A COMPUTATIONAL MODEL OF MOTOR COGNITION FOR ICONIC GESTURE PROCESSING

AMIR SADEGHIPOUR

A doctoral thesis submitted in fulfilment of the requirements for the degree of

Doktor-Ingenieur (Doctor of Engineering)

in the

Social Cognitive Systems group
Cognitive Interaction Technology – Center of Excellence
Faculty of Technology
Bielefeld University
Bielefeld, Germany

Supervisor: Prof. Dr.-Ing. Stefan Kopp

June 2014

Berichte aus der Robotik

Amir Sadeghipour

A computational model of motor cognition for iconic gesture processing

Shaker Verlag
Aachen 2015

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: Bielefeld, Univ., Diss., 2014

Printed in Germany.

ISBN 978-3-8440-3451-6
ISSN 1434-8098

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

ABSTRACT

In the present thesis, we propose a computational cognitive model for processing iconic gestures performed with hands and arms. Among different social behaviors, investigating gestures has gained special attention during the last decades. From a phenomenological perspective, gestures comprise mostly unconventionalized non-verbal behaviors that bring different aspects of social, motor and embodied cognition together. This inherent variability creates considerable challenges during processing, and in particular when the goal is cognitive modeling. In the field of human-computer interaction, advances in motion tracking systems have opened the gateway to the broader application of gesture-based user interfaces. Virtual humanoid agents are been increasingly used in human-computer interaction as user friendly interfaces that allow for intuitive and natural face-to-face communication. Furthermore, they can embody computational models of human social cognition, and provide suitable testbeds for cognitive models and theories.

Against this background, we propose a model that endows artificial humanoid agents with the ability to gesture. This capacity includes, for instance, fast and reliable recognition of highly diverse gesture performances, the ability to learn how to perform gestures through imitation, and to establish gestural alignment during interaction with human users. To this end, we propose a cognitive model of the human sensorimotor system based on neuroscience and psychological empirical evidence, and couched in current cognitive theories. To implement information processing within this cognitive model, we propose two computational approaches with complementary strengths and weaknesses. Empirical Bayesian Belief Update (EBBU) features the fast, incremental and cognitively plausible recognition of gestures during interaction with humans, achieved by combining bottom-up perception with top-down prediction. We also propose a Feature-based Stochastic Context-Free Grammar (FSCFG) to learn discriminative or descriptive grammar models of gestures. Applying this approach to a recorded dataset of iconic gestures demonstrated its considerable generalization capacity over both syntactic structure variabilities, and the statistical spatiotemporal deviations inherent to gesture.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Prof. Stefan Kopp, for his excellent guidance, motivation and enthusiasm, and for allowing me the room to develop and explore my own ideas. His guidance helped me to find my way in the multifaceted research field of cognitive science.

I would like to thank my wonderful colleagues for making it possible to work in an excellent research environment. My special thanks go to Hendrik Buschmeier, Ramin Yaghoubzadeh and Farina Freigang, for the stimulating discussions, and for their helpful suggestions to improve this thesis, as well as for the fruitful collaborations and long working nights. I would also like to thank Prof. Louis-Philippe Morency for offering me a summer internship in his group, and for his guidance in creating a gesture dataset.

Many thanks go to the Cognitive Interaction Technology (CITEC) Center of Excellence at Bielefeld University for providing me with the support and equipment I needed. In addition, my most sincere gratitude to the Deutsche Forschungsgemeinschaft (DFG) for funding this project.

I would never have been able to finish – or even start – my Ph.D. research, without the immeasurable support of my family. I thank my parents for their support and for encouraging me with their best wishes during this long period, and despite the great distance between us. Finally, and most importantly, I would like to thank my dear wife, not only for her support through the good times and bad, but also for the very inspiring discussions we had about my research.

MAIN TOPICS

CONTENTS