

Guest Editorial: Towards Machines Able to Deal with Laughter

Maurizio Mancini, Radoslaw Niewiadomski[✉], Shuji Hashimoto, Mary Ellen Foster[✉],
Stefan Scherer, and Gualtiero Volpe[✉]



1 INTRODUCTION

LAUGHTER is a significant feature of human-human communication. Laughter is characterized by a complex expressive behavior that includes major modalities: auditory, facial expressions, body movements and postural attitudes, and physiological signals. It conveys various meanings and accompanies different emotions, such as amusement, relief, irony, or embarrassment. Laughter has a strong social dimension. For example, it can reduce the sense of threat in a group and facilitate sociability and cooperation. It may also have positive effects on learning, creativity, health, and well-being. Because of its relevance in human-human communication, research on laughter deserves attention from the affective computing community. Several recent initiatives, such as the Special Session on Laughter at the 6th International Conference on Affective Computing and Intelligent Interaction (ACII 2015) and the series of Interdisciplinary Workshops on Laughter and other Non-Verbal Vocalizations in Speech, witness the importance of the topic.

Recent research projects (e.g., [1]) focused on laughter by investigating automatic laughter processing, and by developing proof-of concepts, experiments, and prototypes exploiting laughter for enhancing human-machine interaction in a broad variety of application scenarios. Major research areas where scientific and technological breakthroughs emerged are, for example, those related to the scientific investigation of the different forms of laughter expression and of the mechanisms of laughter elicitation (see e.g., [2], [3]), automated laughter detection from multimodal data sources (see e.g., [4], [5]), and laughter synthesis (e.g., in artificial agents) taking into account multiple sensory modalities and their timing (see e.g., [6], [7]).

The goal of this special section is to gather recent achievements in laughter computing in order to trigger new research directions in this area. Special attention is devoted

to computational models of laughter in human-machine and human-human interaction, capturing the multimodal nature of laughter and its variety of contexts and meanings. The special section ultimately aims at providing an interdisciplinary perspective of ongoing scientific research and ICT developments regarding laughter computing.

2 CONTRIBUTION

Twelve submissions were received for the special section and each of them was reviewed by at least three experts of the related domains (affective computing, emotion psychology, computer graphics, computer vision, signal processing, and so on). Five papers were chosen and finally accepted for publication after revisions, including a second round of reviews. This long process ensured that the papers collected in this special section represent to date the most innovative ideas and techniques related to laughter computing. The five papers provide indeed a broad coverage of the key aspects related to laughter and its computation. Two papers concern the investigation of laughter expressions and their meanings by using different laughter elicitation techniques. Two papers address laughter detection from different perspectives (i.e., by using a wearable device and by analyzing audiovisual input). One paper is devoted to laughter synthesis.

Concerning investigation of laughter and its facets, the work by Hofmann, Platt, and Ruch addresses empirical work about the relation between laughter meaning and facial expressions. The authors recorded and analyzed structured conversations in which participants retold stories involving the 16 positive emotions proposed by Ekman; these emotions include sensory pleasures such as taste and smell, reactions to one's own accomplishments such as pride and relief, as well as responses to the behavior of others such as gratitude or *schadenfreude*. They then looked for presence and intensity of Duchenne ("genuine") smiles, non-Duchenne ("fake") smiles, as well as laughter. Results showed that amusement and *schadenfreude* were related with most laughter, whilst gratitude and contentment resulted in the least intense behaviors. In addition to providing fine-grained empirical evidence on how these behaviors are related to emotions, the results of this study can also be potentially used to improve generation of appropriate facial expressions and behavior of artificial agents.

Whereas Hofmann, Platt, and Ruch used structured conversation to elicit smiles and laughter from their participants,

- M. Mancini, R. Niewiadomski, and G. Volpe are with DIBRIS, University of Genoa, Genoa 16126, Italy. E-mail: {maurizio.mancini, gualtiero.volpe}@unige.it, radoslaw.niewiadomski@dibris.unige.it.
- S. Hashimoto is with Department of Applied Physics, Waseda University, Tokyo 169-8050, Japan. E-mail: shuji@waseda.jp.
- M.E. Foster is with the School of Computing Science, University of Glasgow, Glasgow G12 8QQ, United Kingdom. E-mail: MaryEllen.Foster@glasgow.ac.uk.
- S. Scherer is with the Department of Computer Science, University of Southern California, Los Angeles, CA 90007. E-mail: scherer@ict.usc.edu.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TAFFC.2017.2766938

Fortin and Cooperstock used a different approach: they developed a novel tickling device, which permitted a precisely controlled amount of vibration to be delivered to the bottom of a participant's foot. In an experimental study, they measured the effectiveness of this device at eliciting laughter, either with or without accompanying laughter sounds. The response of the participants was measured through a combination of self-reports, behavioral, and physiological measures. Results showed a clear relationship between the vibration intensity of a stimulus and the self-reported perceived ticklishness. Moreover, hearing laughter had a significant scaling effect on this perception. The properties of the stimulus, however, did not have an effect on the behavioral responses—rather, these were mainly affected by the participants' mood.

Regarding laughter detection, Perusquia-Hernandez, Hirokawa, and Suzuki developed and evaluated a wearable device for analysis of facial expression. They utilized a small device that sits on the user's ears and connects Electromyography (EMG) electrodes with muscles associated with Facial Action Units 12 and 6 to detect genuine smiles. The device enables capturing very subtle and quick facial expressions by using EMG rather than computer vision or signals from other modalities. Further, the work focuses on spontaneous expressions rather than posed facial expressions or actors. As a concrete output of their research the authors present a proof-of-concept of wearable device that can detect subtle facial expressions and validate its accuracy and robustness through a number of studies.

Turker, Yemez, Sezgin, and Erzin approach laughter detection by facing the problem of audiovisual continuous laughter detection, one of the core challenges in affective computing and more broadly in improving our perception of human communication dynamics. Their work does not only introduce a meticulously designed and validated study, seeking to understand the value added of video and audio modalities, but also positions the paper in the tradition of a number of multimodal laughter detection experiments. The proposed work leverages an established affectively colored dataset for their evaluations and achieve state-of-the-art performance in their experiments. They find that both facial expression data and head motion dynamics are relevant for laughter detection performance, especially under difficult and noisy conditions.

In their paper on laughter synthesis, Ding, Huang, and Pelachaud built a laughter behavior controller to generate the upper body movements of a virtual character from laughter audio inputs. They proposed a new continuous-state statistical framework to capture laughter dynamics from human data and constructed a module to infer head and torso movements at each time frame according to the current value of loudness, computed on the audio input, and the output inferred from previous time frames. To train the statistical model parameters the authors recorded a dataset, which contains human laughter behavior and laughter audio. Additionally, they employed a rule-based method to yield shoulder animations from laughter audio.

3 CHALLENGES AND PERSPECTIVES

Papers in this special section at the same time address key challenges in laughter computing and open novel research

directions. A clear example is the development of methods for detection and synthesis of different laughter types and their meaning. Such research challenge emerges, for example, in the paper by Hofmann and colleagues. Whereas their work describes different expressive patterns and their meaning, the process of automatically detecting and synthesizing different laughter types is still largely unexplored. This is nevertheless an important skill for machines that need to establish a socially meaningful interaction with humans, such as, for example, social robots. Misidentification of (human) user's laughter expressions by the machine may have a relevant negative impact on the interaction. For example, confusing an ironic or schadenfreude laughter with a hilarious one might lead to the wrong identification of a user's change of attitude toward the social robot. Detection of different laughter types can be addressed, for example, by considering the moment in which laughter expressions appear during the course of the interaction, by introducing models of laughter contagion, or by considering additional aspects of the verbal content (e.g., joke or irony modeling). Indeed, modeling contextual information (e.g., in the verbal channel) opens new exciting opportunities for building applications that exploit laughter. Moreover, both the papers in this special section and the paradigmatic example of analysis and synthesis of different laughter types, as a future research challenge, point out the strong multidisciplinary characterization of this area. Research requires a deep interaction between psychologists, human scientists, and computer scientists to have at the same time a solid theoretical background corroborated by a strong experimental evidence, and a robust computational approach.

Whereas no paper in this special section explicitly addresses audio synthesis, believable laughter audio overlapped with speech synthesis (i.e., co-occurrence of laughter bursts within speech utterances) remains a crucial challenge for building effective laughter machines.

Future works may also address laughter computing for commercial applications, e.g., according to the principles of Positive Computing. Systems for automatic stress monitoring or depression detection would benefit of algorithms able to detect/monitor hilarious laughter (and distinguish it from other types of laughter). Modeling hilarious laughter can be exploited in systems that explore the positive outcomes of laughter for well-being. For example, they may be a part of an interactive system for training/eliciting positive attitudes. An interactive system may measure user's laughter and use this measurement as a feedback to control such a training/eliciting process.

Finally, laughter computing applied to multimodal interfaces can contribute to develop system that increase creativity and facilitate sociability in educational and artistic contexts. A laughing interface can, for example, be exploited to reinforce the user's perception of pleasantness in performing a joint task, or the degree of funniness of an artistic performance in an audience.

ACKNOWLEDGMENTS

The Guest Editors would like to thank all the authors who have submitted and contributed to this special section and all the reviewers. They would like to thank the IEEE

Transactions on Affective Computing for supporting this special issue, with special thanks to the Editor-in-Chief Prof. Björn Schuller. They would also like to thank the IEEE Administrators Ms. Samantha Jacobs and Ms. Antonia Carl for their technical support.

REFERENCES

- [1] FP7 ILHAIRE Project. [Online]. Available: <http://www.ilhaire.eu/>
- [2] G. McKeown, W. Curran, C. McLoughlin, H. J. Griffin, and N. Bianchi-Berthouze, "Laughter induction techniques suitable for generating motion capture data of laughter associated body movements," in *Proc. 10th IEEE Int. Conf. Workshops Automatic Face Gesture Recognition*, Apr. 2013, pp. 1–5.
- [3] M. Mancini, et al. "Implementing and evaluating a laughing virtual character," *ACM Trans. Internet Technol.*, vol. 17, no. 1, pp. 3:1–3:22, Feb. 2017.
- [4] S. Petridis, B. Martinez, and M. Pantic, "The MAHNOB laughter database," *Image Vis. Comput.*, vol. 31, no. 2, pp. 186–202, Feb. 2013.
- [5] R. Niewiadomski, M. Mancini, G. Varni, G. Volpe, and A. Camurri, "Automated laughter detection from full-body movements," *IEEE Trans. Human-Mach. Syst.*, vol. 46, no. 1, pp. 113–123, Feb. 2016.
- [6] P. C. DiLorenzo, V. B. Zordan, and B. L. Sanders, "Laughing out loud: Control for modeling anatomically inspired laughter using audio," *ACM Trans. Graph.*, vol. 27, no. 5, pp. 125:1–125:8, Dec. 2008.
- [7] J. Urbain, H. Cakmak, A. Charlier, M. Denti, T. Dutoit, and S. Dupont, "Arousal-driven synthesis of laughter," *IEEE J. Selected Topics Signal Process.*, vol. 8, no. 2, pp. 273–284, Apr. 2014.



Maurizio Mancini received the PhD degree in computer science from the University of Paris 8, in 2008. He is assistant professor with DIBRIS (University of Genoa, Italy). His research interests include human-machine interaction, embodied agents, multimodal expressive behavior analysis and synthesis. Since 2001 he worked on EU funded project in FP5 and FP6 (MagiCster, HUMAINE, CALLAS) in the framework of HCI and in particular on the development the Greta ECA research platform. He worked on FP7 (SAME, SIEMPRE, ILHAIRE) and H2020 (DANCE) EU projects, in particular on the development of models and algorithms for expressive behavior analysis for the EyesWeb XMI research platform. He co-authored 13 articles on international journals and more than 80 publications on international conferences and books.



Radoslaw Niewiadomski received the PhD degree in computer science from the University of Perugia, Italy. His research interests are in the area of affective computing and include: recognition and synthesis of nonverbal behaviors, creation of embodied conversational agents and interactive multimodal systems. He participated in several EU Projects and is co-author of more than 50 research publications. Since 2013, he is working as a post-doc researcher with DIBRIS (University of Genoa, Italy) on expressive gesture analysis.



Shuji Hashimoto received the BS, MS, and DrEng degrees in applied physics from Waseda University, Tokyo, Japan, in 1970, 1973, and 1977, respectively. He is a professor in the Department of Applied Physics, Waseda University. He was the director of the Humanoid Robotics Institute, Waseda University for ten years since 2000. He is currently the senior executive vice president for Academic Affairs and Provost of Waseda University since 2010. He is the author of more than 400 technical publications, proceedings, editorials, and books. His research interests include "KANSEI" information processing, image and sound processing, artificial intelligence and robotics. He is a member of major academic societies including IEEE, RSJ (Robotic Society of Japan) and JFace (Japanese Academy of Facial Studies).



Mary Ellen Foster received the PhD degree from the University of Edinburgh, in 2007, and has previously worked at the Technical University of Munich and Heriot-Watt University. She is a lecturer in the School of Computing Science, University of Glasgow. Her primary research interests include human-robot interaction, social robotics, and embodied conversational agents. She is the coordinator of the MuMMER project, a European Horizon 2020 project in the area of socially aware human-robot interaction.



Stefan Scherer received the Dr.rer.nat. degree from the faculty of Engineering and Computer Science, Ulm University in Germany with the grade summa cum laude (i.e., with distinction) in 2011. He is a research assistant professor in the Department of Computer Science, University of Southern California (USC) and the USC Institute for Creative Technologies (ICT) where he leads research projects funded by US National Science Foundation and the Army Research Laboratory. His work is focused on machine learning, multimodal signal processing, and affective computing. He serves as an associate editor of the *IEEE Transactions on Affective Computing* and is the vice chair of the IAPR Technical Committee 9 on Pattern Recognition in Human-Machine Interaction.



Gualtiero Volpe received the MSc degree in computer engineering, in 1999 and the PhD degree in electronic and computer engineering, in 2003 from the University of Genoa. He is an associate professor in the Department of Informatics, Bioengineering, Robotics, and Systems Engineering (DIBRIS), University of Genoa. His research interests include intelligent and affective human-machine interaction, social signal processing, sound and music computing, modeling and real-time analysis and synthesis of expressive content, and multimodal interactive systems. He participated to several EU-funded research projects and he is co-author of more than 150 publications in international journals and conferences.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.