4th Workshop on Social Affective Multimodal Interaction for Health (SAMIH)

Hiroki Tanaka
Nara Institute of Science and Technology, Japan
hiroki-tan@is.naist.jp

Satoshi Nakamura
Nara Institute of Science and Technology, Japan
s-nakamura@is.naist.jp

Jean-Claude Martin
CNRS-LISN, Université Paris-Saclay
France
martin@limsi.fr

Catherine Pelachaud
CNRS-ISIR, Sorbonne University
France
catherine.pelachaud@upmc.fr

ABSTRACT

This workshop discusses how interactive, multimodal technology, such as virtual agents, can measure and train social-affective interactions. Sensing technology now enables analyzing users' behaviors and physiological signals. Various signal processing and machine learning methods can be used for prediction tasks. Such social signal processing and tools can be applied to measure and reduce social stress in everyday situations, including public speaking at schools and workplaces.

CCS CONCEPTS

• **Applied computing → Health care information systems**.

KEYWORDS

cognitive behavioral therapy; motivational interview; social skills training; social signal processing; affective computing; virtual agents; large language models

1 WORKSHOP GOAL

This workshop focuses on how interactive, multimodal technology such as virtual agents or social robots can be used in training, technique, and treatment (motivational interviewing, remote medicine, etc.).

People with socio-affective deficits have difficulties controlling their social behavior and also suffer from interpreting others' social behavior. Behavioral training, such as social skills training, is used in medical settings. Patients are trained by a coach to experience social interaction and reduce social stress. In addition to behavioral training, cognitive behavioral therapy, and motivational interviewing are also useful for understanding better and training social-affective interaction. All these methods are effective but expensive and difficult to access.

This workshop is looking for works describing how multimodal technology can be used in healthcare for measuring and training social-affective interactions [1–3, 7, 11]. Sensing technology analyzes users' behaviors and physiological signals (heart rate, electroencephalography, etc.). Various signal processing and machine learning methods can be used for such prediction tasks [4, 5]. Beyond sensing, it is also important to analyze human behaviors and model and implement training methods (e.g., by virtual agents [6], relevant scenarios, and design appropriate and personalized feedback about social skills performance). Social signal processing and tools can be applied to measure and reduce social stress in everyday situations, including public speaking at schools and workplaces [12]. Target populations include depression, schizophrenia, autism spectrum disorder, and a much larger group of social pathological phenomena. In this workshop, we invite participants from academia, industry, or clinical settings to present and discuss the social-affective design of multimodal training for health. This workshop was organised at ICMI2020, ICMI2021, and ICMI2022 [8–10].

2 WORKSHOP CONTENT

The workshop will cover diverse themes and topics related to perception of virtual agents during public speaking, evaluation of generated reflection for motivational interview, multi-modal emotion expression recognition, and virtual cognitive training for elderly people. We also include broad range of technical terms such as social signal processing, affective computing, virtual agents, and large language models.

3 WORKSHOP ORGANIZATION

3.1 Review process

Invited program committee members represent a broad spectrum of expertise. Each paper received one or two reviews. The selection of accepted papers was based on whether there was sufficient merit of the work to the workshop.

3.2 Organising Committee

- Hiroki Tanaka (NAIST, Japan)
- Satoshi Nakamura (NAIST, Japan)
- Jean-Claude Martin (CNRS-LISN, Université Paris-Saclay, France)
- Catherine Pelachaud, (CNRS-ISIR, Sorbonne University, France)

3.3 Programme Committee

- Yuichiro Fujimoto (NAIST, Japan)
- Kana Miyamoto (NAIST, Japan)
- Kazuhiro Shidara (NAIST, Japan)

4 CONCLUSION

Overall, we are pleased with the diversity of works that will be presented at this workshop. This provides an interesting opportunity to discuss interdisciplinary perspectives on the workshop topics and to promote collaboration further. We strongly believe that a more interdisciplinary perspective (e.g., academia, industry, or clinical settings) will be extremely beneficial in increasing the impact of research on technologies for the understanding and influence of social affective multimodal interaction for health.

5 ACKNOWLEDGMENTS

REFERENCES

[1] S. Burke, Tammy L. Bresnahan, T. Li, K. Epnere, A. Rizzo, Mary Partin, Robert M Ahlness, and M. Trimmer. 2018. Using Virtual Interactive Training Agents (ViTA) with Adults with Autism and Other Developmental Disabilities. *Journal of Autism and Developmental Disorders* 48 (2018), 905–912.

[2] J. Daniels, J. N. Schwartz, C. Voss, N. Haber, A. Fazel, A. Kline, P. Washington, C. Feinstein, T. Winograd, and D. P. Wall. 2018. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. *NPJ Digit Med* 1 (2018), 32.

[3] Ouriel Grynszpan, Julie Bouteiller, Séverine Grynszpan, Florence Le Barillier, Jean-Claude Martin, and Jacqueline Nadel. 2019. Altered sense of gaze leading in autism. *Research in Autism Spectrum Disorders* 67 (2019), 101441. https://doi.org/10.1016/j.rasd.2019.101441

[4] L. Hemamou, G. Felhi, J. Martin, and C. Clavel. 2019. Slices of Attention in Asynchronous Video Job Interviews. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. 1–7. https://doi.org/10.1109/ACII.2019.8925439

[5] L'eo Hemamou, Ghazi Felhi, Vincent Vandenbussche, Jean-Claude Martin, and Chloé Clavel. 2019. HireNet: A Hierarchical Attention Model for the Automatic Analysis of Asynchronous Video Job Interviews. In *AAAI*.

[6] Isabella Poggi, Catherine Pelachaud, F. Rosis, Valeria Carofiglio, and Berardina Carolis. 2005. *Greta. A Believable Embodied Conversational Agent*. 3–25. https://doi.org/10.1007/1-4020-3051-7_1

[7] Kazuhiro Shidara, Hiroki Tanaka, Hiroyoshi Adachi, Daisuke Kanayama, Yukako Sakagami, Takashi Kudo, and Satoshi Nakamura. 2022. Automatic thoughts and facial expressions in cognitive restructuring with virtual agents. *Frontiers in Computer Science* 4 (2022), 762424.

[8] Hiroki Tanaka, Satoshi Nakamura, Jean-Claude Martin, and Catherine Pelachaud. 2020. Social Affective Multimodal Interaction for Health. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 893–894. https://doi.org/10.1145/3382507.3420059

[9] Hiroki Tanaka, Satoshi Nakamura, Jean-Claude Martin, and Catherine Pelachaud. 2021. 2nd Workshop on Social Affective Multimodal Interaction for Health (SAMIH). In *Proceedings of the 2021 International Conference on Multimodal Interaction*. 853–854.

[10] Hiroki Tanaka, Satoshi Nakamura, Kazuhiro Shidara, Jean-Claude Martin, and Catherine Pelachaud. 2022. 3rd Workshop on Social Affective Multimodal Interaction for Health (SAMIH). In *Proceedings of the 2022 International Conference on Multimodal Interaction*. 805–806.

[11] H. Tanaka, H. Negoro, H. Iwasaka, and S. Nakamura. 2017. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. *PLoS ONE* 12, 8 (2017), e0182151.

[12] M. Iftekhar Tanveer, Emy Lin, and Mohammed (Ehsan) Hoque. 2015. Rhema: A Real-Time In-Situ Intelligent Interface to Help People with Public Speaking. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (Atlanta, Georgia, USA) *(IUI '15)*. Association for Computing Machinery, New York, NY, USA, 286–295. https://doi.org/10.1145/2678025.2701386