

ACM Multimedia 2003: Berkeley, CA, USA

Lawrence A. Rowe, [Harrick M. Vin](#), [Thomas Plagemann](#), [Prashant J. Shenoy](#), [John R. Smith](#) (Eds.):
Proceedings of the Eleventh ACM International Conference on Multimedia, Berkeley, CA, USA,
November 2-8, 2003. ACM 2003, ISBN 1-58113-722-2

- [Michael B. Johnson](#):
Building "bows for violinists": designing real digital tools for working artists. 1

Content analysis

- [Liyuan Li](#), [Weimin Huang](#), [Irene Y. H. Gu](#), [Qi Tian](#):
Foreground object detection from videos containing complex background. 2-10
- [Xinguo Yu](#), [Changsheng Xu](#), [Hon Wai Leong](#), [Qi Tian](#), [Qing Tang](#), [Kongwah Wan](#):
Trajectory-based ball detection and tracking with applications to semantic analysis of broadcast soccer video. 11-20
- [Viktor S. Wold Eide](#), [Frank Eliassen](#), [Ole-Christoffer Granmo](#), [Olav Lysne](#):
Supporting timeliness and accuracy in distributed real-time content-based video analysis. 21-32
- [Ling-Yu Duan](#), [Min Xu](#), [Tat-Seng Chua](#), [Qi Tian](#), [Changsheng Xu](#):
A mid-level representation framework for semantic sports video analysis. 33-44

Multimedia streaming and services

- [Mohamed Hefeeda](#), [Ahsan Habib](#), [Boyan Botev](#), [Dongyan Xu](#), [Bharat K. Bhargava](#):
PROMISE: peer-to-peer media streaming using CollectCast. 45-54
- [Philippe de Cuetos](#), [Keith W. Ross](#):
Optimal streaming of layered video: joint scheduling and error concealment. 55-64
- [Ketil Lund](#), [Vera Goebel](#):
Adaptive disk scheduling in a multimedia DBMS. 65-74
- [Roger Zimmermann](#), [Kun Fu](#):
Comprehensive statistical admission control for streaming media servers. 75-85

Demonstration session 1

- [Lisa Amini](#), [Raymond Rose](#), [Chitra Venkatramani](#), [Olivier Verscheure](#), [Peter H. Westerink](#), [Pascal Frossard](#):
ARMS: adaptive rich media secure streaming. 86-87
- [Marc Davis](#), [Jeffrey Heer](#), [Ana Ramirez](#):
Active capture: automatic direction for automatic movies. 88-89
- [Wu-chi Feng](#), [Brian Code](#), [Edward C. Kaiser](#), [Mike Shea](#), [Wu-chang Feng](#):
Panoptes: scalable low-power video sensor networking technologies. 90-91
- [Andreas Girgensohn](#), [Frank M. Shipman III](#), [Lynn Wilcox](#):
Hyper-hitchcock: authoring interactive videos and generating interactive summaries. 92-93
- [Jamey Graham](#), [Berna Erol](#), [Jonathan J. Hull](#), [Dar-Shyang Lee](#):
The video paper multimedia playback system. 94-95
- [Kam-yiu Lam](#), [Calvin K. H. Chiu](#):
Mobile video stream monitoring system. 96-97
- [Ching-Yung Lin](#), [Belle L. Tseng](#), [Milind R. Naphade](#), [Apostol Natsey](#), [John R. Smith](#):
MPEG-7 video automatic labeling system. 98-99
- [Jiazhi Ou](#), [Xilin Chen](#), [Susan R. Fussell](#), [Jie Yang](#):
DOVE: drawing over video environment. 100-101
- [Timothy K. Shih](#), [Liang-Chen Lu](#), [Rong-Chi Chang](#):
An automatic image inpaint tool. 102-103
- [Dezhen Song](#), [Kenneth Y. Goldberg](#):
The co-opticon: shared access to a robotic streaming video camera. 104-105
- [Ming-Yu Wang](#), [Xing Xie](#), [Wei-Ying Ma](#), [HongJiang Zhang](#):
MobiPicture: browsing pictures on mobile devices. 106-107
- [Jiang Yu Zheng](#), [Min Shi](#), [Makoto Kato](#):
Route panoramas for city navigation. 108-109

Music

- [Keiichiro Hoashi](#), [Kazunori Matsumoto](#), [Naomi Inoue](#):
Personalization of user profiles for content-based music retrieval based on relevance feedback. 110-119

- [Victor Lavrenko](#), [Jeremy Pickens](#):
Polyphonic music modeling with random fields. 120-129
- [Richard L. Kline](#), [Ephraim P. Glinert](#):
Approximate matching algorithms for music information retrieval using vocal input. 130-139
- [Lie Lu](#), [HongJiang Zhang](#):
Automated extraction of music snippets. 140-147

Managing images

- [Hao Liu](#), [Xing Xie](#), [Wei-Ying Ma](#), [HongJiang Zhang](#):
Automatic browsing of large pictures on mobile devices. 148-155
- [Kentaro Toyama](#), [Ron Logan](#), [Asta Roseway](#):
Geographic location tags on digital images. 156-166
- [Keiji Yanai](#):
Generic image classification using visual knowledge on the web. 167-176

Student best paper contest

- [Eric P. Bennett](#), [Leonard McMillan](#):
Proscenium: a framework for spatio-temporal video editing. 177-184
- [Sang-Uok Kum](#), [Ketan Mayer-Patel](#), [Henry Fuchs](#):
Real-time compression for dynamic 3D environments. 185-194
- [Beitao Li](#), [Kingshy Goh](#), [Edward Y. Chang](#):
Confidence-based dynamic ensemble for image annotation and semantics discovery. 195-206

Reception and posters

- [Brett Adams](#), [Svetha Venkatesh](#):
Weaving stories in digital media: when Spielberg makes home movies. 207-210
- [Konstantin Biatov](#), [Joachim Köhler](#):
An audio stream classification and optimal segmentation for multimedia applications. 211-214
- [Jeff Boston](#), [Michelle Y. Kim](#), [William L. Luken](#), [Edward So](#), [Steve Wood](#):
Interleaving media data for MPEG-4 presentations. 215-218

- [Rui Cai](#), [Lie Lu](#), [HongJiang Zhang](#):
Using structure patterns of temporal and spectral feature in audio similarity measure. 219-222
- [Wei Chai](#), [Barry Vercoe](#):
Music thumbnailing via structural analysis. 223-226
- [Chris Chambers](#), [Wu-chi Feng](#), [Wu-chang Feng](#), [Debanjan Saha](#):
A geographic redirection service for on-line games. 227-230
- [Won-Sik Cheong](#), [Kyuheon Kim](#), [Gwang Hoon Park](#):
A new scanning method for H.264 based fine granular scalable video coding. 231-234
- [Ludmila Cherkasova](#), [Wenting Tang](#):
Capacity planning tool for streaming media services. 235-238
- [Dawei Ding](#), [Qing Li](#), [Bo Feng](#), [Liu Wenyin](#):
A semantic model for flash retrieval using co-occurrence analysis. 239-242
- [Ling-Yu Duan](#), [Min Xu](#), [Qi Tian](#), [Changsheng Xu](#):
Nonparametric color characterization using mean shift. 243-246
- [Xin Fan](#), [Xing Xie](#), [He-Qin Zhou](#), [Wei-Ying Ma](#):
Looking into video frames on small displays. 247-250
- [Silvia Hollfelder](#), [Peter Fankhauser](#), [Erich J. Neuhold](#):
Observation based vs. model based admission control for interactive multimedia sessions. 251-254
- [Giridharan Iyengar](#), [Harriet J. Nock](#):
Discriminative model fusion for semantic concept detection and annotation in video. 255-258
- [Hang-Bong Kang](#):
Affective content detection using HMMs. 259-262
- [Rainer Lienhart](#), [Igor Kozintsev](#), [Stefan Wehr](#):
Universal synchronization scheme for distributed audio-video capture on heterogeneous computing platforms. 263-266
- [Cheng-Yuan Lin](#), [Jyh-Shing Roger Jang](#), [Mao-Yuan Hsu](#):
An automatic singing voice rectifier design. 267-270
- [Kinji Matsumura](#), [Kazuya Usui](#), [Kenjiro Kai](#), [Koichi Ishikawa](#):
Location-aware data broadcasting: an application for digital mobile broadcasting in Japan. 271-274

- [Florent Monay](#), [Daniel Gatica-Perez](#):
On image auto-annotation with latent space models. 275-278
- [Frank Nack](#), [Amit Manniesing](#), [Lynda Hardman](#):
Colour picking: the pecking prder of form and function. 279-282
- [Chong-Wah Ngo](#):
A robust dissolve detector by support vector machine. 283-286
- [Dinh Q. Phung](#), [Svetha Venkatesh](#), [Chitra Dorai](#):
Hierarchical topical segmentation in instructional films based on cinematic expressive functions. 287-290
- [Juan Carlos Rojas](#), [Miriam Leeser](#):
Programming portable optimized multimedia applications. 291-294
- [Yong Rui](#), [Zicheng Liu](#):
ARTiFACIAL: automated reverse turing test using FACIAL features. 295-298
- [Andrew Salway](#), [Mike Graham](#):
Extracting information about emotions in films. 299-302
- [Yoshinao Takemae](#), [Kazuhiro Otsuka](#), [Naoki Mukawa](#):
Video cut editing rule based on participants' gaze in multiparty conversation. 303-306
- [Chitra Venkatramani](#), [Peter H. Westerink](#), [Olivier Verscheure](#), [Pascal Frossard](#):
Securing media for adaptive streaming. 307-310
- [Kongwah Wan](#), [Xin Yan](#), [Xinguo Yu](#), [Changsheng Xu](#):
Real-time goal-mouth detection in MPEG soccer video. 311-314
- [Feng Wang](#), [Chong-Wah Ngo](#), [Ting-Chuen Pong](#):
Synchronization of lecture videos and electronic slides by video text analysis. 315-318
- [Jun Wang](#), [Mohan S. Kankanhalli](#):
Experience based sampling technique for multimedia analysis. 319-322
- [Yuhang Wang](#), [Fillia Makedon](#):
R-Histogram: quantitative representation of spatial relations for similarity-based image retrieval. 323-326
- [Zhiheng Wang](#), [Sujata Banerjee](#), [Sugih Jamin](#):
Studying streaming video quality: from an application point of view. 327-330
- [Xiaomeng Wu](#), [Wenli Zhang](#), [Shunsuke Kamijo](#), [Masao Sakauchi](#):
Construction of interactive video information system by applying results of object recognition. 331-334

- [Lonce Wyse](#), [Ye Wang](#), [Xinglei Zhu](#):
Application of a content-based percussive sound synthesizer to packet loss recovery in music streaming. 335-338
- [Rong Yan](#), [Alexander G. Hauptmann](#):
The combination limit in multimedia retrieval. 339-342
- [Rong Yan](#), [Alexander G. Hauptmann](#), [Rong Jin](#):
Negative pseudo-relevance feedback in content-based video retrieval. 343-346
- [Satoshi Yonemoto](#), [Hiroshi Nakano](#), [Rin-ichiro Taniguchi](#):
Avatar motion control by user body postures. 347-350
- [Tatsuo Yotsukura](#), [Shigeo Morishima](#), [Satoshi Nakamura](#):
Model-based talking face synthesis for anthropomorphic spoken dialog agent system. 351-354
- [Lei Zhang](#), [Longbin Chen](#), [Mingjing Li](#), [HongJiang Zhang](#):
Automated annotation of human faces in family albums. 355-358
- [Yongwei Zhu](#), [Mohan S. Kankanhalli](#):
Music scale modeling for melody matching. 359-362
- [Lev Manovich](#):
Inventing new media: what we can learn from new media art and media history. 363

Image annotation and video summarization

- [Matthew L. Cooper](#), [Jonathan Foote](#), [Andreas Girgensohn](#), [Lynn Wilcox](#):
Temporal event clustering for digital photo collections. 364-373
- [Yu-Fei Ma](#), [HongJiang Zhang](#):
Contrast-based image attention analysis by using fuzzy growing. 374-381
- [Bin Yu](#), [Wei-Ying Ma](#), [Klara Nahrstedt](#), [HongJiang Zhang](#):
Video summarization based on user log enhanced link analysis. 382-391
- [Frank M. Shipman III](#), [Andreas Girgensohn](#), [Lynn Wilcox](#):
Generation of interactive multi-level video summaries. 392-401

Multimedia coding and security

- [Michael Militzer](#), [Maciej Suchomski](#), [Klaus Meyer-Wegener](#):
Improved p-domain rate control and perceived quality optimizations for MPEG-4 real-

time video applications. 402-411

- [Ye Wang](#), [Ali Ahmaniemi](#), [David Isherwood](#), [Wendong Huang](#):
Content-based UEP: a new scheme for packet loss recovery in music streaming. 412-421
- [Jacob Chakareski](#), [Sangeun Han](#), [Bernd Girod](#):
Layered coding vs. multiple descriptions for video streaming over multiple paths. 422-431
- [Cheng Peng](#), [Robert H. Deng](#), [Yongdong Wu](#), [Weizhong Shao](#):
A flexible and scalable authentication scheme for JPEG2000 image codestreams. 433-441

Demonstration session 2

- [Lalitha Agnihotri](#), [Nevenka Dimitrova](#), [John R. Kender](#), [John Zimmerman](#):
Music videos miner. 442-443
- [Christopher J. C. Burges](#), [John C. Platt](#), [Jonathan Goldstein](#):
Identifying audio clips with *RARE*. 444-445
- [Shu-Ching Chen](#), [Mei-Ling Shyu](#), [Na Zhao](#), [Chengcui Zhang](#):
An affinity-based image retrieval system for multimedia authoring and presentation. 446-447
- [Elaine Chew](#), [Alexandre R. J. François](#):
MuSA.RT: music on the spiral array. real-time. 448-449
- [Wolfgang Hürst](#):
Indexing, searching, and skimming of multimedia documents containing recorded lectures and live presentations. 450-451
- [Jyh-Shing Roger Jang](#), [Yung-Sen Jang](#):
Microcontroller implementation of melody recognition: a prototype. 452-453
- [Andruid Kerne](#), [Vikram Sundaram](#), [Jin Wang](#), [Madhur Khandelwal](#), [J. Michael Mistrot](#):
Human + agent: creating recombinant information. 454-455
- [Alison Lee](#), [Vicki L. Hanson](#):
Enhancing web accessibility. 456-457
- [Florian Pestoni](#), [Clemens Drews](#):
eXtensible content protection. 458-459
- [Claudio S. Pinhanez](#), [Rick Kjeldsen](#), [Lijun Tang](#), [Anthony Levas](#), [Mark Podlaseck](#), [Noi Sukaviriya](#), [Gopal Pingali](#):
Creating touch-screens anywhere with interactive projected displays. 460-461

- [Yong Rui](#), [Zicheng Liu](#):
Excuse me, but are you human? 462-463
- [Jun Shen](#), [Rong Yan](#), [Pei Sun](#), [Song Song](#):
Interactive multimedia messaging service platform. 464-465
- [Naoko Tosa](#), [Seigo Matsuoka](#), [Koji Miyazaki](#):
Interactive storytelling system using behavior-based non-verbal information: ZENetic computer. 466-467
- [Kongwah Wan](#), [Xin Yan](#), [Xinguo Yu](#), [Changsheng Xu](#):
Robust goal-mouth detection for virtual content insertion. 468-469

3D multimedia environments

- [H. Harlyn Baker](#), [Nina T. Bhatti](#), [Donald Tanguay](#), [Irwin Sobel](#), [Dan Gelb](#), [Michael E. Goss](#), [John MacCormick](#), [Kei Yuasa](#), [W. Bruce Culbertson](#), [Thomas Malzbender](#):
Computation and performance issues In coliseum: an immersive videoconferencing system. 470-479
- [Mojtaba Hosseini](#), [Nicolas D. Georganas](#):
Design of a multi-sender 3D videoconferencing application over an end system multicast protocol. 480-489

Multimedia authoring

- [Xian-Sheng Hua](#), [Lie Lu](#), [HongJiang Zhang](#):
AVE: automated home video editing. 490-497
- [Berna Erol](#), [Jonathan J. Hull](#), [Dar-Shyang Lee](#):
Linking multimedia presentations with their symbolic source documents: algorithm and applications. 498-507

Surveillance

- [Daniel DeMenthon](#), [David S. Doermann](#):
Video retrieval using spatio-temporal descriptors. 508-517
- [Cen Rao](#), [Mubarak Shah](#), [Tanveer Fathima Syeda-Mahmood](#):
Invariance in motion analysis of videos. 518-527
- [Gang Wu](#), [Yi-Leh Wu](#), [Long Jiao](#), [Yuan-Fang Wang](#), [Edward Y. Chang](#):
Multi-camera spatio-temporal fusion and biased sequence-data learning for security surveillance. 528-538

Interacting with media

- [Mat C. Hans](#), [Mark T. Smith](#):
Interacting with audio streams for entertainment and communication. 539-545
- [Chunyuan Liao](#), [Qiong Liu](#), [Don Kimber](#), [Patrick Chiu](#), [Jonathan Foote](#), [Lynn Wilcox](#):
Shared interactive video for teleconferencing. 546-554
- [Milton Chen](#):
Visualizing the pulse of a classroom. 555-561

Multimedia for tiny devices

- [Wu-chi Feng](#), [Brian Code](#), [Edward C. Kaiser](#), [Mike Shea](#), [Wu-chang Feng](#), [Louis Bavoil](#):
Panoptes: scalable low-power video sensor networking technologies. 562-571
- [Vikas C. Raykar](#), [Igor Kozintsev](#), [Rainer Lienhart](#):
Position calibration of audio sensors and actuators in a distributed computing platform. 572-581
- [Shivajit Mohapatra](#), [Radu Cornea](#), [Nikil D. Dutt](#), [Alexandru Nicolau](#), [Nalini Venkatasubramanian](#):
Integrated power management for video streaming to mobile handheld devices. 582-591

Video demonstration session

- [Xian-Sheng Hua](#), [Lie Lu](#), [HongJiang Zhang](#):
Photo2Video. 592-593
- [Christos Tryfonas](#), [James Schumacher](#):
Essistants. 594-595
- [Sam Yip](#), [Eugenia Leu](#), [Hunter Howe](#):
The automatic video editor. 596-597
- [John Adcock](#), [Matthew L. Cooper](#), [John Doherty](#), [Jonathan Foote](#), [Andreas Girgensohn](#), [Lynn Wilcox](#):
Managing digital memories with the FXPAL photo application. 598-599
- [John Doherty](#), [Andreas Girgensohn](#), [Jonathan Helfman](#), [Frank M. Shipman III](#), [Lynn Wilcox](#):
Detail-on-demand hypervideo. 600-601
- [Marc Davis](#):
Active capture: automatic direction for automatic movies. 602-603

Content-based retrieval

- [Dongge Li](#), [Nevenka Dimitrova](#), [Mingkun Li](#), [Ishwar K. Sethi](#):
Multimedia content processing through cross-modal association. 604-611
- [Hong Wu](#), [Hanqing Lu](#), [Songde Ma](#):
A practical SVM-based algorithm for ordinal regression in image retrieval. 612-621
- [Kai Yu](#), [Wei-Ying Ma](#), [Volker Tresp](#), [Zhao Xu](#), [Xiaofei He](#), [HongJiang Zhang](#), [Hans-Peter Kriegel](#):
Knowing a tree from the forest: art image retrieval using a society of profiles. 622-631
- [Hui Yang](#), [Lekha Chaisorn](#), [Yunlong Zhao](#), [Shi-Yong Neo](#), [Tat-Seng Chua](#):
VideoQA: question answering on news video. 632-641

Doctoral symposium - session I

- [Mohamed Hefeeda](#):
A framework for cost-effective peer-to-peer content distribution. 642-643
- [David E. Ott](#), [Ketan Mayer-Patel](#):
Transport-level protocol coordination in distributed multimedia applications. 644-645
- [Bin Yu](#), [Klara Nahrstedt](#):
A scalable overlay video mixing service model. 646-647

Doctoral symposium - session II

- [Barbara Barry](#):
The mindful camera: common sense for documentary videography. 648-649
- [A. Viranga Ratnaike](#), [Bala Srinivasan](#), [Surya Nepal](#):
Making sense of video content. 650-651
- [Dezhen Song](#):
Algorithms and systems for shared access to a robotic streaming video camera. 652-653