Stefan Scherer

University of Southern California
Institute for Creative Technologies
12015 Waterfront Dr.
Playa Vista, CA 90094
scherer@ict.usc.edu
http://schererstefan.net

Education

- **Ulm University, Germany**
  Dr. rer. nat. (Degree: *summa cum laude*, with distinction) in Computer Science, Faculty of Engineering and Computer Science, July 2011
- **Ulm University, Germany**
  Dipl.-Inf. in Computer Science, Faculty of Engineering and Computer Science, May 2006

Professional

- **Research Assistant Professor** — July 2014 - present
  University of Southern California, Department of Computer Science — Los Angeles, CA
- **USC Viterbi Engineering Faculty Council Representative** — April 2016 - present
  University of Southern California, Viterbi School of Engineering — Los Angeles, CA
- **Researcher WOC** — April 2014 - present
  San Francisco VA Medical Center — San Francisco, CA
- **Steering Committee Member** — April 2014 - present
  Austrian Scientists & Scholars in North America (ASCINA) - Chapter Pacific South — Los Angeles, CA
- **Research Associate** — October 2013 - July 2014
  USC Institute for Creative Technologies — Los Angeles, CA
- **Postdoctoral Research Associate** — January 2012 - October 2013
  USC Institute for Creative Technologies — Los Angeles, CA
- **Postdoctoral Research Fellow** — September 2011 - December 2011
  Trinity College Dublin — Dublin, Ireland
- **Scientific Employee** — July 2006 - August 2011
  Ulm University — Ulm, Germany
- **Visiting Researcher** — June 2008 - December 2008
  Advanced Telecommunications Research Institute (ATR) — Kyoto, Japan

Teaching Experience

- **Professor at USC** — August 2016
  Multimodal Probabilistic Learning of Human Communication (CSCI 535) — Los Angeles, CA
- **Professor at USC** — January 2015 - present
  Machine Learning for Data Informatics (INF 522) — Los Angeles, CA
- **Professor at USC** — September 2012 - December 2014
  Human Communication and Machine Learning (CSCI 599) — Los Angeles, CA
- **Guest-Lecturer at USC** — April 2014
  Introduction to Computing (CSCI 109) — Los Angeles, CA
- **Guest-Lecturer at USC** — March 2014
  Affective Computing (CSCI 534) — Los Angeles, CA
- **Guest-Lecturer at USC** — October 2013
  Fundamentals of Artificial Intelligence (CSCI 561) — Los Angeles, CA
- **Instructor and Co-Lecturer at Ulm University** — October 2004 - August 2011
  Lectures Neural Networks I/II, Statistical Learning Theory, and Data Mining — Ulm, Germany

---

Supervisory Experience

- **Postdoctoral Fellows**
  - Marcelo Worsley (`http://web.stanford.edu/~mworsley/`) - Topic: Multimodal Learning Analytics *(Now: Faculty at Northeastern University)*
  - Mathieu Chollet (`https://www.linkedin.com/in/mathieuchollet`) - Topic: Cicero - Virtual Human Interpersonal Skill Training Platform
  - Catherine Neubauer (`https://www.linkedin.com/in/catherine-neubauer-11642398`) - Topic: Team Cohesion and Cybernetic Analysis
- **PhD Students**
  - Sayan Ghosh (`https://www.linkedin.com/in/sayanghosh88`) - Topic: Multimodal Deep Learning for User State Recognition (since Fall 2013)
  - Lixing Liu - Topic: Learning Deep Adversarial Representations for Human Behavior Recognition (starting Fall 2016)
  - Michael Tsang - Topic: Human Robot Interaction (since Spring 2016)
  - Sharifa Alghowinem (examiner) - Topic: Multimodal Analysis of Verbal and Nonverbal Behaviour on the Example of Clinical Depression (certification attached to application)
- **Master Students**
  - Verena Venek (`http://www.salzburgresearch.at/person/venek-verena/`) - Topic: Adolescent suicidal risk assessment in clinician-patient interaction: A study of verbal and acoustic behaviors (Summer 2014)
  - Dietmar Wohlbauer - Topic: Automatic audio-visual symptom analysis of parkinson's behavior patterns in clinical settings
  - Lixing Liu (`https://www.linkedin.com/pub/lixing-liu/62/488/338`) - Topic: Robot Eye Gaze Control in Dyadic Interactions (Fall 2014)
  - Maarten Brilman (`https://www.linkedin.com/pub/maarten-brilman/a4/519/a93`) - Topic: A Multimodal Predictive Model of Successful Debaters or How I Learned to Sway Votes
- **Bachelor Students**
  - Torsten Woertwein (`https://interact.anthropomatik.kit.edu/91_326.php`) - Topic: Enabling Interactive Virtual Human Public Speaking Training - Automatic Behavior Sensing and Multimodal Performance Assessment (Summer 2014 & Summer 2015)

- **Undergrad and Summer Interns**
  - Mathieu Chollet, Pooja Voladoddi, Eugene Laksana, Philippa Shoemark, Kalin Stefanov, Talie Massachi, Pankaj Dabade, Hyounghun Kim, Jonathan Chang, Federico Parra, Michelle Morales

---

Awards and Honors

- **Best Paper:** Clinical interviewing by a virtual human agent with automatic behavior analysis, at the 11th International Conference on Disability, Virtual Reality & Associated Technologies, 2016, Los Angeles, CA (Authors: Rizzo, A., Lucas, G., Gratch, J., Stratou, G., Morency, L.-P., Shilling, R., Hartholt, A., and Scherer, S.)
- **Outstanding Paper:** Combining Two Perspectives on Classifying Multimodal Data for Recognizing Speaker Traits, at the ACM International Conference on Multimodal Interaction, 2015, Seattle, WA (Authors: Chatterjee, M., Park, S., Morency, L.-P., and Scherer, S.)
- **Best Paper:** Automatic Behavior Descriptors for Psychological Disorder Analysis, at the IEEE Conference on Automatic Face and Gesture Recognition, 2013, Shanghai, China (Authors: Scherer, S., Stratou, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A., and Morency, L.-P.)
- **Best Paper:** Fuzzy-Input Fuzzy-Output One-Against-All Support Vector Machines, at the International Conference on Knowledge-Based, Intelligent Information and Engineering Systems, 2007, Vietri sul Mare (Authors: Thiel, C., Scherer, S., and Schwenker, F.)
- **USC Collaboration Fund Award:** "Big Data and Human Behavior" Awarded 10k USD as co-PI to organize a conference on Big Data and Human Behavior
- **NetExplo Award (UNESCO):** The SimSensei collaborative effort won the NetExplo Award for being one of the ten most innovative and promising initiatives of the year. This award is organized in partnership with UNESCO.
- **Best Demo:** A Demonstration of the Perception System in SimSensei, a Virtual Human Application for Healthcare Interviews, at the AAAC/IEEE Conference on Affective Computing and Intelligent Interaction, 2015, Xi'An, China (Authors: Giota Stratou, Louis-Philippe Morency, David Devault, Arno Hartholt, Edward Fast, Margaux Lhommet, Gale Lucas, Fabrizio Morbini, Kallirroi Georgila, Stefan Scherer, Jonathan Gratch, Stacy Marsella, David Traum and Albert Rizzo)
- **Nominated for Best Paper:** Verbal indicators of psychological distress in interactive dialogue with a virtual human, at SIGdial 2013, Metz, France (Authors: DeVault, D., Georgilia, K., Artstein, R., Morbini, F., Traum, D., Scherer, S., Rizzo, A., and Morency, L.-P.)
- **Nominated for Best Student Paper:** Mutual Behaviors During Dyadic Negotiation: Automatic Prediction of Respondent Reactions, at Affective Computing and Intelligent Interaction 2013, Geneva, Switzerland (Authors: Park, S., Scherer, S., Gratch, J., Carnevale, P., and Morency, L.-P.)
- **German Exchange Service (DAAD) Scholarship**, sponsored sabbatical research visit to Kyoto, Japan in 2008

Contribution to Science

1. **Multimodal Behavior Analytics and Machine Learning:** The main contribution of my research is to jointly analyze all available modalities (in particular audio and video) within the interactants' observable behavior. I have introduced and spearheaded the integration of perceptual salience in both the applied multimodal machine learning algorithms, which leverage uncertainty in the target classes, and the feature extraction mechanisms, that reflect human characteristics of production and perception. Our algorithms leverage more sources of information and have shown great potential in many of the above-mentioned

areas of research, and have outperformed other state-of-the-art approaches. In particular, I focused on the identification of individuals' multimodal nonverbal behavior in natural and unconstrained environments. *Major publications:* [12, 16, 15, 10]

2. **Behavior Analytics and Its Application to Healthcare:** Unlike laboratory values or radiographic images, clinical material from a mental health encounter consists of complex human interactions between patient and provider characterizing mood, affect, and behavior. Because much of this information is subjective and difficult or impossible to annotate while maintaining affordability of clinical care, it becomes inaccessible to treatment planning among medical team members. Clinical practice suffers from an information deficit that at best compromises and at worst jeopardizes patient care. Within this context we developed machine learning based human behavior analysis technologies that can inform clinicians and healthcare providers with objective and quantified assessments of a person's well being. Exemplarily, the characterization and association of nonverbal behavior with underlying clinical conditions, such as depression or post-traumatic stress, holds transformative potential and could change treatment and the healthcare system's efficiency significantly. *Major publications:* [6, 4, 8, 9]
3. **Automatic and Virtual Human Interpersonal Skill Training:** Virtual humans are used in a wide range of healthcare applications, including psychological disorder assessment and treatment as well as social skills training applications, such as job interview training, public speaking training, conversational skills for autism spectrum disorder, and intercultural communicative skills training. Virtual human social skills training holds several advantages over traditional role-play based or non-experiential training (e.g. classroom lectures). Unlike traditional role-play based social skill training, virtual human based training is less dependent on trainer availability, scheduling, and costs. Further, human role-playing actors might introduce unnecessary and distracting variables to the training (e.g. culture, gender, language, age, etc.), while virtual humans' appearance and behavioral patterns can be precisely programmed, controlled and systematically presented to pace the exposure or interaction, starting at a level that the user is most capable of successfully interacting with and gradually ramping up in difficulty. By incorporating and advancing expertise in the fields of virtual agents, virtual environments for training and nonverbal audiovisual behavior analysis and machine learning, we aim to improve interpersonal skill training within this project. With our National Science Foundation (NSF) funded project focused on interpersonal skill training using multimodal machine learning and virtual humans we are at the forefront of this vibrant and quickly moving field of investigation. *Major publications:* [5, 40, 101, 100, 23]

Current Research Projects

- **Cicero - A Multimodal Virtual Audience Platform for Public Speaking Training (NSF IIS-1421330; 2014-2017); Role: PI:** Public speaking performances are not only characterized by the presentation of the content, but also by the presenters' nonverbal behavior, such as gestures, tone of voice, vocal variety, and facial expressions. Within this project, we seek to identify automatic nonverbal behavior descriptors that correlate with expert-assessments of behaviors characteristic of good and bad public speaking performances. Cicero - named after the great Roman orator Marcus Tullius Cicero - is a virtual audience platform designed for the training of public speaking within a controlled and forgiving environment. Cicero will enable us to conduct a wide variety of experiments reaching from performance assessments to psychological experiments, which would not be possible with a real human audience. NSF funded project (~500k USD).
- **Dyadic Entrainment and Comprehensive Voice Analysis for Virtual Human Interactions (ARL W911NF-14-D-0005; 2014-2017); Role: PI:** Entrainment refers to the tendency of interactants to adapt their communicative behavior to each other. The notion of dyadic entrainment between two interactants is central to successful and engaging interactions. Realistic and virtual character driven simulation and training scenarios has found its way into the training of future competent military and civilian leaders. Virtual humans in such training scenarios often, however, lack the capability to entrain with their human interactants in realtime during continuous interactions rendering the interaction less engaging and effective. Major goals of this effort: (1) data-driven dyadic nonverbal behavior assessment algorithms that detect and measure

entrainment and enable virtual humans to entrain. (2) *SmartVoice* technologies for realtime and robust speech analysis and processing. ARL (Army Research Laboratories) funded project (~540k USD).

- **Modeling Multimodal Communicative Behaviors for Virtual Humans (ARL W911NF-14-D-0005; 2014-2017); Role: PI:** The vision of this research effort is to automatically recognize, model and predict multimodal human nonverbal behavior in the context of interaction with virtual humans, robots and other human participants. For the first specific goal, we envision a new generation of authoring technologies able to efficiently capture the nonverbal identity (i.e., defining how a person or an agent uniquely gestures and vocalizes their sentences) of a person automatically from online multimedia content (e.g., YouTube) and then integrate these behaviors with a specific virtual human. A second specific goal of this project is to improve the learner centric experience with virtual humans by automatically assessing and quantifying the trainee's behaviors during the interaction and reporting multimodal behaviors associated with strong persuasive and learning skills. For example, a virtual human trainer should be able to recognize when a user is confused or disengaged and adjust its behaviors and expectations accordingly (~1.5M USD).
- **Multimodal Computational Framework for the Assessment and Enhancement of Team Cohesion and Performance in Human-System Integration (ARL W911NF-14-D-0005; 2015-2016); Role: PI** High military unit cohesion is a critical factor that enhances unit performance and promotes individual resilience to combat-related trauma. Traditional approaches to quantifying unit cohesion largely rely on questionnaires and cumbersome coding approaches, while the precise behavioral patterns underlying high unit cohesion remain unknown. Advancements in automatic behavior analysis, tracking, and machine learning enable novel approaches to assess complex and dynamic behavior of individuals and help further our understanding of the underlying mechanisms of high team cohesion. Understanding these mechanisms will allow us to predict and control dynamic interaction among humans and enable technology to become an equal member in human-system interaction enhancing and sustaining cohesion as a human team member would. The thorough understanding of the underlying mechanisms of team cohesion will inform the development of technology that tightly cooperates with human team members and has the potential to optimize human performance (~300k USD).
- **BIGDATA: EAGER: Catalyzing Research in Multimodal Learning Analytics (NSF IIS-1548254; 2015-2017); Role: co-PI:** Recent research in Multimodal Learning Analytics, and, more generally, in Multimodal Interaction, has reiterated the importance of capturing and analyzing student learning and cognition through a vast array of modalities (e.g., gesture, speech, pen, physiological sensors and eye gaze). Multimodal Learning Analytics has been particularly informative for understanding and assessing complex learning environments (e.g., Makerspaces, Collaborative Tangible User Interfaces). This two year proposal aims to employ an exploratory, iterative design and test process to develop guidelines, and a prototype for multiparty, multimodal data capture in classroom and laboratory environments. The proposed project is an interdisciplinary initiative, in which computer scientists, data scientists and education researchers collectively identify what is worth measuring (i.e., learning-relevant constructs), how to measure them (i.e., connecting analytic techniques and education theory) and how to capture the necessary data (i.e., the modalities and data quality requirements; ~300k USD).
- **Validation of SimSensei User State Sensing to Advance PTSD Clinical Assessment (ARL W911NF-14-R-0005; 2014-2016); Role: Key Researcher:** This two year project will aim to provide important information for enhancing the accuracy of Post-traumatic Stress Disorder (PTSD) diagnostic methods and provide an objective measure of clinical change in patients who are being treated for this clinical health condition. The proposed project will specifically investigate the use of the SimSensei Virtual Human (VH) Kiosk system to interview patients with PTSD due to Military Sexual Trauma (MST), prior to, during, and at the completion of a course of treatment using Virtual Reality (VR) exposure therapy (VRET). The SimSensei VH Kiosk system is designed to provide participants with the opportunity to engage in a private interview with an intelligent VH character named "Ellie". The approach is expected to provide important and unique assessment data than can be used to supplement the type of information that has been traditionally acquired from self-report psychometric questionnaires and structured interviews.

Past Research Projects

- **ICT Detection and Computational Analysis of Psychological Signals (DCAPS) (2011-2014); Role: Key Researcher:** DARPA funded research project to investigate use of telemedicine and virtual humans to address barriers to care and to provide better care for service members who seek treatment for psychological issues, including post-traumatic stress, depression and suicide risk. Within the DCAPS project I research on the identification and automatic detection of relevant nonverbal psychological signals (e.g. lack of mutual gaze, fidgeting, voice quality, and prosodic trends) in realtime using audiovisual sensors.
- **Prosodic Accommodation in Aviation (PAIA) (2011); Role: co-PI:** Enterprise Ireland (EI) funded Commercialisation Fund project investigating the measurement of speech accommodation as a means of determining effective pilot communication as part of aviation crew resource management (CRM) training. Within this 125k EUR project I function as co-principal investigator together with Dr. Brian Vaughan and Dr. Celine De Looze.

Research Projects in Review/Preparation

- **Health Behavior Informatics; Role: co-PI:** This project introduces our computational framework, named *health behavior informatics*, specifically designed to help clinicians to more effectively incorporate the results of nonverbal behavior analysis into their assessment and treatment of depression. The proposed research will advance this endeavor through the development of new probabilistic algorithms to automatically quantify nonverbal behaviors, salient affective moments of interest indicative of depression severity and nonverbal correlates of patient-clinician rapport. The computational framework will be evaluated in two clinical contexts related to depression: single-session depression screening interviews and multi-session Interpersonal Psychotherapy (IPT) treatments for depression (the latter allowing us to look at longitudinal changes in nonverbal behavior). I act as a co-PI on this project submitted to the NSF/NIH Smart and Connected Health program.
- **Investigating Thought Markers of Depression and Suicidality in Childhood Absence Epilepsy; Role: co-PI:** The present project aims to automatically assess and track acoustic nonverbal behaviors to assess complex psychological conditions like depression and suicidality in adolescents with epilepsy. In particular, Dr. Scherer's effort will further develop sophisticated signal processing techniques and machine learning approaches to better understand the processes and patterns of thought markers within longitudinal and repeated assessments and how they relate to mental and physical health. The analysis will extend across all available modalities, namely verbal, conversational, dyadic, and acoustic data. The automatic analysis of the nonverbal thought markers will help identify and characterize behaviors relevant to the investigated psychological disorders and developed technologies can be utilized to aid identifying other psychological disorders, e.g. post-traumatic stress or schizophrenia, in the future.

Invited Talks and Keynotes

1. Nonverbal Behavior Analytics for Healthcare Applications, Invited talk at 1st International Workshop on Advancements in Social Signal Processing for Multimodal Interaction, Seattle, WA, 2015.
2. Multimodal Behavior Analytics for Healthcare Applications and Interpersonal Skill Training, Keynote Talk at HTC Colloquium on Computational Modelling of Affect and Behaviour, Canberra, Australia, 2015.
3. Modeling communication dynamics: from distress indicators to virtual interviewers, Entertainment Software and Cognitive Neurotherapeutics Society (ESCoNS), San Francisco, 2015.
4. Modeling communication dynamics: from distress indicators to virtual interviewers, University of New South Wales Workshop on Advanced Signal Processing, Sydney, 2015.
5. Computational Behavior Analytics for Healthcare Applications, University of Southern California Alfred E. Mann Institute, Los Angeles, 2014.

6. Computational Behavior Analytics for Healthcare Applications, Veteran Affairs Clinic, University of California San Francisco, San Francisco, 2014.
7. Automatic audiovisual behavior descriptors for psychological disorder analysis within virtual human interaction, Austrian Institute for Artificial Intelligence OFAI, Vienna, 2013.
8. Automatic audiovisual behavior descriptors for psychological disorder analysis within virtual human interaction, Ulm University, Germany, 2013.
9. Automatic audiovisual behavior descriptors for psychological disorder analysis within virtual human interaction, Center for Robotics and Embedded Systems (CRES), University of Southern California, 2013.
10. Multimodal Behavior Analysis: Datasets, Applications, and Tools, D-Meta Challenge International Conference for Multimodal Interaction, Santa Monica, 2012
11. Tools and measures for the computational modeling of human nonverbal behaviors, Institute for Neural Information Processing, Ulm University, Germany, 2012
12. Detecting nonverbal behavioral cues using multimodal and fuzzy machine learning approaches, Signal Analysis and Interpretation Laboratory (SAIL), University of Southern California, April 2012
13. How Partially Supervised Learning can facilitate and enhance user state analysis in naturalistic HCI, International Workshop on Partially Supervised Learning, Ulm University, Ulm Germany, September 2011
14. Analyzing the user's state in HCI: from crisp emotions to conversational dispositions, FastNet Symposium, Speech Communication Laboratory, Trinity College Dublin, Dublin Ireland, March 2011
15. Emotion/Disposition Recognition in Human Computer Interaction, Dipartimento di Ingegneria dell'Informazione, Universita Degli Studi di Siena, Siena Italy, November 2009
16. Conversation Analysis using a Process Engine for Pattern Recognition, Advanced Telecommunications Research Institute (ATR), Kyoto Japan, April 2009

Publications

Journal Articles

[1] Albert RIZZO, Gale LUCAS, Jonathan GRATCH, Giota STRATOU, Louis-Philippe MORENCY, Kenneth CHAVEZ, Russ SHILLING, and Stefan SCHERER. "Automatic Behavior Analysis During a Clinical Interview with a Virtual Human." In: *Medicine Meets Virtual Reality 22: NextMed/MMVR22* 220 (2016), p. 316.

[2] Stefan Scherer. "Multimodal behavior analytics for interactive technologies". In: *KI-Künstliche Intelligenz* 30.1 (2016), pp. 91–92.

[3] V. Venek, S. Scherer, L.-P. Morency, A. Rizzo, and J. P. Pestian. "Adolescent Suicidal Risk Assessment in Clinician-Patient Interaction". In: *IEEE Transactions on Affective Computing* (2016). DOI: 10.1109/TAFFC.2016.2518665.

[4] N. Cummins, S. Scherer, J. Krajewski, S. Schnieder, J. Epps, and T. Quatieri. "A Review of Depression and Suicide Risk Assessment using Speech Analysis". In: *Speech Communication* 17 (2015), pp. 10–49.

[5] S. Park, S. Scherer, J. Gratch, P. Carnevale, and L.-P. Morency. "I Can Already Guess Your Answer: Predicting Respondent Reactions during Dyadic Negotiation". In: *IEEE Transactions on Affective Computing* 6.2 (2015), pp. 86–96.

[6] S. Scherer, G. Lucas, J. Gratch, A. Rizzo, and L.-P. Morency. "Self-reported symptoms of depression and PTSD are associated with reduced vowel space in screening interviews". In: *IEEE Transactions on Affective Computing (in press; doi: 10.1109/TAFFC.2015.2440264)* (2015).

[7] C. De Looze, S. Scherer, B. Vaughan, and N. Campbell. "Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction". In: *Speech Communication* 58 (2014), pp. 11–34.

[8] S. Scherer, G. Stratou, G. Lucas, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo, and L.-P. Morency. “Automatic Audiovisual Behavior Descriptors for Psychological Disorder Analysis”. In: *Image and Vision Computing Journal, Special Issue on Best of Face and Gesture 2013* 32.10 (2014), pp. 648–658.

[9] G. Stratou, S. Scherer, J. Gratch, and L.-P. Morency. “Automatic nonverbal behavior indicators of depression and PTSD: the effect of gender”. In: *Journal on Multimodal User Interfaces* 9.1 (2014), pp. 1–13.

[10] E. Trentin, S. Scherer, and F. Schwenker. “Emotion recognition from speech signals via a probabilistic echo-state network”. In: *Pattern Recognition Letters* (2014). DOI: 10.1016/j.patrec.2014.10.015.

[11] M. Schels, S. Scherer, M. Glodek, H. A. Kestler, G. Palm, and F. Schwenker. “On the Discovery of Events in EEG Data utilizing Information Fusion”. In: *Computational Statistics; Special Issue: Proceedings of Reisensburg 2010* 28.1 (2013), pp. 5–18.

[12] S. Scherer, J. Kane, C. Gobl, and F. Schwenker. “Investigating Fuzzy-Input Fuzzy-Output Support Vector Machines for Robust Voice Quality Classification”. In: *Computer Speech and Language* 27.1 (2013), pp. 263–287. DOI: 10.1016/j.csl.2012.06.001.

[13] J. Gratch, L.-P. Morency, S. Scherer, G. Stratou, J. Boberg, S. Koenig, T. Adamson, and A. Rizzo. “User-State Sensing for Virtual Health Agents and TeleHealth Applications.” In: *Studies in health technology and informatics* 184 (2012), pp. 151–157.

[14] K. Jokinen and S. Scherer. “Embodied Communicative Activity in Cooperative Conversational Interactions - studies in Visual Interaction Management”. In: *Special Issue of Acta Polytechnica Hungarica: CogInfoCom 2011* 9.1 (2012), pp. 19–40.

[15] S. Scherer, M. Glodek, G. Layher, M. Schels, M. Schmidt, T. Brosch, S. Tschechne, F. Schwenker, H. Neumann, and G. Palm. “A Generic Framework for the Inference of User States in Human Computer Interaction: How patterns of low level communicational cues support complex affective states”. In: *Journal on Multimodal User Interfaces, special issue on: Conceptual frameworks for Multimodal Social Signal Processing* 6.3 (2012), pp. 117–141.

[16] S. Scherer, M. Glodek, F. Schwenker, N. Campbell, and G. Palm. “Spotting Laughter in natural multiparty conversations: a comparison of automatic online and offline approaches using audiovisual data”. In: *ACM Transactions on Interactive Intelligent Systems: Special Issue on Affective Interaction in Natural Environments* 2.1 (2012), 4:1–4:31.

Thesis

[17] S. Scherer. “Analyzing the User’s State in HCI: From Crisp Emotions to Conversational Dispositions”. PhD thesis. Ulm University, 2011. URL: http://dx.doi.org/10.18725/OPARU-1766.

Peer Reviewed Conference Articles

[18] Mathieu Chollet, Torsten Wörtwein, Louis-Philippe Morency, and Stefan Scherer. “A Multimodal Corpus for the Assessment of Public Speaking Ability and Anxiety”. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. European Language Resources Association (ELRA), 2016. ISBN: 978-2-9517408-9-1.

[19] Sayan Ghosh, Eugene Laksana, Louis-Philippe Morency, and Stefan Scherer. “Learning Representations of Affect from Speech”. In: *Proceedings of International Conference on Learning Representations (ICLR)*. arXiv preprint arXiv:1511.04747. 2016.

[20] Maarten Brilman and Stefan Scherer. “A Multimodal Predictive Model of Successful Debaters or How I Learned to Sway Votes”. In: *Proceedings of the 23rd ACM international conference on Multimedia*. ACM. 2015, pp. 149–158.

[21] M. Chatterjee, S. Park, L.-P. Morency, and S. Scherer. “Combining Two Perspectives on Classifying Multimodal Data for Recognizing Speaker Traits”. In: *to appear in International Conference Multimodal Interaction (ICMI 2015)*. 2015.

[22] Moitreya Chatterjee, Sunghyun Park, Louis-Philippe Morency, and Stefan Scherer. “Combining Two Perspectives on Classifying Multimodal Data for Recognizing Speaker Traits”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM. 2015, pp. 7–14.

[23] M. Chollet, T. Wortwein, L.-P. Morency, A. Shapiro, and S. Scherer. “Exploring Feedback Learning Strategies to Improve Public Speaking: An Interactive Virtual Audience Framework”. In: *to appear in Proceedings of ACM UbiComp 2015*. 2015.

[24] Mathieu Chollet, Kalin Stefanov, Helmut Prendinger, and Stefan Scherer. “Public Speaking Training with a Multimodal Interactive Virtual Audience Framework”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM. 2015, pp. 367–368.

[25] Mathieu Chollet, Torsten Wörtwein, Louis-Philippe Morency, Ari Shapiro, and Stefan Scherer. “Exploring feedback strategies to improve public speaking: an interactive virtual audience framework”. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM. 2015, pp. 1143–1154.

[26] Sayan Ghosh, Eugene Laksana, Stefan Scherer, and Louis-Philippe Morency. “A multi-label convolutional neural network approach to cross-domain action unit detection”. In: *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE. 2015, pp. 609–615.

[27] Gale M Lucas, Jonathan Gratch, Stefan Scherer, Jill Boberg, and Giota Stratou. “Towards an affective interface for assessment of psychological distress”. In: *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE. 2015, pp. 539–545.

[28] S. Scherer, L.-P. Morency, J. Gratch, and J. P. Pestian. “REDUCED VOWEL SPACE IS A ROBUST INDICATOR OF PSYCHOLOGICAL DISTRESS: A CROSS-CORPUS ANALYSIS”. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2015, pp. 4789–4793.

[29] H. S. Shim, S. Park, M. Chatterjee, S. Scherer, K. Sagae, and L.-P. Morency. “ACOUSTIC AND PARAVERBAL INDICATORS OF PERSUASIVENESS IN SOCIAL MULTIMEDIA”. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing*. 2015.

[30] Marcelo Worsley, Stefan Scherer, Louis-Philippe Morency, and Paulo Blikstein. “Exploring Behavior Representation for Learning Analytics”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM. 2015, pp. 251–258.

[31] Torsten Wörtwein, Mathieu Chollet, Boris Schauerte, Louis-Philippe Morency, Rainer Stiefelhagen, and Stefan Scherer. “Multimodal public speaking performance assessment”. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM. 2015, pp. 43–50.

[32] Torsten Wörtwein, Louis-Philippe Morency, and Stefan Scherer. “Automatic assessment and analysis of public speaking anxiety: A virtual audience case study”. In: *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*. IEEE. 2015, pp. 187–193.

[33] M. Chatterjee, G. Stratou, S. Scherer, and L.-P. Morency. “Context-based Signal Descriptors of Heart-Rate Variability for Anxiety Assessment”. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2014, pp. 3659–3663.

[34] M. Chollet, G. Stratou, A. Shapiro, L.-P. Morency, and S. Scherer. “An Interactive Virtual Audience Platform for Public Speaking Training”. In: *Proceedings of International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2014, pp. 1657–1658.

[35] D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgilia, J. Gratch, A. Hartholt, M. Lhommet, G. Lucas, S. Marsella, F. Morbini, A. Nazarian, S. Scherer, G. Stratou, A. Suri, D. Traum, R. Wood, Y. Xu, A. Rizzo, and L.-P. Morency. “SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support”. In: *Proceedings of Autonomous Agents and Multiagent Systems (AAMAS)*. 2014, pp. 1061–1068.

[36] G. Degottex, J. Kane, T. Drugman, T. Raitio, and S. Scherer. “COVAREP - A collaborative voice analysis repository for speech technologies”. In: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014)*. 2014, pp. 960–964.

[37] J. Gratch, R. Artstein, G. Lucas, S. Scherer, A. Nazarian, R. Wood, J. Boberg, D. DeVault, S. Marsella, D. Traum, A. Rizzo, and L.-P. Morency. “The Distress Analysis Interview Corpus of human and computer interviews”. In: *Proceedings of Language Resources and Evaluation Conference (LREC)*. 2014, pp. 3123–3128.

[38] S. Scherer, Z. Hammal, Y. Yang, L.-P. Morency, and J. F. Cohn. “Dyadic Behavior Analysis in Depression Severity Assessment Interviews”. In: *Proceedings of International Conference on Multimodal Interaction*. ACM, 2014.

[39] V. Venek, S. Scherer, L.-P. Morency, A. Rizzo, and J. P. Pestian. “Adolescent Suicidal Risk assessment in clinician-patient interaction: A Study of verbal and acoustic behaviors”. In: *Proceedings of Spoken Langugae Technologies (SLT)*. 2014.

[40] L. Batrinca, G. Stratou, A. Shapiro, L.-P. Morency, and S. Scherer. “Cicero - Towards a Multimodal Virtual Audience Platform for Public Speaking Training”. In: *Proceedings of Intelligent Virtual Agents (IVA) 2013*. Springer, 2013, pp. 116–128.

[41] D. DeVault, K. Georgilia, R. Artstein, F. Morbini, D. Traum, S. Scherer, A. Rizzo, and L.-P. Morency. “Verbal indicators of psychological distress in interactive dialogue with a virtual human”. In: *Proceedings of SigDial 2013*. Association for Computational Linguistics, 2013, pp. 193–202.

[42] H. Iseman, B. Ilari, H.-Y. Pan, S. Scherer, and A. Habibi. “Synchronization, empathy, and prosocial behavior in children”. In: *Proceedings of the meeting of the Society for Music Perception and Cognition*. 2013.

[43] J. Kane, S. Scherer, M. Aylett, L.-P. Morency, and C. Gobl. “Speaker and Language Independent Voice Quality Classification Applied to Unlabelled Corpora of Expressive Speech”. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2013, pp. 7982–7986.

[44] J. Kane, S. Scherer, L.-P. Morency, and C. Gobl. “A comparative study of glottal open quotient estimation techniques”. In: *Proceedings of Interspeech 2013*. ISCA, 2013, pp. 1658–1662.

[45] S. Marsella, A. Shapiro, A. Feng, Y. Xu, M. Lhommet, and S. Scherer. “Towards Higher Quality Character Performance in Previz”. In: *Proceedings of Digital Production Symposium 2013 (DigiPro2013)*. ACM, 2013, pp. 31–35.

[46] S. Marsella, Y. Xu, M. Lhommet, A. Feng, S. Scherer, and A. Shapiro. “Virtual Character Performance From Speech”. In: *Proceedings of Symposium on Computer Animation 2013 (SCA2013)*. ACM, 2013, pp. 25–35.

[47] E. Nouri, S. Park, S. Scherer, J. Gratch, P. Carnevale, L.-P. Morency, and D. Traum. “Prediction of Strategy and Outcome as Negotiation Unfolds by Using Basic Verbal and Behavioral Features”. In: *Proceedings of Interspeech 2013*. ISCA, 2013, pp. 1458–1461.

[48] S. Park, S. Scherer, J. Gratch, P. Carnevale, and L.-P. Morency. “Mutual Behaviors During Dyadic Negotiation: Automatic Prediction of Respondent Reactions”. In: *Proceedings of Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2013. (**Nominated Best Student Paper Award**.)

[49] S. Scherer, J. P. Pestian, and L.-P. Morency. “Investigating the Speech Characteristics of Suicidal Adolescents”. In: *Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, 2013, pp. 709–713.

[50] S. Scherer, G. Stratou, J. Gratch, and L.-P. Morency. “Investigating Voice Quality as a Speaker-Independent Indicator of Depression and PTSD”. In: *Proceedings of Interspeech 2013*. ISCA, 2013, pp. 847–851.

[51] S. Scherer, G. Stratou, M. Mahmoud, J. Boberg, J. Gratch, A. Rizzo, and L.-P. Morency. “Automatic Behavior Descriptors for Psychological Disorder Analysis”. In: *Proceedings of IEEE Conference on Automatic Face and Gesture Recognition*. IEEE, 2013.

[52] S. Scherer, G. Stratou, and L.-P. Morency. “Audiovisual Behavior Descriptors for Depression Assessment”. In: *Proceedings of International Conference on Multimdoal Interaction*. ACM, 2013.

[53] G. Stratou, S. Scherer, J. Gratch, and L.-P. Morency. “Automatic Nonverbal Behavior Indicators of Depression and PTSD: Exploring Gender Differences”. In: *Proceedings of International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2013.

[54] Y. Zhou, S. Scherer, D. Devault, J. Gratch, G. Stratou, L.-P. Morency, and J. Cassell. “Multimodal Prediction of Psychological Disorder: Learning Nonverbal Commonality in Adjacency Pairs”. In: *Proceedings of Workshop Series on the Semantics and Pragmatics of Dialogue*. 2013.

[55] C. De Looze, A. Ghio, S. Scherer, G. Pouchoulin, and F. Viallet. “Automatic analysis of prosodic variations in Parkinson read and semi-spontaneous speech”. In: *Proceedings of Speech Prosody 2012*. ISCA, 2012.

[56] C. De Looze, S. Scherer, B. Vaughan, and N. Campbell. “Mesure automatique de l’adaptation prosodique en interaction conversationnelle”. In: *Proceedings of Journees d’Etudes sur la Parole (JEP 2012)*. 2012.

[57] S. Finkelstein, S. Scherer, A. Ogan, L.-P. Morency, and J. Cassell. “Investigating the Influence of Virtual Peers as Dialect Models on Students’ Prosodic Inventory”. In: *Proceedings of Workshop on Child, Computer and Interaction (WOCCI’12)*. ISCA, 2012.

[58] S. Scherer, J. Kane, C. Gobl, and F. Schwenker. “The Effect of Fuzzy Training Targets on Voice Quality Classification”. In: *Proceedings of First IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human Computer Interaction*. 2012.

[59] S. Scherer, G. Layher, J. Kane, H. Neumann, and N. Campbell. “An audiovisual political speech analysis incorporating eye-tracking and perception data”. In: *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC’12)*. ELRA, 2012, pp. 1114–1120.

[60] S. Scherer, S. Marsella, G. Stratou, Y. Xu, F. Morbini, A. Egan, A. Rizzo, and L.-P. Morency. “Perception Markup Language: Towards a Standardized Representation of Perceived Nonverbal Behaviors”. In: *Proceedings of Intelligent Virtual Agents (IVA’12)*. LNAI 7502. Springer, 2012, pp. 455–463.

[61] S. Scherer, N. Weibel, S. L. Oviatt, and L.-P. Morency. “Multimodal prediction of expertise and leadership in learning groups”. In: *Proceedings of International Conference on Multimodal Interaction*. 2012.

[62] E. Szekely, K. Kane, S. Scherer, and J. Carson-Berndsen. “Detecting a targeted voice style in an audiobook using voice quality features”. In: *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2012)*. IEEE, 2012, pp. 4593–4596.

[63] J. Esparza, S. Scherer, and F. Schwenker. “Studying self- and active-training methods for multi-feature set emotion recognition”. In: *Proceedings of the Workshop on Partially Supervised Learning (PSL’11)*. Springer, 2011, pp. 19–31.

[64] M. Glodek, S. Scherer, F. Schwenker, and G. Palm. “Conditioned Hidden Markov Model Fusion for Multimodal Classification”. In: *Proceedings of Interspeech 2011*. ISCA, 2011, pp. 2269–2272.

[65] M. Glodek, S. Tschechne, G. Layher, M. Schels, T. Brosch, S. Scherer, M. Kächele, M. Schmidt, H. Neumann, G. Palm, and F. Schwenker. “Multiple Classifier Systems for the Classification of Audio-Visual Emotional States”. In: *Affective Computing and Intelligent Interaction*. Ed. by Sidney D’Mello, Arthur Graesser, Björn Schuller, and Jean-Claude Martin. Vol. 6975. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2011, pp. 359–368.

[66] G. Layher, S. Tschechne, S. Scherer, T. Brosch, C. Curio, and H. Neumann. “Social Signal Processing in Companion Systems - Challenges Ahead”. In: *Proceedings of Informatik 2011*. Ed. by H.-U. Heiss, P. Pepper, H. Schlingloff, and J. Schneider. Springer, 2011, pp. 239–239.

[67] C. Oertel, C. De Looze, S. Scherer, A. Windmann, and P. Wagner. “Towards the Automatic Detection of Involvement in Conversation”. In: *Proceedings of the SSPnet-COST 2102 PINK International Conference*. Ed. by A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, and A. Nijholt. Springer, 2011, pp. 163–170.

[68] C. Oertel, S. Scherer, P. Wagner, and N. Campbell. “On the use of multimodal cues for the prediction of involvement in spontaneous conversation”. In: *Proceedings of Interspeech 2011*. ISCA, 2011, pp. 1541–1544.

[69] S. Scherer, M. Schels, and G. Palm. “How low level observations can help to reveal the user’s state in HCI”. In: *Proceedings of the 4th international conference on Affective Computing and Intelligent Interaction (ACII’11)*. Ed. by Sidney D’Mello, Arthur Graesser, Björn Schuller, and Jean-Claude Martin. Vol. 2. Springer, 2011, pp. 81–90.

[70] S. Walter, S. Scherer, M. Schels, M. Glodek, D. Hrabal, M. Schmidt, R. Böck, K. Limbrecht, H. C. Traue, and F. Schwenker. “Multimodal Emotion Classification in Naturalistic User Behavior”. In: *Proceedings of 14th International Conference on Human-Computer Interaction (HCII’11)*. Ed. by J. A. Jacko. Vol. 3. Springer, 2011, pp. 603–611.

[71] W. N. Campbell and S. Scherer. “Comparing Measures of Synchrony and Alignment in Dialogue Speech Timing with respect to Turn-taking Activity”. In: *Proceedings of Interspeech 2010*. ISCA, 2010, pp. 2546–2549.

[72] S. Egger, R. Schatz, and S. Scherer. “It Takes Two to Tango - Assessing The Impact of Delay on Conversational Interactivity on Perceived Speech Quality”. In: *Proceedings of Interspeech 2010*. ISCA, 2010, pp. 1321–1324.

[73] V. Fritzsch, S. Scherer, and F. Schwenker. “An Open Source Process Engine Framework for Realtime Pattern Recognition and Information Fusion Tasks”. In: *Proceedings of the Seventh International conference on Language Resources and Evaluation (LREC’10)*. Ed. by N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, M. Piperidis S.and Rosner, and D. Tapias. Valletta, Malta: European Language Resources Association (ELRA), 2010, pp. 880–884. ISBN: 2-9517408-6-7.

[74] S. Scherer, I. Siegert, L. Bigalke, and S. Meudt. “Developing an Expressive Speech Labeling Tool Incorporating the Temporal Characteristics of Emotion”. In: *Proceedings of the Seventh International conference on Language Resources and Evaluation (LREC’10)*. Ed. by N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, M. Piperidis S.and Rosner, and D. Tapias. Valletta, Malta: European Language Resources Association (ELRA), 2010, pp. 1172–1175. ISBN: 2-9517408-6-7.

[75] F. Schwenker, S. Scherer, M. Schmidt, M. Schels, and M. Glodek. “Multiple Classifier Systems for the Recognition of Human Emotions”. In: *9th International Workshop on Multiple Classifier Systems (MCS 2010)*. Ed. by N. El Gayar, J. Kittler, and F. Roli. Springer, 2010, pp. 315–324.

[76] P.-M. Strauss, S. Scherer, G. Layher, and H. Hoffmann. “Evaluation of the PIT Corpus Or What a Difference a Face Makes?” In: *Proceedings of the Seventh International conference on Language Resources and Evaluation (LREC’10)*. Ed. by N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odjik, S. Piperidis, M. Rosner, and D. Tapias. Valletta, Malta: ELRA, 2010, pp. 3470–3474. ISBN: 2-9517408-6-7.

[77] E. Trentin, S. Scherer, and F. Schwenker. “Maximum Echo-State Likelihood Networks for Emotion Recognition”. In: *Proceedings of Artificial Neural Networks in Pattern Recognition (ANNPR’10)*. Springer, 2010, pp. 61–72.

[78] S. Scherer, V. Fritzsch, and F. Schwenker. “Multimodal real-time conversation analysis using a novel process engine”. In: *Proceedings of the 3rd international conference on Affective Computing and Intelligent Interaction (ACII’09)*. IEEE, 2009, pp. 253–255.

[79] S. Scherer, V. Fritzsch, F. Schwenker, and N. Campbell. “Demonstrating laughter detection in natural discourses”. In: *Proceedings of the Interdisciplinary Workshop on Laughter and other Interactional Vocalisations in Speech*. 2009.

[80] S. Scherer, F. Schwenker, W. N. Campbell, and G. Palm. “Multimodal laughter detection in natural discourses”. In: *Proceedings of 3rd International Workshop on Human-Centered Robotic Systems (HCRS’09)*. Ed. by H. Ritter, G. Sagerer, R. Dillmann, and M. Buss. Cognitive Systems Monographs. Springer, 2009, pp. 111–121.

[81] S. Scherer, E. Trentin, F. Schwenker, and G. Palm. “Approaching emotion in human computer interaction”. In: *International Workshop on Spoken Dialogue Systems (IWSDS’09)*. 2009, pp. 156–168.

[82] F. Schwenker, S. Scherer, Y. Magdi, and G. Palm. “The GMM-SVM Supervector Approach for the Recognition of the Emotional Status from Speech”. In: *19th International Conference on Artificial Neural Networks 2009 Part I*. Ed. by C. Alippi. Vol. LNCS 5768. Berlin Heidelberg: Springer, 2009, pp. 894–903.

[83] S. Scherer and W. N. Campbell. “Automatic Laughter Detection for Measuring Discourse Engagement”. In: *Autumn Meeting of the Acoustical Society of Japan 2008 (ASJ’08)*. 2008, pp. 265–266.

[84] S. Scherer, H. Hofmann, M. Lampmann, M. Pfeil, S. Rhinow, F. Schwenker, and G. Palm. “Emotion Recognition from Speech: Stress Experiment”. In: *Proceedings of the Sixth International Language Resources and Evaluation (LREC’08)*. Marrakech, Morocco: European Language Resources Association (ELRA), 2008, pp. 1325–1330.

[85] S. Scherer, M. Oubbati, F. Schwenker, and G. Palm. “Real-Time Emotion Recognition Using Echo State Networks”. In: *Proceedings of the 4th IEEE tutorial and research workshop on Perception and Interactive Technologies for Speech-Based Systems (PIT’08)*. Kloster Irsee, Germany: Springer, 2008, pp. 200–204. ISBN: 978-3-540-69368-0.

[86] S. Scherer, M. Oubbati, F. Schwenker, and G. Palm. “Real-Time Emotion Recognition from Speech Using Echo State Networks”. In: *Proceedings of the 3rd IAPR workshop on Artificial Neural Networks in Pattern Recognition (ANNPR’08)*. Paris, France: Springer, 2008, pp. 205–216. ISBN: 978-3-540-69938-5.

[87] S. Scherer and P.-M. Strauss. “A Flexible Wizard of Oz Environment for Rapid Prototyping”. In: *Proceedings of the Sixth International Language Resources and Evaluation (LREC’08)*. Marrakech, Morocco: ELRA, 2008, pp. 958–961.

[88] P.-M. Strauss, H. Hoffmann, W. Minker, H. Neumann, G. Palm, S. Scherer, H. C. Traue, and U. Weidenbacher. “The PIT corpus of German multi-party dialogues”. In: *Proceedings of the Sixth International Language Resources and Evaluation (LREC’08)*. Marrakech, Morocco: ELRA, 2008, pp. 2442–2445.

[89] H. K. Maganti, S. Scherer, and G. Palm. “A novel feature for emotion recognition in voice based applications”. In: *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction (ACII’07)*. Springer, 2007, pp. 710–711.

[90] S. Scherer, F. Schwenker, and G. Palm. “Classifier fusion for emotion recognition from speech”. In: *3rd IET International Conference on Intelligent Environments 2007 (IE’07)*. IEEE, 2007, pp. 152–155.

[91] P.-M. Strauss, H. Hoffmann, and S. Scherer. “Evaluation and user acceptance of a dialogue system using Wizard-of-Oz recordings”. In: *3rd IET International Conference on Intelligent Environments 2007 (IE’07)*. IEEE, 2007, pp. 521–524.

[92] C. Thiel, S. Scherer, and F. Schwenker. “Fuzzy-Input Fuzzy-Output One-Against-All Support Vector Machines”. In: *11th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems (KES’07)*. Vol. 3. Lecture Notes in Artificial Intelligence. Springer, 2007, pp. 156–165.

[93] P.-M. Strauss, H. Hoffmann, H. Neumann, W. Minker, G. Palm, S. Scherer, F. Schwenker, H. C. Traue, and U. Weidenbacher. “Wizard-of-Oz data collection for perception and interaction in multi-user environments”. In: *Proceedings of the Fifth International Language Resources and Evaluation (LREC’06)*. ELRA, 2006, pp. 2014–2017.

Book Chapters

[94] Albert Rizzo, Russell Shilling, Eric Forbell, Stefan Scherer, Jonathan Gratch, and Louis-Philippe Morency. “Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing”. In: *Artificial Intelligence in Behavioral and Mental Health Care* (2015), p. 53.

[95] M. Schels, M. Glodek, S. Meudt, S. Scherer, M. Schmidt, G. Layher, S. Tschechne, T. Brosch, D. Hrabal, S. Walter, H. C. Traue, G. Palm, H. Neumann, and F. Schwenker. “Multi-Modal Classifier-Fusion for the Recognition of Emotions”. In: ed. by M. Rojc and N. Campbell. Vol. Coverbal Synchrony in Human-Machine Interaction. CRC Press, 2013. Chap. 4, pp. 73–97.

[96] H. C. Traue, F. Ohl, A. Brechmann, F. Schwenker, H. Kessler, K. Limbrecht, H. Hoffmann, S. Scherer, M. Kozyba, A. Scheck, and S. Walter. “A Framework for Emotions and Dispositions in Man-Companion Interaction”. In: ed. by M. Rojc and N. Campbell. Vol. Coverbal Synchrony in Human-Machine Interaction. CRC Press, 2013. Chap. 5, pp. 98–140.

[97] S. Scherer, F. Schwenker, and G. Palm. “Classifier fusion for emotion recognition from speech”. In: ed. by W. Minker, M. Weber, H. Hagras, V. Callagan, and A. D. Kameas. Advanced Intelligent Environments. Springer, 2009. Chap. 5, pp. 95–117.

[98] S. Scherer, F. Schwenker, and G. Palm. “Emotion recognition from speech using multi-classifier systems and RBF-ensembles”. In: ed. by B. Prasad and S.R.M. Prasanna. Vol. 83. Studies in Computational Intelligence. Springer, 2008. Chap. 3, pp. 49–70.

Patents

[99] S. Scherer, J.P. Pestian, and L.P. Morency. *System and Method for Assessing Suicide Risk of a Patient Based Upon Non-Verbal Characteristics of Voice Data*. US Patent App. 14/893,253. 2016. URL: https://www.google.com/patents/US20160099011.

Media Coverage

- Dormehl, L. (2016): *Machine Learning Algorithm Spots Depression in Speech Patterns*; Digital Trends. http://tinyurl.com/digitalTrends-machine-learning
- Byrne, M. (2016): *Machine Learning Algorithm Spots Depression in Speech Patterns*; Motherboard, VICE. http://tinyurl.com/machine-learning-depression
- Cremin, G. (2016): *Robots Are Learning to Fake Empathy*; Motherboard, VICE. http://tinyurl.com/mb-empathy
- Robinson, A. (2015): *Meet Ellie, the machine that can detect depression*; The Guardian. http://tinyurl.com/MeetEllieGuardian
- Kaplan, M. (2015): *Sounds bad: Computers can spot symptoms of depression psychiatrists may miss*; The Economist. http://tinyurl.com/soundsBad-2015
- Farrell, L. (2015): *Dr Roboto will see you now*; Science Node. http://tinyurl.com/DrRoboto-2015
- Lafrance, A. (2015): *Machines That Can See Depression on a Person’s Face*; The Atlantic. http://tinyurl.com/Depression-TheAtlantic
- Carrasquillo, R. (2015): *Dr. Stefan Scherer from USC talks qualitative research, virtual humans, and the future of research tools.* Clevermind IO Blog. http://tinyurl.com/clevermind-blog-scherer
- Belman, O. (2014): *A New Way to Practice What You Preach*; USC Viterbi magazine. http://tinyurl.com/kuykdwz
- Firth, N. (2013): *Digital shrinks find depressed faces and body language*; New Scientist, Issue 2910; http://tinyurl.com/l9x9rdn
- Hall, T. (2014): *How virtual therapy could help fight PTSD* [Television series episode]. In MSNBC Los Angeles, California http://tinyurl.com/la5e4jt
- Alvarado, G. (2013): *Virtual human “Ellie”* [Television series episode]. In CNN en Español. Los Angeles, California http://tinyurl.com/m3sdhsn
- Leithead, A. (2013): *Consulting California’s ’virtual shrink’* [Radio series episode]. In Today programme. London, England: BBC. http://tinyurl.com/mmrdmn5
- Spiegel, A. (2013). *If your shrink is a bot, how do you respond?* [Radio series episode]. In Morning Edition. Los Angeles, California: National Public Radio (NPR) http://tinyurl.com/lq7h38e

---

Professional Service Activity

Chairing and Organizing Activities

- **Program Chair** International Conference on Intelligent Virtual Agents 2016

- **Organizer** Speaker Series in Big Data and Human Behavior (Sponsored by USC with 6k USD per year)
- **Tutorial Organizer** A review of depression and suicide risk assessment using speech analysis at Interspeech 2015 in Dresden, Germany.
- **Demonstration and Exhibit Chair** of the ACM International Conference on Multimodal Interaction (ICMI 2015) in Seattle, WA.
- **Demonstration and Exhibit Chair** of the ACM International Conference on Multimodal Interaction (ICMI 2013) in Sydney, Australia.
- **Co-organizer** of the first international Workshop on Multimodal Learning Analytics co-located to the International Conference on Multimodal Interaction (ICMI 2012) in Santa Monica, California.
- **Co-organizer** of the second international Workshop on Multimodal Learning Analytics co-located to the International Conference on Multimodal Interaction (ICMI 2013) in Sydney, Australia.
- **Co-organizer** of the first international Workshop on Multimodal Pattern Recognition of Social Signals (MPRSS) in Human Computer Interaction co-located to the International Conference on Pattern Recognition (ICPR 2012) in Tokyo, Japan.
- **Co-organizer** of the second international Workshop on Multimodal Pattern Recognition of Social Signals (MPRSS) in Human Computer Interaction co-located to the IEEE International Conference on Cybernetics (CYBCONF 2013) in Lausanne, Switzerland.

Editorial Activities

- **Associate Editor** of IEEE Transactions on Affective Computing
- **Guest Editor** of Pattern Recognition Letters Journal Special Issue on Pattern Recognition in Human Computer Interaction
- **Guest Editor** of IEEE Transactions on Affective Computing Special Issue on Laughter
- **Review Editor** of Frontiers in Computational Neuroscience

Scientific Committee and Reviewing Activities

- **Award Review Panelist** for NSF Information & Intelligent Systems Division, 2016

- **Scientific Committee member** of the Language Resources and Evaluation Conference (LREC)
- **Program Committee member** of the conference on Affective Computing and Intelligent Interaction (ACII)
- **Program Committee member** of the conference on Intelligent Virtual Agents (IVA)
- **Program Committee member** of the conference on International Audiovisual Emotion Challenge (AVEC)
- **Program Committee member** of the conference on Autonomous Agents and Multiagent Systems (AAMAS)
- **Program Committee member** of the conference on Multimodal Machine Learning Workshop in conjunction with NIPS 2015 (MMML)
- **Reviewer** for ACM Transactions on Interactive Intelligent Systems (TiiS)
- **Reviewer** for Elsevier Computer Speech and Language (CSL)
- **Reviewer** for Elsevier Pattern Recognition Letters (PRL)
- **Reviewer** for Elsevier Image and Vision Computing Journal (IVCJ)
- **Reviewer** for Elsevier Neurocomputing
- **Reviewer** for Springer Journal on Computational Statistics (COST)
- **Reviewer** for IEEE Transactions on Affective Computing (TAC)
- **Reviewer** for IEEE Signal Processing Letters (SPL)
- **Reviewer** for Clinical Pharmacology and Therapeutics (Nature Publishing)

Professional Memberships

- **Member** IEEE, ACM, ISCA