Improvement of Public Speaking Skills using Virtual Reality: Development of a Training System

Sarah Saufnay
QuantOM, HEC Liège
University of Liège
Liège, Belgium
sarah.saufnay@uliege.be

Elodie Etienne
QuantOM, HEC Liège
University of Liège
Liège, Belgium
elodie.etienne@uliege.be

Michaël Schyns
QuantOM, HEC Liège
University of Liège
Liège, Belgium
m.schyns@uliege.be

***Abstract*—The *Public Speaking* project is a Virtual Reality (VR) training system dedicated to the improvement of public speaking skills. It provides users with the opportunity to rehearse their speech in a realistic virtual environment which displays an interactive audience. The experience is customized based on the chosen training scenario, which is determined by the room, participant's role, audience behaviors, and timer. A companion website provides comprehensive feedback reports on speakers' performance, based on their verbal and nonverbal behavior, as well as replay functionalities. In addition to its educational objective, this system is used as a tool to advance research into the development of effective public speaking training environments.**

***Index Terms*—Public speaking, Training, Virtual Reality, Virtual Agents**

I. Introduction

Mastering public speaking skills is essential in both our personal and professional lives. However, a significant portion of the population still struggles with such social interactions, which in turn affects their daily lives [4] [6]. Besides, people interested in attending training programs to purely improve their speaking capabilities face a lack of dedicated solutions.

In this context has emerged the *Public Speaking* project, a Virtual Reality (VR) system developed to provide individuals, especially students, with an accessible and comprehensive training solution. By immersing the user in an environment displaying a responsive virtual audience, VR provides speakers with the opportunity to practice their speech in lifelike conditions [1] [5], which consequently fosters skills improvement and stress reduction [8] [10].

Other training systems exist, such as *Ovation* [7] and *Virtual Speech* [9], but do not disclose the models used to evaluate speakers' performance and do not allow the implementation of specific performance indicators. The development of a transparent and comprehensive system was then preferred.

II. The Public Speaking application

The application provides a multitude of training opportunities, allowing users to engage in VR simulations of classic public speaking scenarios. Although currently composed of six training rooms (*i.e., the Auditorium, the Boardroom, the Classroom, the Courtroom, the Meeting Room*, and *the Office*), the application can be repurposed in a multitude of contexts given the broad spectrum of the public speaking challenge. The project is nevertheless constantly evolving, regularly integrating new rooms, functionalities, and the results of the authors' latest scientific research. A demonstration of the VR application is available here: https://youtu.be/PIETuo4_UHs.

A. Devices

The application has been developed to run on autonomous Meta Quest VR headsets, leveraging hand tracking technology and enabling users to move around freely, as they would in real-life presentations. An internet connection is however required for recording and linking the session with the companion website developed along with the application (see *Section II-G*).

B. Methodology

The application was developed by a team of 3D artists, animators and technical artists from the SIG AR/VR Lab from the University of Liège, employing a variety of software, including *Unity, Blender*, and *Maya*. In particular, the creation and animation of the virtual agents were carried out using the *Reallusion* suite. In addition, the rooms and their associated functionalities have been designed in collaboration with field experts from a range of related disciplines in order to perfectly grasp the challenges inherent to the considered use cases.

C. System design

Through a dedicated interface, users are able to tailor the training experience according to their needs by selecting:

- The **room**, which corresponds to a specific public speaking context.
- The **role**, which corresponds to the assigned role in the chosen public speaking context.
- The **training scenario**, which is based on the educational objectives associated with the public speaking task. The scenarios vary in terms of:
  - The **virtual audience attitudes**, either associated with a positive, negative or neutral perception of agents' nonverbal behavior.
  - Specific **sequences of events**, which may comprise a combination of audience verbal and nonverbal behavior, as well as disturbance sounds.
  - The **allotted time** for speaking.

979-8-3315-1645-1/24/$31.00

Whenever the training requires further customization, scenario codes corresponding to specific sequence of audience reactions can be generated via the companion website. These codes can then be entered during the training setup process.

A final interface allows users to enter their personal identification code to link the training session with their profile on the website. In this way, the experience can be enriched by incorporating the users' slides and notes uploaded via the dedicated web page into the VR experience. Personal session feedback reports can also be viewed on their profile after the training (see *Section II-F*).

D. The virtual audience

The audience displayed in the present application is constituted by photo-realistic avatars of different gender, ethnic origin, and age, whose behaviors are derived from a perceptual study carried out in the scope of this project. The attitudes and behaviors of the system's audience have indeed been meticulously crafted based on the findings of [3], which investigated the perception of the virtual agents' valence and arousal. This study resulted in a comprehensive library of virtual agents' non-verbal behaviors, classified as either positive, negative, or neutral perceived attitudes, and validated by 125 participants in VR. The implemented audience attitudes were previously investigated in [2] but using non-immersive devices. They combine facial expressions, head movements, and postures.

The influence of realism was investigated as well in [3]. The use of photo-realistic agents did not result in the uncanny valley effect. Furthermore, the participants exhibited greater confidence in characterizing the attitudes of photo-realistic agents compared to cartoon ones. Photo-realistic agents were therefore preferred in the present system.

It is noteworthy to highlight that the environment presented a high level of presence as presented in [3], but should be evaluated as well under other training configurations.

E. Speaker's performance indicators

With the user's consent, training sessions are recorded. Specifically, the audio as well as the hands, head, and eye movements are tracked, allowing for an analysis of the speaker's performance and providing them with personalized feedback for further improvement.

Specific Key Performance Indicators (KPIs), including both verbal and non-verbal cues, were included as part of the proposed feedback. They currently encompass head and hand movements, gaze, fundamental frequency, and voice intensity. Multimodal cues have been preferred since recognized as best predictors of speakers' performance, when compared to solely verbal or nonverbal signals, as shown by [11].

The list of KPIs incorporated into the system is currently limited, but will be expanded based on previous research conducted in this field to provide users with comprehensive feedback reports. Although fundamental to analyze performance and detect stress, the currently implemented KPIs are not enough for the creation of an accurate performance assessment model. Context-related KPIs will also be added to further enrich the feedback.

F. Feedback and replay

Provided that the user has accepted the training session to be recorded, a performance report is available on the associated website after each training session. These reports display the evolution of the speaker's KPIs throughout the training session, along with a transcription of the speech. Additionally, the recorded audio can be played back.

Users have the opportunity to replay their speaking task as well, either on the website or directly in the VR environment by putting the headset back on. They can decide the location within the room from which they will observe their presentation, as well as the gender of the avatar that will assume their role. The audio recording is played back simultaneously with the gestures made during the presentation, and mapped onto the avatar. Extensive work has been dedicated to the optimization of the Inverse Kinematic (IK) system, in order to accurately reproduce users' movements onto the avatar.

G. Associated website

The website is accessible to all registered users, who may be either be students or teachers, although certain functionalities are reserved for teachers alone. As a student, notes and slides can be uploaded using the website, to be later synchronized with the VR experience. Personal feedback reports are accessible as well. In addition to the aforementioned functionalities, teachers are able to create groups of students. Consequently, they will have access to the details of students' training sessions, along with their respective feedback reports. Moreover, specific training scenarios can be created to better meet the students' needs by creation of specific codes.

III. Contributions and future work

As the project is in continuous development, improvements and additional features will be regularly incorporated in the system. A current limitation identified is the restricted number of scenarios provided, which results in repeated training situations. To address this, Artificial Intelligence and Machine Leaning algorithms will be integrated to obtain an autonomous system, where virtual agents automatically react to the speaker's speech performance, evaluated at runtime. The development of complex audience and performance assessment models are therefore required. Additionally, the effectiveness of the training system is currently being investigated, with comparisons being made to other training settings. A multi-user functionality is under development as well and will be incorporated into the system to enable users to embody a virtual avatar and directly influence the training scenario. This encompass further optimization of the IK system. In addition, the feeling of presence and immersion will continue to be assessed in the different training configurations of the system.

To conclude, this project provides a solid foundation for research, and has already been employed to conduct a perceptual study and validate the associated audience. Moreover, the environment has been tailored to meet the authors' research objectives, thus facilitating further investigations.

REFERENCES

[1] M. Bachmann, A. Subramaniam, J. Born, and D. Weibel, "Virtual reality public speaking training: effectiveness and user technology acceptance", Frontiers in Virtual Reality, vol. 4, September 2023.

[2] M. Chollet and S. Scherer, "Perception of virtual audiences", IEEE Computer Graphics and Applications, vol. 37, no. 4, pp. 50-59, July 2017.

[3] E. Etienne, A.-L. Leclercq, A. Remacle, L. Dessart, and M. Schyns, "Perception of avatars nonverbal behaviors in virtual reality", Psychology & Marketing, vol. 40, pp. 2464–2481, July 2023.

[4] T. Furmark, M. Tillfors, P.-O. Everz, I. Marteinsdottir, O. Gefvert, and M. Fredrikson, "Social phobia in the general population: prevalence and sociodemographic profile", Social Psychiatry and Psychiatric Epidemiology, vol. 34, pp. 416–424, September 1999.

[5] Y. Glémarec, J.-L. Lugrin, A.-G. Bosser, C. Buche, and M. E. Latoschik, "Controlling the stage: a high-level control system for virtual audiences in virtual reality", Frontiers in Virtual Reality, vol. 3, May 2022.

[6] P. Jefferies and M. Ungar, "Social anxiety in young people: a prevalence study in seven countries", PLoS ONE, vol. 15, no. 9, September 2020.

[7] "Ovation VR." https://www.ovationvr.com/ (accessed Jul. 20, 2024).

[8] R. Reeves, D. Curran, A. Gleeson, and D. Hanna, "A meta-analysis of the efficacy of virtual reality and in vivo exposure therapy as psychological interventions for public speaking anxiety", Behavior Modification, vol. 46, no. 4, pp. 937-965, July 2022.

[9] "Virtual Speech." https://virtualspeech.com/ (accessed Jul. 20, 2024).

[10] T. Wörtwein, L.-P. Morency, and S. Scherer, "Automatic assessment and analysis of public speaking anxiety: a virtual audience case study", 2015 International Conference on Affective Computing and Intelligent Interaction, China, pp. 187-193, December 2015.

[11] T. Wörtwein, M. Chollet, B. Schauerte, L.-P. Morency, R. Stiefelhagen, and S. Scherer, "Multimodal public speaking performance assessment", Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, USA, pp. 43–50, November 2015.