

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Alfred Kobsa

*University of California, Irvine, CA, USA*

Friedemann Mattern

*ETH Zurich, Switzerland*

John C. Mitchell

*Stanford University, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

Oscar Nierstrasz

*University of Bern, Switzerland*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*TU Dortmund University, Germany*

Madhu Sudan

*Microsoft Research, Cambridge, MA, USA*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Gerhard Weikum

*Max Planck Institute for Informatics, Saarbruecken, Germany*

Emanuele Salerno A. Enis Çetin  
Ovidio Salvetti (Eds.)

# Computational Intelligence for Multimedia Understanding

International Workshop, MUSCLE 2011  
Pisa, Italy, December 13-15, 2011  
Revised Selected Papers

## Volume Editors

Emanuele Salerno  
Ovidio Salvetti  
ISTI  
National Research Council of Italy  
Via Moruzzi, 1  
56124 Pisa, Italy  
E-mail: {emanuele.salerno, ovidio.salvetti}@isti.cnr.it

A. Enis Çetin  
Bilkent University  
Electrical and Electronics Engineering  
06800 Bilkent, Ankara, Turkey  
E-mail: cetin@bilkent.edu

ISSN 0302-9743 e-ISSN 1611-3349  
ISBN 978-3-642-32435-2 e-ISBN 978-3-642-32436-9  
DOI 10.1007/978-3-642-32436-9  
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012946551

CR Subject Classification (1998): H.5.1, I.4, I.2.10, I.5, I.2.4, I.2.6, H.2.8, G.1.2

LNCS Sublibrary: SL 6 – Image Processing, Computer Vision, Pattern Recognition, and Graphics

© Springer-Verlag Berlin Heidelberg 2012

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.

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

*Typesetting:* Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Preface

We live in a world where multimedia technology and resources improve the quality of our daily experiences. Multimedia data, in the form of still pictures, graphics, 3D models, audio, speech, video and a combination of these are in fact playing an increasingly important role in our life. Multimedia offers awe-inspiring possibilities to meet modern needs by enabling the distribution of content through several means. This information can be processed to create new knowledge. Therefore, we need to develop algorithms for computational interpretation and processing of such data, but also to share them across the network in an intelligent manner. Networked multimedia communications, such as video and audio on demand, video conferencing and e-learning, give us an idea of the growing diffusion of these data. Multimedia communications produce a growing demand for systems and tools to satisfy the most sophisticated requirements for storing, searching, accessing, retrieving, managing and sharing complex resources with many different formats of media types. A multimedia system is generally made of different components: database, multimedia storage server, network and client systems, often in a mobile environment. New initiatives are trying to bind these components together in open multimedia frameworks for an interoperable use of multimedia data in a distributed environment. Finding multimedia objects by content in a distributed database means searching for them on the basis of content descriptions and similarity measures. Looking for events in a video library requires not only searching a database for objects related to the specified events, but also considering bandwidth constraints. Enabling automatic classification and segmentation of digital resources is necessary to produce information that can be then extracted from multimedia also considering annotations. The above processes involve multimedia producers, but also several scientific communities (*e.g.*, cultural heritage, education, medicine, etc.) who want to merge multimedia description, management, and processing for their domain-specific requirements. This entails a proper understanding of the semantic relationships between the multimedia requisites in different domains. Research in multimedia analysis and understanding ranges from low-level data processing to automatic recognition. Current research domains involve methods and techniques for pattern recognition, content-based indexing, image understanding, object/subject-based retrieval and representation, segmentation, multimodal and multisource signal fusion, as well as knowledge processing and tracking. Multimedia understanding has a wide range of applications. Fundamental research and new technologies give rise to the growth of new understanding and applications outcomes.

March 2012

A. Enis Çetin  
Ovidio Salvetti

# Introduction

This book contains the contributions presented at the International Workshop on Computational Intelligence for Multimedia Understanding, organized by the ERCIM working group on Multimedia Understanding Through Semantics, Computation, and Learning (MUSCLE) as an open forum to discuss the most recent advances in this field. The workshop took place in Pisa, Italy, during December 13–15, 2011.

An open call-for-papers was issued in May 2011. After a first review, the contributions that scored more than 4/5 (43% of the original submissions) were accepted for publication. The papers that scored no more than 3/5 were rejected. Among the other papers, those that scored 4/5 were accepted, but their second versions were subjected to scrutiny by the editors, and the ones that scored between 3/5 and 4/5 were given the possibility to be presented at the workshop, but their second versions underwent a new full review before final acceptance. In this way, with the help of more than 20 external reviewers, the Program Committee managed to ensure both a high scientific level and an acceptance rate of more than 80%. I am highly indebted to the members of the Program Committee and the reviewers for their hard work. I must also thank all the contributors, whose presentations made the workshop extremely interesting and fruitful, and whose authoritative papers give value to this book. Two greatly appreciated invited lectures were given by Sanni Siltanen of VTT, Finland, and Bülent Sankur of Bogaziçi University, Turkey. I thank them very much on behalf of the entire audience.

The papers presented here are both theoretical and application-oriented, covering a wide range of subjects. As far as the media are concerned, most contributions deal with still images, but there are some on video and text. Two papers also use results from human perception experiments for image and video analysis and synthesis. The organization of the volume follows, where possible, both criteria of media/methodology and application. The first three papers are all related to meaning and semantics. The papers by Perner and Colantonio et al. are devoted to ontologies, a very important subject for knowledge-based media analysis, also devising the possibility of including the processing algorithms in the ontological reference. Then, the only paper dealing with text analysis, by Malandrakis et al., describes a technique to detect the affective meaning of text fragments to expand specialized lexica. The papers by Ruiz et al., Kayabol et al., and Szirányi and Szolgay deal with classification, with different applications and approaches. Classification also emerges as one of the most important issues in media understanding, since, with segmentation, it is probably an essential prerequisite for the extraction of meaning. Another often essential step to extraction of meaning is shape reconstruction. The papers by Kim and Dahyot and Kovács are devoted to this topic. Keskin and Çetin improve on fractional

wavelets to build a powerful signal analysis tool. Han et al. treat spatial registration, another essential task when dealing with multimodal data. They adopt a geometric approach, pointing out that raw data or low-level features may not be appropriate to match images that are different in nature. Four papers from the same research team (Vácha and Haindl, Haindl and V. Havlíček, Haindl and M. Havlíček, and Filip et al.) deal with texture modeling, synthesis, and recognition, which are relevant to multimedia understanding and visualization, and also affect the fixation/attention of human observers. Carcel et al., Martienne et al., and Gallego et al. deal with the video media, describing annotation, labeling, and segmentation strategies, respectively. Magrini et al. present a real-time image analysis system deployed in an infomobility scenario. Finally, Szalai et al. describe a statistical model of human observers' behavior that helps to estimate possible tracking paths and regions of interest of the human visual system presented with video data.

Just to add a general remark to this conceptual map, I would say that there is a ubiquitous quest in this volume, which is mentioned explicitly in the Preface: “information can be processed to create new knowledge.” This is not immediate, and trying to work the details out would drive us to ancient philosophical questions. As often happens, great poetry helps us:<sup>1</sup>

*Where is the wisdom we have lost in knowledge?  
Where is the knowledge we have lost in information?*

Well, these papers show us how information processing can bring us to *meaning*. Perhaps meaning processing could bring us nearer to *knowledge*. Also, knowledge processing could bring us to wisdom, but this is out of the scope of this volume.

March 2012

Emanuele Salerno

---

<sup>1</sup> T. S. Eliot, *The Rock*, 1934

# Organization

MUSCLE 2011 was organized by ERCIM, the European Research Consortium in Informatics and Mathematics, Working Group on Multimedia Understanding through Semantics, Computation and Learning.

## Workshop Chairs

### General Chairs

|                 |                                    |
|-----------------|------------------------------------|
| A. Enis Çetin   | Bilkent University, Turkey         |
| Ovidio Salvetti | National Research Council of Italy |

### Program Chair

|                  |                                    |
|------------------|------------------------------------|
| Emanuele Salerno | National Research Council of Italy |
|------------------|------------------------------------|

## Program Committee

|                            |  |
|----------------------------|--|
| Rozenn Dahyot              | Trinity College Dublin, Ireland              |
| Patrick Gros               | INRIA, France                                |
| Michal Haindl              | UTIA, Czech Republic                         |
| Nahum Kiryati              | Tel Aviv University, Israel                  |
| Ercan E. Kuruoğlu          | CNR, Italy                                   |
| Marie-Colette van Lieshout | CWI, The Netherlands                         |
| Rafael Molina              | University of Granada, Spain                 |
| Montse Pardás              | Universitat Politècnica de Catalunya, Spain  |
| Petra Perner               | IBAI, Germany                                |
| Béatrice Pesquet-Popescu   | Telecom Paristech, France                    |
| Ioannis Pitas              | Aristotle University of Thessaloniki, Greece |
| Alex Potamianos            | Technical University of Crete, Greece        |
| Sanni Siltanen             | VTT, Finland                                 |
| Tamas Sziranyi             | SZTAKI, Hungary                              |
| Anna Tonazzini             | CNR, Italy                                   |
| Simon Wilson               | Trinity College Dublin, Ireland              |
| Josiane Zerubia            | INRIA, France                                |

## Local Committee

|                      |  |
|----------------------|--|
| Francesca Martinelli | CNR, Italy                             |
| Davide Moroni        | CNR, Italy, Chair                      |
| Gabriele Pieri       | CNR, Italy                             |
| Marco Tampucci       | CNR, Italy                             |
| Ettore Ricciardi     | COCES and ISTI-CNR, Italy, Secretariat |

## Referees

|               |               |                |
|---------------|---------------|----------------|
| G. Amato      | P. Gros       | E. Salerno     |
| M. Belkhatir  | M. Haindl     | O. Salvetti    |
| G. Bianco     | L. Havasi     | P. Savino      |
| A. E. Çetin   | K. Kilic      | F. Sebastiani  |
| G. Charpiat   | N. Kiryati    | J. Shi         |
| S. Colantonio | R. Molina     | S. Siltanen    |
| M. D'Acunto   | D. Moroni     | C. Strapparava |
| L. Dempere    | S. Pankhanti  | T. Sziranyi    |
| R. Dahyot     | M. Pardás     | A. Tonazzini   |
| J. Estévez    | G. Pieri      | A. Utasi       |
| G. Franzè     | I. Pitas      | D. Vitulano    |
| C. Gennaro    | P. Perner     | S. Wilson      |
| C. Germain    | A. Potamianos | B. Zitová      |

## Sponsoring Institutions

European Research Consortium in Informatics and Mathematics - ERCIM



National Research Council - CNR, Italy



*Consiglio Nazionale delle Ricerche*

National Institute for Research in Informatics and Automation - INRIA, France





# Table of Contents

|  |     |
|--|-----|
| Learning an Ontology for Visual Tasks . . . . .  | 1   |
| <i>Petra Perner</i>  |     |
| Ontology and Algorithms Integration for Image Analysis . . . . .   | 17  |
| <i>Sara Colantonio, Massimo Martinelli, and Ovidio Salvetti</i>  |     |
| EmotiWord: Affective Lexicon Creation with Application to Interaction<br>and Multimedia Data . . . . .                   | 30  |
| <i>Nikos Malandrakis, Alexandros Potamianos, Elias Iosif, and<br/>Shrikanth Narayanan</i>                                |     |
| A Bayesian Active Learning Framework for a Two-Class Classification<br>Problem . . . . .                                 | 42  |
| <i>Pablo Ruiz, Javier Mateos, Rafael Molina, and<br/>Aggelos K. Katsaggelos</i>  |     |
| Unsupervised Classification of SAR Images Using Hierarchical<br>Agglomeration and EM . . . . .                           | 54  |
| <i>Koray Kayabol, Vladimir A. Krylov, and Josiane Zerubia</i>  |     |
| Geometrical and Textural Component Separation with Adaptive Scale<br>Selection . . . . .                                 | 66  |
| <i>Tamás Szirányi and Dániel Szolgay</i>   |     |
| Bayesian Shape from Silhouettes . . . . .  | 78  |
| <i>Donghoon Kim and Rozenn Dahyot</i>  |     |
| Shape Retrieval and Recognition on Mobile Devices . . . . .  | 90  |
| <i>Levente Kovács</i>  |     |
| Directionally Selective Fractional Wavelet Transform Using a 2-D<br>Non-separable Unbalanced Lifting Structure . . . . . | 102 |
| <i>Furkan Keskin and A. Enis Çetin</i>   |     |
| Visible and Infrared Image Registration Employing Line-Based<br>Geometric Analysis . . . . .                             | 114 |
| <i>Jungong Han, Eric Pauwels, and Paul de Zeeuw</i>  |     |
| Texture Recognition Using Robust Markovian Features . . . . .  | 126 |
| <i>Pavel Vácha and Michal Haindl</i>   |     |
| A Plausible Texture Enlargement and Editing Compound Markovian<br>Model . . . . .  | 138 |
| <i>Michal Haindl and Vojtěch Havlíček</i>  |     |

|   |            |
|---|------------|
| Bidirectional Texture Function Simultaneous Autoregressive Model . . . . .                          | 149        |
| <i>Michal Haindl and Michal Havlíček</i>  |            |
| Analysis of Human Gaze Interactions with Texture and Shape . . . . .                                | 160        |
| <i>Jiří Filip, Pavel Vácha, and Michal Haindl</i>   |            |
| Rich Internet Application for Semi-automatic Annotation of Semantic<br>Shots on Keyframes . . . . . | 172        |
| <i>Elisabet Carcel, Manuel Martos, Xavier Giró-i-Nieto, and<br/>Ferran Marqués</i>                  |            |
| Labeling TV Stream Segments with Conditional Random Fields . . . . .                                | 183        |
| <i>Emmanuelle Martienne, Vincent Claveau, and Patrick Gros</i>                                      |            |
| Foreground Objects Segmentation for Moving Camera Scenarios Based<br>on SCGMM . . . . .             | 195        |
| <i>Jaime Gallego, Montse Pardàs, and Montse Solano</i>  |            |
| Real Time Image Analysis for Infomobility . . . . .   | 207        |
| <i>Massimo Magrini, Davide Moroni, Gabriele Pieri, and<br/>Ovidio Salvetti</i>                      |            |
| Tracking the Saliency Features in Images Based on Human Observation<br>Statistics . . . . .         | 219        |
| <i>Szilard Szalai, Tamás Szirányi, and Zoltan Vidnyanszky</i>                                       |            |
| <b>Author Index . . . . .</b>   | <b>235</b> |