Social and Affective Robotics Tutorial

Maja Pantic[1,2], Vanessa Evers[2], Marc Deisenroth[1], Luis Merino[3] and Bjoern Schuller[4,1]
[1]Imperial College London, Exhibition Road, London SW7 2AZ, UK
[2]University of Twente, Drienerlolaan 5, 7522 NB Enschede, Netherlands
[3]Pablo de Olavide University, 41013 Sevilla, Spain
[4]University of Passau, Innstrasse 43, 94032 Passau, Germany
m.pantic@imperial.ac.uk; v.evers@utwente.nl; m.deisenroth@imperial.ac.uk; lmercab@upo.es; bs@fim.uni-passau.de

ABSTRACT
Social and Affective Robotics is a growing multidisciplinary field encompassing computer science, engineering, psychology, education, and many other disciplines. It explores how social and affective factors influence interactions between humans and robots, and how affect and social signals can be sensed and integrated into the design, implementation, and evaluation of robots. With talks by renowned researchers in this area, Social and Affective Robotics Tutorial will help both new and experienced researchers to identify trends, concepts, methodologies and applications in this field, identified as a technological megatrend driving the fourth industrial revolution.

CCS Concepts
• **Computing methodologies ⊙ Artificial Intelligence**
• **Computing methodologies ⊙ Machine Learning • Human-centered computing.**

Keywords
Affective computing, Social robotics, Machine Learning, Signal Processing, Robotic Planning.

1. INTRODUCTION
Until recently, the use of robots was confined to tightly controlled tasks in specific industries such as automotive. Today, however, robots are increasingly used across all sectors and for a wide range of tasks from precision manufacturing to nursing. Rapid progress in robotics makes naturalistic interaction between humans and machines an everyday reality. Advances in sensors, computer vision, audio processing, and machine learning, are enabling robots to understand their environment and their users better and, in turn, to become more adaptive and flexible, responding better and engaging in a broader variety of tasks such as household chores, tourists guiding, education, autism research, healthcare, etc. As robots become capable of sensing and responding to the context in which they (and their human users) operate, the next generation of robots emerges, where the emphasis will be on human-machine collaboration. Collaborative Robotics are a technological megatrend expected to reach its tipping point – the moment when it hits mainstream society – in the next ten years [1], shaping our AI-empowered future and robot-aided personal ecosystem. Social and Affective Robotics Tutorial will identify concepts, methodologies, and trends in this burgeoning field and explore some of the ethical and psychological questions raised by human-machine relations.

The tutorial is divided into seven main parts: five talks on the main aspects of Affective and Social Robotics including sensing affect and social signals, machine learning, and socially-adaptive navigation and planning, an interactive demo session exhibiting a number of socially-aware and affective robots, and a panel discussion revolving around technological, ethical, and psychological questions raised by this robotic technology driving the fourth industrial revolution.

2. TALKS

2.1 The Rise of Social Robotics
Robotics and AI are expected to leave factory floors and to co-inhabit our homes, our schools, our hospitals and our streets. These environments are social and if the technology put in these environments isn't socially intelligent, it is likely that it will not be accepted by human users. The talk by Prof. Vanessa Evers explains the core idea behind making machines socially intelligent – that sentient beings need social referencing to learn – and identifies technical challenges that need to be addressed in order to realize the full potential of this technology.

2.2 Robot Social Navigation
Robots acting among humans require social awareness when performing their tasks. This has to be considered in each step, from high-level planning and supervision, to motion planning and execution. The talk by Dr. Luis Merino identifies different techniques employed for robot social navigation, including techniques for human-aware motion planning, learning users' navigation intention models and social cost functions from observations, planning based on the learned models, and how all these different techniques can be combined to achieve social navigation.

2.3 Machine Learning for Robotics
One of the fundamental challenges in fully autonomous robots is learning from data directly. This requires data-driven statistical methods for modeling, predicting, and decision making, while taking uncertainty into account, e.g., due to measurement noise, sparse data, or stochasticity in the environment. A practical challenge that comes with autonomous robots is that learning needs to be efficient – performing millions of experiments with robots is time consuming and wears out the hardware. The talk by Dr. Marc Deisenroth identifies practical algorithms for data-efficient learning in robotics and discusses further challenges and opportunities in the field.

MM'16, October 15–19, 2016, Amsterdam, The Netherlands.
ACM ISBN 978-1-4503-3603-1/16/10.
DOI: http://dx.doi.org/10.1145/2964284.2986914

2.4 Visual Affect & Social Signals Sensing

While the physiognomy (and the dynamics) of the face serves to identify other members of our species, facial behavior is our preeminent means to communicating affective and social signals. The talk by Prof. Maja Pantic summarizes a number of aspects of the human face and facial behavior and how they can be automatically sensed and analyzed by computers. Past research in the field and how far are we from enabling robots to identify and verify human users and sense and recognize human facial expressions and behavior are discussed in greater detail.

2.5 Auditory Affect & Social Signals Sensing

Speech processing technology is the crux to making human-robot interaction more intuitive, more efficient, and more naturalistic. To reach this goal, rich speaker traits and personal states, carried by the tone of the voice and the spoken words, must be reliably identified by machines. The talk by Prof. Bjoern Schuller explains novel holistic and time-evolving techniques for multi-task and weakly supervised learning of universal speaker characteristics, which have been considered only in isolation so far. The talk will also discuss overcoming today's sparseness of annotated realistic speech data by large-scale speech and meta-data mining from public sources (e.g., social media, crowd-sourcing for labelling and quality control, and shared semi-automatic annotation).

3. DEMOS

Attendees will be able to take part in an interactive exhibition were the Tutorial organizers showcase examples of the Social and Affective robotics technology in action.

3.1 FROG Robot

Imagine the local Zoo guide or the info kiosk in Disneyland being replaced by an attentive robot. This is exactly what FROG Robot is! There are two FROG robots, the red one and the green one. The red one is the info bot and messenger between buildings at the University of Twente, Netherlands. The green one is tourist guide bot at Royal Alcazar in Seveille, Spain. FROG is pretty much autonomous – it spots and approached people, shows them around based on its automatic estimation of their interest level, and then moves on to the next group. The red FROG will be shown in Amsterdam in October 2016. For a video of FROG see: https://www.facebook.com/FrogEU/videos/1515431328741253/.

3.2 TERESA Robot

TERESA is a tele-presence robot built upon the commercially available Giraff robot, which consists of a video screen attached to a base with wheels. The robot is controlled by a user at a remote location, whose face is displayed on the robot's video screen. The robot has a camera, microphone and a speaker to enable video conferencing. The user remotely interacts with people by guiding a remotely located robot, allowing the user to be more physically present than with standard teleconferencing. TERESA is socially intelligent in that it navigates semi-autonomously among groups of people, maintains face-to-face contact during conversations, and displays appropriate body-pose behavior. For a video synopsis see: https://www.youtube.com/watch?v=eaDYa6TIMQM. For more info see: http://teresaproject.eu

3.3 "Mimic Me" with NAO Robot

"Mimic-Me" is an interactive game played with the NAO robot. It is designed to engage both typically developing and autistic children, to raise their interest in information and communication technologies (ICT), and in the later case, to help autistic kids to practice the recognition and display of facial expressions of emotions. The game involves the robot 'mimicking' the player's expression using a combination of body gestures and audio cues. The robot interacts with players in a naturalistic way by automatically recognizing their speech, head movements, and facial expressions [2]. The "Mimic-Me" game is implemented as a modular, loosely coupled, software system using the HCI^2 Framework [3]. As a result, the system and its modules can be easily reused, or extended, for the purposes of other studies on human-robot interactions. For a video synopsis of "Mimic-Me" game see: http://www.youtube.com/watch?v=i49KUjFz-nM. For video synopsis of the performed facial expression tracking, see: https://www.facebook.com/maja.pantic.758/videos/vb.100003177678822/773346092781306/?type=2. For info on human-robot interaction studies with autistic children, see: http://de-enigma.eu.

4. PANEL DISCUSSION

Social and Affective Robots and the future Collaborative Robots are set to transform many services we use each day and to co-inhabit our personal ecosystems, but what exactly does this mean? In future, humanoid robots could improve security, healthcare, education, and revolutionize industry. But despite its reach, this powerful technology raises many questions. Would we consult an AI-driven robot doctor with a perfect or near-perfect diagnostic success rate or stick with the human GP who has known us for years? If our own behavior in any context becomes predictable by robot (or AI), how much personal freedom would we have or feel that we have to deviate from the prediction? How do we maintain our individuality, the very source of our diversity, in the digital age? Our panel of speakers will discuss these and other related issues raised by this powerful technology driving the fourth industrial revolution.

5. ACKNOWLEDGMENTS

This work has been funded by the European Community 7th Framework Programme [FP7/2007-2013] under grant agreement no. 611153 (TERESA) and the European Community Horizon 2020 Programme [H2020] under grant agreement no. 688835 (De-Enigma). Previous work on the topic was also funded in part by the European Community 7th Framework Programme [FP7/2007-2013] under grant agreement no. 288235 (FROG).

6. REFERENCES

[1] Schwab, K. 2016. *The Fourth Industrial Revolution*. World Economic Forum Press, Geneva, Switzerland.

[2] Shen, J., Rudovic, O., Cheng, S., and Pantic, M. 2015. Sentiment Apprehension in Human-Robot Interaction with NAO. *IEEE Proc. Int'l Conf. Affective Computing and Intelligent Interaction*. 1st International Workshop on Automatic Sentiment Analysis in the Wild (WASA'15). DOI=10.1109/ACII.2015.7344676.

[3] Shen, J., and Pantic, M. 2013. HCI^2 Framework: A Software Framework for Multimodal Human-Computer Interaction Systems. IEEE Trans. Cybernetics, 43, 6 (Dec. 2013), 1593-1606. DOI=10.1109/TCYB.2013.2271563.