



# Multimodal Processing and Interaction

## Audio, Video, Text

**Multimodal Processing and Interaction: Audio, Video and Text** presents high quality, state-of-the-art research ideas and results from theoretic, algorithmic and application viewpoints. This edited volume contains both state-of-the-art reviews and original contributions by leading experts in the scientific and technological field of multimedia. It grew out of a four-year collaboration among research groups participating in the European Network of Excellence on Multimedia Understanding, Semantics, Computation and Learning (MUSCLE).

**Multimodal Processing and Interaction: Audio, Video and Text** covers a broad spectrum of novel perspectives, analytic tools, algorithms, design practices and applications in multimedia science and engineering with emphasis on multimodal integration and modality fusion. This volume also contains contributions in the area of interaction with multimedia, especially multimodal interfaces for accessing multimedia content.

**Multimodal Processing and Interaction: Audio, Video and Text** is designed for a professional audience composed of practitioners and researchers in industry and academia. This book is suitable for advanced-level students in computer science and engineering as well.

Maragos · Potamianos · Gross



Petros Maragos  
Alexandros Potamianos  
Patrick Gros



Multimodal Processing and Interaction  
Audio, Video, Text

# Multimodal Processing and Interaction

## Audio, Video, Text

ISBN 978-0-387-76315-6



9 780387 763156

springer.com

 Springer

---

# **Multimodal Processing and Interaction: Audio, Video, Text**

Edited by

Petros Maragos, Alexandros Potamianos and Patrick Gros

---

## Contents

### **Multimodal Processing and Interaction: Audio, Video, Text**

*Edited by Petros Maragos, Alexandros Potamianos and Patrick Gros . . . V*

---

#### **Part I Review of the State-of-the-Art**

---

##### **1 Cross-Modal Integration for Performance Improving in Multimedia: A Review**

*Petros Maragos, Patrick Gros, Athanassios Katsamanis, George  
Papandreou . . . . . 3*

##### **2 Human-Computer Interfaces to Multimedia Content: a Review**

*Alexandros Potamianos, Manolis Perakakis . . . . . 49*

---

#### **Part II INTEGRATED MULTIMEDIA ANALYSIS AND RECOGNITION**

---

##### **3 Stochastic Models for Multimodal Video Analysis**

*Manolis Delakis, Guillaume Gravier, Patrick Gros . . . . . 91*

##### **4 Adaptive Multimodal Fusion by Uncertainty Compensation with Application to Audio-Visual Speech Recognition**

*George Papandreou, Athanassios Katsamanis, Vassilis Pitsikalis,  
Petros Maragos . . . . . 111*

##### **5 Action Recognition in Multimedia Streams**

*Rozenn Dahyot, François Pitié, Daire Lennon, Naomi Harte, Anil  
Kokaram . . . . . 127*

<b>6 Surveillance Using Both Video and Audio</b>	
<i>Yigithan Dedeoglu, B. Ugur Toreyin, Ugur Gudukbay, A. Enis Cetin</i>	143
<b>7 Movie Analysis with Emphasis to Dialogue and Action Scene Detection</b>	
<i>Emmanouil Benetos, Spyridon Siatras, Constantine Kotropoulos, Nikos Nikolaidis, Ioannis Pitas</i>	157
<b>8 Audiovisual Attention Modeling and Salient Event Detection</b>	
<i>Georgios Evangelopoulos, Konstantinos Rapantzikos, Petros Maragos, Yannis Avrithis, Alexandros Potamianos</i>	179
<b>9 Toward the Integration of Natural Language Processing and Automatic Speech Recognition: Using Morpho-Syntax and Pragmatics for Transcription</b>	
<i>Stéphane Huet, Gwénoùlé Lecorvé, Guillaume Gravier, Pascale Sébillot</i>	201
<hr/>	
<b>Part III SEARCHING MULTIMEDIA CONTENT</b>	
<hr/>	
<b>10 Interactive Image Retrieval Using a Hybrid Visual and Conceptual Content Representation</b>	
<i>Marin Ferecatu, Nozha Boujemaa, Michel Crucianu</i>	221
<b>11 Multimodal Analysis of Text and Audio Features for Music Information Retrieval</b>	
<i>Robert Neumayer, Andreas Rauber</i>	241
<b>12 Intelligent Search for Image Information on the Web through Text and Link Structure Analysis</b>	
<i>Euripides G.M. Petrakis</i>	259
<hr/>	
<b>Part IV INTERFACES TO MULTIMEDIA CONTENT</b>	
<hr/>	
<b>13 Design Principles for Multimodal Spoken Dialogue Systems</b>	
<i>Alexandros Potamianos, Manolis Perakakis</i>	279
<b>14 Eye Tracking: A New Interface for Visual Exploration</b>	
<i>Oyewole K. Oyekoya, Fred W. M. Stentiford</i>	297
<b>15 User Interaction for Mobile Devices</b>	
<i>Sanni Siltanen, Charles Woodward, Seppo Valli, Petri Honkamaa, Andreas Rauber</i>	311
<b>References</b>	331