GENEA Workshop 2022: The 3rd Workshop on Generation and Evaluation of Non-verbal Behaviour for Embodied Agents

Pieter Wolfert
pieter.wolfert@ugent.be
IDLab, Ghent University – imec
Ghent, Belgium

Taras Kucherenko
tkucherenko@ea.com
SEED – Electronic Arts (EA)
Stockholm, Sweden

Youngwoo Yoon
youngwoo@etri.re.kr
ETRI
Daejeon, Republic of Korea

Carla Viegas
cviegas@andrew.cmu.edu
CMU & NOVA University Lisbon
Lisbon, Portugal

Zerrin Yumak
z.yumak@uu.nl
Utrecht University
Utrecht, The Netherlands

Gustav Eje Henter
ghe@kth.se
KTH Royal Institute of Technology
Stockholm, Sweden

ABSTRACT

Embodied agents benefit from using non-verbal behavior when communicating with humans. Despite several decades of non-verbal behavior-generation research, there is currently no well-developed benchmarking culture in the field. For example, most researchers do not compare their outcomes with previous work, and if they do, they often do so in their own way which frequently is incompatible with others. With the GENEA Workshop 2022, we aim to bring the community together to discuss key challenges and solutions, and find the most appropriate ways to move the field forward.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **Computer systems organization** → Robotics.

KEYWORDS

behavior synthesis, gesture generation, datasets, evaluation

1 WORKSHOP MOTIVATION

The field of non-verbal behaviour generation for embodied agents has seen many improvements in the recent years. Researchers have been proposing methods and models to generate non-verbal behaviors for conversational agents in the form of facial expressions, gestures, gaze and posture [2, 3, 5, 10]. Embodied Social AI starts to play a larger role in human-machine interaction. In 2020 the first gesture generation challenge was organized at IVA [6]. The corresponding results from both the challenge, as well as later work

that builds on that challenge, shows that there is a vibrant community of researchers willing to work on non-verbal behaviour generation for embodied agents. The topic of non-verbal behaviour generation has attracted the attention of different communities such as HCI, HRI and 3D graphics/animation, as well as social and behavioral scientists. Yet, these embodied agents are still far from having non-verbal behaviors synthesized on-the-fly in interactive settings, based on factors such as personality, emotions and the social context. Non-verbal behavior synthesis for groups as well as plausibility and controlability of the generated motions remains an important topic of research. In order to advance the field of non-verbal behavior generation, there needs to be clear methods for evaluating and benchmarking the outcomes [6].

The 3rd edition of this workshop aims to bring together researchers that have different motivations and approaches for non-verbal behavior generation, with the intent to stimulate and advance discussions on how to improve the generation methods and the evaluation of their results. The workshop is a continuation of a successful workshop held at the ACM International Conference on Intelligent Virtual Agents (IVA) 2020 and the ACM International Conference on Multimodal Interaction (ICMI) 2021; see genea-workshop.github.io/2021/. The webpage for the 2022 iteration of the workshop (the one described in this paper) can be found at: genea-workshop.github.io/2022/workshop/

This year's edition of the workshop is combined with the GENEA Challenge 2022, an official Grand Challenge of ICMI 2022. More details are given in the main challenge paper [11] and on the challenge webpage at genea-workshop.github.io/2022/challenge/.

2 EXPECTED OUTCOMES AND IMPACT

- The communication among different communities will lead to a better understanding of the challenges and opportunities in non-verbal behaviour generation and evaluation.
- New insights on research challenges of generating non-verbal behaviours in different social and emotional contexts, for different individuals, identities, and applications.
- New approaches and ideas to evaluate and benchmark the results of non-verbal behaviour synthesis approaches for virtual agents and social robots.
- Gathering the community
- Contributing toward a benchmarking culture

3 WORKSHOP CONTRIBUTIONS

Five contributions were accepted for the workshop, covering a range of topics in HCI. These are listed below:

- *Can you tell that I'm confused? An overhearer study for German backchannels by an embodied agent* [7] by Isabel Donya Meywirth and Jana Götze.
- *Automatic facial expressions, gaze direction and head movements generation of a virtual agent* [1] by Alice Delbosc, Stéphane Ayache and Magalie Ochs.
- *Emotional Respiration Speech Dataset* [8] by Rozemarijn Roes, Francisca Pessanha and Almila Akdag.
- *Understanding Interviewees' Perceptions and Behaviour towards Verbally and Non-verbally Expressive Virtual Interviewing Agents* [9] by Jinal Thakkar, Pooja S. B. Rao, Kumar Shubham, Vaibhav Jain and Dinesh Babu Jayagopi.
- *ReCell: replicating recurrent cell for auto-regressive pose generation* [4] by Vladislav Korzun, Anna Beloborodova and Arkady Ilin.

4 LIST OF ORGANIZERS

- **Pieter Wolfert** (PhD Student, IDLab Ghent University – imec, Belgium)
- **Taras Kucherenko** (Research Scientist, SEED – Electronic Arts, Stockholm, Sweden)
- **Youngwoo Yoon** (Principal Researcher, ETRI, Daejeon, South Korea)
- **Carla Viegas** (PhD Student CMU, Pittsburgh, PA, USA, & NOVA University Lisbon, Lisbon, Portugal)
- **Zerrin Yumak** (Assistant Professor, Utrecht University, The Netherlands)
- **Gustav Eje Henter** (Assistant Professor, KTH Royal Institute of Technology, Stockholm, Sweden)

5 KEYNOTE SPEAKERS

The workshop also invited two keynote speakers:

- Judith Holler (Associate Professor, Radboud University, The Netherlands)
- Carlos Ishi (Group Leader, ATR, Osaka, Japan)

These researchers were selected to bring perspectives from both research on language and interaction (Judith Holler), as well as on virtual agents and gestures (Carlos Ishi).

6 PROGRAM COMMITTEE

The names of the program committee members (in alphabetical order) are the following: Chaitanya Ahuja (CMU); Sean Andrist (Microsoft Research); Jonas Beskow (KTH Royal Institute of Technology); Uttaran Bhattacharya (University of Maryland); Carlos Busso (University of Texas at Dallas); Yingruo Fan (The University of Hong Kong); David Greenwood (University of East Anglia); Dai Hasegawa (Hokkai Gakuen University); Carlos Ishi (ATR); Minsu Jang (ETRI); Naoshi Kaneko (Aoyama Gakuin University); James Kennedy (Futronics NA Corporation); Vladislav Korzun (Moscow Institute of Physics and Technology); Nikhil Krishnaswamy (Colorado State University); Rachel McDonnell (Trinity College Dublin); Tan

Viet Tuyen Nguyen (King's College London); Kunkun Pang (Guangdong Institute of Intelligent Manufacturing); Catherine Pelachaud (Sorbonne University); Tiago Ribeiro (Soul Machines); Carolyn Saund (University of Glasgow); Hiroshi Shimodaira (University of Edinburgh); Jonathan Windle (University of East Anglia); and Bowen Wu (Osaka University).

7 CONCLUSIONS

We are confident that the accepted contributions in combination with our keynote speakers will lead the floor for fruitful discussions about the research progress in the field. We hope that this workshop will bring together researchers working on gesture generation in avatars and robotics from different angles, and that this will lead to cross pollination in terms of methodology and evaluation.

REFERENCES

[1] Alice Delbosc, Stéphane Ayache, and Magalie Ochs. 2022. Automatic facial expressions, gaze direction and head movements generation of a virtual agent. In *Companion Publication of the 2022 International Conference on Multimodal Interaction (ICMI '22 Companion)*. ACM.

[2] Patrik Jonell, Taras Kucherenko, Gustav Eje Henter, and Jonas Beskow. 2020. Let's Face It: Probabilistic multi-modal interlocutor-aware generation of facial gestures in dyadic settings. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*. Article 31, 8 pages.

[3] Alex Klein, Zerrin Yumak, Arjen Beij, and A. Frank van der Stappen. 2019. Data-driven gaze animation using recurrent neural networks. In *Proceedings of the ACM SIGGRAPH Conference on Motion, Interaction and Games*. Article 4, 11 pages.

[4] Vladislav Korzun, Anna Beloborodova, and Arkady Iliin. 2022. ReCell: replicating recurrent cell for auto-regressive pose generation. In *Companion Publication of the 2022 International Conference on Multimodal Interaction (ICMI '22 Companion)*. ACM.

[5] Taras Kucherenko, Dai Hasegawa, Naoshi Kaneko, Gustav Eje Henter, and Hedvig Kjellström. 2021. Moving fast and slow: Analysis of representations and post-processing in speech-driven automatic gesture generation. *International Journal of Human-Computer Interaction* 37, 14 (2021), 1300–1316.

[6] Taras Kucherenko, Patrik Jonell, Youngwoo Yoon, Pieter Wolfert, and Gustav Eje Henter. 2021. A large, crowdsourced evaluation of gesture generation systems on common Data: The GENEA Challenge 2020. In *Proceedings of the 26th International Conference on Intelligent User Interfaces*. 11–21.

[7] Isabel Donya Meywirth and Jana Götze. 2022. Can you tell that I'm confused? An overhearer study for German backchannels by an embodied agent. In *Companion Publication of the 2022 International Conference on Multimodal Interaction (ICMI '22 Companion)*. ACM.

[8] Rozemarijn Roes, Francisca Pessanha, and Almila Akdag. 2022. Emotional respiration speech dataset. In *Companion Publication of the 2022 International Conference on Multimodal Interaction (ICMI '22 Companion)*. ACM.

[9] Jinal Thakkar, Pooja S. B. Rao, Kumar Shubham, Vaibhav Jain, and Dinesh Babu Jayagopi. 2022. Understanding interviewees' perceptions and behaviour towards verbally and non-verbally expressive virtual interviewing agents. In *Companion Publication of the 2022 International Conference on Multimodal Interaction (ICMI '22 Companion)*. ACM.

[10] Youngwoo Yoon, Bok Cha, Joo-Haeng Lee, Minsu Jang, Jaeyeon Lee, Jaehong Kim, and Geehyuk Lee. 2020. Speech gesture generation from the trimodal context of text, audio, and speaker identity. *ACM Transactions on Graphics (TOG)* 39, 6, Article 222 (2020), 16 pages.

[11] Youngwoo Yoon, Pieter Wolfert, Taras Kucherenko, Carla Viegas, Teodor Nikolov, Mihail Tsakov, and Gustav Eje Henter. 2022. The GENEA Challenge 2022: A large evaluation of data-driven co-speech gesture generation. In *Proceedings of the ACM International Conference on Multimodal Interaction (ICMI '22)*. ACM.