AudienceAR - Utilising Augmented Reality and Emotion Tracking to Address Fear of Speech

Philipp Hartl
University of Regensburg
Regensburg, Germany
philipp1.hartl@stud.uni-regensburg.de

Thomas Fischer
University of Regensburg
Regensburg, Germany
thomas1.fischer@stud.uni-regensburg.de

Andreas Hilzenthaler
University of Regensburg
Regensburg, Germany
andreas.hilzenthaler@stud.uni-regensburg.de

Martin Kocur
University of Regensburg
Regensburg, Germany
martin.kocur@ur.de

Thomas Schmidt
University of Regensburg
Regensburg, Germany
thomas.schmidt@ur.de

ABSTRACT

With Augmented Reality (AR) we can enhance the reality by computer-generated information about real entities projected in the user's field of view. Hence, the user's perception of a real environment is altered by adding (or subtracting) information by means of digital augmentations. In this demo paper we present an application where we utilise AR technology to show visual information about the audience's mood in a scenario where the user is giving a presentation. In everyday life we have to talk to and in front of people as a fundamental aspect of human communication. However, this situation poses a major challenge for many people and may even go so far as to lead to fear and and avoidance behaviour. Based on findings in previous work about fear of speech, a major cause of anxiety is that we do not know how the audience judges us. To eliminate this feeling of uncertainty, we created an AR solution to support the speaker while giving a speech by tracking the audience's current mood and displaying this information in real time to the speaker's view: *AudienceAR*. By doing so we hypothesise to reduce the speaker's tension before and during presentation. Furthermore, we implemented a small web interface to analyse the presentation based on the audience mood after the speech is given. Effects will be tested in future work.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; *Displays and imagers.*

KEYWORDS

affective computing, augmented reality, emotion, facial, recognition, hololens

1 INTRODUCTION

The fear of public speaking has a lifetime prevalence of 15 to 30 % in general population [16]. Burnley et al. states that, "Approximately 85 percent of the general population report experiencing some level of anxiety about speaking in public" [6]. This frequently occurring disorder as a part of social phobia make those affected having a feeling of intense anxiety not only immediately before a fearful situation, but also at the mere imagination of such a situation [9]. Due to the rapid progress of technology and the affordability of hardware in recent years, immersive computer-mediated interventions have been on the rise in the domain of eMental-health. Shown in previous work, VR and AR applications are evidence based therapy tools for different kind of phobias and disorders [4, 7, 17]. Hence, "Clinical VR" [14] is used as treatment modality for fear of public speaking by immersing the patients in a virtual environment and confront them virtually with fearful stimuli in a exposure therapy. In [9] participants give a speech to a virtual audience. The auditorium is filled gradually with avatars intensifying the fearful stimuli in a controlled manner. Anderson et al. do not change the number of listeners, however they vary the reaction of the audience (e.g. interested, bored, applause) in a virtual auditorium scenario [1]. Results from both indicate that the in-virtuo exposure therapy shows significant decrease in public speaking anxiety. Although the patients are

confronted with virtual stimuli results imply that VR exposures provoke similar anxiety levels like in vivo exposures [8]. Furthermore, related work propose exposure therapies with AR as treatment modality. One study demonstrated an AR therapy application where participants with insect phobia are exposed to cockroaches and improved significantly after treatment [3]. Similar results are shown in a study where AR is applied to treat arachnophobia [10]. Findings indicate that AR can evoke a sufficient sense of presence as a requirement for triggering the same emotions and reactions as a real experience would [13]. In contrast to VR, we can utilise AR to let the user interact with real fearing stimuli having a natural sense of presence in the real world augmented with digital entities. Hence, the patients are exposed to real entities (in our case, real people) like in in-vivo exposure enhanced with digital information [4]. As shown in [18], real communication partners may lead to different outcome of exposure therapy.

In this demo paper we address the fear of public speaking by taking advantages of the possibilities of AR to enhance the reality with additional information. We track the emotional state of an audience and display an emotional distribution in the user's viewport. As the major cause of anxiety and nervousness is that we do not know how the audience evaluates us [11, 12], we want to alleviate this feeling of uncertainty by first tracking the facial expressions of each individual listener in a customisable interval and calculating an overall audience emotional distribution in real time. Inspired by [15], we hypothesise to reduce the presenter's anxiety before and during a presentation by providing continuous information about the emotional state of the audience.

2 METHODS

In the following section we briefly describe which user groups we have identified and how our application AudienceAR is structured and implemented.

Use Cases

While it is conceivable for almost everybody to use our application, because presenting information is an important task in everyday life, two groups of this rather heterogeneous group are particularly relevant. The first one being professional presenters such as managers or consultants who are interested in optimising their presentation performance. Our second target group consists of novice speakers who give speeches very irregularly thus having little experience and often suffering from a fear of public speaking. They are primarily interested in enhancing their self-confidence and alleviate their existing anxiety during a presentation or speech. By giving them an objective measurement of the crowd's current emotions we want to fight their negative bias towards their own current performance.

Application Structure

Our application is divided into two parts. The first one being a live representation of the audience's emotions on the Microsoft Hololens[1], while the second one is a web application for analysing the presentation afterwards (see figure 3).

Live presentation. Before presentation the user has a few options to customise his/her experience. It is either possible to attach the live emotions to the user's viewport or place it inside the AR environment which corresponds with a real place inside the presentation room. This enables the user to either be always aware of the audience's emotions or have the opportunity to utilise it on demand. Additionally, it is possible to manually take a snapshot by using the air tap gesture of the MS Hololens, however this might be distracting and get tiresome quickly, so there is an option to automatically analyse the emotions in three different intervals (see figure 1).
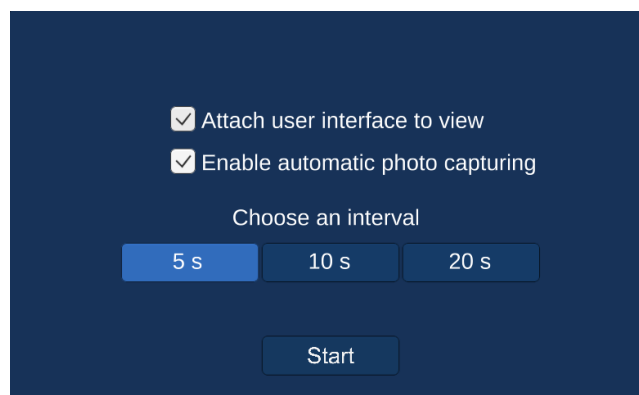

Figure 1: Live interface configuration

To minimise potential distraction while presenting we decided to use a rather simple method to represent the measured emotions - a stacked bar chart (see figure 2). Before we present the emotions to the user we calculate the arithmetic mean, if applicable, to get an average representation of all individuals in the audience. By using the double tap gesture it is also possible to hide the user interface during a presentation completely if desired.

While every emotion has a simple colour representation we decided to reduce the user's cognitive load by adding a small legend indicating which two emotional dimensions are present the most at the moment. To prevent a heavily split chart we also decided to combine all emotions detected between 0 - 5% into the category *other* emotions.

---

[1]https://www.microsoft.com/en-us/hololens (note: all URLs mentioned in this article were last visited June 21, 2019)
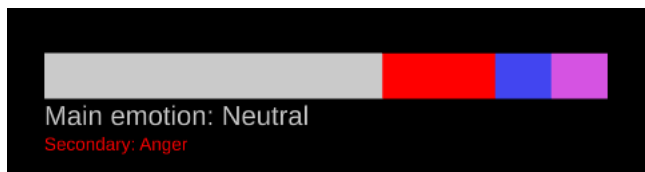
Figure 2: Live emotional distribution

Follow-up. After a presentation we provide a web page to further explore the emotional course throughout the presentation. This documentation enables to pinpoint critical moments which can be either in a positive (e.g. peak in happiness) or negative (e.g. peak in anger) way.
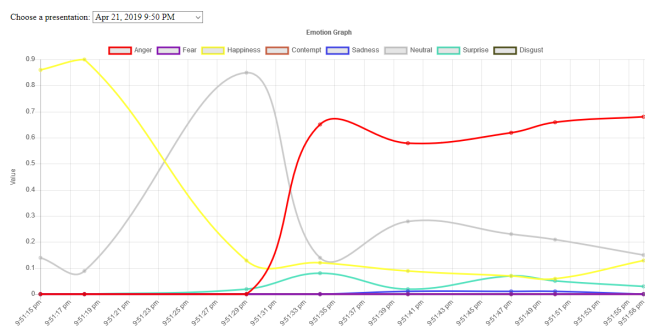

Figure 3: Web application

Our web tool provides live feedback during presentation. This feature is especially useful when used by a therapist who can either investigate while a presentation is running (the chart is updated during a presentation) or afterwards in combination with his observations.

Implementation

We created the application by using the Unity3D game engine[2] in combination with the AR platform Vuforia[3] and deployed it specifically for the Mircosoft Hololens. During a presentation we use Hololens's main standard camera to take a picture of the present audience either automatically or on demand depending on the users preferences. We upload this picture to our own server which then forwards it to Mircosoft Cognitive Services[4]. Microsoft Cognitive Services performs the facial emotion recognition and delivers the metrics we use in our application. We store the JSON values received from Microsoft Cognitive Services in a database and plot them into a graph for further investigation during or after a presentation. The communication is done by using HTTPS to ensure no sensitive data can be leaked by third-parties.

---

[2]https://unity.com

[3]https://www.ptc.com/en/products/augmented-reality

[4]https://azure.microsoft.com/en-us/services/cognitive-services/

Furthermore no images or individual emotional data is stored because of privacy concerns. Only averaged audience scores are available.

3 LIMITATIONS

Since the MS Hololens provides a two megapixel video camera without flash, lightning conditions play an important role for capturing the audience. We found natural daylight to suite best. The application cannot be used properly in dark rooms (or only to a very limited extent).

Based on those limitations, the distance between audience and presenter has to be considered. We recommend a distance of max. five meters to track the audience. Generally, a smaller audience size is preferred as the field the of view is restricted to 67 degrees maximum.

However, our implementation does not rely on detecting all persons at all times as we only calculate the average emotion in the audience. We informally tested the prototype with ten participants in a teacher-student classroom scenario in which the participants (sitting approximately three meters away from the potential speaker) play with facial expressions to test the emotion detection performance. We found a crowd of about five to ten people within a distance of a few meters to be working best, so that the application could detect the overall emotion of the audience (happiness, neutral, surprise, sadness or a combination of those). We have yet to investigate the application in a realistic presentation scenario with a podium and a real audience to provide information about the usability and the performance.

4 CONCLUSION

We presented the motivation and implementation for AudienceAR, an AR-application that tracks the emotional state of an audience for presenters. Please note, since this is a work-in-progress project we have yet to evaluate the overall idea and the effect on potential users. We are currently planning first studies for our prototype about the acceptance of the idea and to gather feedback on how to design and to place the AR-elements as advantageous as possible. We also need to include potential audiences into the evaluation process since it is important to evaluate how the usage of our application changes the impression of a presenter and if this usage is at all accepted. We also work close with psychologists to design studies to evaluate if this tool can indeed reduce the fear and nervousness during presentations or if it does indeed worsen the situation.

Furthermore, we also plan to broaden (1) our potential target group and (2) the potential use cases. We are currently focused on the area of E-Health and especially persons with anxiety problems during presentations. Another group we hypothesise can be supported by our tool are persons with social-emotional agnosia [5] . People with this disorder are

not able to perceive and understand human emotions expressed via body language, the face or the voice. The disorder is also known to co-occur with autism [2]. AudienceAR can support those people to understand human emotion also outside of the presentation context in their everyday life. Outside of the use case of E-Health, we also suggest that presenters without any major anxiety problems during presentations can profit from our tool since they can better assess the current state of their audience e.g. if they are bored or excited. The presenter can use this information to react during the presentation on this feedback.

Please note, that we are also aware of the ethical dimensions of this project. Of course filming and analysing an audience during presentation cannot be done without their clear consent, so it is necessary for potential users of this application to keep this in my mind. In the same context, it might also not be possible to use this application in everyday life although we see benefits for the target group of persons with social-emotional agnosia. However, we also want to continue the research for this user group and see potential to use our application in specific use cases, in which they can acquire the consent of the persons filmed.

Overall, we envision a lot of use cases for our application idea and we will continue our research to optimise our prototype and address further user groups.

REFERENCES

[1] Page L. Anderson, Elana Zimand, Larry F. Hodges, and Barbara O. Rothbaum. 2005. Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and Anxiety* 22, 3 (2005), 156–158. https://doi.org/10.1002/da.20090

[2] George Mihai Bădescu, Mădălina Fîlfan, Raluca Elena Sandu, Roxana Surugiu, Ovidiu Ciobanu, and Aurel Popa-Wagner. 2016. Molecular mechanisms underlying neurodevelopmental disorders, ADHD and autism. *Rom J Morphol Embryol* 57, 2 (2016), 361–366.

[3] Cristina Botella, Juana Bretón-López, Soledad Quero, Rosa Baños, and Azucena García-Palacios. 2010. Treating Cockroach Phobia With Augmented Reality. *Behavior Therapy* 41, 3 (2010), 401 – 413. https://doi.org/10.1016/j.beth.2009.07.002

[4] Cristina Botella, M. Ángeles Pérez-Ara, Juana Bretón-López, Soledad Quero, Azucena García-Palacios, and Rosa María Baños. 2016. In Vivo versus augmented reality exposure in the treatment of small animal phobia: A randomized controlled trial. *PLoS ONE* 11, 2 (2016), 1–22. https://doi.org/10.1371/journal.pone.0148237

[5] Paula Brown. 2017. Facial expressions: understanding the social information code. *Early Years Educator* 19, 4 (2017), 39–44.

[6] M. Caroline E. Burnley, Patricia A. Cross, and Nicholas P. Spanos. 1993. The Effects of Stress Inoculation Training and Skills Training on the Treatment of Speech Anxiety. *Imagination, Cognition and Personality* 12, 4 (1993), 355–366. https://doi.org/10.2190/N6TK-AR8Q-L4E9-0RJ0 arXiv:https://doi.org/10.2190/N6TK-AR8Q-L4E9-0RJ0

[7] Tom K.J. Craig, Mar Rus-Calafell, Thomas Ward, Julian P. Leff, Mark Huckvale, Elizabeth Howarth, Richard Emsley, and Philippa A. Garety. 2017. AVATAR therapy for auditory verbal hallucinations in people with psychosis: A single-blind, randomised controlled trial. *The Lancet Psychiatry* 0366, 17 (2017). https://doi.org/10.1016/S2215-0366(17)30427-3

[8] P Emmelkamp, M Krijn, H Hulsbosch, and S de Vries. 2002. Virtual Reality Treatment versus exposure in vivo: A Comparative Evaluation in Acrophobia. *Behaviour Research & Therapy* 40 (2002), 509–516.

[9] Sandra R. Harris, Robert L. Kemmerling, and Max M. North. 2003. Brief Virtual Reality Therapy for Public Speaking Anxiety. *CyberPsychology & Behavior* 5, 6 (2003), 543–550. https://doi.org/10.1089/109493102321018187

[10] M. Carmen Juan and Jérôme Calatrava. 2011. An augmented reality system for the treatment of phobia to small animals viewed via an optical see-through HMD: Comparison with a similar system viewed via a video see-through HMD. *International Journal of Human-Computer Interaction* 27, 5 (2011), 436–449. https://doi.org/10.1080/10447318.2011.552059

[11] James C McCroskey. 1984. The communication apprehension perspective. *Avoiding communication: Shyness, reticence, and communication apprehension* (1984), 13–38.

[12] James C. McCroskey. 2006. Oral Communication Apprehension: A Summary of Recent Theory and Research. *Human Communication Research* 4, 1 (03 2006), 78–96. https://doi.org/10.1111/j.1468-2958.1977.tb00599.x arXiv:http://oup.prod.sis.lan/hcr/article-pdf/4/1/78/22344472/jhumcom0078.pdf

[13] Giuseppe Riva and Fabrizia Mantovani. 2012. Being There: Understanding the Feeling of Presence in a Synthetic Environment and its Potential for Clinical Change. *Virtual Reality in Psychological, Medical and Pedagogical Applications* (2012), 3–34. https://doi.org/10.5772/46411

[14] Albert Skip Rizzo. 2017. Is Clinical Virtual Reality Ready for Primetime? *American Psychological Association* (2017), 1–42. https://doi.org/10.1037/neu0000405

[15] Albrecht Schmidt, Philipp Eiglsperger, Mariam Hassib, Stefan Schneegass, Niels Henze, and Florian Alt. 2017. EngageMeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography. (2017), 5114–5119. https://doi.org/10.1145/3025453.3025669

[16] Vickram Tejwani, Duc Ha, and Carlos Isada. 2016. Public Speaking Anxiety in Graduate Medical Education—A Matter of Interpersonal and Communication Skills? *Journal of Graduate Medical Education* 8, 1 (2016), 111–111. https://doi.org/10.4300/JGME-D-15-00500.1 arXiv:https://doi.org/10.4300/JGME-D-15-00500.1

[17] Sara Ventura, Rosa M Baños, and Cristina Botella. 2018. Virtual and Augmented Reality: New Frontiers for Clinical Psychology. In *State of the Art Virtual Reality and Augmented Reality Knowhow*. IntechOpen.

[18] Helene S. Wallach, Marilyn P. Safir, and Margalit Bar-Zvi. 2009. Virtual Reality Cognitive Behavior Therapy for Public Speaking Anxiety: A Randomized Clinical Trial. *Behavior Modification* 33, 3 (2009), 314–338. https://doi.org/10.1177/0145445509331926 arXiv:https://doi.org/10.1177/0145445509331926 PMID: 19321811.