Life-size Sequential Photography in a Mixed Reality Environment

Shunji Muto
Shibaura Institute of Technology
Tokyo, Japan

Takashi Ijiri
Shibaura Institute of Technology
Tokyo, Japan

Figure 1: Life-size sequential photography in the MR environment. The observer with the video-see-through head-mounted device records a sports motion with a camera (a). Our system places life-size photographs in the MR space. Our system supports multiple visualizations such as sequential photography (b) and video with afterimages (c).

ABSTRACT

We visualize life-size sequential photographs of sports activities in a mixed reality (MR) environment. Wearing a video-see-through head-mounted display, an observer records the motions of a player using a handheld camera. Our system then places billboards in the three-dimensional MR space, on which the sequential photographs of the player's motion are presented at life-size. In a user study, we found that the observers perceived the size of the motions more accurately than when viewing sequential photographs on a monitor display.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**.

KEYWORDS

Life-size Sequential Photography, sports motion, Mixed Reality

ACM Reference Format:
Shunji Muto and Takashi Ijiri. 2020. Life-size Sequential Photography in a Mixed Reality Environment. In *26th ACM Symposium on Virtual Reality Software and Technology (VRST '20), November 1–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3385956.3422086

1 INTRODUCTION

Sequential photography is an imaging technique that acquires multiple photographs of a moving object with a constant time interval, and combines them into a single image. Because sequential photography illustrates one motion per image, it is widely used in sports visualization. If the sequential photographs are displayed on a paper or monitor, the motions are smaller than life-size. Consequently, observers cannot easily recognize the actual size of the motion.

Many sports-visualization methods have been explored. Examples are sequential photography with a handheld camera [3], and a free-viewpoint video on a tabletop [4]. However, motions visualised with these methods are smaller than life-size. A smart mirror has also been used in sports visualization [1], but this method prevents the observer from freely approaching the visualized motion.

Our goal is to develop a sports replay system in which observers can easily perceive the sizes of sports motions. We propose life-size visualization of sequential photography in a mixed reality (MR) environment. The players' motions are presented as life-sized sequential photographs or video recordings with afterimage effects in the MR space (Figure 1). The player's life-size motions can be observed from a close viewpoint for easy size perception. The effectiveness of our method was tested in a user study. Participants of the study recognized the height of a jump more accurately when using our system than when using a monitor display. The feasibility of our system was further demonstrated in visualization examples of multiple sports.

2 PROPOSED METHOD

When using our system, an observer observes a video-transparent MR environment through a head-mounted display (Figure 2a), and holds a camera-mounted controller (Figure 2b). The observer first captures the motion of a player. When a controller button is pressed, the camera records a video until the button is released (Figure 1a). Our system then provides a life-size sequential visualization of the photography. We present three different visualizations: the sequential photographs, the video, and the video with afterimages (Figure 1b, c).

VRST '20, November 1–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-7619-8/20/11.
https://doi.org/10.1145/3385956.3422086

To realize the life-size visualization, we place planner billboards in the three-dimensional (3D) MR space. The human region in the video are extracted by segmenting each video frame using Mask R-CNN [2]. We next generate two rays by connecting the camera position and foot positions in the first and last frames. The two rays are projected onto the floor, creating points $\mathbf{p}_0$ and $\mathbf{p}_1$ (Figure 2c). Each of the segmented video frames is then placed as a planner billboard in the 3D space, and the depth of each billboard is determined by interpolating between $\mathbf{p}_0$ and $\mathbf{p}_1$ (Figure 2c).

3 RESULTS AND DISCUSSION

The effectiveness of our method was compared with that of traditional visualization in a user study. Twelve undergraduate students participated in the study. The participants observed sequential photographs of standing long jumps, and were tasked with indicating the highest ankle height of the jumper. One ankle of each jumper was attached with markers for easy identification of the ankle positions. Each participant performed the task eight times under four conditions: (1) using our method (Figure 3a), (2) using a monitor (Figure 3b), (3) using our method with a time limit, and (4) using a monitor with a time limit. Under the time-limited conditions, the sequential photographs were observed for three seconds. By imposing time-limited conditions, we tested our hypothesis that life-size visualization affects the time of recognizing motion size.

Figure 3c summarizes the average absolute differences between the participants' answers and the ground truths under the four conditions. Under the time-free conditions, the errors in the ankle-height estimates significantly differed (p<0.01) between our method and traditional viewing on a monitor (paired one-sided t-test). When using our method, most of the participants measured the height by using their bodies as rulers. For instance, one participant approached the billboards and marked a body part corresponding to the height of the jumper's ankle. We believe that this measurement improved the accuracy of the estimation. No clear differences were observed between the time-free and time-limited conditions. To illustrate the usefulness of our method, we also visualized various sports motions such as running, badminton swing, tennis swing, and baseball swing (see supporting video).

4 CONCLUSION AND FUTURE WORK

We proposed a life-size visualization of sequentially photographed sports motions in an MR environment. In a comparative study of our method and traditional visualization on a monitor, we confirmed that our method improves the accuracy of estimating sizes of motions. The heavy time demands of segmenting the video frames is one limitation of our present system. In future, we would like to accelerate this process by using a graphics processing unit. A more detailed user study on the time of recognizing the size of a motion is also planned in future work.

ACKNOWLEDGMENTS

We thank Daisuke Niino who graduated from Shibaura Inst. Tech. for his support in implementing our prototype. This work was supported by JSPS Grant-in-Aid for Scientific Research (C)[18K11606].

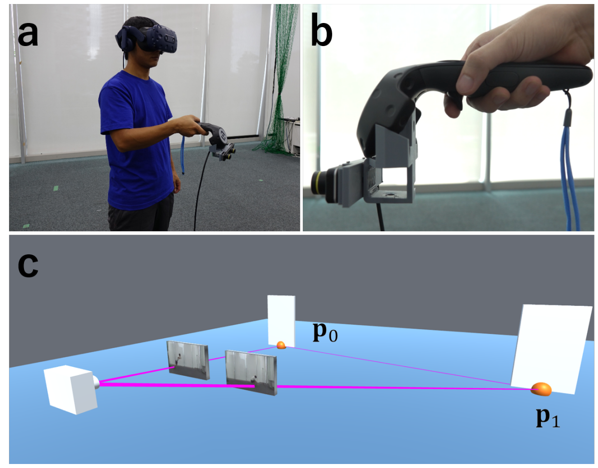

Figure 2: The observer wears the head-mounted display and records sports motions (a) with a handheld camera (b). The depths of the frames are determined by generating two rays passing through the camera and the foot positions of the first and last frames (c).

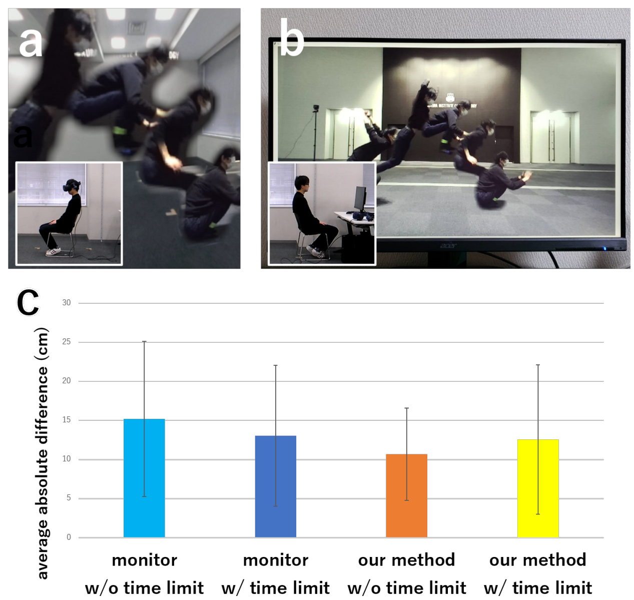

Figure 3: Setup (a, b) and results (c) of the user study.

REFERENCES

[1] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: Enhancing Movement Training with an Augmented Reality Mirror. In *Proc. UIST.* 311–320.

[2] He Kaiming, Gkioxari Georgia, Dollár Piotr, and Girshick Ross. 2017. Mask R-CNN. In *Proc ICCV.* 2980–2988.

[3] Kentaro Matsui, Masaki Iwase, Masato Agata, Toshimi Tsu Tanaka, and Noboru Ohnishi. 1998. Soccer image sequence computed by a virtual camera. In *Proc. CVPR.* IEEE, 860–865.

[4] Konstantinos Rematas, Ira Kemelmacher-Shlizerman, Brian Curless, and Steve Seitz. 2018. Soccer on your tabletop. In *Proc. CVPR.* 4738–4747.