Social Affective Multimodal Interaction for Health

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-LIMSI, 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, Jean-Claude Martin, and Catherine Pelachaud. 2020. Social Affective Multimodal Interaction for Health. In *Proceedings of the 2020 International Conference on Multimodal Interaction (ICMI '20), October 25–29, 2020, Virtual event, Netherlands.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3382507.3420059

1 WORKSHOP GOAL

Social Skill Training 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.

People with social affective deficits have difficulties controlling their own social behavior and also suffer from interpreting others' social behavior. Behavioral therapy (eg with a clinician) and Social Skill Training (SST) are used in medical settings [2]. Patients are trained with a coach to experience social interaction and to reduce social stress. SST includes a role-play of a simulation of actual situations [3, 11, 14]. In addition to behavioral training, cognitive therapy [6] and motivational interviewing are also useful to better understand and train social-affective interaction [8]. All these methods applied with a human clinician are effective but expensive and difficult to access.

This workshop is looking for works describing how interactive, multimodal technology such as virtual agents [13] can be used in social skills training for measuring and training social-affective interactions [4, 5, 7]. 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 [9, 10, 12]. Beyond sensing, it is also important to analyze human behaviors, to model and to 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 [15]. 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].

In this workshop, we invite participants from academia, industry or clinical settings to present and discuss social-affective design of multimodal training for health.

2 WORKSHOP CONTENT

2.1 Call for papers

The following keywords describe the topics of the workshop. These themes and topics will be discussed at the workshop.

- Cognitive behavioral therapy
- Motivational interview
- Social skills training

We also include broad range of technical terms such as:

- Social signal processing
- Affective computing
- Physiological signal processing
- Virtual agents
- Social robotics

ICMI '20, October 25–29, 2020, Virtual event, Netherlands

ACM ISBN 978-1-4503-7581-8/20/10…$15.00
https://doi.org/10.1145/3382507.3420059

2.2 Invited Speakers

In addition to the submitted papers, we also have an invited speaker. The researcher was selected to bring perspectives from both the social skills training and the affective computing fields.

3 WORKSHOP ORGANIZATION

3.1 Review Process

Invited program committee members represent a broad spectrum of expertise. Each paper received at least 2 reviews. Selection of accepted papers will be based on whether there is sufficient merit of the work judged by the reviewers. The accepted papers will be then subdivided for (virtual) oral presentation.

3.2 Organising Committee

- Hiroki Tanaka (Nara Institute of Science and Technology, Japan)
- Satoshi Nakamura (Nara Institute of Science and Technology, Japan)
- Jean-Claude Martin (CNRS-LIMSI, Université Paris-Saclay, France)
- Catherine Pelachaud (CNRS-ISIR, Sorbonne University, France)

3.3 Program Committee

- Mohamed Chetouani (University Pierre and Marie Curie)
- Mathieu Chollet (IMT Atlantique)
- Yuichiro Fujimoto (Nara Institute of Science and Technology)
- Shogo Okada (Japan Advanced Institute of Science and Technology)
- Shiro Kumano (NTT)
- Theodora Chaspari Texas A&M University)

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 (eg 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.

REFERENCES

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

[2] A.S. Bellack, K.T. Mueser, S. Gingerich, and J. Agresta. 2013. *Social Skills Training for Schizophrenia, Second Edition: A Step-by-Step Guide.* Guilford Publications.

[3] Merijn Bruijnes, Jeroen Linssen, and Dirk Heylen. 2019. Special issue editorial: Virtual Agents for Social Skills Training. *Journal on Multimodal User Interfaces* 13, 1 (01 Mar 2019), 1–2. https://doi.org/10.1007/s12193-018-00291-7

[4] 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.

[5] 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.

[6] K. K. Fitzpatrick, A. Darcy, and M. Vierhile. 2017. Delivering Cognitive Behavior Therapy to Young Adults With Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Ment Health* 4, 2 (Jun 2017), e19.

[7] 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

[8] Rahul Gupta, Nishant Nath, Taruna Agrawal, Panayiotis Georgiou, David C. Atkins, and Shrikanth S. Narayanan. 2016. Laughter Valence Prediction in Motivational Interviewing Based on Lexical and Acoustic Cues. In *Interspeech 2016*. 505–509. https://doi.org/10.21437/Interspeech.2016-184

[9] 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

[10] 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*.

[11] Mohammed (Ehsan) Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W. Picard. 2013. MACH: My Automated Conversation Coach. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 697–706. https://doi.org/10.1145/2493432.2493502

[12] 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.

[13] 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

[14] 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.

[15] 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