DEIM Forum 2017 I3-1

A VR Application for Rehearsing and Improving Presentation

Yixuan He† Hayato Yamana† ‡

† Graduate School of Fundamental Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku Tokyo, Japan ‡ Faculty of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku Tokyo, Japan

‡ National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan

E-mail: {pandaxuan, yamana}@yama.info.waseda.ac.jp

Abstract: In this paper we propose a presentation rehearsal application in Virtual Reality (VR) that could be used to help users practice, improve their presentation skills, and overcome stage fright. In this application, users can import their own materials and thus they can see the slides during their speech in virtual environment. When they finish the rehearsal, a score will be given as a feedback. The whole presentation is evaluated by the user's stage usage and interaction with audience, such as eye contacts. Additionally, an eye tracking heat map will also be generated. It will allow users to see where and how they interact within the virtual environment. Recently, in Virtual Reality Applications, there are several domains show great potentials: Video Games, Tourism, Education, etc. Especially, Interactive Learning in VR gradually becomes more useful and approachable compared to traditional education ways. It is proved to be more effective to enable users to get more involved in the learning process. In this research, our goal is not only to provide the user a place to practice presentation despite the limitations of time and space, but also to help them overcome the performance anxiety and improve the quality. To the best of our knowledge, our paper is the first to research on the "Visual" impact of presentation in Virtual Reality. "Visual" means how the presenter uses the space and interacts with the audience. In order to compare users' pre and post-practice, we evaluate our system by self-assessment and expert-assessment.

Keyword: Virtual Reality, Application, Presentation, Public Speaking, Stage Fright

1. INTRODUCTION

Presentations gradually become common and important in our everyday life. Besides the context and layout of the materials, nonverbal communication, such as volume and speed of voice, eye contacts, stage usage, are key elements of a successful presentation. However, millions of people suffer from performance anxiety, commonly called "stage fright"[1], which causes trembling hands and voice, lacking eye contacts and interactions. One of the best ways to deal with stage fright is to practice more[2]. However, practice is limited to the space and time. Because of lacking of proper places and time to practice, people are seeking to find a solution using technology. Unable to simulate nervousness and presence, traditional platforms, such as computers and phones, failed to give a satisfying solution. With the ability to provide an immersion, virtual reality is therefore being expected to solve this kind of problems[3]. Virtual reality gives a solution, which works well to provide users a presence. And at the same time, unlike other[1] ways, practicing in VR is not dependent on people and scheduling, place availability and costs [1].

According to the research, there are mainly three aspects when defining a good presentation: "Verbal," "Vocal," and "Visual"[4]. Verbal represents words people choose, vocabulary, grammar, etc. Vocal includes tone of voice, pitch range, loudness, etc. Visual includes postures, gestures, eye contacts, etc.[5]. Researchers also identify communication as 55% Visual, 38% Vocal, and 7% Verbal.[5] It indicates that when we make speaking, the words we choose only have a small impact on the presentation quality, the way we convey the message matters much more. Especially the visual effect, for example how we look when we say it, is a dominant element for a good presentation.

1 http://www.webmd.com/anxiety-panic/guide/stage-fright-performance-anxiety#2

2 https://blog.bufferapp.com/what-happens-to-our-brains-when-we-have-stage-fright-the-science-of-public-speaking

3 http://www.vrphobia.com/therapy.htm

4 http://successfully-speaking.com/blog/the-three-vs-of-communication

5 http://www.nickols.us/presenta/ons.pdf

In previous research, there existed many research focusing on "Verbal", using natural language processing to analyze the context of the speech, using image processing to analyze the layout, color, and quantity of the contents on the presentation materials [2]. Some of them are focusing on "Vocal". Researchers do audio analysis to categorize utterance duration, pitch, filled pause, and so on by using the input from headset microphone. They also build a real-time alarm system as a feedback to the user [3]. Due to the difficulty of collecting data, such as stage usage data, eye tracking data, there is a few research focusing on "Visual." In [4], they integrate signals obtained from variety of sources, such as the gestures and voice. Facial expressions are collected as videos by using cameras. The data are analyzed using image processing.

In our research, HTC Vive head-mounted headset is used. Lighthouse, which is HTC Vive's tracking system, is used to collect position data. For eye tracking data, there existed several ways: 1) using eye-tracker glasses, 2) using eye-tracker for VR, and 3) estimation. However, glasses cannot be fit into VR, while eye-tracker (SMI)[1] for VR is very expensive. There is also one specialized VR, FOVE[2], which is said to be the world's first eye tracking VR headset. However, it is still under developing. Therefore, we will collect eye-tracking data by using an estimation method, which we will introduce in Chapter 3.

In our research, we are mainly focusing on Visual. We evaluate the presentation based on the usage of stage, and eye contacts, and how well the presenter interacts with the audience. To our best knowledge, our proposal is the state-of-the-art approach to build presentation rehearsal application in VR.

In the reminder of the paper, Section 2 describes the prior works related to virtual reality and evaluation of good presentation. We present our approach in Section 3. And we conclude the paper and make a discussion on future work in Section 4.

2. RELATED WORK

2.1. Research on Virtual Reality development

In jujunjun110's work[3], they build an A-Frame application, which is a web framework for building virtual reality experiences. It enables users to have highly immersion when viewing in VR headset. Space 3D Scan Camera and Omnidirectional Camera are used. They managed to provide a good immersion. However, it only enables the user to see from one fixed point. Also, interaction between environment and the user is not possible in their work.

In Tarid's work [5], they discussed several technologies in a forensic psychiatry context. They present a work in the development of a multimodal brain-computer interface (BCI) combining neuroimaging and eye tracking with VR devices. In their work, SensoMotric Instrument (SMI) Eye-Tracking Glasses V2.0 is used to track the gaze in the 3D environment. The Volfoni Glasses are synchronized with the VR System's display using a Volfoni ActiveHub IR100. They do not use any VR device, instead, a modified version of Volfoni Active Eyes Glasses is set on top of the SMI ETG glasses to enable stereoscopic vision.

In Nikolaos's work [6], they present a gaze-controlled Multimedia User Interface and developed six applications including mail composing, multimedia viewing for VR headsets. They mentioned that now various companies, such as SensoMotoric Instruments, provide eye-tracking add-on to the Oculus Rift Development Kit 2 (DK2) HMD. Within this, immersive User Interface paradigms embedded in a VR setup controlled via eye tracking can be designed, implemented and evaluated. In their approach, they used twin-CCD binocular eye-tracker by Arrington Research to record the eye-tracking data. And to model good 3D Characters that meet their needs, they developed ARViPL Character Designer, which is an in-house Unity 3D plugin. Together with Singular Inversion FaceGen SDK, it can generate full bodies.

2.2. Research on evaluating presentation

In [2], Kazutaka proposed an automated basic quantitative evaluation of presentation materials. They evaluated by analyzing the quantity of the contents, font size, and how well the font color matches with the background color. The score of font size is given as:

$$\frac{H_m f_p l_{min}}{H_p l f_{min}} \geq 1$$

(1)

where H_m is the Height/Length of the display, H_p is the Vertical resolution of the display, f_p represents the number of pixel of the font, l represents the distance between audience and display, l_{min} equals to 0.4m, and F_{min} equals to 0.003m.

They provide two ways[2] to evaluate how well the font color and background color matches. The first is to compare the Contrast, and the formula is given as follows:

$$0.299 \times R_d + 0.587 \times G_d + 0.114 \times B_d \geq 125 \quad (2)$$

The other is to calculate the color-difference:

1 http://www.smivision.com/en/gaze-and-eye-tracking-systems/products/eye-tracking-htc-vive.html
2 https://www.getfove.com/
3 http://vr-lab.voyagegroup.com

$$\max(R_1, \quad R_2) - \min(R_1, \quad R_2) + \max(G_1, \; G_2) - \min(G_1, \; G_2) + \max(B_1, \; B_2) - \min(B_1, \; B_2) \geq 500 \quad (3)$$

Mathieu et al. [4] compared three feedback strategies to improve public speaking: a non-interactive virtual audience, direct visual feedback, and nonverbal feedback from an interactive virtual audience. They take self-assessment questionnaires, expert assessments, eye contacts, and avoidance of pause fillers into account when evaluate the speaking. And they found that among these three conditions, interactive virtual audience works the best in: (1) increasing engagement and challenges, (2) improving public speaking.

As described above, there exist several research working on helping users improve their speaking based on "Verbal" and "Volume", but there is only a few research on evaluating "Visual", eye contacts, interactions, and usage of the stage due to the difficulties of colleting such data. On the other hand, the innovation of VR, which has the ability to provide an immersion, has a great potential.

In our paper, we will present a novel approach and an application that could help users practice, evaluate, and improve their presentation based on their "Visual" performance.

3. METHODOLOGY

The final goal of this research is to build a virtual reality application that can help users practice and improve their presentation by giving them practical advice based on their "Visual" performance.

The task of our research can be divided as the following steps: 1) Construct a place that meets normal presentation room standards, and build it into Virtual Reality, 2) in order to make it similar to real situations, how to import user's materials into Virtual environment and how to make it to be possible to interact with, 3) since there is yet little research on eye tracking in Virtual Reality, how to estimate eyes' gazing data, 4) how to present an eye tracking heat map, which is aimed to give the users a reference and an advice on eye contacts performance during the presentation.

In this paper, we present our approach solving problems 1) to 3), and we will continue our research on 4). The limitation of this research comes from the inaccurate eye tracking data, since in this paper, we ignore eye movements, and we assume that the head orientation equals to the eyes orientation.

3.1. Device and Platform

In this research, our experiments are implemented on HTC Vive head-mounted headset. Resolution is 2160×1200. Eye sight range is 110°. Game Engine Unity3D 5.5 is used to develop the application. We assume that the classroom, developed by AutoCAD, in VR as shown in Fig.1.

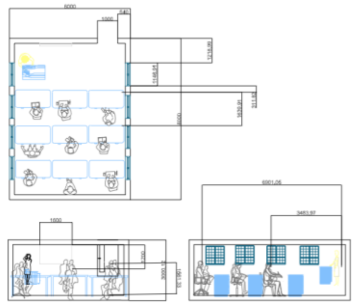

Fig.1 Structure of Classroom in AutoCAD

3.2. Interaction with Materials in VR

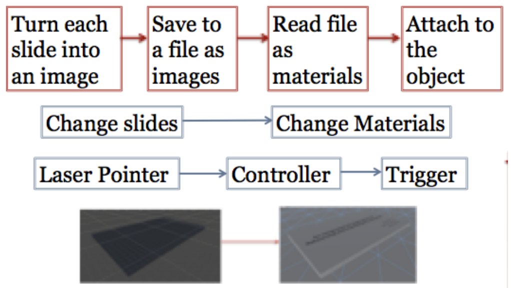

Fig.2 Whole process of importing materials

Users' presentation materials are imported into VR environment as shown in Fig.2. Firstly, we convert all the materials, such as slides, into images, then each image is saved as one file. Then we attach the materials' files to the board object. Secondly, we change materials of the board object to realize the slide-change.

Controllers, which are packaged in HTC Vive, are used to imitate the functions of laser pointer. User can change slides in virtual environment by simply pushing down the triggers on controllers. However, there has one limitation, which is that animations cannot be used any more due to the conversion into pictures.

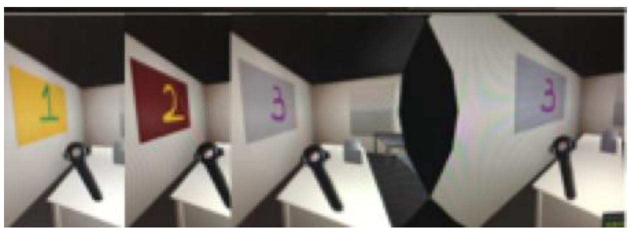

Fig.3 Slide-Change in virtual environment

We could find that the left screen and the right screen are not exactly same. It is due to the phenomenon called “parallax,” based on which our brains can perceive depth.

3.3. Estimation of Eye Tracking Data

Position Data of user’s head can be easily collected by using lighthouse. For eye-tracking data, to simplify the estimation, we assume that the position of the screen center, which means the position of head, is close to the position of eyes, i.e., eye gaze.

The following steps are adopted for eye-tracking.

1) Adding a cursor at the center of the screen in VR as shown in Fig.5 (could be transparent, blue here for explanation).

2) Collecting all the positions where the cursor and the objects in VR collide, which is similar to the collision of eye gaze and objects. For example, if the cursor (user’s eye gaze) collides with the square, which is one of the objects, the position of the square is collected as shown in Fig.4.

3) Refining the results by further dividing each object into several parts, i.e., fine grain level. When the cursor (user’s eye gaze) collides with any part of an object, this time instead of collecting the position of the whole objects, now we only collect the position data we of the part being watched. For example, as shown in Fig.5, we divide the square into three parts, and the user is looking at the left side. Then we can obtain and collect the position of the left side part,

4) Dividing the objects further into more parts until units in Unity3D.

```
Show over state
UnityEngine.Debug:Log(Object)
(-0.5, 1.5, 3.0)
UnityEngine.Debug:Log(Object)
```

Fig.4 Mirror screen from VR

As shown in Fig.4, if there is a collision of cursor (eye gaze) and block (object)

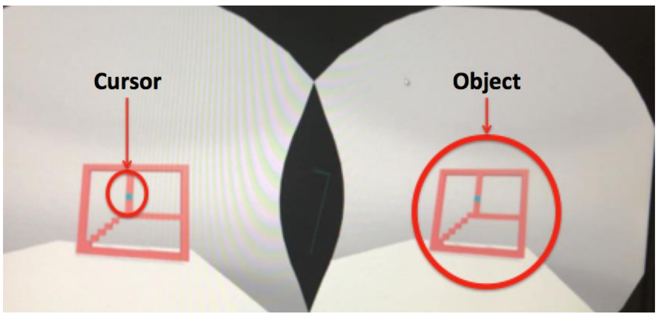

Fig.5 Output of collision positions

3.4. Interaction with audience

In this paper, we use balls to represent audience, temporary. Our goal is to examine whether the presenter could give enough eye contacts with each audience. Currently, when the cursor, which means his/her eye gaze, and the ball collide, the ball will be changed to another material to show him/her that the system detect your eye-gaze. This is for the situation for example when there is some student sitting in the corner sleeping, whether the professor could make a notice or give a warn is a very important element when evaluating his teaching performance. In the future, we will change balls to human models, and change materials to animations, such as sleeping, listening, and so on.

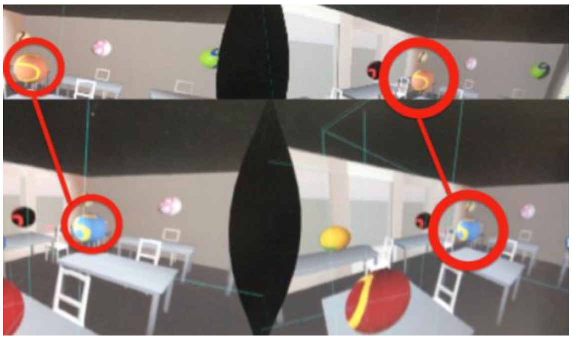

Fig.6 Material change when collisions occur

3.5. Eye Tracking Heat Map

To generate an eye-tracking heat-map, we are now considering two approaches: 1) using a third-party heat map, 2) changing the color and transparency of the transparent blocks when collision occurs. The former may require some additional work. While if adopting the latter, we may need to make many more blocks but it could be a pretty good real-time feedback.

3.6. Evaluation of the presentation

We evaluate the whole presentation by analyzing the eye tracking data, whether the presenter gives enough eye contacts to every corner of the room, and stage usage, whether the presenter has made good use of the space. However, we have not decided the weight of the parameters.

4. CONCLUSION and FUTURE WORK

In this paper, we present a VR application for presentation rehearsal and improvement. We manage to import users’ presentation materials into virtual environment, and we also manage to detect the approximate position that the user is looking at. In order to improve the system, there are also some future works to do. We need to do more survey and try to implement 3D eye-tracking heat map in Virtual Reality. We also need to do some test and assign proper value of weight to each parameter when evaluating the presentation. Lastly, instead of using balls, we need to find or make 3D audience models and animations.

REFERENCE

[1] Mathieu Chollet, Torsten Wortwein, Louis-Philippe Morency, Ari Shapiro, Stefan Scherer, *Exploring Feedback Strategies to Improve Public Speaking: An Interactive Virtual Audience & Framework,* UbiComp'15, 2015

[2] 栗原 一貴, 加藤 公一, 大浦 弘樹, *SlideChecker: プレゼンテーション資料の基礎的な 定量的自動評価手法*, 2009

[3] Kazutaka Kurihara, Masataka Goto, Jun Ogata, Yosuke Matsusaka, Takeo Igarashi, *Presentation Sensei: A Presentation Training System using Speech and Image Processing,* ICMI'07

[4] Mathieu Chollet, Torsten Wortwein, Louis-Philippe Morency, Ari Shapiro, Stefan Scherer, *Exploring Feedback Strategies to Improve Public Speaking: An Interactive Virtual Audience Framework,* UbiComp'15, 2015

[5] Tarid Boukhalfi, Christian Joyal, Stephane Bouchard, Sarah Michelle Neveu, Patrice Renaud, *Tools and Techniques for Real-time Data Acquisition and Analysis in Brain Computer Interface studies using qEEG and Eye Tracking in Virtual Reality Environment*, International Federation of Automatic Control Hosting by Elsevier Ltd. 2015

[6] Nikolaos Sidorakis, George Alex Koulieris, Katerina Mania, *Binocular Eye-Tracking for the Control of a 3D Immersive Multimedia User Interface*, IEEE, 2015s as References. Instead, if you want to refer such web contents, just put them at footnote.