Yannis Avrithis
Yiannis Kompatsiaris
Steffen Staab
Noel E. O'Connor (Eds.)

LNCS 4306

Semantic Multimedia

First International Conference on Semantic and Digital Media Technologies, SAMT 2006
Athens, Greece, December 2006, Proceedings

Springer

Yannis Avrithis   Yiannis Kompatsiaris
Steffen Staab   Noel E. O'Connor (Eds.)

Semantic Multimedia

First International Conference on Semantic
and Digital Media Technologies, SAMT 2006
Athens, Greece, December 6-8, 2006
Proceedings

Springer

Volume Editors

Yannis Avrithis
Image, Video & Multimedia Systems Laboratory
School of Electrical and Computer Engineering
National Technical University of Athens
9, Iroon, Polytechniou Str., 15773 Zographou, Athens, Greece
E-mail: iavr@image.ntua.gr

Yiannis Kompatsiaris
Informatics and Telematics Institute
Centre for Research and Technology Hellas
6th Km Thessaloniki-Thermi Road, 57001 Thermi-Thessaloniki, Greece
E-mail: ikom@iti.gr

Steffen Staab
University of Koblenz-Landau
Universitaetsstrasse 1, 56016 Koblenz, Germany
E-mail: staab@uni-koblenz.de

Noel E. O'Connor
Centre for Digital Video Processing
Dublin City University
Collins Avenue, Dublin 9, Ireland
E-mail: oconnorn@eeng.dcu.ie

Library of Congress Control Number: 2006936895

CR Subject Classification (1998): H.5.1, H.4, H.5, I.2.10, I.4, I.7, C.2

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

ISSN 0302-9743
ISBN-10 3-540-49335-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-49335-8 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: 11930334 06/3142 5 4 3 2 1 0

Preface

We are delighted to welcome you to the proceedings of the 1st International Conference on Semantic and Digital Media Technologies held in Athens.

SAMT 2006 aims to narrow the large disparity between the low-level descriptors that can be computed automatically from multimedia content and the richness and subjectivity of semantics in user queries and human interpretations of audiovisual media — The Semantic Gap. SAMT started out as two workshops, EWIMT 2004 and EWIMT 2005, that quickly achieved success in attracting high-quality papers from across Europe and beyond. This year EWIMT turned into the full-fledged conference SAMT, bringing together forums, projects, institutions and individuals investigating the integration of knowledge, semantics and low-level multimedia processing, and linking them with industrial engineers who exploit the underlying emerging technology.

In total, 68 papers were submitted to the SAMT 2006 conference and each was reviewed by at least two independent reviewers. We are grateful to the members of the Technical Program Committee who completed these reviews and allowed us to put together a very strong technical program of 17 papers. The selection process was very competitive with only 25% of papers being selected for oral presentation. The program also included two invited keynote talks from Alan Smeaton and Guus Schreiber, and we are very grateful to them for their insightful presentations.

We thank all our colleagues in the Organization Committee who helped us tremendously with putting together an interesting and rewarding number of events including four workshops, three tutorials, three special sessions, poster, demo and projects sessions, the SAMT 2006 industry day, and two invited talks from Roberto Cencioni and Luis Rodriguez-Rosello featuring the launch of the 7th Framework ICT Program. Our Steering Committee was of great assistance in all critical decisions, and the staff of the Image, Video and Multimedia Systems Laboratory of the National Technical University of Athens provided invaluable help in making this conference happen. Special thanks go to Marios Phinikettos and Themis Zervou for the tremendous amount of work they put into supporting SAMT 2006.

The SAMT 2006 conference was held in cooperation with the European Commission, EURASIP, COST 292 and the IET, and we are grateful to these organizations for promoting the event. We are also indebted to K-Space Network of Excellence, of which SAMT was envisioned as the flagship event, and all sponsors for their contributions and financial support.

December 2006

Y. Avrithis
Y. Kompatsiaris
S. Staab
N.E. O'Connor

Printing: Mercedes-Druck, Berlin
Binding: Stein+Lehmann, Berlin

International Conference on Semantic and Digital Media Technologies

Organizing Committee

General Chairs	Yannis Avrithis *National Technical University of Athens, Greece* Yiannis Kompatsiaris *Informatics and Telematics Institute, Greece*
Program Chairs	Noel O'Connor *Dublin City University, Ireland* Steffen Staab *University of Koblenz-Landau, Germany*
Special Sessions	José Martinez *Universidad Autonoma de Madrid, Spain*
Tutorials and Workshops	Raphaël Troncy *CWI, Netherlands* Vassilis Tzouvaras *National Technical University of Athens, Greece*
EU Liaison	Paola Hobson *Motorola Labs, UK*

Programme Committee

Bruno Bachimont, INA, France
Wolf-Tilo Balke, University of Hannover, Germany
Jenny Benois-Pineau, University of Bordeaux, France
Jesús Bescós, GTI-UAM, Spain
Nozha Boujemaa, INRIA, France
Patrick Bouthemy, IRISA, France
Pablo Castells, NETS-UAM, Spain
Andrea Cavallaro, Queen Mary University of London, UK
Stavros Christodoulakis, Technical University of Crete, Greece
Fabio Ciravegna, University of Sheffield, UK
Thierry Declerck, DFKI, Germany
Anastasios Delopoulos, Aristotle University of Thessaloniki, Greece
Edward Delp, Purdue University, USA
Martin Dzbor, Knowledge Media Institute, UK
Touradj Ebrahimi, Swiss Federal Institute of Technology, Switzerland
Jerome Euzenat, INRIA, France

Borko Furht, Florida Atlantic University, USA
Moncef Gabbouj, Tampere University of Technology, Finland
Christophe Garcia, France Telecom, France
William I. Grosky, University of Michigan, USA
Werner Haas, Joanneum Research, Austria
Christian Halaschek-Wiener, University of Maryland, USA
Siegfried Handschuh, DERI Galway, Ireland
Alan Hanjalic, TU Delft, Netherlands
Lynda Hardman, CWI, Netherlands
Andreas Hotho, University of Kassel, Germany
Jane Hunter, University of Queensland, Australia
Antoine Isaac, Vrije Universiteit Amsterdam, Netherlands
Ebroul Izquierdo, Queen Mary University of London, UK
Franciska de Jong, University of Twente, Netherlands
Joemon Jose, University of Glasgow, UK
Aggelos K. Katsaggelos, Northwestern University, USA
Hyoung Joong Kim, Kangwon National University, Korea
Joachim Koehler, Fraunhofer IMK, Germany
Stefanos Kollias, National Technical University of Athens, Greece
Paul Lewis, University of Southampton, UK
Petros Maragos, National Technical University of Athens, Greece
Ferran Marques, Technical University of Catalonia, Spain
Jose Martinez, GTI-UAM, Spain
Adrian Matellanes, Motorola Labs, UK
Bernard Merialdo, EURECOM, France
Frank Nack, V2_ Institute for the Unstable Media and CWI, Netherlands
Jacco van Ossenbruggen, CWI, Netherlands
Jeff Pan, University of Manchester, UK
Thrasyvoulos Pappas, Northwestern University, USA
Ioannis Pitas, Aristotle University of Thessaloniki, Greece
Dietrich Paulus, University of Koblenz, Germany
Eric Pauwels, CWI, Netherlands
Dennis Quan, IBM, USA
Keith van Rijsbergen, University of Glasgow, UK
Andrew Salway, University of Surrey, UK
Mark Sandler, Queen Mary, University of London, UK
Simone Santini, University of California, San Diego, USA
Guus Schreiber, Vrije Universiteit Amsterdam, Netherlands
Timothy Shih, Tam Kang University, Tawain
Sergej Sizov, University of Koblenz, Germany
Alan Smeaton, Dublin City University, Ireland
John R. Smith, IBM Research, USA
Giorgos Stamou, National Technical University of Athens, Greece
Fred S. Stentiford, University College London, UK
Michael Strintzis, ITI, Greece

Rudi Studer, University of Karlsruhe, Germany
Vojtech Svatek, UEP, Czech Republic
Murat Tekalp, University of Rochester, USA
Raphaël Troncy, CWI, Netherlands
Paulo Villegas, Telefonica I+D, Spain
Gerhard Widmer, University of Linz, Austria
Li-Qun Xu, British Telecom, UK

Sponsors

SAMT 2006 was organized by the Image, Video and Multimedia Systems Laboratory at the National Technical University of Athens. The event was co-sponsored by:

Publicity Sponsors

Information Society Technologies (IST), European Commission

European Association for Signal, Speech and Image Processing (EURASIP)

European Cooperation in the field of Scientific and Technical Research, COST 292 Action — Semantic Multimodal Analysis and Digital Media

Visual Information Engineering Professional Network (VIE PN), Institution of Engineering and Technology (IET)

Emerald Sponsor

K-Space Network of Excellence — Knowledge Space of Semantics Inference for Automatic Annotation and Retrieval of Multimedia Content

Gold Sponsor

Muscle Network of Excellence — Multimedia Understanding through Semantics, Computation and Learning

Silver Sponsors

aceMedia Integrated Project — Integrating Knowledge, Semantics and Content for User-Centred Intelligent Media Services

Motorola UK Ltd

Industry Day Sponsors

Athens Technology Center (ATC)

Deutsche Welle

MESH Integrated Project — Multimedia Semantic Syndication for Enhanced News Services

Best Paper Award Sponsor

X-Media Integrated Project — Large Scale Knowledge Sharing and Reuse Across Media

Lecture Notes in Computer Science

For information about Vols. 1–4216

please contact your bookseller or Springer

Vol. 4313: T. Margaria, B. Steffen (Eds.), Leveraging Applications of Formal Methods. IX, 197 pages. 2006.

Vol. 4312: S. Sugimoto, J. Hunter, A. Rauber, A. Morishima (Eds.), Digital Libraries: Achievements, Challenges and Opportunities. XVIII, 571 pages. 2006.

Vol. 4311: K. Cho, P. Jacquet (Eds.), Technologies for Advanced Heterogeneous Networks II. XI, 253 pages. 2006.

Vol. 4306: Y. Avrithis, Y. Kompatsiaris, S. Staab, N.E. O'Connor (Eds.), Semantic Multimedia. XII, 241 pages. 2006.

Vol. 4302: J. Domingo-Ferrer, L. Franconi (Eds.), Privacy in Statistical Databases. XI, 383 pages. 2006.

Vol. 4300: Y.Q. Shi (Ed.), Transactions on Data Hiding and Multimedia Security I. IX, 139 pages. 2006.

Vol. 4296: M.S. Rhee, B. Lee (Eds.), Information Security and Cryptology – ICISC. XIII, 358 pages. 2006.

Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), MICAI 2006: Advances in Artificial Intelligence. XXVIII, 1232 pages. 2006. (Sublibrary LNAI).

Vol. 4292: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part II. XXXII, 906 pages. 2006.

Vol. 4291: G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. Nefian, G. Meenakshisundaram, V. Pascucci, J. Zara, J. Molineros, H. Theisel, T. Malzbender (Eds.), Advances in Visual Computing, Part I. XXXI, 916 pages. 2006.

Vol. 4290: M. van Steen, M. Henning (Eds.), Middleware 2006. XIII, 425 pages. 2006.

Vol. 4284: X. Lai, K. Chen (Eds.), Advances in Cryptology – ASIACRYPT 2006. XIV, 468 pages. 2006.

Vol. 4283: Y.Q. Shi, B. Jeon (Eds.), Digital Watermarking. XII, 474 pages. 2006.

Vol. 4281: K. Barkaoui, A. Cavalcanti, A. Cerone (Eds.), Theoretical Aspects of Computing - ICTAC 2006. XV, 371 pages. 2006.

Vol. 4280: A.K. Datta, M. Gradinariu (Eds.), Stabilization, Safety, and Security of Distributed Systems. XVII, 590 pages. 2006.

Vol. 4279: N. Kobayashi (Ed.), Programming Languages and Systems. XI, 423 pages. 2006.

Vol. 4278: R. Meersman, Z. Tari, P. Herrero (Eds.), On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops, Part II. XLV, 1004 pages. 2006.

Vol. 4277: R. Meersman, Z. Tari, P. Herrero (Eds.), On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops, Part I. XLV, 1009 pages. 2006.

Vol. 4276: R. Meersman, Z. Tari (Eds.), On the Move to Meaningful Internet Systems 2006: CoopIS, DOA, GADA, and ODBASE, Part II. XXXII, 752 pages. 2006.

Vol. 4275: R. Meersman, Z. Tari (Eds.), On the Move to Meaningful Internet Systems 2006: CoopIS, DOA, GADA, and ODBASE, Part I. XXXI, 1115 pages. 2006.

Vol. 4273: I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, L. Aroyo (Eds.), The Semantic Web - ISWC 2006. XXIV, 1001 pages. 2006.

Vol. 4272: P. Havinga, M. Lijding, N. Meratnia, M. Wegdam (Eds.), Smart Sensing and Context. XI, 267 pages. 2006.

Vol. 4271: F.V. Fomin (Ed.), Graph-Theoretic Concepts in Computer Science. XIII, 358 pages. 2006.

Vol. 4270: H. Zha, Z. Pan, H. Thwaites, A.C. Addison, M. Forte (Eds.), Interactive Technologies and Sociotechnical Systems. XVI, 547 pages. 2006.

Vol. 4269: R. State, S. van der Meer, D. O'Sullivan, T. Pfeifer (Eds.), Large Scale Management of Distributed Systems. XIII, 282 pages. 2006.

Vol. 4268: G. Parr, D. Malone, M. Ó Foghlú (Eds.), Autonomic Principles of IP Operations and Management. XIII, 237 pages. 2006.

Vol. 4267: A. Helmy, B. Jennings, L. Murphy, T. Pfeifer (Eds.), Autonomic Management of Mobile Multimedia Services. XIII, 257 pages. 2006.

Vol. 4266: H. Yoshiura, K. Sakurai, K. Rannenberg, Y. Murayama, S. Kawamura (Eds.), Advances in Information and Computer Security. XIII, 438 pages. 2006.

Vol. 4265: L. Todorovski, N. Lavrač, K.P. Jantke (Eds.), Discovery Science. XIV, 384 pages. 2006. (Sublibrary LNAI).

Vol. 4264: J.L. Balcázar, P.M. Long, F. Stephan (Eds.), Algorithmic Learning Theory. XIII, 393 pages. 2006. (Sublibrary LNAI).

Vol. 4263: A. Levi, E. Savas, H. Yenigün, S. Balcisoy, Y. Saygin (Eds.), Computer and Information Sciences – ISCIS 2006. XXIII, 1084 pages. 2006.

Vol. 4261: Y. Zhuang, S. Yang, Y. Rui, Q. He (Eds.), Advances in Multimedia Information Processing - PCM 2006. XXII, 1040 pages. 2006.

Vol. 4260: Z. Liu, J. He (Eds.), Formal Methods and Software Engineering. XII, 778 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. (Sublibrary LNAI).

Vol. 4257: I. Richardson, P. Runeson, R. Messnarz (Eds.), Software Process Improvement. XI, 219 pages. 2006.

Vol. 4256: L. Feng, G. Wang, C. Zeng, R. Huang (Eds.), Web Information Systems – WISE 2006 Workshops. XIV, 320 pages. 2006.

Vol. 4255: K. Aberer, Z. Peng, E.A. Rundensteiner, Y. Zhang, X. Li (Eds.), Web Information Systems – WISE 2006. XIV, 563 pages. 2006.

Vol. 4254: T. Grust, H. Höpfner, A. Illarramendi, S. Jablonski, M. Mesiti, S. Müller, P.-L. Patranjan, K.-U. Sattler, M. Spiliopoulou (Eds.), Current Trends in Database Technology – EDBT 2006. XXXI, 932 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. (Sublibrary LNAI).

Vol. 4252: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXIII, 1335 pages. 2006. (Sublibrary LNAI).

Vol. 4251: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXVI, 1297 pages. 2006. (Sublibrary LNAI).

Vol. 4249: L. Goubin, M. Matsui (Eds.), Cryptographic Hardware and Embedded Systems - CHES 2006. XII, 462 pages. 2006.

Vol. 4248: S. Staab, V. Svátek (Eds.), Managing Knowledge in a World of Networks. XIV, 400 pages. 2006. (Sublibrary LNAI).

Vol. 4247: T.-D. Wang, X. Li, S.-H. Chen, X. Wang, H. Abbass, H. Iba, G. Chen, X. Yao (Eds.), Simulated Evolution and Learning. XXI, 940 pages. 2006.

Vol. 4246: M. Hermann, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIII, 588 pages. 2006. (Sublibrary LNAI).

Vol. 4245: A. Kuba, L.G. Nyúl, K. Palágyi (Eds.), Discrete Geometry for Computer Imagery. XIII, 688 pages. 2006.

Vol. 4244: S. Spaccapietra (Ed.), Journal on Data Semantics VII. XI, 267 pages. 2006.

Vol. 4243: T. Yakhno, E.J. Neuhold (Eds.), Advances in Information Systems. XIII, 420 pages. 2006.

Vol. 4242: A. Rashid, M. Aksit (Eds.), Transactions on Aspect-Oriented Software Development II. IX, 289 pages. 2006.

Vol. 4241: R.R. Beichel, M. Sonka (Eds.), Computer Vision Approaches to Medical Image Analysis. XI, 262 pages. 2006.

Vol. 4239: H.Y. Youn, M. Kim, H. Morikawa (Eds.), Ubiquitous Computing Systems. XVI, 548 pages. 2006.

Vol. 4238: Y.-T. Kim, M. Takano (Eds.), Management of Convergence Networks and Services. XVIII, 605 pages. 2006.

Vol. 4237: H. Leitold, E. Markatos (Eds.), Communications and Multimedia Security. XII, 253 pages. 2006.

Vol. 4236: L. Breveglieri, I. Koren, D. Naccache, J.-P. Seifert (Eds.), Fault Diagnosis and Tolerance in Cryptography. XIII, 253 pages. 2006.

Vol. 4234: I. King, J. Wang, L. Chan, D. Wang (Eds.), Neural Information Processing, Part III. XXII, 1227 pages. 2006.

Vol. 4233: I. King, J. Wang, L. Chan, D. Wang (Eds.), Neural Information Processing, Part II. XXII, 1203 pages. 2006.

Vol. 4232: I. King, J. Wang, L. Chan, D. Wang (Eds.), Neural Information Processing, Part I. XLVI, 1153 pages. 2006.

Vol. 4231: J. F. Roddick, R. Benjamins, S. Si-Saïd Cherfi, R. Chiang, C. Claramunt, R. Elmasri, F. Grandi, H. Han, M. Hepp, M. Hepp, M. Lytras, V.B. Mišić, G. Poels, I.-Y. Song, J.D. Trujillo, C. Vangenot (Eds.), Advances in Conceptual Modeling - Theory and Practice. XXII, 456 pages. 2006.

Vol. 4230: C. Priami, A. Ingólfsdóttir, B. Mishra, H.R. Nielson (Eds.), Transactions on Computational Systems Biology VII. VII, 185 pages. 2006. (Sublibrary LNBI).

Vol. 4229: E. Najm, J.F. Pradat-Peyre, V.V. Donzeau-Gouge (Eds.), Formal Techniques for Networked and Distributed Systems - FORTE 2006. X, 486 pages. 2006.

Vol. 4228: D.E. Lightfoot, C.A. Szyperski (Eds.), Modular Programming Languages. X, 415 pages. 2006.

Vol. 4227: W. Nejdl, K. Tochtermann (Eds.), Innovative Approaches for Learning and Knowledge Sharing. XVII, 721 pages. 2006.

Vol. 4226: R.T. Mittermeir (Ed.), Informatics Education – The Bridge between Using and Understanding Computers. XVII, 319 pages. 2006.

Vol. 4225: J.F. Martínez-Trinidad, J.A. Carrasco Ochoa, J. Kittler (Eds.), Progress in Pattern Recognition, Image Analysis and Applications. XIX, 995 pages. 2006.

Vol. 4224: E. Corchado, H. Yin, V. Botti, C. Fyfe (Eds.), Intelligent Data Engineering and Automated Learning – IDEAL 2006. XXVII, 1447 pages. 2006.

Vol. 4223: L. Wang, L. Jiao, G. Shi, X. Li, J. Liu (Eds.), Fuzzy Systems and Knowledge Discovery. XXVIII, 1335 pages. 2006. (Sublibrary LNAI).

Vol. 4222: L. Jiao, L. Wang, X. Gao, J. Liu, F. Wu (Eds.), Advances in Natural Computation, Part II. XLII, 998 pages. 2006.

Vol. 4221: L. Jiao, L. Wang, X. Gao, J. Liu, F. Wu (Eds.), Advances in Natural Computation, Part I. XLI, 992 pages. 2006.

Vol. 4220: C. Priami, G. Plotkin (Eds.), Transactions on Computational Systems Biology VI. VII, 247 pages. 2006. (Sublibrary LNBI).

Vol. 4219: D. Zamboni, C. Kruegel (Eds.), Recent Advances in Intrusion Detection. XII, 331 pages. 2006.

Vol. 4218: S. Graf, W. Zhang (Eds.), Automated Technology for Verification and Analysis. XIV, 540 pages. 2006.

Vol. 4217: P. Cuenca, L. Orozco-Barbosa (Eds.), Personal Wireless Communications. XV, 532 pages. 2006.

Table of Contents

Content vs. Context for Multimedia Semantics: The Case of SenseCam Image Structuring

Invited Keynote Paper

Alan F. Smeaton

Adaptive Information Cluster
& Centre For Digital Video Processing,
Dublin City University,
Ireland
Alan.Smeaton@dcu.ie

Abstract. Much of the current work on determining multimedia semantics from multimedia artifacts is based around using either context, or using content. When leveraged thoroughly these can independently provide content description which is used in building content-based applications. However, there are few cases where multimedia semantics are determined based on an *integrated* analysis of content and context. In this keynote talk we present one such example system in which we use an integrated combination of the two to automatically structure large collections of images taken by a SenseCam, a device from Microsoft Research which passively records a person's daily activities. This paper describes the post-processing we perform on SenseCam images in order to present a structured, organised visualisation of the highlights of each of the wearer's days.

1 Introduction

When we think of multimedia information retrieval and multimedia semantics we tend to think of fairly standard multimedia artifacts such as still images, music, video and maybe 3D objects. When we think of how to determine the semantic content of such multimedia artifacts we do so because we want to perform a variety of content-based operations on such information including browsing, searching, summarisation, linking, etc. And finally, when we look at *how* we might determine semantics of multimedia objects we find that there are generally two approaches, namely:

1. use the context of the objects such as information gathered at the time of object creation or capture, to help determine some content features;
2. extract information directly from within the content of the objects in order to determine some content aspects.

Trying to determine and then usefully use a user's *context* is a fairly hot topic in information retrieval at the moment with lots of attempts to capture and then

Y. Avrithis et al. (Eds.): SAMT 2006, LNCS 4306, pp. 1–10, 2006.

apply such context in retrieval [11]. Determining a document or a multimedia object's context has also been explored for a long time and this forms the basis for many current systems for multimedia object management. For example such basic metadata as date and time of creation form the essential content representation for many tools which manage personal photos. Examples of such popular photoware includes Photoshop Album [2], PhotoFinder [16], ACDSee [1], Picasa [9] and others. Other metadata created at the time of photo capture such information as shutter speed and lens aperture, whether a flash was used or not can also be used to support automatic grouping of photos [14]. Finally, there are emerging online photoware systems such as Flickr [5] and Yahoo 360 [19] which support user-supplied context information to help with photo organisation.

What all these applications have in common, apart from the fact that they are all used to manage personal photos, is that they all use semantic information to describe multimedia objects (photos) which are derived from the context of the photo ... either directly from the capture process, or provided by an end-user afterwards.

To complement semantics derived from context we also use semantics derived from content in helping to manage our multimedia objects. Returning to the example of personal photos, this corresponds to extracting features directly from the image contents. An example system which does this is MediAssist which automatically determines whether a picture was taken indoors or outdoors, whether it is of a built or of a natural environment, whether a picture contains faces and if so whether those faces are faces of known individuals [14]. While this is a limited set of descriptive semantics, automatically determining the presence or absence of a larger number of medium and high level semantic features in visual media is notoriously difficult as is shown repeatedly in the TRECVid benchmarking evaluation campaign [17].

Once we have determined some level of semantic representation for multimedia objects we can then use these for content-based operations such as retrieval and we find that those derived from content and from context are almost always used either independently of each other or collaboratively with each other, but rarely are they truly integrated with each other. In other words, because these semantics are derived from different primary sources they maintain and retain their differing heritages when they are used subsequently.

To illustrate this let us examine the different ways in which video shots can be retrieved. In [18] we presented a classification of five different experimental approaches to video shot retrieval, namely:

1. Use metadata determined at the time of video capture/creation to access video by date, time, title, genre, actors, popularity rating, etc. as in [12];
2. Use one or more example query images to match against shot keyframes using whole-image matching approaches based on colour, texture or edges, as shown by many systems in [17];
3. Use text queries to match against text derived from transcriptions of the spoken dialogue of text determined from video OCR, also as shown by many systems in [15];

4. Use video objects, semi-automatically determined from shot keyframes and from user query images, and match these video objects based on shape, colour and/or texture, also as shown by many systems in [17];
5. Use the presence or absence of semantic video features such as indoor, outdoor, beach, sky, boats, motor vehicles, certain named persons, etc. to narrow the scope of shot retrieval to only those shots likely to contain such features;

Many systems have been developed to support video shot retrieval using one, two or perhaps three of the above but none have been developed to support all of them and for those that support multiple modalities for shot retrieval, the user is normally left with responsibility for combining and integrating them.

In this paper we argue for a more integrated approach to using semantic features determined from content and from context, and we illustrate what is possible with a novel application based around sets of images taken with a SenseCam. In the next section we introduce the SenseCam and its possible range of applications and in section 3 we present a summary of our work on structuring SenseCam images based on an integrated combination of content and context features. Section 4 concludes the paper.

2 The SenseCam

A SenseCam is a device developed by Microsoft Research in Cambridge, UK, for recording visual images of a wearer's day. It passively captures images through a fisheye lens and stores them on-board for subsequent download to a personal computer [8]. In addition to being a camera, a SenseCam also has other sensors including a light meter, a passive infra-red sensor and a 3-axis accelerometer and sensor readings from all these devices are also stored for later download. However, in addition to recording some elements of the SenseCam environment, the additional sensors are also used in a semi-intelligent way to trigger when photos are to be taken. For example when a person walks in front of the wearer this can be picked up by the passive infra-red sensor to trigger a photo to be taken. Similarly, when the user moves by standing up, or moves from indoor to outdoor or vice-versa, these are picked up by the accelerometer and light level sensors respectively and also trigger taking of photos. As a default, without an explicit triggering from the sensors, or from a user-controlled button on the SenseCam, the device will take a new photo every 45 seconds anyway. In this way a typical day can have up to 3,000 photos taken, which could add up to almost a million images in a year. A SenseCam being worn around a wearer's neck is shown in Figure 1 and a set of sample images taken from the author's use of a SenseCam is shown in Table 1.

The SenseCam device has been used extensively in the MyLifeBits project at Microsoft Research [6], [7] as well as being used in other, exploratory projects at Microsoft Research in Cambridge [10]. Like many other sensor devices, the SenseCam is great at capturing raw data, up to a million images per year for each user, and the main challenge is to effectively manage this huge volume of personal data. This requires automatic analysis and structuring in order to

Table 1. Sample SenseCam images