Toward Affective Social Interaction in VR

I first encountered VR in the late 1990s, as a researcher looking at how it provided engineers and designers an environment for prototyping. After that I became more interested in looking at how to augment reality and our surrounding environment. However, although VR had been around for decades by that point, there were many aspects, in particular from an interaction point of view, that I felt still deserved investigation. VR is enjoying renewed interest thanks to the recent proliferation of consumer products, content, and applications. This is accompanied by unprecedented interest from consumers and by the maturity of VR, now considered less and less to be hype and to be more market ready. However, important challenges remain, associated with dizziness and the limits of current wearable displays, as well as interaction techniques. Despite these limitations, application fields are flourishing in training, therapy, and well-being beyond the more traditional VR fields of games and military applications.

One of the most ambitious research goals for interactive systems is to be able to recognize and influence emotions. Affect plays an important role in all we do—it is in fact an essential aspect of social interaction. So the study of *affective social interaction* in VR can be important to the above-mentioned fields to support mediated communication. For example, in mental or psychological disorders, VR can be used for interventions and training to monitor patient engagement and emotional responses to stimuli, providing feedback and correction on particular behaviors. Moreover, VR is increasingly accepted as a research platform in psychology, social science,

and neuroscience as an environment that helps introduce contextual and dynamic factors in a controllable way. In such disciplines, affect recognition and synthesis are critical to many investigated phenomena.

MULTIMODAL SYNTHETIC AFFECT IN VR

In social interaction, emotions can be expressed through gestures, posture, facial expressions, speech and its acoustic features, and touch. Our sense of touch plays a large role in establishing how we feel and act toward another person, considering, for example, hostility, nurturance, dependence, and affiliation.

Having done work on physiological and affective computing and haptics separately, I saw a unique way to combine these techniques to develop synthetic affect in VR, combining different modalities. For example, the emotional interpretation of touch can vary depending on cues from other modalities that provide a social context. Facial expressions have been found to modulate the touch-perception [1] and post-touch-orienting response. Such multimodal affect stimuli are perceived differently according to individual differences of gender and behavioral inhibition. For example, behavioral inhibition sensitivity in males was associated with stronger affective touch perception [2].

Taking facial expressions and touch as modalities for affective interaction, we can uncover different issues in their production. Currently, emotional expression on avatars can be produced using off-the-shelf software that analyzes the facial movements of an actor modeling basic expressions, head orientation, and gaze. Subsequently, the descriptions are used to animate virtual

characters. Emotional expressions in avatars are often the result of a process involving several steps ensuring that these relate to the intended emotions. The expressions are recorded first by capturing the live presentation from a professional actor, using a facial-tracking software that also animates a virtual character. Expressions can then be manually adjusted to last for the same amount of time and end with a neutral expression. Different animations are created for each distinct emotion type. The expressions can then be validated by measuring the recognition accuracy of participants who watch and classify the animations. This process works well to customize the facial expression to the intended use in replicable experiments. But this is resource-intensive and does not scale well for other uses where facial expression might need to be generated in greater variations (for expressing nuances) or for generalizability, since every expression is unique. While mediated touch has been proven to affect emotion [1,3], behavior research into the deployability, resolution, and fidelity of haptics is ongoing. However, in our recent studies, we compared several techniques to simulate a virtual hand of a character touching the hand of a participant [4].

Emotion tracking is more challenging in the case of a wearable VR headset, as facial expressions cannot be easily tracked through recent computer-vision software [5]. For recognizing change in psychophysiological states or to assess emotional responses to particular stimuli, physiological sensors can be used. These sensors are also being integrated into more and more off-the-shelf consumer products, as in the case of electrodermal activity (EDA) or electroencephalogram (EEG). While EDA provides a way to track

Figure 1. Bringing it all together: hand tracking of the user through a glass. Wearable haptics, an EEG cap, and an HMD for VR allow the simulation of a situation in which a person sitting in front of the user touches her hand with different facial expressions [1,2,4].

arousal among other states and is easy to use unobtrusively, it is suited for changes in the order of minutes, not for time-sensitive events in seconds and milliseconds. EEG, on the other hand, increasingly provided in commercial devices, is better suited for time-precise measurements. For example, the study

of how emotional expressions modulate the processing of touch can be done by event-related potential (ERP) in EEG resulting from touch. Studies show that the use of EEG is compatible with commonly available HMDs. Eye tracking, which recently appeared commercially to be used

inside HMDs, can be used both to identify whether users attend to a particular stimulus to track its emotional response, and to track psychophysiological phenomena such as cognitive load and arousal. As an example, the setup in Figure 1 includes VR, haptics, and physiological sensors.

Figure 2. RelaWorld using VR and physiological sensors (EEG) for a neuroadaptive meditation environment [6].

Association for
Computing Machinery

ACM Conference Proceedings Now Available via Print-on-Demand!

*Did you know that you can
now order many popular
ACM conference proceedings
via print-on-demand?*

Institutions, libraries and individuals can choose from more than 100 titles on a continually updated list through Amazon, Barnes & Noble, Baker & Taylor, Ingram and NACSCORP: CHI, KDD, Multimedia, SIGIR, SIGCOMM, SIGCSE, SIGMOD/PODS, and many more.

For available titles and
ordering info, visit:
librarians.acm.org/pod

► BLOG@IX

It can be used to simulate a social interaction at a table where mediated multimodal affect can be studied while an avatar touches the user's hand, at the same time delivering a facial expression. The user recognizes the virtual hand in Figure 1A as her own hand, as it is synchronized in real time.

This setup can be used for a number of training, entertainment, or therapy purposes. For example, a recent product applies VR for treating anxiety patients. Recent studies have evaluated the impact in training autism spectrum disorder patients to apply this to dealing with anxiety. In our own recent study, we used the same setup for an air hockey game. The haptics simulated the hitting of the puck, and the emotional expression of the avatar allowed us to study effects on players' performance and experience of the game.

FUTURE STEPS

VR devices, applications, and content are emerging quickly. An important feature in the future will be the affective capability of the environment, including the recognition and synthesis of emotions.

A variety of research challenges exist for affective interaction:

- Techniques in recognizing users' emotions from easily deployable sensors, including the fusion of signals. Physiological computing is advancing fast in research and in commercial products. Vision-based solutions that track facial expression have recently seen success; physiological-based sensors could soon follow suit.

- Synthesis of affect utilizing multiple modalities, as exemplified here. For example, combining touch and facial expression, but considering also speech and its acoustic features and other nonverbal cues. How to ensure that these multimodal expressions are generally valid and can be generated uniquely.

While these are challenges, the potential application fields are numerous and replete with emerging evidence of their relevance:

- *Training* in, for example, emergency or disaster situations, but in principle in any setting where a learner needs to simulate a task in an environment where

she needs to attend to numerous features and social interaction.

- In certain training situations, affective capabilities are essential to carrying out the task, such as in *therapy*, which can be more physical, as in limb injuries, or more mental, as in autism disorder and social phobias, or both, as in cases such as stroke rehabilitation.

In several of these situations—for example, mental disorders such as autism, anxiety, and social phobias—the patient practices social interaction while monitoring how they recognize or respond to emotional situations.

- *Well-being* examples such as physical exercise and meditation (Figure 2). Affective interaction here can motivate physical exercise or monitor psychophysiological states such as engagement or relaxation.

ENDNOTES

1. Ravaja, N., Harjunen, V., Ahmed, I., Jacucci, G., and Spapé, M.M. Feeling touched: Emotional modulation of somatosensory potentials to interpersonal touch. *Scientific Reports* 7 (2017), 40504.
2. Harjunen, V.J., Spapé, M., Ahmed, I., Jacucci, G., and Ravaja, N. Individual differences in affective touch: Behavioral inhibition and gender define how an interpersonal touch is perceived. *Personality and Individual Differences* 107 (2017), 88–95.
3. Spapé, M.M., Hoggan, E.E., Jacucci, G., and Ravaja, N. The meaning of the virtual Midas touch: An ERP study in economic decision making. *Psychophysiology* 52, 3 (2015), 378–387.
4. Ahmed, I., Harjunen, V., Jacucci, G., Hoggan, E., Ravaja, N., and Spapé, M.M. Reach out and touch me: Effects of four distinct haptic technologies on affective touch in virtual reality. *Proc. of the 18th ACM International Conference on Multimodal Interaction*. ACM, 2016, 341–348.
5. Affectiva; <http://www.affectiva.com/>
6. Kosunen, I., Salminen, M., Järvelä, S., Ruonala, A., Ravaja, N., and Jacucci, G. RelaWorld: Neuroadaptive and immersive virtual reality meditation system. *Proc. of the 21st International Conference on Intelligent User Interfaces*. ACM, 2016, 208–217.

✎ **Giulio Jacucci** is a professor in computer science at the University of Helsinki and founder of MultiTaction (www.multitaction.com). His research interests include multimodal interaction; physiological, tangible, and ubiquitous computing; search and information discovery; as well as behavioral change.
→ giulio.jacucci@helsinki.fi