

T4 (Full day)

Media Semantics and the Statistical Foundations for Understanding It

AM: Marc Davis (UC Berkeley),
Chitra Dorai (IBM Research),
Frank Nack (CWI, Netherlands)

PM: Edward Chang (UC Santa Barbara)

Part 1 (AM):

Focus: Understanding Media Semantics

Presentors: Marc Davis, University of California at Berkeley
Chitra Dorai, IBM T.J. Watson Research Center
Frank Nack, CWI

Introduction

The current goal of multimedia research is to make multimedia information pervasively accessible and useable. Achieving this goal requires bridging “the semantic gap,” which describes the gulf between the meaningful descriptions that users expect systems to associate with media and the low-level features that systems actually compute. One promising approach to bridging the semantic gap and building high-level semantic descriptions of media content is founded on understanding the semantics of various media within a computationally informed and systematic study of media production and reception. The purpose of this tutorial is to provide an understanding of the role, technologies, and applicability of media semantics to the tasks of managing and reusing various media. Understanding media semantics is essential for progress in media content analysis and applications and will be of great interest to many attendees at the ACM Multimedia 2005. In our tutorial we address a number of specific issues, including some recent advances and new areas:

- The challenge of the semantic gap and strategies for bridging it
- The basic communication, semiotic, and media theories underlying media semantics
- The particular semantics of audio-visual media
- The semantics of media combination (especially in video sequencing and editing)
- The applicability of media semantics to capturing, representing, processing, managing, repurposing, and personalizing media
- The future of media semantics
- Emerging new areas in multimedia and the role of media semantics

Due to the widespread interest and participation from the community in the tutorial we organized at the 2003 ACM Multimedia conference in Berkeley and the 2004 ACM Multimedia conference in New York centred on the same theme, we suggest this tutorial be offered again this year at ACM MM. Our 2003 tutorial in Berkeley, California was successful and attended by a total of 18 participants, with 12 of them as registered attendees. Our tutorial at New York in 2004 was again well-attended, with 22 registered attendees (*the most popular of all the tutorials*). From the surveys filled out by the participants, the ranking of our New York tutorial was 1.7 on a scale from 1 (exceptional) to 5 (extremely bad).

Audience / Format / Requirements/Schedule

This tutorial is designed for researchers and practitioners who would like to learn how to understand the semantics of various media, how to describe them, and how to make use of such descriptions in the whole value chain of media creation, management, distribution, delivery, and reuse. While the tutorial is focused on Media Semantics, given Media Semantics' fundamental import to multimedia systems, the tutorial should also be of interest to people working in: Wireless Multimedia Systems, XML and Multimedia, Content-Based Indexing and Retrieval, Desktop Video Editing, Multimedia on the Web, Digital Asset Management, Media-Rich Homes / Digital Entertainment, Practical Digital Libraries, Multimedia User Interface Design, Multimedia in Collaboration, TREC Video Retrieval, and Multimedia Information Retrieval. The tutorial will include lectures, demonstrations and audience participation. The preliminary but proven schedule is as follows:

Outline:

Welcome

Welcome participants, provide overview of tutorial goals, schedule, and ground rules

Sections:

Welcome Participants

Provision with tutorial material – printout of all slides plus annotated bibliography (suggestions of books, journals, conferences and other fields of interest) plus additional information on media standards (annotated list of relevant industrial, W3C and ISO standards).

Introduce Tutorial Organizers

Overview of Tutorial (goals, schedule, ground rules)

Bridging the semantic gap (20 minutes) Presenter: Chitra Dorai

Explain the limits of signal-based media analysis and the need for media semantics (and syntax)

Sections:

Media Signal Analysis Promises and Problems

The Semantic Gap

The Sensory Gap

Bridging the Semantic Gap

The Need for Media Semantics

cs and Computation

Fundamentals of Media Semantics (30 minutes) Presenter: Marc Davis

Explain the basic concepts of communication theory and semiotics as foundations for media semantics

Sections:

Communications Theory

Introduction to Semiotics

Semiotics for Media

Formalist and Cognitivist Media Theory

Computational Media Theory

Semantics of Images (30 minutes) Presenter: Frank Nack

Introduce the main semantic concepts of static visuals and how these can be computationally described

Examples are taken from existing applications

Sections:

Image Semantics

Mise-en-scene Properties

Cinematographic Properties

Semantic Descriptions of Images (Media Streams, W3C and ISO standards as applied to images) Introduction to Metadata Structures (hierarchy, inheritance, facets)

Semantics of Audio (45 minutes) Presenter: Chitra Dorai

Provide an overview of audio semantics and representational structures for audio metadata
Examples are taken from existing applications

Sections:

Audio Analysis (sounds, sound effects, music)

Audio Semantics (describing affect)

Standards for Audio Content Description

Semantics of Video (45 minutes) Presenter: Marc Davis

Introduce the main semantic concepts of video and how these can be computationally described

Examples are taken from existing applications

Sections:

Relationship to Image Semantics

Semantics of Time and Motion

Montage, or the Semantics and Syntax of Editing and Sequencing

Representing Space, Time, People, and Actions in Video

Metadata Systems for Temporal Media (Media Streams, MPEG-7, SMIL)

Final Discussion

Provide participants the possibility to reflect once again on the various concepts introduced in the Tutorial and allow for feedback, discussion, and questions and answers on the impact on their own research directions. Provide an outlook on the future of Media Semantics. We outline a few exciting emerging topics in multimedia in which media semantics becomes a central underpinning and describe new challenges and open problems

Sections:

Tutorial Summary

Questions and Answers

The equipment comprises: a video projector and audio speakers to be used with a laptop, a projection screen or surface.

Organizing Committee and Backgrounds

Marc Davis

University of California at Berkeley

Tel.: (510) 643-2253, Fax: (510) 642-5814

Email: marc@sims.berkeley.edu, Web: <http://www.sims.berkeley.edu/~marc>

Marc Davis is an Assistant Professor at the School of Information Management and Systems (SIMS) at the University of California at Berkeley where he directs Garage Cinema Research (<http://garage.sims.berkeley.edu>). His work is focused on creating the technology and applications that will enable daily media consumers to become daily media producers. Prof. Davis' research and teaching encompass the theory, design, and development of digital media systems for creating and using media metadata to automate media production, sharing, and reuse. Garage Cinema Research is researching and developing media metadata frameworks, smart cameras, automatic video editing systems, mobile media metadata technology and applications, and the social uses of personal media. Prof. Davis earned his B.A. in the College of Letters at Wesleyan University, his M.A. in Literary Theory and Philosophy at the University of Konstanz in Germany, and his Ph.D. in Media Arts and Sciences at the Massachusetts Institute of Technology Media Laboratory. As part of his doctoral dissertation at the MIT Media Laboratory, he developed Media Streams, an iconic visual language for annotating, retrieving, and repurposing digital video. At the MIT Media Laboratory, Marc Davis co-founded the Narrative Intelligence Reading Group, which innovated interdisciplinary discourse at the intersection of literary and media theory, artificial intelligence, and media technology and design. From 1993 to 1998 at Interval Research Corporation, he led research and development teams in automatic media production technology for which a patent was awarded in 2001. In 1997, he was an invited contributor to the 50th Anniversary Edition of the Communications of the ACM, for which he wrote a vision piece about the next 50 years of media technology. From 1999 to 2002, Marc Davis was Chairman and Chief Technology Officer of Amova, Inc., a developer of media automation and personalization technology. At UC Berkeley, Prof. Davis is a Co-Founder and Executive Committee Member of the new interdisciplinary Center for New Media (CNM), an Advisory Board Member of the Art, Technology, and Culture Colloquium (ATC), and an Affiliated Faculty Member of the Berkeley Institute of Design (BiD). Prof. Davis is on the editorial boards of leading journals in the field of multimedia: IEEE MultiMedia and ACM Transactions on Multimedia Computing, Communications, and Applications.

Chitra Dorai

IBM T.J. Watson Research Center

Tel.: (914) 784-6049, Fax: (914) 784-7455

Email: dorai@watson.ibm.com, Web: <http://www.research.ibm.com/people/d/dorai>

Dr. Chitra Dorai is a Research Staff Member and Manager at the IBM T.J. Watson Research Center, New York, where she leads the Media Semantics and e-Learning Media projects. Her research interests are in the areas of multimedia systems and digital video analysis, computer vision, pattern recognition, and machine learning. Her current research focuses on developing technologies for digital media analysis and content management in various domains, such as education and training media and motion pictures that are useful in content-based structuring, annotation and search and smart browsing. Dr. Dorai created the new research approach, jointly with Svetha Venkatesh, called "Computational Media Aesthetics" to address the problem of the semantic gap in automatic content annotation and management systems. Computational Media Aesthetics is defined as the algorithmic study of a variety of image and aural elements in media, founded on their patterns of use in film grammar, and the computational analysis of the principles that have emerged underlying their manipulation, individually or jointly, in the creative art of clarifying, intensifying, and interpreting some event for the audience. This media production-guided semantic analysis approach has been well received in the research community with awards. She recently served as a Guest-Editor of a special issue of IEEE Multimedia dedicated to this theme in spring 2003 which includes an article by Prof. Marc Davis. She also edited a volume titled *Media Computing: Computational Media Aesthetics* published in June 2002 by Kluwer Academic

Publishers, which includes a chapter by Dr. Frank Nack. She was an author and contributor of the Videotext Multimedia Description Scheme, which is now part of the MPEG-7 standard. Chitra received her Ph.D. from the Department of Computer Science at Michigan State University, where she was a recipient of the Distinguished Academic Achievement Award from the College of Engineering. She is a senior member of the IEEE and a member of the ACM.

Frank Nack

Centrum voor Wiskunde en Informatica (CWI)

Tel.: +31 20 592 4223, Fax: +31 20 592 4312

Email: Frank.Nack@cwi.nl, Web: <http://homepages.cwi.nl/~nack>

Dr. Frank Nack is a senior researcher at CWI, currently working in the Multimedia and Human-Computer Interaction group. He obtained his Ph.D. with a thesis on “The Application of Video Semantics and Theme Representation for Automated Film Editing,” at Lancaster University, UK. The main thrust of his research is on the representation, retrieval and reuse of media in distributed hypermedia systems, educational hypermedia systems that enhance human communication and creativity, computational assistance for the development, maintenance and usage of hypermedia systems and distributed hypermedia systems, computational applications of media theory & semiotics, automated video editing, interactive storytelling, and computational humour theory. He was member of the MPEG-7 standardization group where he served as the editor of the Context and Objectives Document and the Requirements Document, and chaired the MPEG-7 DDL development group. Frank is on the editorial board of IEEE Multimedia, where he serves as associated editor in chief and edits the Media Impact column.