2016 IEEE International Symposium on Mixed and Augmented Reality Adjunct Proceedings

First Deployment of Diminished Reality for Anatomy Education

Naoto Ienaga* Keio University Yokohama, Japan
Felix Bork† Technische Universitt Mnchen Munich, Germany
Siim Meerits‡ Keio University Yokohama, Japan
Shohei Mori§ Keio University Yokohama, Japan
Pascal Fallavollita¶ University of Ottawa Ottawa, Canada
Nassir Navab‖ Technische Universitt Mnchen Munich, Germany
Hideo Saito** Keio University Yokohama, Japan

Abstract

Understanding the anatomy of the human body is vital for everyone working in the medical domain. Augmented reality (AR) systems for anatomy teaching, which display virtual information directly on top of a users' body, have proven to facilitate mental mapping compared to traditional teaching paradigms. In this paper, we explore the potential of diminished reality (DR) in the context of anatomy education. As a first necessary step to achieving a DR anatomy education system, parts of the human body have to be extracted and diminished from the video stream. Our system diminishes either the arm or head of the user by projecting a background image recovered using RGB-D cameras. Such a system, if combined with an accurate overlay of virtual counterparts, could potentially improve the learning effect by attracting the users' attention to the virtual information and improve visual perception by avoiding the well-known floating effect of AR.

Index Terms: Mixed Reality, Diminished Reality, Anatomy Education

1 Introduction

Classes on human anatomy have always been the basis for a sound medical education. Paradigms for teaching anatomy to students have evolved significantly over the centuries. The most widespread techniques include textbooks, three-dimensional (3D) models, cadaver studies, and more recently, web-based learning platforms, mobile apps for smartphones or tablet devices, and virtual reality environments. One key disadvantage of such techniques is that the user cannot see the information in relation to his or her own body.

Recently, augmented reality (AR) systems have been developed to overcome this limitation and improve mental mapping of anatomical information. Blum *et al*. proposed *mirracle*, an AR system specifically developed for medical education, enabled by Microsoft Kinect skeleton tracking, which overlays virtual organs of the abdomen and thorax on top of a color image, alongside medical slices from different modalities registered to the users' body [2, 8]. Based on this work, Ma *et al*. proposed a personalized augmented reality system which further increases the accuracy of the AR overlay in anatomy teaching scenarios [7, 5]. Another system was developed by Juan *et al*., who employed a head-mounted display (HMD) to overlay human anatomy onto a phantom [4]. Davis *et al*. used a similar setup for simulation purposes to overlay virtual airways [3].

*e-mail:ienaga@hvrl.ics.keio.ac.jp
†e-mail:felix.bork@tum.de
‡e-mail:meerits@hvrl.ics.keio.ac.jp
§e-mail:mori@hvrl.ics.keio.ac.jp
¶e-mail:fallavol@in.tum.de
‖e-mail:nassir.navab@tum.de
**e-mail:hs@keio.jp

Compared to AR, diminished reality (DR) aims at modifying real-world scenes by detecting certain objects in a video stream and subsequently removing them. DR has been applied in various areas, such as layout visualization in industry [10], robot manipulation safety [9], and automobile safety [1]. To the best of our knowledge, there are no published works exploring the potential of DR for anatomy education. Our proposed system can diminish the head or the arm of the user by capturing the background and reconstructing it in real time. The purpose is to improve mental mapping and visual perception compared to a simple AR overlay.

2 Materials & Methods

The proposed system consists of two Microsoft Kinect RGB-D cameras and a large TV screen. In terms of DR, the system has to handle 3D dynamic scenes assuming a moving user and audience behind the user. To deal with this situation, we use two RGB-D cameras, one for the user perspective and tracking, C_u, and the other to capture 3D dynamic background structures, C_b.

2.1 Calibration

We first calibrate two RGB-D cameras C_u and C_b using a fiducial checkerboard. The checkerboard is captured by the two cameras simultaneously. Given known 3D points on the board and corresponding 2D points on images, transformation matrices from the board to each camera are calculated by solving Perspective-n-Point problems. Let $\mathbf{T}_u$ denote a 4×4 transformation matrix from the board to C_u and $\mathbf{T}_b$ denote the transformation matrix for C_b.

2.2 Background Reconstruction

The structure of the background scene is reconstructed in real time as a textured 3D mesh using Meerits and Saito's method [6]. This procedure takes RGB-D camera C_b depth and color image frames as input. The mesh is generated by connecting neighboring depth pixels of the depth map to form triangles.

To achieve viewpoint change, vertices $\mathbf{p}_b \in \mathbb{R}^3$ of the generated 3D mesh are transformed to $\mathbf{u}_u \in \mathbb{R}^2$ in the user reference frame using the transformations $\mathbf{T}_u$ and $\mathbf{T}_b$ introduced in the previous subsection and perspective projection $\pi(\cdot)$,

$$\mathbf{u}_u = \pi(\mathbf{T}_u \mathbf{T}_b^{-1} \tilde{\mathbf{p}}_b) \quad (1)$$

Finally, the textured triangle mesh is rendered with conventional computer graphics methods.

The method was chosen mainly for its speed and simplicity. A more complex reconstruction method that can fill in missing geometry is unlikely to benefit us as we only have a single color image source for the background. In other words, even if there was a more complete background geometry model we would not have the texture map to completely color it.

2.3 Background Overlay

The projected background image is masked and overlaid on the user image to diminish an arm or head of the user. Depending on selected computer-generated (CG) skeleton models, pre-defined joint

978-1-5090-3740-7/16 $31.00 © 2016 IEEE
DOI 10.1109/ISMAR-Adjunct.2016.90

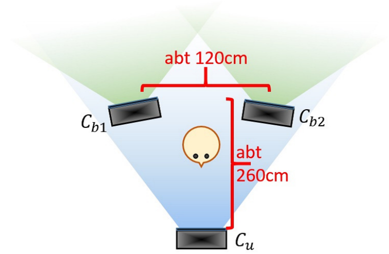

Figure 1: Setup illustration. We used two RGB-D cameras $C_{\mathrm{b}1}$ and $C_{\mathrm{b}2}$.

Figure 2: (Top) Input, (bottom) resulting DR of arm.

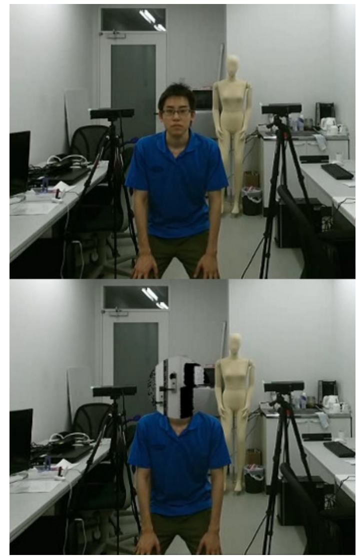

Figure 4: DR of head.

Figure 5: Example of failed DR.

Figure 3: (Top left) Full size version of the upper image in Figure 2; (top right) background reconstructed by $C_{\mathrm{b}2}$, (bottom left) background reconstructed by $C_{\mathrm{b}2}$, (bottom right) full size version of the lower image in Figure 2.

pairs of the user and thresholds are used to determine diminishing regions. Thresholds are radius of a circle and width of a rectangle. We mask the user's arm area by combining rectangles and circles and the head area by using a circle.

3 Results & Discussion

Figure 1 shows the setup of this experiment. We used two RGB-D cameras (Microsoft Kinect v2) to expand an area within which capturing background is possible. C_{u} and C_{b} should not be separated enough to overlap their fields of view. The fields of view of the two C_{b}s should be set to capture as wide an area as possible without overlap. Reconstructed background, input image and the DR result of an arm are shown in Figure 3. We determine which background should be overlaid only by the value of the horizontal direction. Figure 2 shows an expanded view of the input image and result in Figure 3, and Figure 4 shows the input image and the DR result for the head.

One of the limitations is the angle of view of C_{b}. We of course can reconstruct only backgrounds within the angle of view of C_{b}. We cannot mask when the user's head or arm areas deviate from the angle of view of C_{b}. An example of such a failure is shown in Figure 5.

In future work, virtual models of the head and arm will be overlaid on top of the diminished video stream. An evaluation has to be performed in order to test whether such a DR visualization can yield improvements in terms of anatomy learning. Both accuracy and perception are important topics in this sense: on the one hand, virtual models have to be placed accurately with regard to their real-world counterparts. One the other hand, simply overlaying such virtual models will lead to perceptual challenges. In the literature, various ways to deal with such perceptual issues of AR overlays have been described. A sound evaluation has to be performed on this end as well.

4 Conclusion

In this paper, we have brought diminished reality technology to the field of anatomy teaching. Our dual RGB-D camera system captures dynamic backgrounds and overlays such regions in the regular video camera stream displayed to the user of the system. Both the head and the arm of a user can be diminished, paving the way for mixed reality applications in the future, incorporating overlays of virtual information on top of the diminished video stream.

References

[1] P. Barnum, Y. Sheikh, A. Datta, and T. Kanade. Dynamic seethroughs: Synthesizing hidden views of moving objects. In *Mixed and Augmented Reality, 2009. ISMAR 2009. 8th IEEE International Symposium on*, pages 111–114. IEEE, 2009.

[2] T. Blum, V. Kleeberger, C. Bichlmeier, and N. Navab. Mirracle: An augmented reality magic mirror system for anatomy education. *Proc. - IEEE Virtual Real.*, pages 115–116, mar 2012.

[3] L. Davis, F. G. Hamza-Lup, J. Daly, Y. Ha, S. Frolich, C. Meyer, G. Martin, J. Norfleet, K.-C. Lin, C. Imielinska, et al. Application of augmented reality to visualizing anatomical airways. In *AeroSense 2002*, pages 400–405. International Society for Optics and Photonics, 2002.

[4] C. Juan, F. Beatrice, and J. Cano. An augmented reality system for learning the interior of the human body. In *2008 Eighth IEEE International Conference on Advanced Learning Technologies*, pages 186–188. IEEE, 2008.

[5] M. Ma, P. Fallavollita, I. Seelbach, A. M. Von Der Heide, E. Euler, J. Waschke, and N. Navab. Personalized augmented reality for anatomy education. *Clin. Anat.*, dec 2015.

[6] S. Meerits and H. Saito. Real-time diminished reality for dynamic scenes. In *Mixed and Augmented Reality Workshops (ISMARW), 2015 IEEE International Symposium on*, pages 53–59. IEEE, 2015.

[7] M. Meng, P. Fallavollita, T. Blum, U. Eck, C. Sandor, S. Weidert, J. Waschke, and N. Navab. Kinect for interactive AR anatomy learning. In *2013 IEEE Int. Symp. Mix. Augment. Reality, ISMAR 2013*, pages 277–278, Adelaide, Australia, oct 2013.

[8] N. Navab, T. Blum, L. Wang, A. Okur, and T. Wendler. First deployments of augmented reality in operating rooms. *Computer (Long. Beach. Calif).*, 45(7):48–55, 2012.

[9] K. Sugimoto, H. Fujii, A. Yamashita, and H. Asama. Half-diminished reality image using three rgb-d sensors for remote control robots. In *2014 IEEE International Symposium on Safety, Security, and Rescue Robotics (2014)*, pages 1–6. IEEE, 2014.

[10] S. Zokai, J. Esteve, Y. Genc, and N. Navab. Multiview paraperspective projection model for diminished reality. In *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*, pages 217–226. IEEE, 2003.