International Journal of Virtual and Augmented Reality
Volume 5 • Issue 1

VR Presentation Training System Using Machine Learning Techniques for Automatic Evaluation

Yuto Yokoyama, Nagoya University, Japan
Katashi Nagao, Nagoya University, Japan

ABSTRACT

In this paper, the authors build an immersive training space using building-scale VR, a technology that makes a virtual space based on an entire building existing in the real world. The space is used for presentations, allowing students to self-train. The results of a presentation are automatically evaluated by using machine learning or the like and fed back to the user. In this space, users can meet their past selves (more accurately, their avatars), so they can objectively observe their presentations and recognize weak points. The authors developed a mechanism for recording and reproducing activities in virtual space in detail and a mechanism for applying machine learning to activity records. With these mechanisms, a system for recording, reproducing, and automatically evaluating presentations was developed.

KEYWORDS

3D Avatars, Automatic Evaluation of Presentations, Building-Scale VR, Machine Learning, Presentation Training, Virtual Environment, VR Playback, VR Recording

1. INTRODUCTION

The current use of VR is biased towards entertainment such as games, and much hardware and software development is for that purpose (Grubert et al., 2018; Gugenheimer et al., 2017; Huang et al., 2017). However, 6 degrees of freedom and controllers for interacting with the virtual world are also suitable for training, learning, and meetings, and we think that usage in these directions should be increased. In fact, the medical and athletic fields use VR simulation systems for surgical training and rehabilitation (Hamzeheinejad et al., 2018; Zahedi et al., 2017). In addition, research has been conducted on visualizing the movement of players in VR (Kaplan et al., 2018) and on improving the efficiency of VR training (Loreto et al., 2018; Isogawa et al., 2018).

Similarly, companies and universities are introducing VR conferences and VR classes, and their effectiveness has been verified (Clifford et al., 2018; Slavova et al., 2018; Volonte et al., 2018; Wienrich et al., 2018). In addition to being in a VR space remotely, VR-specific learning methods

DOI: 10.4018/IJVAR.290044

have been devised that display interactive 3D models or use 360-degree content for effective learning (Hussein et al., 2011).

However, VR usage related to meetings and classes mainly involves participation from remote locations with only a sense of reality, such as presence communication. Therefore, the VR feature of tracking the movement of parts of the body or the whole body is not faithfully taken advantage of.

Nonverbal communication channels such as gestures are also important in VR, and they are known to be particularly effective in presentations (Hostetter, 2011). Therefore, the body movements of VR users should be considered.

In an actual meeting, recording mainly involves video and audio. For this reason, extra work is required to obtain knowledge from these records, such as estimating the body movement of a speaker from images. However, in VR, the user always wears a tracking device, so data can be acquired more easily. In other words, a conference using VR is more likely to uncover the characteristics and customs of the participants than an in-person conference would.

Several systems for evaluating real-world presentations already exist (Kurihara et al., 2007). However, a methodology for improving the problems discovered and the specific solutions to the problems has not been established. The threshold of the evaluation indices used in the systems is also determined by intuition. Furthermore, gestures are not evaluated much during presentations.

In this paper, we propose a VR presentation training system that automatically evaluates and feeds back a user's own presentations.

In this system, a presentation venue (conference hall) similar to one in reality is reproduced in VR, the presenter and audience are expressed as VR avatars, and the presentation is recorded in detail and evaluated automatically in real time. Speech/gaze are evaluated by Kurihara et al. (2007) and Wolf In Motion Ltd. (2016) in the same manner, and we also evaluate gestures in the same way. To evaluate gestures, we record the movements of presenters during presentations made in VR space and perform subjective evaluations with multiple subjects.

We perform machine learning with this data and develop an automatic presentation-evaluation and feedback system. One of the functions of the VR presentation training system is a VR playback function that browses past presentations in VR and encourages users to reflect on their own presentation performance. However, at present, there is no appropriate means of recording and analyzing past presentations and motions.

In this research, we developed VRec, a library that records activities in VR in detail, and VR2ML, a library that analyzes recorded data using machine learning (Yokoyama & Nagao, 2020), as a supplemental system for the VR presentation training system.

VRec is a library for recording activities in VR, and it can record and play back 3D model movements such as human avatars and objects and events such as sound and effect generation.

VR2ML is a library that extends the functions of VRec and reconstructs the content recorded by VRec into a form suitable for machine learning and statistical processing. In addition, we provide a system that performs machine learning on a VR game engine and feeds back the results, although it is still limited to learning functions.

Both libraries are designed so that they can be used for purposes other than this research, and they have been released to the public and are being distributed.

Using these two libraries, we implemented and evaluated a VR presentation training system. In addition, short-term and long-term experiments were performed using this system, and the effectiveness of the system was verified on the basis of changes in evaluation scores and verbal questionnaires.

Our main contributions are three-fold: (1) a methodology for realizing and operating a VR presentation training system, (2) a method for analyzing body movements including gestures in presentations, which has not been considered much in the past, based on machine learning, and (3) the results of evaluation experiments done on the system using subjects.