Lecture Notes in Computer Science 4299

Steve Renals Samy Bengio
Jonathan G. Fiscus (Eds.)

Machine Learning for Multimodal Interaction

Third International Workshop, MLMI 2006
Bethesda, MD, USA, May 1-4, 2006
Revised Selected Papers

Springer

Volume Editors

Steve Renals
University of Edinburgh
The Centre for Speech Technology Research
Edinburgh EH8 9LW, UK
E-mail: s.renals@ed.ac.uk

Samy Bengio
IDIAP Research Institute
CP 592, Rue du Simplon, 4, 1920 Martigny, Switzerland
E-mail: bengio@idiap.ch

Jonathan G. Fiscus
National Institute of Standards and Technology (NIST)
100 Bureau Drive, Gaithersburg, MD 20899-8940, USA
E-mail: jfiscus@nist.gov

Library of Congress Control Number: 2006938909

CR Subject Classification (1998): H.5.2-3, H.5, I.2.6, I.2.10, I.2, I.7, K.4, I.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743
ISBN-10 3-540-69267-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-69267-6 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11965152 06/3142 5 4 3 2 1 0

Preface

This book contains a selection of refereed papers presented at the 3rd Workshop on Machine Learning for Multimodal Interaction (MLMI 2006), held in Bethesda MD, USA during May 1–4, 2006.

The workshop was organized and sponsored jointly by the US National Institute for Standards and Technology (NIST), three projects supported by the European Commission (Information Society Technologies priority of the sixth Framework Programme)—the AMI and CHIL Integrated Projects, and the PASCAL Network of Excellence—and the Swiss National Science Foundation national research collaboration, IM2.

In addition to the main workshop, MLMI 2006 was co-located with the 4th NIST Meeting Recognition Workshop. This workshop was centered on the Rich Transcription 2006 Spring Meeting Recognition (RT-06) evaluation of speech technologies within the meeting domain. Building on the success of previous evaluations in this domain, the RT-06 evaluation continued evaluation tasks in the areas of speech-to-text, who-spoke-when, and speech activity detection.

The conference program featured invited talks, full papers (subject to careful peer review, by at least three reviewers), and posters (accepted on the basis of abstracts) covering a wide range of areas related to machine learning applied to multimodal interaction—and more specifically to multimodal meeting processing, as addressed by the various sponsoring projects. These areas included human–human communication modeling, speech and visual processing, multimodal processing, fusion and fission, human–computer interaction, and the modeling of discourse and dialog, with an emphasis on the application of machine learning. Out of the submitted full papers, about 50% were accepted for publication in the present volume, after authors had been invited to take review comments and conference feedback into account. The workshop featured invited talks from Roderick Murray-Smith (University of Glasgow), Tsuhan Chen (Carnegie Mellon University) and David McNeill (University of Chicago), and a special session on projects in the area of multimodal interaction including presentations on the VACE, CHIL and AMI projects.

Based on the successes of the first three MLMI workshops, and to strengthen and broaden the base of this workshop series, the MLMI standing committee was formed. The initial membership comprises Samy Bengio (IDIAP), Hervé Bourlard (IDIAP and EPFL), Tsuhan Chen (Carnegie Mellon University), John Garofolo (NIST), Mary Harper (Purdue University), Sharon Oviatt (Natural Interaction Systems), Steve Renals (Edinburgh University), Rainer Stiefelhagen (Universität Karlsruhe), and Alex Waibel (Carnegie Mellon University and Universität Karlsruhe). The committee will provide a permanent link across MLMI workshops. MLMI 2007, the fourth workshop in the series, will take place in

Brno, Czech Republic during June 28–30, 2007, directly after ACL–2007, which takes place in Prague.

Finally, we take this opportunity to thank our Programme Committee members, the sponsoring projects and funding agencies, and those responsible for the excellent management and organization of the workshop and the follow-up details resulting in the present book.

October 2006

Steve Renals
Samy Bengio
Jonathan Fiscus

Organization

Organizing Committee

Samy Bengio	IDIAP Research Institute
Hervé Bourlard	IDIAP Research Institute
Jonathan Fiscus (Co-chair)	NIST
John Garofolo	NIST
Steve Renals (Co-chair)	University of Edinburgh
Vincent Stanford (Demonstrations)	NIST
Alex Waibel (Special Sessions)	CMU and Universität Karlsruhe

Workshop Organization

Patrice Boulanger	NIST
Caroline Hastings	University of Edinburgh
Avril Heron	University of Edinburgh
Jonathan Kilgour	University of Edinburgh
Teresa Vicente	NIST

Program Committee

Marc Al-Hames	Munich University of Technology
Tilman Becker	DFKI
Jean Carletta	University of Edinburgh
Dan Ellis	Columbia University
Corinne Fredouille	University of Avignon
Thomas Hain	University of Sheffield
Mary Harper	Purdue University
Thomas Huang	University of Illinois
Alejandro Jaimes	Fuji Xerox
Samuel Kaski	University of Helsinki
Stephane Marchand-Maillat	University of Geneva
Nelson Morgan	ICSI
Andrei Popescu-Belis	University of Geneva
Mubarak Shah	University of Central Florida
Rainer Stiefelhagen	Universität Karlsruhe
Jean-Philippe Thiran	EPFL
Victor Tom	BAE Systems
Pierre Wellner	IDIAP Research Institute

Sponsoring Projects and Institutions

Institutions:

- US National Institute of Standards and Technology (NIST), http://www.nist.gov/speech/
- European Commission, through the Multimodal Interfaces objective of the Information Society Technologies (IST) priority of the sixth Framework Programme
- Swiss National Science Foundation, through the National Center of Competence in Research (NCCR) program

Projects:

- AMI, Augmented Multiparty Interaction, http://www.amiproject.org/
- CHIL, Computers in the Human Interaction Loop, http://chil.server.de/
- PASCAL, Pattern Analysis, Statistical Modeling and Computational Learning, http://www.pascal-network.org/
- IM2, Interactive Multimodal Information Management, http://www.im2.ch/

Table of Contents

MLMI'06

I Invited Paper

II Multimodal Processing

III Image and Video Processing

IV HCI and Applications

V Discourse and Dialogue

VI Speech and Audio Processing

VII NIST Meeting Recognition Evaluation