

Challenges in portraying emotion in generated sign language

Rosalee Wolfe, Institute for Language and Speech Processing, Athena RC, Athens Greece
Fabrizio Nunnari, German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany

In collaboration with deaf communities, we are developing avatars with the goal of portraying sign languages in a legible, natural manner. This presentation describes ongoing work and identifies challenges in portraying emotion via a signing avatar.

Previous research and feedback have provided valuable insights that inform emotive sign language generation. For example, a facial posture can convey different meanings depending on the intent of the signer. Raised eyebrows might indicate the posing of a polar question, but in a different context, raised eyebrows might indicate surprise [Kimmelman, 2020]. Further, grammatical non-manual signals can co-occur on the face simultaneously with emotion, but the onset and duration of each process may differ [Weast, 2011].

A layered approach for organizing the animation preserves the intent and integrity of individual processes. Although this approach has proved capable of portraying several co-occurring processes, there are open questions, including resolving the competing influences of emotion and mouthing on the lower part of the face, and integrating the actions of manual and nonmanual channels. Although researchers have made initial forays into analyzing the effect of emotion on the manual channel [Hietanen, 2004], an open question remains as to how to incorporate this into sign language generation.

Much of the research in computer science, addressing the recognition and manifestation of emotions, rely on the established (and often abused) list of Ekman universal facial expressions (happiness, sadness, fear, surprise, anger, disgust, contempt) [Ekman, 1971]. Instead of relying on a discrete model which is limited to seven classes, we experiment with the description and manifestation of emotions expressed in the PAD (Pleasure, Arousal, Dominance) representation [Bakker, 2014]. The PAD model does not use descriptive labels such as sadness or anger, but instead defines emotions via the quantification of the P, A, and D values in a closed range. It provides a more flexible method to identify emotion via automatic recognition or to produce emotion via automatic generation. (See figure 1.) In a recent project (EASIER, <https://www.project-easier.eu>), this representation formed the basis for conveying emotion content recognized in spoken language to a signing avatar.

However, there are still many open questions on the methodology used to map PAD values into avatar motion. We currently rely on research from psychology, however, when person-to-person communication relies on sign language, rather than on spoken language, assumptions based on previous research might not hold. It is imperative that next steps include refining avatar portrayal in collaboration with deaf researchers and deaf communities.

In conclusion, there are many interesting and promising avenues to explore, but only the ones deemed worthy by deaf communities are worth pursuing. Thus, the most important open question is one of collaboration. How can we encourage more deaf researchers and deaf community members to get involved in avatar research?

References

- I. Bakker, T. Van Der Voordt, P. Vink, and J. De Boon, "Pleasure, Arousal, Dominance: Mehrabian and Russell revisited," *Curr Psychol*, vol. 33, no. 3, pp. 405–421, Sep. 2014, doi: 10.1007/s12144-014-9219-4.
- P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion.," *Journal of personality and social psychology*, vol. 17, no. 2, p. 124, 1971.
- J. Hietanen, J. Leppänen, and U. Lehtonen. "Perception of emotions in the hand movement quality of Finnish sign language", *Journal of nonverbal behavior*, vol. 28, pp. 53-64, 2004
- V. Kimmelman, A. Imashev, M. Mukushev, and A. Sandygulova. "Eyebrow position in grammatical and emotional expressions in Kazakh-Russian Sign Language: A quantitative study," *PloS one*, vol. 15, no. 6, pp e0233731, 2020.
- T. Weast, "American Sign Language tone and intonation: A phonetic analysis of eyebrow properties," *Formational units in sign languages*, pp. 203-226, 2011.

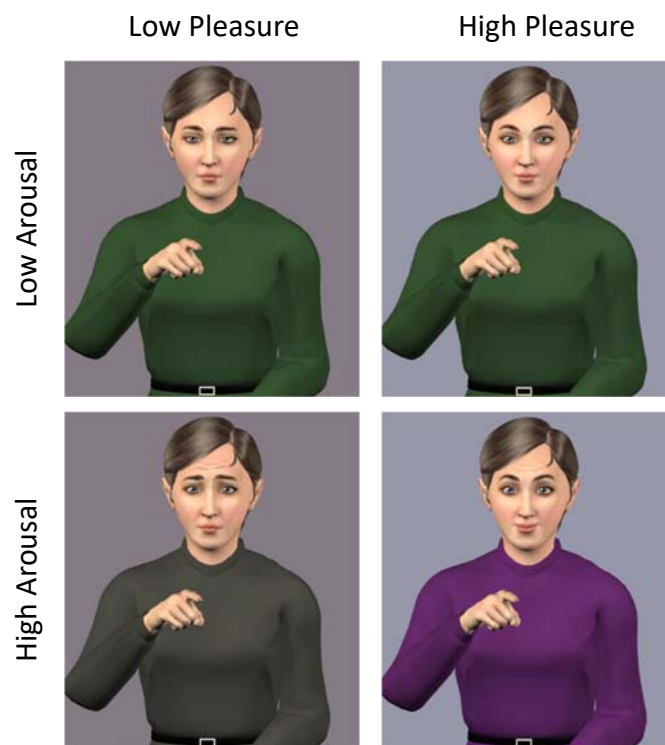


Figure 1: Utilizing the PAD model