3rd 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

Kazuhiro Shidara
Nara Institute of Science and Technology, Japan
shidara.kazuhiro.sc5@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 be used in social skills training for measuring and training social-affective interactions. Sensing technology now enables analyzing user's behaviors and physiological signals. Various signal processing and machine learning methods can be used for such 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; physiological signal processing; virtual agents; social robotics

ACM Reference Format:
Hiroki Tanaka, Satoshi Nakamura, Kazuhiro Shidara, Jean-Claude Martin, and Catherine Pelachaud. 2022. 3rd Workshop on Social Affective Multimodal Interaction for Health (SAMIH). In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '22), November 7–11, 2022, Bengaluru, India.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3536221.3564030

1 WORKSHOP GOAL

Social Skills Training (SST) is often used in the multimodal Interaction research community as an umbrella term for systems that aim at training social skills: managing appropriately verbal and nonverbal behaviors when interacting with one or more persons, in relation with various communicative functions such as turn taking and emotions.

This workshop is looking for works describing how interactive, multimodal technology such as virtual agents [10] can be used in social skills training for measuring and training social-affective interactions [2, 3, 6]. Sensing technology now enables analyzing user's behaviors and physiological signals (heart-rate, EEG, etc.). Various signal processing and machine learning methods can be used for such prediction tasks [7–9]. Beyond sensing, it is also important to analyze human behaviors and model and implement training methods (e.g., by virtual agents, social robots, relevant scenarios, design appropriate and personalized feedback about social skills performance). 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 [16]. Target populations include depression, Social Anxiety Disorder (SAD), Schizophrenia, Autism Spectrum Disorder (ASD), but also a much larger group of different social pathological phenomena [1]. This workshop was organised at ICMI2020 and ICMI2021 [14, 15].

2 WORKSHOP CONTENT

The workshop will cover themes and topics related to cognitive behavioral therapy, motivational interview and social skills training. We also include broad range of technical terms such associal signal processing, affective computing, physiological signal processing, virtual agents, social robotics. Here is a brief description of the accepted papers:

- Richter et al. [11] analyzed multimodal information for remote schizophrenia detection. Data were collected using a virtual agent, and experimental participants conducted multiple tasks. The results showed significant differences in speech and facial behaviors between healthy control and schizophrenia groups.
- Schütze et al. [13] proposed a visual programming software for enabling non-technical users to create robot-assisted therapy scenarios. The usability and performance of the system have been evaluated by non-technical domain expert users.
- Groß et al. [5] present a software architecture for robot-assisted configurable and autonomous scenarios of Joint-Attention-Training, one of therapeutic strategies to support autism therapy. The preliminary evaluation showed the high usability of the system, and even if the non-computer science students could complete all tasks regarding the training.

ICMI '22, November 7–11, 2022, Bengaluru, India

ACM ISBN 978-1-4503-9390-4/22/11.
https://doi.org/10.1145/3536221.3564030

- Filippou et al. [4] proposed a predictive framework that leverages multimodal physiological signals to detect alexithymia. Their proposed method outperformed more complex classification models, including deep neural networks.
- Roesler et al. [12] compared two facial landmark predictors in combination with four different normalization methods with respect to their effect on the utility of facial metrics obtained through a multimodal assessment platform with a virtual agent. For the experiment, they used Parkinson's disease group and healthy controls.

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 is sufficient merit of the work judged by the reviewers.

3.2 Organising Committee

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

3.3 Programme Committee

- Shiro Kumano (NTT, Japan)
- Yuichiro Fujimoto (NAIST, Japan)
- Kana Miyamoto (NAIST, Japan)
- Liu Yang (CNRS-ISIR, Sorbonne University, France)
- Jennifer Hamet Bagnou (CNRS-LISN, Université Paris-Saclay, France)

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 further promote collaboration. 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

We would like to thank ANR-CREST-TAPAS Japan - France project. This workshop was funded by the JST CREST Grant Number JPMJCR19M5-1, Japan and ANR-19-JSTS-0001, France (TAPAS).

REFERENCES

[1] American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed. ed.). Autor, Washington, DC.

[2] 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.

[3] 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.

[4] Valeria Filippou, Nikolas Theodosiou, Mihalis Nicolaou, Elena Constantinou, Panayiotou Georgia, and Marios Theodorou. 2022. A Wavelet-based Approach for Multimodal Prediction of Alexithymia from Physiological Signals. In *Companion Publication of the 2022 International Conference on Multimodal Interaction.*

[5] André Groß, Christian Schütze, Britta Wrede, and Birte Richter. 2022. An Architecture Supporting Configurable Autonomous Multimodal Joint-Attention-Therapy for Various Robotic Systems. In *Companion Publication of the 2022 International Conference on Multimodal Interaction.*

[6] 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

[7] 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

[8] 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.*

[9] H. Maki, H. Tanaka, S. Sakti, and S. Nakamura. 2018. Graph Regularized Tensor Factorization for Single-Trial EEG Analysis. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* 846–850.

[10] 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

[11] Vanessa Richter, Michael Neumann, Hardik Kothare, Oliver Roesler, Jackson Liscombe, David Suendermann-Oeft, Sebastian Prokop, Anzalee Khan, Christian Yavorsky, Jean-Pierre Lindenmayer, and Vikram Ramanarayanan. 2022. Towards Multimodal Dialog-Based Speech Facial Biomark- ers of Schizophrenia. In *Companion Publication of the 2022 International Conference on Multimodal Interaction.*

[12] Oliver Roesler, Hardik Kothare, William Burke, Michael Neumann, Jackson Liscombe, Andrew Cornish, and Doug Habbersta. 2022. Exploring Facial Metric Normalization For Within- and Between-Subject Comparisons in a Multimodal Health Monitoring Agent. In *Companion Publication of the 2022 International Conference on Multimodal Interaction.*

[13] Christian Schütze, André Groß, Britta Wrede, and Birte Richter. 2022. Enabling Non-Technical Domain Experts to Create Robot-Assisted Therapeutic Scenarios via Visual Programming. In *Companion Publication of the 2022 International Conference on Multimodal Interaction.*

[14] 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

[15] 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.

[16] 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