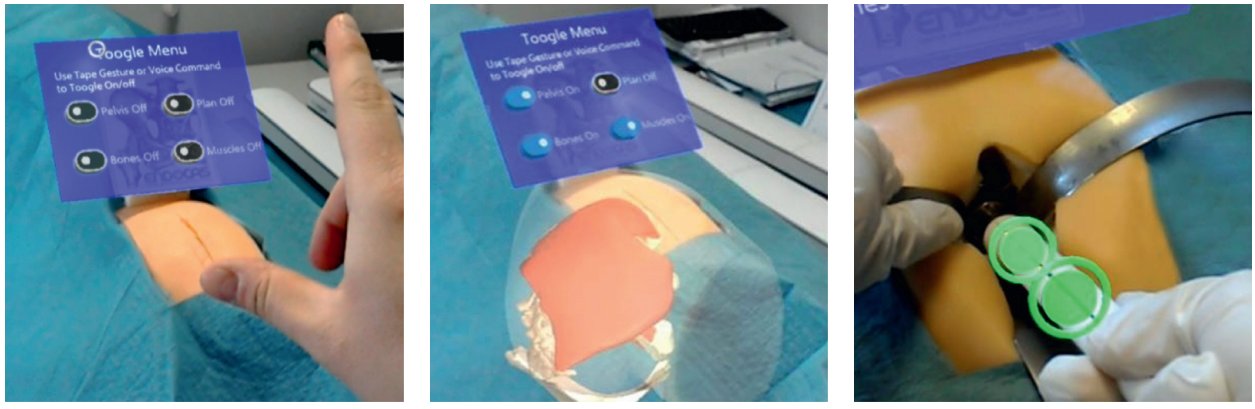
(ii) the rigid components of the mannequin, without the synthetic soft tissue;

(iii) the Vuforia SDK Image Target for tracking and registration;



(a) (b)

Figure 5: The designed Image Target that was printed with a size of 12 × 14 cm (a) and the image features detected by the Vuforia SDK (b). This Image Target obtained a 5/5-star rating: star rating defines how well an image can be detected and tracked using the Vuforia SDK, and this rating is displayed in the Target Manager and returned for each uploaded Image Target via the Vuforia web API.



(a) (b) (c)

Figure 6: Examples of AR images captured during the simulated surgical procedure: (a) the mannequin, positioned on a surgical table and covered with a surgical drape to enhance the realism of simulation, and the virtual AR menu for the selection of the virtual anatomical components to be visualized; (b) the surgeon can visualize in AR mode the virtual anatomy before performing the surgical incision; (c) with the help of the virtual viewfinder, the surgeon can orient the surgical instrument, so that the acetabulum reaming can proceed in the direction of the planned implant.

(iv) the NDI Aurora electromagnetic tracking system (V2 System); and

(v) the NDI Aurora calibrated 6 degrees of freedom (DOF) digitizer.

The mannequin and the Aurora EM emitter were fixed in a stable position to avoid relative movement during the targeting trials.

The rigid transformation ${}^{A}\mathbf{T}_{V}$ between the Aurora reference system and virtual environment reference frame was derived with a point-based registration algorithm: the positions of three landmarks (three corners of the simulator) were acquired in the CAD environment; the positions of the same landmarks were then acquired in the Aurora reference system with the digitizer; and then the transformation was derived with a least-squares error minimization algorithm [40]. Finally, the root mean squared registration error (RMSE) and the maximum registration error (MR) were computed and saved.

The official HoloLens app was used to calibrate the HMD for each user before the targeting session. The tracking and registration functionalities supported by the Vuforia SDK were used for the real-time registration of the virtual targets and the real mannequin.

The subjects were asked to use the digitizer to point at the perceived position of the four (4) virtual targets displayed

(a) (b) (c)

Figure 7: Experimental setup for the quantitative study: (a) planned position of targets (red points) in the CAD environment; (b) mannequin used for the test with the Vuforia Image Target; (c) a user wearing the HoloLens during a targeting task and using the Aurora digitizer to point at the perceived position of one target.

through the HMD (Figure 7). Each target was acquired 3 times by each user, for a total of 12 targeting trials per person (60 in total). Target positions, acquired in the Aurora reference frame, were then expressed in the virtual environment reference frame by means of the ${}^{A}\mathbf{T}_{V}$ rigid transformation.

Targeting accuracy was measured as the average Euclidean distance between the perceived (digitized) position and the actual position of each target. The maximum and minimum error (Euclidean distance), as well as the standard deviation, were also calculated for each target.

2.7. Qualitative Study. Twenty (20) subjects with 10/10 or corrected (lenses) to 10/10 vision were recruited from technical employees (engineers) and personnel with medical background (medical students, orthopaedic resident surgeons, orthopaedic surgeons) of the University of Pisa (see Table 1 for detailed demographics).

The qualitative study includes: subjective workload assessments with the NASA Task Load Index (NASA-TLX) Questionnaire and a Likert Questionnaire to evaluate visual and audio perception, and interaction and ergonomics issues. NASA-TLX is a multidimensional rating procedure that provides an overall workload score, between 0 and 100, based on a weighted average of ratings on six subscales [41]:

(i) mental demands ("How mentally demanding was the task?"),

(ii) physical demands ("How physically demanding was the task?"),

(iii) temporal demands ("How hurried or rushed was the pace of the task?"),

(iv) own performance ("How successful were you in performing the task?"),

(v) effort ("How hard did you have to work to achieve your level of performance?"), and

Table 1: Demographics of participants in the qualitative evaluation study.

Profession/Position Held (engineers; med. staff: students, orthop. residents, and orthop. surgeons)	10; 10 (6, 1, 3)
Gender (male, female, nonbinary)	13, 7, 0
Age (min, max, mean, STD)	23, 48, 32, 7
Handedness (left, right, ambidextrous)	2, 18, 0
Vision (10/10 naked eyes, corrected to 10/10 with lenses)	10, 10
Experience with AR (none, limited, familiar, experienced)	8, 5, 5, 2
Experience with HoloLens (none, limited, familiar, experienced)	16, 3, 1, 0
Colour Blindness (no, yes)	20, 0
English Reading (none, limited, familiar, experienced)	0, 0, 12, 8
English Speaking (none, limited, familiar, experienced)	0, 2, 11, 6

(vi) frustration ("How insecure, discouraged, irritated, stressed, and annoyed were you?").

NASA-TLX Questionnaire was administrated to identify the primary source of workload during the execution of the proposed AR-based simulation and to investigate workload levels of users with differing characteristics ("Profession/Position Held," "Experience with AR" etc.).

The Likert Questionnaire, which is reported in Table 2, comprises 14 items, each evaluated using a 5-points Likert scale (from 1 = strongly disagree, to 5 = strongly agree).

The experimental setup is depicted in Figure 6. The mannequin was positioned on a fixed height surgical table. The study protocol for each participant included the following steps:

(1) The participant fills out a Consent Form and a Demographic Form (Table 1) including information

Table 2: Spatial accuracy evaluation.

	Accuracy (mean error)	Max. error	Min. error	STD
Target 1	2.1	4.4	1.0	1.1
Target 2	1.7	3.3	0.9	0.8
Target 3	1.7	3.3	0.7	0.7
Target 4	2.5	5.2	0.8	1.4
Total	2.0	5.2	0.7	1.1

about his/her previous experience with AR and HoloLens.

(2) The subject calibrates the HoloLens using the Calibration app (by Microsoft).

(3) The subject learns how to interact with HoloLens by means of head movements, gestures, and voice commands, using the Learn Gestures app (by Microsoft).

(4) The subject fills out the NASA-TLX Questionnaire (part 1, weights form).

(5) The HipSim app is launched and the subject has to perform a series of tasks (Figure 8).

(6) The subject fills out the NASA-TLX Questionnaire (part 2, rating form).

(7) The subject fills out the Likert Questionnaire.

(8) The total time of the study was recorded for every participant.

Statistical analysis of data was performed using the SPSS® Statistics Base 19 software.

Results of the NASA-TLX Questionnaire are summarized in terms of means and standard deviation. Data were processed using the analysis of variance (ANOVA) to examine possible relationships between individual characteristics and workload.

As for the Likert Questionnaire, the central tendencies of responses to a single Likert item were summarized by using median, with dispersion measured by interquartile range. The Mann–Whitney U test and Kruskal–Wallis test were used to understand whether the answering tendencies (with respect to each Likert item) differ based on "Profession/Position Held" and "Experience with AR"/"Experience with HoloLens". A p value <0.05 was considered statistically significant.

3. Results

3.1. Quantitative Evaluation Results. The obtained RMSE and MR are, respectively, 0.6 mm and 0.8 mm. Table 2 reports the accuracy obtained for each target, as well as the maximum error, minimum error and the standard deviation. The maximum error is compatible with values declared by HoloLens developers: Klein G. reported [42] a maximum static registration error <10 mrad, which results in an error of about 5 mm at a distance of 50 cm from the user (the approximate working distance in our setup).

3.2. Qualitative Evaluation Results. The average time for the completion of the study was 40 minutes.

Figure 9 shows the results of the subjective workload scores from the NASA-TLX Questionnaire. No statistically significant differences were found between personnel with medical background and engineers (Mental Demand $p = 0.741$; Physical Demand $p = 0.079$; Temporal Demand $p = 0.246$; Frustration Demand $p = 0.297$; Effort $p = 0.445$; Performance Evaluation $p = 0.826$; Overall Workload $p = 0.825$). Moreover, no statistically significant differences were found between groups with different experience with AR (Mental Demand $p = 0.418$; Physical Demand $p = 0.539$; Temporal Demand $p = 0.524$; Frustration Demand $p = 0.912$; Effort $p = 0.218$; Performance Evaluation $p = 0.709$; Overall Workload $p = 0.931$); and HoloLens (Mental Demand $p = 0.419$; Physical Demand $p = 0.800$; Temporal Demand $p = 0.718$; Frustration Demand $p = 0.831$; Effort $p = 0.530$; Performance Evaluation $p = 0.704$; Overall Workload $p = 0.905$).

The overall workload obtained (30.65) can be considered low giving that the average overall score observed in the literature for medical task is 50.60 (min 9.00; max 77.35) and for computer activities is 54.00 (min 7.46; max 78.00) [43]. Performance induced the highest workload indicating the overall satisfaction with self-assessed performance.

Table 3 summarizes the results of the Likert Questionnaire. Results show no statistically significant differences in answering tendencies between engineers and clinicians with an exception for the postural discomfort during the application and the ease of aligning the surgical instrument to the AR viewfinders.

As for the postural discomfort, clinicians expressed a neutral opinion (median 3), while engineers agreed that they did not experience postural discomfort (median 4). Moreover, clinician also expressed a neutral opinion (median 3) regarding the ease of aligning the surgical instrument, while engineers strongly agreed that this task is easy (median 5).

Overall, participants agreed/strongly agreed that the virtual content is correctly aligned to the real objects (median 5), it is easy to perceive the spatial relationships between real and virtual objects (median 5), they did not notice motion of virtual content (median 4), they did not notice latency (median 4), they did not notice jitter (median 4), they did not experience double vision (median 5), they did not notice colour separation (median 5), the field of view is adequate for the application (median 4), the spatial sounds make the experience more immersive (median 4.5), the gesture interaction is easy and intuitive (median 5), and the voice interaction is easy and intuitive (median 4.5). The overall median opinion regarding the experience of visual fatigue is neutral (median 3.5).

4. Conclusions

As suggested by a recent literature review on orthopaedic surgery simulation [11], *"an ideal simulator should be multimodal, combining haptic, visual and audio technology to create an immersive training environment."* In this work, we present an innovative multimodal simulation tool, which takes advantage from patient-specific modelling to improve the realism of the simulated surgical case; rapid prototyping

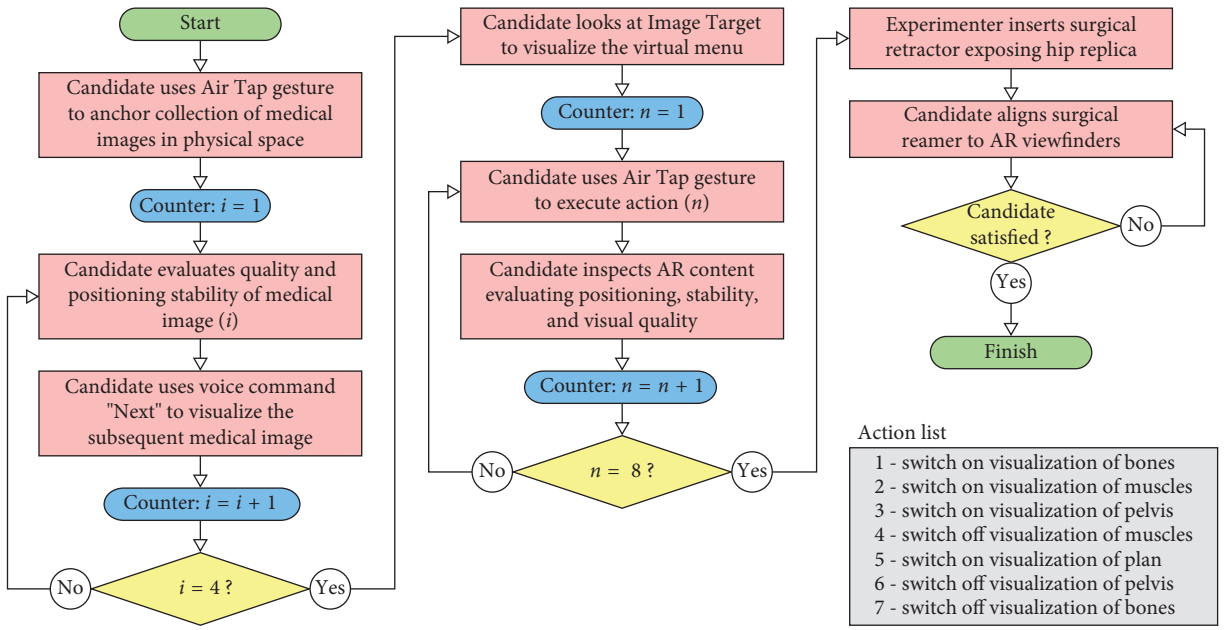


Figure 8: Flowchart of the experiment using the HipSim app.

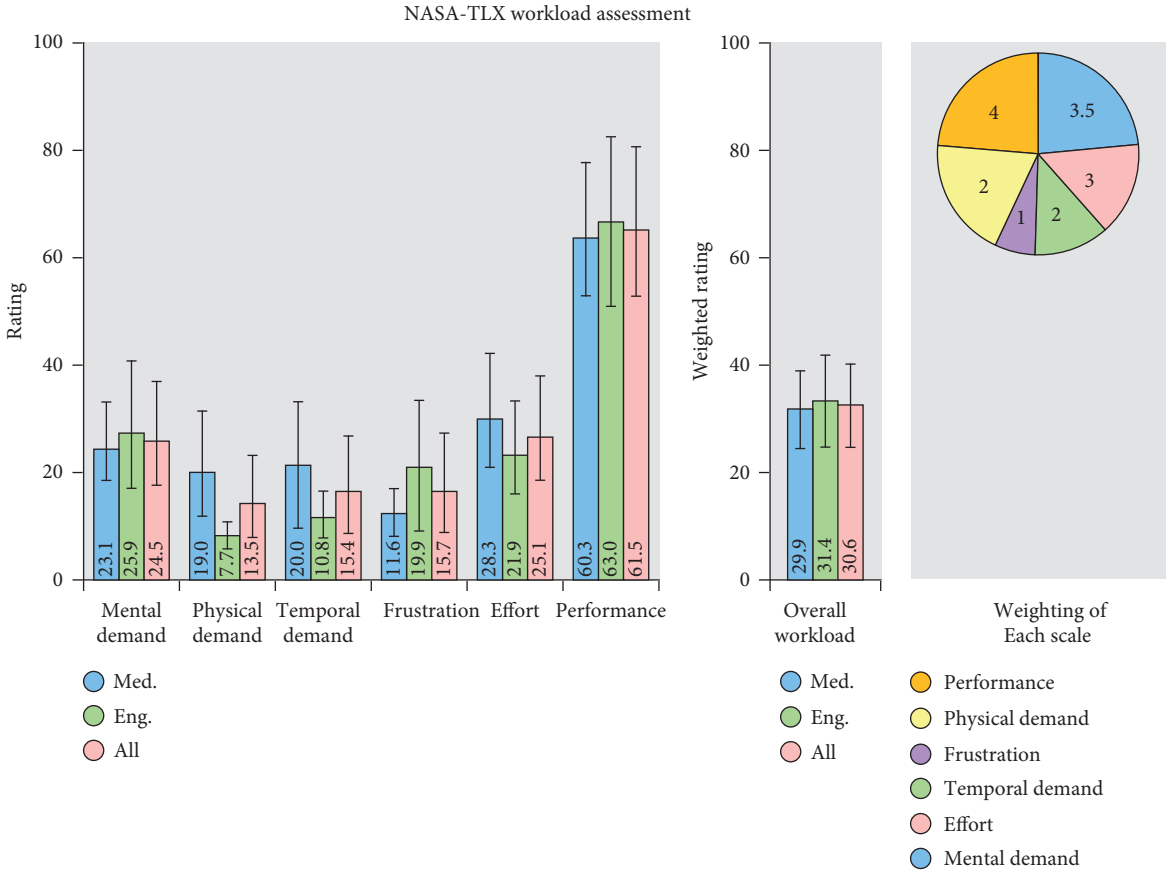


Figure 9: The bar charts show the mean rating with the standard deviation of each subscale and the overall weighted workload for personnel with medical background, engineers, and all participants. The pie chart shows the averaged (across all participants) weighting for each subscale (the total weighting is 15).

Table 3: Qualitative evaluation using a 5-point Likert questionnaire. Central tendency summarized using median with dispersion measured by interquartile range (25°~75°).

	Item	Questionnaire items	Median (25°~75°)			P value
			Engineers	Clinicians	All	(Eng. vs Clin.)
Visual and audio perception	A	The virtual content is correctly aligned to the real objects.	5 (5~4)	4 (5~3.75)	5 (5~4)	0.280
	B	It is easy to perceive the spatial relationships between real and virtual objects.	5 (5~4)	4.5 (5~4)	5 (5~4)	0.739
	C	I did not notice motion of virtual content.	4 (5~4)	4 (5~3.75)	4 (5~4)	0.436
	D	I did not notice latency (lag, delay) between virtual content and objects real.	4.5 (5~4)	4 (5~4)	4 (5~4)	0.353
	E	I did not notice jitter (high-frequency shaking of the virtual content).	4 (5~2.75)	4 (4.75~3.75)	4 (5~3)	0.912
	F	I did not experience double vision.	4.5 (5~4)	5 (5~4)	5 (5~4)	0.481
	G	I did not notice colour separation.	5 (5~3.75)	5 (5~4.75)	5 (5~4)	0.393
	H	The field of view (FOV) is adequate for the application.	4 (4.25~2.75)	3.5 (4~2.0)	4 (4~2.25)	0.579
	I	Spatial sounds make the experience more immersive.	4 (5~4)	5 (5~3.75)	4.5 (5~4)	0.796
Interaction and ergonomics	J	I did not experience postural discomfort during the application.	4 (4.25~3.75)	3 (4 ~ 2)	4 (4~2.25)	**0.029**
	K	I did not experience visual fatigue (eyestrain, dried mucus or tears around the eyelids, discomfort when the eyes are open, hot eyes, and headaches).	4 (4.25~2.75)	2.5 (4.25~2)	3.5 (4~2)	0.393
	L	Gesture interaction is easy and intuitive.	4.5 (5~4)	5 (5~4)	5 (5~4)	0.631
	M	Voice interaction is easy and intuitive.	4 (5~4)	5 (5~4)	4.5 (5~4)	0.481
	N	It is easy to aligning the surgical instrument to the AR viewfinders.	5 (5~4)	3 (4~2)	4 (5~3)	**0.023**

No statistically significant differences were found between groups with different experience with AR (Item A $p = 0.126$; Item B $p = 0.219$; Item C $p = 0.789$; Item D $p = 0.653$; Item E $p = 0.590$; Item F $p = 0.085$; Item G $p = 0.204$; Item H $p = 0.466$; Item I $p = 0.196$; Item J $p = 0.204$; Item K $p = 0.246$; Item L $p = 0.469$; Item M $p = 0.284$; Item N $p = 0.193$) and HoloLens (Item A $p = 0.606$; Item B $p = 0.662$; Item C $p = 0.772$; Item D $p = 0.326$; Item E $p = 0.986$; Item F $p = 0.986$; Item G $p = 0.772$; Item H $p = 0.499$; Item I $p = 0.364$, item J $p = 0.470$; Item K $p = 0.508$; Item L $p = 0.739$; Item M $p = 0.187$; Item N $p = 0.760$).

for the manufacturing of synthetic models, which guarantees a realistic haptic feedback; AR to enrich the simulated scenario and guide the learner during the surgical procedure; and HoloLens functionalities for an interactive and immersive simulation experience.

Results of quantitative and qualitative study encourage the usage of HoloLens technology for the implementation of a hybrid simulator for orthopaedic open surgery. The perceived positioning accuracy matches the requirements of the target application. Moreover, the perceived overall workload can be considered low, and subjects participating in this study expressed satisfaction with self-assessed performance. A positive feedback was obtained on visual and audio perception, and gesture and voice interaction independently of the level of previous experience with AR and HoloLens, and education backgrounds (medical or technical). As regards postural discomfort during the application and the experience visual fatigue, obtained results show a non-negative opinion for a simulation experience with duration of 40 minutes (enough for the specific purposes). A more prolonged usage could negatively impact the comfort because of an increase of the visual fatigue. An optimal design of the simulation tasks and the simulation setup (time for each task, height of the surgical table, distance of user interaction) are required to minimize the user discomfort, so that the virtual content appears in the optimal/comfort zone for most of the time of the simulation period, and the head tilt is sustainable. Moreover, attention should be paid to the design of AR viewfinders (optimal shape, colour, transparency level) to ease the alignment task, which is already impaired by the focus rivalry between the physical and virtual content.

Hip arthroplasty, a surgical procedure which could take great advantage from simulation, was selected as a benchmark for this study. Primary and revision total HA indeed were ranked third and fourth among the orthopaedic interventions accounted for the greatest share of adverse events and excess hospital stay [44] and, as showed by several studies [45, 46], the risk of complications after HA is strongly related to the surgeon's case volume. In this context, surgical simulation could play a pivotal role, offering novices an opportunity to practice skills outside the operating theatre, in a safe controlled environment.

Future work will include Face Validity, Content Validity, and Construct Validity for a complete assessment of the proposed simulator for this specific orthopaedic intervention. Additionally, in the future, our system could integrate novel haptic equipment and able to simulate high-magnitude force feedback. However, in this case, the usage of haptic interfaces will be limited to the simulation of the reamer-bone interactions, whereas the direct interactions between the surgeon hands and the soft

tissue will be still simulated using the current synthetic mannequin.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was partially supported by the European Project VOSTARS (H2020 Call ICT-29-2016 G.A. (731974)).

References

[1] D. M. Gaba, "The future vision of simulation in health care," *Quality and Safety in Health Care*, vol. 13, no. 1, pp. i2–i10, 2004.

[2] M. Morgan, A. Aydin, A. Salih, S. Robati, and K. Ahmed, "Current status of simulation-based training tools in orthopedic surgery: a systematic review," *Journal of Surgical Education*, vol. 74, no. 4, pp. 698–716, 2017.

[3] S. Misra, K. T. Ramesh, and A. M. Okamura, "Modeling of tool-tissue interactions for computer-based surgical simulation: a literature review," *Presence*, vol. 17, no. 5, p. 463, 2008.

[4] Sawbones® Orthopaedic Models, 2017, https://www.sawbones.com/products/orthopaedic-models.html.

[5] M. P. Gardner, A. C. Chong, A. G. Pollock, and P. H. Wooley, "Mechanical evaluation of large-size fourth-generation composite femur and tibia models," *Annals of Biomedical Engineering*, vol. 38, no. 3, pp. 613–620, 2010.

[6] A. D. Heiner, "Structural properties of fourth-generation composite femurs and tibias," *Journal of Biomechanics*, vol. 41, no. 15, pp. 3282–3284, 2008.

[7] A. C. Chong, E. A. Friis, G. P. Ballard, P. J. Czuwala, and F. W. Cooke, "Fatigue performance of composite analogue femur constructs under high activity loading," *Annals of Biomedical Engineering*, vol. 35, no. 7, pp. 1196–1205, 2007.

[8] P. D. Parchi, V. Ferrari, N. Piolanti et al., "Computer tomography prototyping and virtual procedure simulation in difficult cases of hip replacement surgery," *Surgical Technology International*, vol. 23, pp. 228–234, 2013.

[9] P. Tack, J. Victor, P. Gemmel, and L. Annemans, "3D-printing techniques in a medical setting: a systematic literature review," *Biomedical Engineering Online*, vol. 15, no. 1, p. 115, 2016.

[10] S. M. Botden and J. J. Jakimowicz, "What is going on in augmented reality simulation in laparoscopic surgery?," *Surgical Endoscopy*, vol. 23, no. 8, pp. 1693–1700, 2009.

[11] N. Vaughan, V. N. Dubey, T. W. Wainwright, and R. G. Middleton, "A review of virtual reality based training simulators for orthopaedic surgery," *Medical Engineering and Physics*, vol. 38, no. 2, pp. 59–71, 2016.

[12] M. Carbone, S. Condino, L. Mattei, P. Forte, V. Ferrari, and F. Mosca, "Anthropomorphic ultrasound elastography phantoms—characterization of silicone materials to build breast elastography phantoms," in *Proceedings of Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Conference*, vol. 2012, pp. 492–494, San Diego, CA, USA, August 2012.

[13] J. B. Nysaether, E. Dorph, I. Rafoss, and P. A. Steen, "Manikins with human-like chest properties—a new tool for chest compression research," *IEEE Transactions on Bio-Medical Engineering*, vol. 55, no. 11, pp. 2643–2650, 2008.

[14] J. L. Sparks, N. A. Vavalle, K. E. Kasting et al., "Use of silicone materials to simulate tissue biomechanics as related to deep tissue injury," *Advances in Skin & Wound Care*, vol. 28, no. 2, pp. 59–68, 2015.

[15] M. Carbone, S. Condino, V. Ferrari, M. Ferrari, and F. Mosca, "Surgical simulators integrating virtual and physical anatomies," in *Proceedings of CEUR Workshop*, pp. 13–18, Amsterdam, Netherlands, 2011.

[16] S. Condino, M. Carbone, V. Ferrari et al., "How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simulators," *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 7, no. 2, pp. 202–213, 2011.

[17] R. Viglialoro, S. Condino, M. Gesi, M. Ferrari, and V. Ferrari, "Augmented reality simulator for laparoscopic cholecystectomy training," in *Augmented and Virtual Reality: Lecture Notes in Computer Science*, L. T. De Paolis and A. Mongelli, Eds., pp. 428–433, Springer International Publishing, New York, NY, USA, 2014.

[18] V. Ferrari, R. M. Viglialoro, P. Nicoli et al., "Augmented reality visualization of deformable tubular structures for surgical simulation," *International Journal of Medical Robotics + Computer Assisted Surgery: MRCAS*, vol. 12, no. 2, pp. 231–240, 2015.

[19] R. Viglialoro, S. Condino, C. Freschi et al., "AR visualization of "synthetic Calot's triangle" for training in cholecystectomy," in *Proceedings of 12th IASTED International Conference on Biomedical Engineering, BioMed 2016*, pp. 85–89, Innsbruck, Austria, February 2016.

[20] F. Cutolo, U. Fontana, M. Carbone, R. D. Amato, and V. Ferrari, "[POSTER] hybrid video/optical see-through HMD," in *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 52–57, Nantes, France, October 2017.

[21] ©Microsoft Hololens, 2017, https://www.microsoft.com/en-us/hololens.

[22] I. D. Learmonth, C. Young, and C. Rorabeck, "The operation of the century: total hip replacement," *The Lancet*, vol. 370, no. 9597, pp. 1508–1519, 2007.

[23] P. Parchi, S. Condino, M. Carbone et al., "Total hip replacement simulators with virtual planning and physical replica for surgical training and reharsal," in *Proceedings of 12th IASTED International Conference on Biomedical Engineering, BioMed 2016*, pp. 97–101, Innsbruck, Austria, February 2016.

[24] W. Barfield and T. Caudell, *Fundamentals of Wearable Computers and Augmented Reality*, Taylor and Francis, Milton Park, UK, 2001.

[25] M. Carbone, S. Condino, F. Cutolo et al., "Proof of concept: wearable augmented reality video see-through display for neuro-endoscopy," in *Augmented and Virtual Reality: Lecture Notes in Computer Science*, L. T. De Paolis and A. Mongelli, Eds., pp. 95–104, Springer International Publishing, New York, NY, USA, 2018.

[26] J. P. Rolland and H. Fuchs, "Optical versus video see-through mead-mounted displays in medical visualization," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 3, pp. 287–309, 2000.

[27] D. Nicholson, "Advances in human factors in cybersecurity," in *Proceedings of AHFE 2017 International Conference on Human Factors in Cybersecurity*, Springer International Publishing, Los Angeles, CA, USA, July 2017.

[28] L. Qian, A. Barthel, A. Johnson, G. Osgood, P. Kazanzides, N. Navab et al., "Comparison of optical see-through head-mounted displays for surgical interventions with object-anchored 2D-display," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 6, pp. 901–910, 2017.

[29] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, no. 3, p. 33, 2008.

[30] G. Kramida, "Resolving the vergence-accommodation conflict in head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 7, pp. 1912–1931, 2016.

[31] J. Kim, D. Kane, and M. S. Banks, "The rate of change of vergence-accommodation conflict affects visual discomfort," *Vision Research*, vol. 105, pp. 159–165, 2014.

[32] D. W. Kim, J. S. Yoo, and Y. H. Seo, "Qualitative analysis of individual and composite content factors of stereoscopic 3D video causing visual discomfort," *Displays*, vol. 34, no. 3, pp. 223–240, 2013.

[33] V. Ferrari, C. Cappelli, G. Megali, and A. Pietrabissa, "An anatomy driven approach for generation of 3D models from multi-phase CT images," *International Journal of Computer Assisted Radiology and Surgery*, vol. 3, no. 1, pp. 271–273, 2008.

[34] P. A. Yushkevich, J. Piven, H. C. Hazlett et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[35] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "MeshLab: an open-source mesh processing tool," in *Eurographics Italian Chapter Conference*, Salerno, Italy, July 2008.

[36] Foundation B Blender free and open source 3D creation suite, 2017, https://www.blender.org/.

[37] G. Turini, S. Condino, S. Sinceri et al., "Patient specific virtual and physical simulation platform for surgical robot movability evaluation in single-access robot-assisted minimally-invasive cardiothoracic surgery," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 10325, Springer, Berlin, Germany, 2017.

[38] J. Grubert, Y. Itoh, K. R. Moser, and J. E. Swan II, "A survey of calibration methods for optical see-through head-mounted displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 9, pp. 2649–2662, 2017.

[39] Inc. P Vuforia SDK, 2017, https://www.vuforia.com.

[40] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 698–700, 1987.

[41] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load Index): results of empirical and theoretical research," in *Advances in Psychology*, P. A. Hancock and N. Meshkati, Eds., vol. 52, pp. 139–183, 1988.

[42] G. Klein, "Keynotes," in *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Nantes, France, October 2017.

[43] A. G. Rebecca, "How high is high? A meta-analysis of NASA-TLX global workload scores," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 59, no. 1, pp. 1727–1731, 2015.

[44] B. R. Wolf, X. Lu, Y. Li, J. J. Callaghan, and P. Cram, "Adverse outcomes in hip arthroplasty: long-term trends," *Journal of Bone and Joint Surgery American*, vol. 94, no. 14, p. e103, 2012.

[45] Y. Hasegawa and T. Amano, "Surgical skills training for primary total hip arthroplasty," *Nagoya Journal of Medical Science*, vol. 77, no. 1-2, pp. 51–57, 2015.

[46] R. N. de Steiger, M. Lorimer, and M. Solomon, "What is the learning curve for the anterior approach for total hip arthroplasty?," *Clinical Orthopaedics and Related Research*, vol. 473, no. 12, pp. 3860–3866, 2015.