

Thomas S. Huang
Anton Nijholt
Maja Pantic
Alex Pentland (Eds.)

Artificial Intelligence for Human Computing

ICMI 2006 and IJCAI 2007 International Workshops
Banff, Canada, November 2006 and
Hyderabad, India, January 2007
Revised Selected and Invited Papers



Springer

TP11-53

I16
2006

Thomas S. Huang Anton Nijholt
Maja Pantic Alex Pentland (Eds.)

Artificial Intelligence for Human Computing

ICMI 2006 and IJCAI 2007 International Workshops
Banff, Canada, November 3, 2006 and
Hyderabad, India, January 6, 2007
Revised Selected and Invited Papers



Springer



E2007003205

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Thomas S. Huang
Beckman Institute, University of Illinois at Urbana-Champaign
405 N. Mathews Ave., Urbana, IL 61801, USA
E-mail: huang@ifp.uiuc.edu

Anton Nijholt
University of Twente
Faculty of Electrical Engineering, Mathematics and Computer Science
Postbus 217, 7500 AE Enschede, The Netherlands
E-mail: a.nijholt@ewi.utwente.nl

Maja Pantic
Imperial College, Computing Department
180 Queens Gate, London SW7 2AZ, U.K.
E-mail: m.pantic@imperial.ac.uk

Alex Pentland
MIT Media Laboratory, Massachusetts Institute of Technology
Building E15, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA
E-mail: pentland@media.mit.edu

Library of Congress Control Number: 2007926703

CR Subject Classification (1998): I.2, H.5.2, I.2.10, I.4, I.5

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743

ISBN-10 3-540-72346-3 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-72346-2 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12059386 06/3180 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence

4451

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Preface

This volume in the *Lecture Notes of Artificial Intelligence* represents the first book on human computing. We introduced the notion of human computing in 2006 and organized two events that were meant to explain this notion and the research conducted worldwide in the context of this notion.

The first of these events was a Special Session on Human Computing that took place during the Eighth International ACM Conference on Multimodal Interfaces (ICMI 2006), held in Banff, Canada, on November 3, 2006. The theme of the conference was multimodal collaboration and our Special Session on Human Computing was a natural extension of the discussion on this theme. We are grateful to the organizers of ICMI 2006 for supporting our efforts to organize this Special Session during the conference.

The second event in question was a Workshop on AI for Human Computing organized in conjunction with the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), held in Hyderabad (India), on January 6, 2007. The main theme of IJCAI 2007 was AI and its benefits to society. Our workshop presented a vision of the future of computing technology in which AI, in particular machine learning and agent technology, plays an essential role. We want to thank the organizers of IJCAI 2007 for their support in the organization of the Workshop on AI for Human Computing.

A large number of the contributions in this book are updated and extended versions of the papers presented during these two events. In order to obtain a more complete overview of research efforts in the field of human computing, a number of additional invited contributions are included in this book on AI for human computing.

One of the contributions in this volume starts with the observation that humans are social beings. Unfortunately, it is exceptional when we can say that a particular computer system, a computer application, or a human – computer interface has been designed from this point of view. Rather, we talk about users that have to perform tasks in a way that is prescribed by the computer. However, when we take the point of view of designing systems for social beings, we should talk rather about partners or participants instead of users, and when we do so, it is also the computer or a computer-supported environment that plays the role of a partner or a participant.

Human computing, as advocated and illustrated in this volume, aims at making computing devices and smart environments social partners of humans interacting with these devices or inhabiting these environments. These devices and environments need to understand what exactly the specifics of the current interaction flow and the surrounding environment are. This understanding allows for anticipatory and proactive feedback and real-time, unobtrusive support of human activities in the environment.

The *LNAI* volume on *AI for Human Computing* consists of three parts: a part on foundational issues of human computing, a part on sensing humans and their activities, and a part on anthropocentric interaction models.

Sensing humans and understanding their behavior is a core issue in the research on human computing. Numerous papers presented in this volume can be considered from

this point of view. Human behavioral cues like facial expressions and body gestures are sensed by the computer / environment and then interpreted. Unfortunately, this interpretation is often carried out while taking into account only a limited portion of the information (if any) about the context (sensed or explicitly provided) in which the observed behavioral patterns have occurred. However, as accurate interpretation of human behavior cannot be achieved in a context-free manner, several papers in this volume argue that the realization of automatic context sensing and context-dependant analysis of human behavioral cues are the two challenging issues that need immediate attention of researchers in the field.

Sensing humans and interpreting their behavior should be followed by proactive support of their current activities. The support is to be provided by the environment and its 'inhabitants' in a suitable form. For example, the environment, as perceived by the user, can provide the support by turning on the air-conditioning when the user gets sweaty and his or her face reddens. On the other hand, human users being a part of a smart environment can be supported by artificial partners while conducting a task or a leisure-oriented activity within the environment. That is, smart digital devices within the environment including robots, virtual humans displayed within the environment, and hidden software agents of the environment can all provide support to human users within the environment and can cooperate with them and each other to provide this support. From a global point of view, realizing a smart environment able to support human activities simply means adapting the environment including its human and artificial inhabitants in such a way that the implicitly or explicitly communicated habits, preferences, desires, questions, and commands of a particular human user within the environment are anticipated, supported, and executed as well as possible.

Several papers in this book also emphasize the importance of having multimodal corpora available for research purposes. These are necessary for training and testing machine learning methods for human behavior analysis. In addition, analysis of a multimodal corpus makes us aware of properties of human behavior, correlations between behavioral signals and types (categories) of human behavior, and (causal) relations between the different categories and between the context specifics and these categories. This allows us to modify and extend existing theories (in computer science as well as in cognitive sciences) and also to refine existing algorithms to enable more accurate / suitable analysis of the data. In a research context it is not always possible to study or fully simulate realistic situations. Financial or ethical reasons will often make it impossible to have fully unobtrusive recordings of human behavior in naturalistic contexts. Hence, the large majority of currently conducted experiments relate to scripted behavior, often deliberately displayed behavior that follows specific scenarios designed by researchers. Behavioral data obtained in this way can be useful in the start-up phase of the research on human computing. However, such data and the methods trained and tested using such data are not applicable in real-life situations, where subtle changes in expression of behavior are typical rather than the exaggerated changes that typify deliberately displayed behavior. Hence, the focus of the research in the field started to shift to automatic analysis of spontaneous behavior (produced in a reflex-like manner). Several works presented in this volume address machine analysis of human spontaneous behavior and present efforts towards collecting data in realistic settings.

In what follows, we shortly summarize each of the contributions to this volume.

Foundations of Human Computing

In his contribution to this volume, Cohn surveys ways to infer emotions from expressive human behavior, in particular facial expressions. He discusses several key issues that need to be considered when designing interfaces that approach the naturalness of human face-to-face interaction. Among them are the differences between the judgment-based approach (inferring the underlying affective state) and the sign-based approach (labeling facial muscle actions) to measurement of facial behavior, the timing and overall dynamics of facial behavior, and individual differences in facial expressions. A less researched issue discussed in this paper, which is of particular importance for natural interaction, is that of synchrony of affect display in face-to-face interaction.

“Instinctive Computing” by Cai argues that for genuine intelligence and natural interaction with humans, computers must have the ability to recognize, understand, and even have primitive instincts. The message, supported by the literature, is that instincts influence how we look, feel, think, and act. Foraging, vigilance, reproduction, intuition, and learning are distinguished as the human basic instincts. These basic instincts need to be addressed and have their metaphors in the computer systems of the future. The paper discusses different case studies and observations on nontraditional physiological sensors, sensors related to reproductive esthetics, and those related to updating of instincts (learning). Instinctive computing is seen as the foundation for ambient intelligence and empathetic computing, where the latter includes the detection, understanding, and reacting to human expressions of pain, illness, depression, and anomaly.

Sensing Humans for Human Computing

Automatic sensing and understanding of human behavior in computer-supported and smart environments are discussed by the editors of this volume in their position paper on human computing. They summarize the current state of the art, challenges, and opportunities facing the researchers in intertwined research areas of context sensing, human affect sensing, and social signaling analysis. Context sensing is discussed from the W5+ (Who, Where, What, When, Why, How) perspective and it is argued that the most promising way to achieve accurate context sensing in naturalistic settings is to realize multimodal, multi-aspect context sensing. In this approach, the key is to automatically determine whether observed behavioral cues share a common cause [e.g., whether the mouth movements and audio signals complement to indicate an active known or unknown speaker (How, Who, Where) and whether his or her focus of attention is another person or a computer (What, Why)]. In addition, the paper also argues that the problem of context-constrained analysis of multimodal behavioral signals shown in temporal intervals of arbitrary length should be treated as one complex problem rather than a number of detached problems in human sensing, context sensing, and human behavior understanding.

Natural settings, in which we can infer the users’ emotional state by combining information from facial expression and speech prosody, are discussed in several contributions to this volume. As remarked in the relevant papers, it is well known that posed (deliberately displayed) expressions of emotions differ in appearance and timing from those shown in a natural setting. Audiovisual recognition of spontaneous expressions of human positive and negative affective feedback in a natural

human – human conversation setting has been investigated in the work of Zeng et al. Facial expressions extracted from the video signal and prosodic features (pitch and energy) extracted from the speech signal were treated separately and in a combination. It was shown that bimodal data provide more effective information for human affect recognition than single-modal data and that linear bimodal fusion is sub-optimal for the task at hand.

The paper of Karpouzis et al. discusses audiovisual emotion recognition (positive vs. negative and active vs. passive) in naturalistic data obtained from users interacting with an artificial empathetic agent. This sensitive artificial listener (SAL) agent gives the impression of sympathetic understanding. The ‘understanding’ component is keyword driven but suffices to induce states that are genuinely emotional and involve speech. The SAL agent can be given different personalities to elicit a range of emotions. Facial expressions and hand gestures extracted from the video signal and a range of acoustic features extracted from the speech signal were treated subsequently by means of recurrent neural networks trained for recognition of target affective states. Similar to Zeng et al., Karpouzis et al. show that multimodal data provide more effective information for human affect recognition than single-modal data.

Human – robot interaction in settings where both the human and the robot can show affective behavior has been investigated in the work of Broekens. Part of the presented setup (i.e., the robot) has been simulated. The aim of the simulated robot is to survive in a grid-world environment. It uses reinforcement learning to find its way to food. Two versions of the robot are presented: a nonsocial robot that learns its task without the affective feedback of a human observer and a social robot that uses the communicated affective feedback as the social reward. Affective feedback is obtained from the facial expressions of the human observer who monitors the robot’s actions. The results show that the “social robot” learns its task significantly faster than its “nonsocial sibling” and Broekens argues that this presents strong evidence that affective communication with humans can significantly enhance the reinforcement learning loop.

The paper of Oikonomopoulos et al. presents the research on trajectory-based representation of human actions. In this work, human actions are represented as collections of short trajectories that are extracted by means of a particle filtering tracking scheme that is initialized at points that are considered salient in space and time. Improving the detection of spatio-temporal salient points and enhancing the utilized tracking scheme by means of an online background estimation algorithm is also discussed. By defining a distance metric between different sets of trajectories corresponding to different actions using a variant of the longest common subsequence algorithm and a relevance vector machine, the authors obtained promising results for recognition of human body gestures such as aerobic exercises.

Modeling the communication atmosphere is the main topic of the paper by Rutkowski and Mandic. They identify a 3D communication space for face-to-face communication, where the dimensions are environmental (i.e., related to the ambient conditions like noise and visual activity rather than to the communication itself), communicative (related to the audiovisual behavior of the communicators like feedback provision and turn taking), and emotional (related to coupled emotional states of the participants). The main focus of the paper is on the dynamics of nonverbal communication as the basis for estimating communication atmosphere. The

ability to model this aspect of face-to-face communication implies the ability to manipulate (some of) the relevant conditions in order to adjust the atmosphere. Experiments using cameras and microphones to capture communicators' face-to-face situations have been conducted.

Dong and Pentland discuss the problem of combining evidence from different dynamic processes. For example, evidence about the context of a particular user can be obtained from different sensors (possibly connected in a sensor network). The proposed approach to multisensorial data fusion and interpretation introduces an 'influence model' in which experts with different knowledge can consult each other about their understanding of the data in order to decide about the classification. The authors illustrate their approach by means of two examples. The first one is a situation where data are collected from several wearable sensors (accelerometers, an audio recorder, and a video recorder) and where a team of four experts need to recognize various types of wearer context (locations, audio contexts, postures, and activities). The second situation relates to a social network where participants have mobile phones that record various data based on which the participants' social circles and individual behaviors are to be determined.

Anthropocentric Interaction Models for Human Computing

Humans have social skills. These skills allow them to manage relationships with other people in a given social structure. In ambient intelligence and virtual community environments, we need models of social intelligence in order to understand, evoke, and anticipate social behavior and to generate social intelligent behavior by the environment and its physical and synthetic agents performing in the environment. Here, generating socially intelligent behavior means performing actions, providing feedback, taking the initiative in interactions, displaying verbal and nonverbal affective and social signals, and amplifying social intelligence in such a way that there is a smooth, natural, but also effective embedding in the socio-cultural structure of the virtual, human, or augmented-reality community. Social intelligence design, as advocated by Nishida, aims at understanding and augmentation of social intelligence. Three perspectives are distinguished. The first one relates to social interaction in face-to-face and multi-party interactive settings in small groups where traditional discourse modeling and verbal and nonverbal interaction issues play important roles in displaying socially intelligent interactive behavior. The second perspective relates to social interaction in the large, possibly multimedial, chat and game environments, where sociological and socio-psychological models of multi-party interaction, large-scale collaboration, and social attitudes are of importance. The third perspective relates to the design of social artifacts that embody social intelligence and therefore facilitate social interaction. Among these artifacts are embodied agents, interactive robots, as well as collaboration technologies.

Within the human computing framework one can argue that interactive systems need to have the same communicative capabilities that humans have. In other words, human – computing technologies (i.e., interactive systems or environments) need to be based on theories and models of human – human interaction. The paper by Op den Akker and Heylen explores this view on human computing. Presently, computational models of human – human interaction are very limited and hardly take into account subtleties of verbal, let alone nonverbal, interaction. Nevertheless, there are tools such

as annotation schemes that enable the researchers in the field to analyze multimodal interaction corpora and come up with new, enhanced models of human - human interplay. In the paper by Op den Akker and Heylen these issues are discussed and illustrated by analyzing conversations constituting the multimodal augmented multi-party interaction (AMI) corpus.

Human computing applications are based on displayed human behavior such as affective and social signals. Being able to understand human behavior and behavioral cues is far beyond traditional human – computer interaction research. Moreover, within human computing applications in the home, office, and public spaces, human behavior to be supported by the environment does not necessarily have to be task-oriented and efficiency may not be an issue at all. For example, how should a smart environment support leisure activities, how can it increase social interaction among inhabitants, and how can it act as a social and entertaining space for its inhabitants? Clearly, designing the environment as a social and entertaining partner is an open and complex research question where various perspectives on environmental characteristics such as efficiency, related to computing and human resources, should be considered. Poppe et al. discuss trends for emerging human computing applications. They identify where challenges remain when it comes to evaluating emerging applications of human computing technologies. The paper aims to create awareness that business-as-usual will not work for these applications, it stresses the fact that the current evaluation practices are inappropriate, and it proposes partial solutions to some of the challenges.

Maat and Pantic discuss an agent-based, adaptive interface in which context sensing and context modeling are integrated within a context-aware application. The human part of the user's context is sensed using face recognition, eye tracking, speech recognition and, as usual, key strokes and mouse movements. Contextual questions related to who the current user is, what his or her task is, how he or she feels, and when a certain user's (re)action occurred are answered in an automated manner. Case-based reasoning is used to match events against user preferences. Important parts of the paper are the reasoning about the user's preferences, the interaction adaptation in accordance to these preferences, and the conducted usability study. In the latter an assessment of the system's usability was made in terms of issues like effectiveness, usability, usefulness, affective quality, and a number of ethical issues relevant to the system.

Sonntag et al. present a study on an environment where the user employs a context-aware handheld device with which it can interact in a natural (multimodal) way using speech and gestures. More specifically, Sonntag et al. investigated development of a context-aware, multimodal mobile interface to Web-based information services. User expressions (i.e., his or her speech and pointing gestures) are processed with respect to the situational and discourse context. Question clarification, resolution of referring expressions, and resolution of elliptical expressions are handled by the dialogue manager. A media ontology guides the presentation of results to the user, using text, image, and speech as possible modalities.

We started this volume with the paper by Cohn on the human face, the recognition of facial expressions, and how to use them in helping to infer an underlying affective state. The paper of Blanz revisits the topic of the human face. Blanz presents a model-based approach to reconstruct, animate, modify, and exchange faces in images or in

3D. Human computing addresses computer-supported environments inhabited by humans and human-like agents. Faces, displaying affective and social signals, need to be understood, both for understanding and for generating purposes. Autonomous and human-controlled agents need faces, for example, to represent their ‘owners.’ We also need tools to generate and exchange faces (for example, by reconstructing them from images), to manipulate them, and to do facial animation, for example, to generate subtle facial expressions and visual speech. The model-based tools proposed by Blanz allow semantically meaningful manipulation of faces.

Humans and human-like agents (robots, virtual humans) can interact with each other in smart-, virtual-, and mixed-reality environments. While Op den Akker and Heylen take human – human conversations as the starting point for modeling natural interactive systems, Reidsma et al. require not only that the interactive system or environment interacts with its human partner in a natural and intuitive way, but that it does so by means of a human-like, embodied, virtual human. This virtual human may represent a particular functionality of the environment (e.g., a doorman, a butler, a financial adviser, a fitness trainer, or a virtual friend) or it may represent another user of the environment with which it is possible to interact. The paper emphasizes the necessity of being able to model all kinds of subtleties and ‘imperfections’ in human – human communication. For example, designing a virtual human that is reluctant to answer the user’s questions may seem a waste of time, but if this virtual human is a virtual tutor in an educational virtual environment, or a negotiation partner within a virtual auction, or an opponent in a game environment, then this is rather acceptable behavior. The paper also discusses applications and (individual) user characteristics that require modeling of preferences and peculiarities of humans and human – human interactions.

Virtual humans, but controlled by human actors, are the topic of the research presented by Zhang et al. The ‘avatars’ discussed in the paper represent human actors that have mediated interaction in a storytelling environment. One of the avatars in the environment is controlled by a human director who takes care through her avatar that the story develops appropriately. More specifically, the affective cues present in the natural language interaction of the avatars help the director to make the emerging story interesting. The aim of Zhang et al. is to automate the role of the director. This requires the detection of the affective content of the utterances. One of the issues that are researched is the metaphorical conveyance of affect. Rather than aiming at fully linguistic analyses of the utterances mediated by the avatars, Zhang et al. aim at building robust representations of the affective connotations. One obvious reason to do so is that the utterances are often ungrammatical, they borrow from language used in text-messaging and chat rooms, and often the literal meaning hardly gives a clue about the intended meaning. Obviously, being able to detect the affective state of a user-controlled virtual human makes it also possible to generate a suitable emotional animation.

March 2007

Tom Huang
Anton Nijholt
Maja Pantic
Sandy Pentland

Organization

Editorial Board

Thomas S. Huang, Beckman Institute, University of Illinois at Urbana-Champaign, USA

Anton Nijholt, University of Twente, Faculty of EEMCS, The Netherlands

Maja Pantic, Imperial College London, Computing Department, UK

Alex Pentland, MIT Media Laboratory, Massachusetts Institute of Technology, USA

Program Committee

Jan Alexandersson, German Research Centre for Artificial Intelligence (DFKI), Germany

Marian S. Bartlett, University of California at San Diego, USA

Ira Cohen, Hewlett Packard, USA

Jeffrey Cohn, University of Pittsburgh, USA

Silvia Coradeschi, Orebro University, Sweden

Daniel Gatica Perez, IDIAP Research Institute, Switzerland

Dennis Hofs, University of Twente, The Netherlands

Seong-Whan Lee, Korea University, Korea

Alexis Martinez, Ohio State University, USA

Nikos Paragios, Ecole Centrale de Paris, France

Ioannis Patras, Queen Mary University, UK

Vladimir Pavlovic, Rutgers University, USA

Catherine Pelachaud, University of Paris 8, France

Bjoern Schuller, University of Munich, Germany

Nicu Sebe, University of Amsterdam, The Netherlands

Phoebe Sengers, Cornell University, USA

Ivo Swartjes, University of Twente, The Netherlands

Mariet Theune, University of Twente, The Netherlands

Matthew Turk, University of California at Santa Barbara, USA

Yaser Yacoob, University of Maryland, USA

Ingrid Zukerman, Monash University, Australia

Editorial Assistants

Hendri Hondorp, University of Twente, The Netherlands

Antonis Oikonomopoulos, Imperial College, UK

Lynn E. Packwood, University of Twente, The Netherlands

Lecture Notes in Artificial Intelligence (LNAI)

- Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), *Logic Programming and Nonmonotonic Reasoning*. IX, 327 pages. 2007.
- Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*. XIV, 585 pages. 2007.
- Vol. 4481: J.T. Yao, P. Lingras, W.-Z. Wu, M. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), *Rough Sets and Knowledge Technology*. XIV, 576 pages. 2007.
- Vol. 4452: M. Fasli, O. Shehory (Eds.), *Agent-Mediated Electronic Commerce*. VIII, 249 pages. 2007.
- Vol. 4451: T.S. Huang, A. Nijholt, M. Pantic, A. Pentland (Eds.), *Artificial Intelligence for Human Computing*. XVI, 359 pages. 2007.
- Vol. 4438: L. Maicher, A. Sigel, L.M. Garshol (Eds.), *Leveraging the Semantics of Topic Maps*. X, 257 pages. 2007.
- Vol. 4429: R. Lu, J.H. Siekmann, C. Ullrich (Eds.), *Cognitive Systems*. X, 161 pages. 2007.
- Vol. 4426: Z.-H. Zhou, H. Li, Q. Yang (Eds.), *Advances in Knowledge Discovery and Data Mining*. XXV, 1161 pages. 2007.
- Vol. 4411: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), *Programming Multi-Agent Systems*. XII, 249 pages. 2007.
- Vol. 4410: A. Branco (Ed.), *Anaphora: Analysis, Algorithms and Applications*. X, 191 pages. 2007.
- Vol. 4399: T. Kovacs, X. Llorà, K. Takadama, P.L. Lanzi, W. Stolzmann, S.W. Wilson (Eds.), *Learning Classifier Systems*. XII, 345 pages. 2007.
- Vol. 4390: S.O. Kuznetsov, S. Schmidt (Eds.), *Formal Concept Analysis*. X, 329 pages. 2007.
- Vol. 4389: D. Weyns, H.V.D. Parunak, F. Michel (Eds.), *Environments for Multi-Agent Systems III*. X, 273 pages. 2007.
- Vol. 4384: T. Washio, K. Satoh, H. Takeda, A. Inokuchi (Eds.), *New Frontiers in Artificial Intelligence*. IX, 401 pages. 2007.
- Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), *Computational Logic in Multi-Agent Systems*. X, 315 pages. 2007.
- Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), *Declarative Programming for Knowledge Management*. X, 229 pages. 2006.
- Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), *Theory and Applications of Relational Structures as Knowledge Instruments II*. X, 373 pages. 2006.
- Vol. 4335: S.A. Brueckner, S. Hassas, M. Jelasity, D. Yamins (Eds.), *Engineering Self-Organising Systems*. XII, 212 pages. 2007.
- Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), *Verification of Object-Oriented Software*. XXIX, 658 pages. 2007.
- Vol. 4333: U. Reimer, D. Karagiannis (Eds.), *Practical Aspects of Knowledge Management*. XII, 338 pages. 2006.
- Vol. 4327: M. Baldoni, U. Endriss (Eds.), *Declarative Agent Languages and Technologies IV*. VIII, 257 pages. 2006.
- Vol. 4314: C. Freksa, M. Kohlhase, K. Schill (Eds.), *KI 2006: Advances in Artificial Intelligence*. XII, 458 pages. 2007.
- Vol. 4304: A. Sattar, B.-H. Kang (Eds.), *AI 2006: Advances in Artificial Intelligence*. XXVII, 1303 pages. 2006.
- Vol. 4303: A. Hoffmann, B.-H. Kang, D. Richards, S. Tsumoto (Eds.), *Advances in Knowledge Acquisition and Management*. XI, 259 pages. 2006.
- Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), *MICA 2006: Advances in Artificial Intelligence*. XXVIII, 1232 pages. 2006.
- Vol. 4289: M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenić, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, M. van Someren (Eds.), *Semantics, Web and Mining*. X, 197 pages. 2006.
- Vol. 4285: Y. Matsumoto, R.W. Sproat, K.-F. Wong, M. Zhang (Eds.), *Computer Processing of Oriental Languages*. XVII, 544 pages. 2006.
- Vol. 4274: Q. Huo, B. Ma, E.-S. Chng, H. Li (Eds.), *Chinese Spoken Language Processing*. XXIV, 805 pages. 2006.
- Vol. 4265: L. Todorovski, N. Lavrač, K.P. Jantke (Eds.), *Discovery Science*. XIV, 384 pages. 2006.
- Vol. 4264: J.L. Balcázar, P.M. Long, F. Stephan (Eds.), *Algorithmic Learning Theory*. XIII, 393 pages. 2006.
- Vol. 4259: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H.S. Nguyen, R. Słowiński (Eds.), *Rough Sets and Current Trends in Computing*. XXII, 951 pages. 2006.
- Vol. 4253: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part III*. XXXII, 1301 pages. 2006.
- Vol. 4252: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), *Knowledge-Based Intelligent Information and Engineering Systems, Part II*. XXXIII, 1335 pages. 2006.

- Vol. 4251: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXVI, 1297 pages. 2006.
- Vol. 4248: S. Staab, V. Svátek (Eds.), Managing Knowledge in a World of Networks. XIV, 400 pages. 2006.
- Vol. 4246: M. Hermann, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIII, 588 pages. 2006.
- Vol. 4223: L. Wang, L. Jiao, G. Shi, X. Li, J. Liu (Eds.), Fuzzy Systems, and Knowledge Discovery. XXVIII, 1335 pages. 2006.
- Vol. 4213: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Knowledge Discovery in Databases: PKDD 2006. XXII, 660 pages. 2006.
- Vol. 4212: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Machine Learning: ECML 2006. XXIII, 851 pages. 2006.
- Vol. 4211: P. Vogt, Y. Sugita, E. Tuci, C.L. Nehaniv (Eds.), Symbol Grounding and Beyond. VIII, 237 pages. 2006.
- Vol. 4203: F. Esposito, Z.W. Raś, D. Malerba, G. Semeraro (Eds.), Foundations of Intelligent Systems. XVIII, 767 pages. 2006.
- Vol. 4201: Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, E. Tomita (Eds.), Grammatical Inference: Algorithms and Applications. XII, 359 pages. 2006.
- Vol. 4200: I.F.C. Smith (Ed.), Intelligent Computing in Engineering and Architecture. XIII, 692 pages. 2006.
- Vol. 4198: O. Nasraoui, O. Zaïane, M. Spiliopoulou, B. Mobasher, B. Masand, P.S. Yu (Eds.), Advances in Web Mining and Web Usage Analysis. IX, 177 pages. 2006.
- Vol. 4196: K. Fischer, I.J. Timm, E. André, N. Zhong (Eds.), Multiagent System Technologies. X, 185 pages. 2006.
- Vol. 4188: P. Sojka, I. Kopeček, K. Pala (Eds.), Text, Speech and Dialogue. XV, 721 pages. 2006.
- Vol. 4183: J. Euzenat, J. Domingue (Eds.), Artificial Intelligence: Methodology, Systems, and Applications. XIII, 291 pages. 2006.
- Vol. 4180: M. Kohlhase, OMDoc – An Open Markup Format for Mathematical Documents [version 1.2]. XIX, 428 pages. 2006.
- Vol. 4177: R. Marín, E. Onaindía, A. Bugarín, J. Santos (Eds.), Current Topics in Artificial Intelligence. XV, 482 pages. 2006.
- Vol. 4160: M. Fisher, W. van der Hoek, B. Konev, A. Lisitsa (Eds.), Logics in Artificial Intelligence. XII, 516 pages. 2006.
- Vol. 4155: O. Stock, M. Schaerf (Eds.), Reasoning, Action and Interaction in AI Theories and Systems. XVIII, 343 pages. 2006.
- Vol. 4149: M. Klusch, M. Rovatsos, T.R. Payne (Eds.), Cooperative Information Agents X. XII, 477 pages. 2006.
- Vol. 4140: J.S. Sichman, H. Coelho, S.O. Rezende (Eds.), Advances in Artificial Intelligence - IBERAMIA-SBIA 2006. XXIII, 635 pages. 2006.
- Vol. 4139: T. Salakoski, F. Ginter, S. Pyysalo, T. Pahikkala (Eds.), Advances in Natural Language Processing. XVI, 771 pages. 2006.
- Vol. 4133: J. Gratch, M. Young, R. Aylett, D. Ballin, P. Olivier (Eds.), Intelligent Virtual Agents. XIV, 472 pages. 2006.
- Vol. 4130: U. Furbach, N. Shankar (Eds.), Automated Reasoning. XV, 680 pages. 2006.
- Vol. 4120: J. Calmet, T. Ida, D. Wang (Eds.), Artificial Intelligence and Symbolic Computation. XIII, 269 pages. 2006.
- Vol. 4118: Z. Despotovic, S. Joseph, C. Sartori (Eds.), Agents and Peer-to-Peer Computing. XIV, 173 pages. 2006.
- Vol. 4114: D.-S. Huang, K. Li, G.W. Irwin (Eds.), Computational Intelligence, Part II. XXVII, 1337 pages. 2006.
- Vol. 4108: J.M. Borwein, W.M. Farmer (Eds.), Mathematical Knowledge Management. VIII, 295 pages. 2006.
- Vol. 4106: T.R. Roth-Berghofer, M.H. Göker, H.A. Güvenir (Eds.), Advances in Case-Based Reasoning. XIV, 566 pages. 2006.
- Vol. 4099: Q. Yang, G. Webb (Eds.), PRICAI 2006: Trends in Artificial Intelligence. XXVIII, 1263 pages. 2006.
- Vol. 4095: S. Nolfi, G. Baldassarre, R. Calabretta, J.C.T. Hallam, D. Marocco, J.-A. Meyer, O. Miglino, D. Parisi (Eds.), From Animals to Animats 9. XV, 869 pages. 2006.
- Vol. 4093: X. Li, O.R. Zaïane, Z. Li (Eds.), Advanced Data Mining and Applications. XXI, 1110 pages. 2006.
- Vol. 4092: J. Lang, F. Lin, J. Wang (Eds.), Knowledge Science, Engineering and Management. XV, 664 pages. 2006.
- Vol. 4088: Z.-Z. Shi, R. Sadananda (Eds.), Agent Computing and Multi-Agent Systems. XVII, 827 pages. 2006.
- Vol. 4087: F. Schwenker, S. Marinai (Eds.), Artificial Neural Networks in Pattern Recognition. IX, 299 pages. 2006.
- Vol. 4068: H. Schärfe, P. Hitzler, P. Øhrstrøm (Eds.), Conceptual Structures: Inspiration and Application. XI, 455 pages. 2006.
- Vol. 4065: P. Perner (Ed.), Advances in Data Mining. XI, 592 pages. 2006.
- Vol. 4062: G.-Y. Wang, J.F. Peters, A. Skowron, Y. Yao (Eds.), Rough Sets and Knowledge Technology. XX, 810 pages. 2006.
- Vol. 4049: S. Parsons, N. Maudet, P. Moraitis, I. Rahwan (Eds.), Argumentation in Multi-Agent Systems. XIV, 313 pages. 2006.
- Vol. 4048: L. Goble, J.-J.C. Meyer (Eds.), Deontic Logic and Artificial Normative Systems. X, 273 pages. 2006.
- Vol. 4045: D. Barker-Plummer, R. Cox, N. Swoboda (Eds.), Diagrammatic Representation and Inference. XII, 301 pages. 2006.
- Vol. 4031: M. Ali, R. Dapoigny (Eds.), Advances in Applied Artificial Intelligence. XXIII, 1353 pages. 2006.

¥538.00元

Table of Contents

Part I: Foundations of Human Computing

Foundations of Human Computing: Facial Expression and Emotion	1
<i>Jeffrey F. Cohn</i>	
Instinctive Computing	17
<i>Yang Cai</i>	

Part II: Sensing Humans for Human Computing

Human Computing and Machine Understanding of Human Behavior: A Survey	47
<i>Maja Pantic, Alex Pentland, Anton Nijholt, and Thomas S. Huang</i>	
Audio-Visual Spontaneous Emotion Recognition	72
<i>Zhihong Zeng, Yuxiao Hu, Glenn I. Roisman, Zhen Wen, Yun Fu, and Thomas S. Huang</i>	
Modeling Naturalistic Affective States Via Facial, Vocal, and Bodily Expressions Recognition	91
<i>Kostas Karpouzis, George Caridakis, Loic Kessous, Noam Amir, Amaryllis Raouzaiou, Lori Malatesta, and Stefanos Kollias</i>	
Emotion and Reinforcement: Affective Facial Expressions Facilitate Robot Learning	113
<i>Joost Broekens</i>	
Trajectory-Based Representation of Human Actions	133
<i>Antonios Oikonomopoulos, Ioannis Patras, Maja Pantic, and Nikos Paragios</i>	
Modelling the Communication Atmosphere: A Human Centered Multimedia Approach to Evaluate Communicative Situations	155
<i>Tomasz M. Rutkowski and Danilo P. Mandic</i>	
Modeling Influence Between Experts	170
<i>Wen Dong and Alex Pentland</i>	

Part III: Anthropocentric Interaction Models for Human Computing

Social Intelligence Design and Human Computing	190
<i>Toyoaki Nishida</i>	

Feedback Loops in Communication and Human Computing 215
 Rieks op den Akker and Dirk Heylen

Evaluating the Future of HCI: Challenges for the Evaluation of
Emerging Applications..... 234
 Ronald Poppe, Rutger Rienks, and Betsy van Dijk

Gaze-X: Adaptive, Affective, Multimodal Interface for Single-User
Office Scenarios 251
 Ludo Maat and Maja Pantic

SmartWeb Handheld — Multimodal Interaction with Ontological
Knowledge Bases and Semantic Web Services 272
 *Daniel Sonntag, Ralf Engel, Gerd Herzog, Alexander Pfalzgraf,
 Norbert Pfleger, Massimo Romanelli, and Norbert Reithinger*

A Learning-Based High-Level Human Computer Interface for Face
Modeling and Animation 296
 Volker Blanz

Challenges for Virtual Humans in Human Computing 316
 Dennis Reidsma, Zsófia Ruttkay, and Anton Nijholt

Affect Detection and an Automated Improvisational AI Actor in
E-Drama 339
 *Li Zhang, Marco Gillies, John A. Barnden, Robert J. Hendley,
 Mark G. Lee, and Alan M. Wallington*

Author Index 359

Foundations of Human Computing: Facial Expression and Emotion*

Jeffrey F. Cohn

Department of Psychology, University of Pittsburgh, 3137 SQ, 210 S. Bouquet Street,
Pittsburgh, PA 15260 USA
jeffcohn@cs.cmu.edu
www.pitt.edu/~jeffcohn

Abstract. Many people believe that emotions and subjective feelings are one and the same and that a goal of human-centered computing is emotion recognition. The first belief is outdated; the second mistaken. For human-centered computing to succeed, a different way of thinking is needed. Emotions are species-typical patterns that evolved because of their value in addressing fundamental life tasks. Emotions consist of multiple components, of which subjective feelings may be one. They are not directly observable, but inferred from expressive behavior, self-report, physiological indicators, and context. I focus on expressive facial behavior because of its coherence with other indicators and research. Among the topics included are measurement, timing, individual differences, dyadic interaction, and inference. I propose that design and implementation of perceptual user interfaces may be better informed by considering the complexity of emotion, its various indicators, measurement, individual differences, dyadic interaction, and problems of inference.

Keywords: Emotion, measurement, facial expression, automatic facial image analysis, human-computer interaction, temporal dynamics.

1 Introduction

How can computers recognize human emotions? Is this even the correct question? By emotion, people often think of subjective feelings, but emotions are more than that and subjective feeling is in no sense essential. There is no *sin qua non* for emotion. Emotions are species-typical patterns consisting of multiple components that may include intentions, action tendencies, appraisals, other cognitions, neuromuscular and physiological changes, expressive behavior, and subjective feelings. None of these alone is necessary or sufficient for any given situation. In human-human interaction, intentions and action tendencies often are more important than what an individual may be feeling. People may or may not be aware of what they're feeling, and feelings often come about some time late in the temporal unfolding of an emotion.

* A previous version of this paper was originally published in the *Proceedings of the ACM International Conference on Multimodal Interfaces*, Banff, Canada, 2006 (Copyright © ACM Press).