



Recording and replaying psychomotor user actions in VR

Manos Kamarianakis

kamarianakis@uoc.gr

FORTH - ICS, University of Crete, ORamaVR
Heraklion, Greece

Mike Kentros

mike.kentros@oramavr.com

FORTH - ICS, University of Crete, ORamaVR
Heraklion, Greece

Ilias Chrysovergis

ilias.chrysovergis@oramavr.com

ORamaVR
Heraklion, Greece

George Papagiannakis

papagian@ics.forth.gr

FORTH - ICS, University of Crete, ORamaVR
Heraklion, Greece

KEYWORDS

Virtual Reality, Recording, Replay, Data Collection, Lip Sync

ACM Reference Format:

Manos Kamarianakis, Ilias Chrysovergis, Mike Kentros, and George Papagiannakis. 2022. Recording and replaying psychomotor user actions in VR. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '22 Posters)*, August 07-11, 2022. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3532719.3543253>

1 INTRODUCTION

Nowadays, session recording and playback of a single or multi-user VR session has become an increasingly market-required asset. The need for effective VR recording and replaying (VRRR) is especially highlighted in virtual training applications, as replaying user actions can serve as an additional and powerful educational tool. Despite the effort, achieving VRRR is a task not natively undertaken by modern game engines and therefore most VR applications do not include such a feature by default.

Current bibliography contains numerous studies of how the VR record and replay features can enhance the learning impact that VR educational-oriented applications provide, by mainly measuring the performance of users [Lahanas et al. 2015]. Usually, the data are captured in video format [Zia et al. 2016] and a post process of this high-dimensional data is required to obtain any further analysis. Since video data is not sufficient for reasons we explain in Section 2, our approach is close to [Kloiber et al. 2020], where user's motion were analysed by recording their hands and head trajectories. Current ongoing research also explores the proper methods and data structures that must be employed to achieve real-time logging while keeping the required data storage manageable and allowing effective replay.

2 OUR APPROACH

Our proposed VRRR functionality, already implemented in the Unity3D MAGES SDK 4.0 ([Papagiannakis et al. 2020]), publicly

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

SIGGRAPH '22 Posters, August 07-11, 2022, Vancouver, BC, Canada

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9361-4/22/08.

<https://doi.org/10.1145/3532719.3543253>

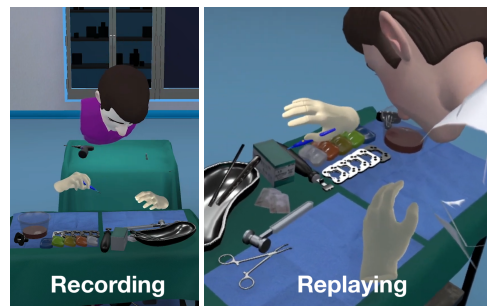


Figure 1: VR Record (Left) and Replay (Right) functionality. The user is able to replay the recorded session from any perspective, as well as pause the replay in order to continue the session on his own.

available for free), enables a) experts to record and replay their sessions, b) novices to learn how to correctly perform an operation by watching the expert's recording and reviewing their own sessions, and c) evaluators to assess the learning outcomes of the apprentices, utilizing the VR Replay feature (see Figure 1).

VR Recorder: Existing methods dictate that accurate recording of a VR session can be achieved via two methods, depending on the logged data. The first method records all users' inputs whereas, the second method focuses on the effects that these inputs have on the virtual scene. Our research revolves around a VR logger that records raw user input, as such an approach fits well with the current growth direction of virtual reality environments. In this respect, we avoid collecting high-dimensional data, such as the state of the entire virtual world. Instead, we collect and store low-dimensional information, such as the users' inputs and triggers, that may be used to obtain valuable analytics even by using simple, machine learning (ML) based, processing algorithms. Finally, the recorded data can be utilized to easily reproduce the user session by applying the world's mechanics, while the second method would demand sophisticated computer vision algorithms.

Our VR Recorder allows, for the first time, to log the user's sessions within a virtual environment in the form of positions, rotations, user-object interactions and training scenegraph nodes [Zikas et al. 2022], resulting in improved accuracy without compromising generalization, while requiring minimal storage space, approximately 1MB per minute per user. As the objects which the user interacts with might come into contact or interaction with other

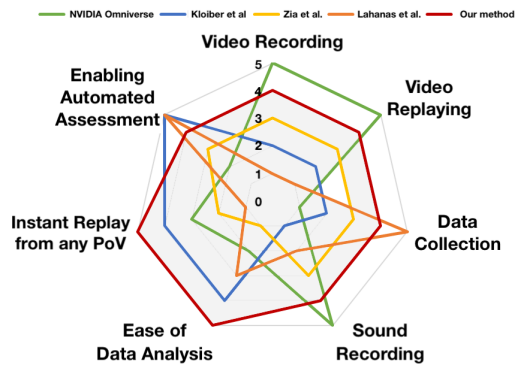


Figure 2: Comparison chart amongst existing VR recording and/or replaying methods (NVIDIA Omniverse, [Kloiber et al. 2020; Lahanas et al. 2015; Zia et al. 2016]) and our proposed method. Various features are graded from 0 (non-existent, inner layers) to 5 (exceptional, outer layers).

objects, i.e., changing their location or status, we are also recording the transformations of all subsequently affected objects. The voice of each player is recorded individually, including incoming voice in cases of multi-player sessions.

VR Replay: Via VRRR, we can replay a VR session, in both single and multi player modes. These recordings can be synchronized with the cloud and also be replayed on any device regardless of the original hardware they were recorded on. VRRR allows a full reproduction of the session as it happened when it was recorded, rather than simply video replaying the in-game view. While replaying the session, the users are free to move around the virtual world, watch the scene from any angle, or act simultaneously with the various recorded interactions and events. Such a functionality allows the creation of high fidelity VR replays that can be used creatively to increase the pedagogical benefits of various simulations.

Audio-Video Synchronization: Audio Video synchronization (AV-sync) is a common issue when one tries to replay graphics with audio. According to the European Broadcasting Union, the relative timing between audio and vision components must be less than 40 ms for sound before graphics, and less than 60ms for sound after graphics. The proposed AV-sync method is similar to the one presented by Zhang et. al [Zhang et al. 1705], which uses timestamps to eliminate the lip-sync error. In this regard, two issues need to be tackled a) the graphics replay times and b) the remote user’s sound reception for multiplayer sessions. The first issue is caused by the different frame rate between recording and replaying a session. As a result, a desynchronization between graphics and the respective sounds is created. The *sound after graphics* issue is resolved by waiting the amount of frames needed to visualize graphical changes. On the other hand, the *sound before graphics* is tackled by skipping visual changes. The second issue, caused by the reception time difference between transformation and sound data of a remote user, is eliminated by adding noise or removing samples respectively, to match the expected sound sample rate, i.e., 48000 samples per second.

Results: A comparison chart of existing VR recording or/and replaying methods against our proposed method is presented in Figure 2. Moreover, Table 1 shows a negligible performance overhead in terms of FPS in the use of our proposed VR Recorder.

Table 1: Measuring the FPS burden of a VR application due to the Recorder feature. The results were obtained using a PC with an i7-11375H, 16 GB of RAM and an RTX 3060.

Metric	Session without VR Recording	Session with VR Recording	Difference
Average FPS	89.56	85.13	4.43
Minimum FPS	76.56	68.78	7.78
Maximum FPS	93.29	92.57	0.72

3 CONCLUSIONS AND RELATED WORK

We introduce a novel method that describes the functionality and characteristics of an efficient VR recorder with replay capabilities, implemented in a modern game engine, publicly available for free.

Relating to this work, we have also developed a Convolution Neural Network (CNN) that can digest the stored data to assess in real-time the user’s achievements compared to prerecorded actions. In the future, we aim to create intelligent agents trained using experts’ session data, that are able to complete tasks on their own. Furthermore, we plan to develop a no-code VR authoring tool, that would allow the virtual environment designers to develop new training modules. This tool will also fuse all recorded data and train ML algorithms in order to understand the type of task the designer intends to develop.

ACKNOWLEDGMENTS

The project was partially funded by the European Union’s Horizon 2020 research and innovation programme under grant agreements No 871793 (ACCORDION) and No 101016509 (CHARITY). We would like to thank Antonis Protosaltis for his valuable comments.

REFERENCES

- Simon Kloiber, Volker Settgast, Christoph Schinko, Martin Weinzerl, Johannes Fritz, Tobias Schreck, and Reinhold Preiner. 2020. Immersive analysis of user motion in VR applications. *The Visual Computer* 36, 10-12 (Aug. 2020), 1937–1949.
- Vasileios Lahanas, Constantinos Loukas, Nikolaos Smailis, and Evangelos Georgiou. 2015. A novel augmented reality simulator for skills assessment in minimal invasive surgery. *Surg. Endosc.* 29, 8 (2015), 2224–2234.
- George Papagiannakis, Paul Zikas, Nick Lydatakis, Steve Kateros, Mike Kentros, Efstratios Geronikolakis, Manos Kamarianakis, Ioanna Kartsonaki, and Giannis Evangelou. 2020. MAGES 3.0: Tying the Knot of Medical VR. In *ACM SIGGRAPH 2020 Immersive Pavilion*. Association for Computing Machinery, Article 6, 2 pages.
- Min Zhang, Yu-Hua Wu, Ji Li, and Shi-Jun Li. 2017/05. Research on Audio and Video Synchronization Algorithm Based on AVI Format. In *Proceedings of the 2017 2nd International Conference on Materials Science, Machinery and Energy Engineering (MSMEE 2017)*. Atlantis Press, 959–962.
- Aneeq Zia, Yachna Sharma, Vinay Bettadapura, Eric L Sarin, Thomas Ploetz, Mark A Clements, and Irfan Essa. 2016. Automated video-based assessment of surgical skills for training and evaluation in medical schools. *Int. J. Comput. Assist. Radiol. Surg.* 11, 9 (Sept. 2016), 1623–1636.
- Paul Zikas, Steve Kateros, Nick Lydatakis, Mike Kentros, Efstratios Geronikolakis, Manos Kamarianakis, Giannis Evangelou, Ioanna Kartsonaki, Achilles Apostolou, Tanja Birrenbach, Aristomenis K. Exadaktylos, Thomas C. Sauter, and George Papagiannakis. 2022. Virtual Reality Medical Training for COVID-19 Swab Testing and Proper Handling of Personal Protective Equipment: Development and Usability. *Frontiers in Virtual Reality* 2 (2022).