

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,

PETITIONER,

v.

POINTWISE VENTURES LLC,

PATENT OWNER.

Case No. IPR2025-01102

U.S. Patent No. 8,471,812

Declaration of Lauren Gluckman

I, Lauren Gluckman, state and declare as follows:

I. Introduction

1. I have prepared this Declaration in connection with a petition for inter partes review of U.S. Patent No. 8,471,812 which I understand will be filed concurrently with this Declaration.

2. I am currently a Senior Research Analyst at Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, 901 New York Avenue, NW, Washington, DC 20001-4413.

3. I am over eighteen years of age and am competent to make this Declaration. I make this Declaration based on my own personal knowledge and based on my knowledge of library science practices.

4. I earned a Master of Science in Library Science (“MLS”) from Rutgers University in 1993, and a Bachelor of Arts in Political Science from Rutgers University in 1991. I have worked as a librarian for over thirty years. I have been employed at Finnegan, Henderson, Farabow, Garrett & Dunner, LLP since August 2021. My previous position, from 2006-2021, was as the Head of Collection Management and Reference Librarian at the Montague Law Library at Penn State Law of the Pennsylvania State University. From 2002-2006, I was employed as a Reference Librarian at Sutherland Asbill & Brennan. I was employed as a Senior Reference Librarian at the United States Senate Library from

1996-2002 and from 1993-1995 I was an Assistant Librarian at Debevoise & Plimpton.

5. I am a member of the American Association of Law Libraries.

II. Standard Library Practice for Receiving, Cataloging, and Making Materials, including Serial and Monograph Publications, Publicly Available

6. I have knowledge of and experience with standard library practices regarding the receipt, cataloging, shelving, and making materials, including serial and monograph publications, available to the public. I have knowledge of and experience with the Machine-Readable Cataloging (MARC) system, an industry-wide standard that libraries use to catalog materials. I also have knowledge of and experience with OCLC (previously called Online Computer Library Center, Inc.) and OCLC Control Numbers, which are unique accession numbers assigned by the OCLC system when a bibliographic record is catalogued and made available to the public through WorldCat. OCLC control numbers are often used by the public to search for specific records within WorldCat, which is an online database of 586,648,074 bibliographic records from libraries throughout the world as of December 2024 (Source: <https://www.oclc.org/en/worldcat/inside-worldcat.html>).

7. Based on standard library practice, when a library receives a print item, the cataloguer or cataloguer's assistant will examine the item and date-stamp and/or affix a label to the item with the library name and/or barcode with the date it

was received. Next, the library will catalog the item within a few days of receiving it. As a general practice, cataloguing is centralized and performed by a cataloguing department within a library or university setting.

8. Normally, after an item is cataloged, the public may access the item by searching a library online catalog, browsing the physical library shelves, searching a physical card catalog, or by seeking assistance from a Librarian to identify the item. Standard library practice is to make an item available to the public within several days of cataloging. If an item requires original cataloguing, the item will be made available to the public up to and including one week from the date of receipt.

III. Serial Publications

9. A serial publication, often known as a “journal” or “periodical,” is a resource issued in successive parts and has no predetermined conclusion. These successive parts are commonly referred to as “issues,” and each issue is usually chronologically numbered and dated. The presence of enumeration, years of coverage, and/or other chronological information also indicates a serial publication.

10. There are significant differences between cataloging finite resources (books/monographs) and continuing resources (serials). For serials, the catalog record provides information about the serial as a whole, including the first or

earliest available issue. It also provides information as to holdings – the volumes and issues, with dates, received by the library and made available to the public. In serials cataloging, there are identifying characteristics unique to serials that are slightly different from monographs (books). The issue date for a print serial publication, for example, generally appears on the cover (front or back), the masthead page, the title page (if any), the table of contents page(s), or on the pages of the individual articles contained in the issue. The initial periodicals cataloging record will sometimes not reflect all subsequent changes in publication details (including minor variations in title, etc.). More information regarding the unique aspects of cataloguing serials can be found at this link:

<https://www.loc.gov/aba/pcc/conser/scctppt/Basic-2014/Basic-Trainee-Mannual.pdf>.

IV. Monographs versus Serials Cataloguing of Conference Proceedings

11. Conference proceedings often consist of a collection of individual papers presented at a meeting, symposium or conference held during a specific time-period in a designated geographic place. Although some ongoing conference publications are catalogued as serials (if a conference publication exhibits evidence of seriality after the first issue and it does not contain a distinctive title or series number), cataloguing a conference publication as a monograph can provide the public a way to easily locate and access the

conference resource, and its papers, by searching by a unique title, topic, or other data points based on descriptive elements of the conference proceeding. Monographs are efficient for public searching and offer the public an easier way to discover what an institution holds, while changes to the record are not an issue.

V. Ownership and Date-Stamp

12. Libraries have different policies on whether to date-stamp received publications. It is becoming increasingly common for libraries to no longer date-stamp the publications they receive. In the past, it was common practice to affix an ownership and date stamp somewhere within the publication, usually on the cover page, verso of the cover page, or a designated page within the publication. The timing of when an ownership/date stamp is entered on the monograph or periodical can vary from one library to another. It could occur when the item is received in acquisitions after shipment to the library, or at the time of cataloging. Therefore, it is possible for there to be instances when the date of receipt precedes the cataloging date or vice versa.

VI. MARC Records

13. The MARC system was developed during the 1960s to standardize bibliographic catalog records so they could be read by computers and shared among libraries. By the mid-1970s, MARC had become the international

standard for cataloging bibliographic materials and is still used today. Many libraries provide public access to their MARC records via their electronic cataloging systems at the library. In a MARC record, each field provides specific information about the cataloged item, including how materials are held and when they were made available to the public.

14. The MARC record system uses specific three-digit numeric codes (“field tags”) (from 001-999) to identify each field in a catalog record. For example, field tag 008 provides the six-digit date an item was catalogued (Date entered on file). The first six characters of field tag 008 are always in the “YYMMDD” format. Descriptions and definitions of all the character positions of field tag 008 are outlined here:

<https://www.loc.gov/marc/bibliographic/bd008a.html>.

15. As is relevant to this Declaration, MARC field tag 022 provides the International Standard Serial Number (ISSN), a unique identification number assigned to a serial publication. MARC field tag 245 sets forth the Title and statement of responsibility for the work. MARC field tag 260 sets forth information relating to the place of publication, printing, distribution, and copyright date.

VII. Open Access Resources and SPIE Digital Library

16. Open Access (OA) refers to all electronic resources available online without any licensing or copyright restrictions. A publication is considered Open Access when there are no financial, legal, or technical barriers to accessing it. This means anyone can read, download, copy, distribute, print, search for, and search within the information or use it for education or any other legal purpose. Examples of Open Access resources include articles, journals, books, conference proceedings, theses, videos, music, and more. Open Access has become increasingly popular in scholarly research and communication as academic resource prices have significantly increased over the past decade, and academic libraries' budgets have not kept pace with these new expenses.

17. The SPIE Digital Library is a publicly accessible online platform on optics and photonics, providing access to technical publications from SPIE Journals, Conference Proceedings and Presentations, and eBooks from SPIE Press from 1962 to the present. It is located at the following website:

<https://www.spiedigitallibrary.org>. The SPIE Digital Library provides the ability to search and browse features or to view tables of contents and abstracts for free. Access to its full text content generally requires a subscription or a pay-per-view purchase unless the content is freely available through Open Access.

18. New papers from SPIE's Journals and Proceedings are published as they receive approvals and complete production. Print publications follow completion of the online volume. For e-First publications, the electronic version is the publication of record, and citation information indicates the online publication date according to details provided from the SPIE Digital Library, (<https://www.spiedigitallibrary.org/about-the-dl>).

VIII. U.S. Copyright Office Records

19. The United States Copyright Office, a part of the Library of Congress, is the official U.S. government body which maintains copyright certificates and copyright records of registration in the United States, including providing public access to an online catalog of copyright records held by the Copyright Office. The copyright records contained in the public catalog are routinely used by copyright title researchers who are attempting to clear a chain of title for copyrighted works. The records contained in the public catalog also denote the date of creation and the date of publication of the copyrighted work.

IX. The Internet Archive

20. The Internet Archive, a 501(c)(3) non-profit and founded in 1996, is a publicly available digital library of Internet sites and other cultural artifacts that can be searched and accessed in digital form. Its mission is to provide Universal Access to All Knowledge including books, films, magazines, websites, videos,

software programs, audio recordings, and television. Like a physical library, it provides free access to everyone. The Internet Archive is available here:

<https://archive.org/>.

21. The Internet Archive created a resource known as the “Wayback Machine” (<https://web.archive.org/>). The Wayback Machine is named in reference to the famous Mr. Peabody’s WABAC (pronounced way-back) machine from the Rocky and Bullwinkle cartoon show. The Wayback Machine makes it possible to search and browse more than 450 billion publicly available web pages stored in the Internet Archive's web archive. Visitors to the Wayback Machine can search the Archive by keyword or a website URL. If archived records for a URL or the key words are available, the visitor will be presented with a display of available dates. The visitor may select one of those dates and be taken to the collection of web captures associated with the specific web crawl and point in time the capture came from.

22. This archived data made viewable by the Wayback Machine is obtained using web archiving software that automatically stores copies of files available via the Internet, each file preserved as it existed at a particular point in time.

23. The Internet Archive assigns a URL to all archived files within the Wayback Machine in the following format: [http://web.archive.org/web/\[Year in](http://web.archive.org/web/[Year in)

yyyy][Month in mm][Day in dd][Time code in hh:mm:ss]/[Archived URL]. This is known as an “extended URL” format.

X. Public Availability of Bove

24. This Declaration relates to the date of receipt, accessibility, and public availability of the following reference: Bove, V. Michael, et al. “Adding hyperlinks to digital television.” *SMPTE journal* 108.11 (1999): 795-801. I understand the Bove reference is being submitted as Exhibit 1012.

25. As detailed below, I have reviewed the print reference of *SMPTE Journal* containing Bove which was retained by Wisconsin TechSearch (WTS) from the supplier Linda Hall Library- University of Missouri on May 2, 2025, to determine the date of public accessibility and availability for this reference.

26. **Appendix 1** to this Declaration is a true and accurate copy of the cover page, table of contents and the reference for the print version of Bove, V. Michael, et al. “Adding hyperlinks to digital television.” *SMPTE journal* 108.11 (1999): 795-801. **Appendix 1** also sets forth the reference was received by Linda Hall Library NOV 17 1999, as indicated with a date stamp.

27. **Appendix 2** to this Declaration is a true and accurate copy of the Linda Hall Library public catalog record for its copy of *SMPTE journal* containing Bove which was downloaded from <https://catalog.lindahall.org/discovery/fulldisplay?docid=alma992679233405961&>

[context=L&vid=01LINDAHALL_INST:LHL&lang=en&search_scope=MyInstitution&adaptor=Local%20Search%20Engine&tab=LibraryCatalog&query=issn,contains,0036-1682,AND&mode=advanced&offset=0](https://catalog.lindahall.org/discovery/sourceRecord?vid=01LINDAHALL_INST:LHL&lang=en&search_scope=MyInstitution&adaptor=Local%20Search%20Engine&tab=LibraryCatalog&query=issn,contains,0036-1682,AND&mode=advanced&offset=0) on May 2, 2025.

28. The Linda Hall Library public catalog record, **Appendix 2** sets forth the publication, contents, and holdings information for its copy of *SMPTE journal* containing *Bove*. **Appendix 2** also denotes *SMPTE journal* containing *Bove* may be requested from the closed stacks in the Linda Hall Library.

29. **Appendix 3** contains the MARC record for its holdings of *SMPTE journal* containing *Bove*. **Appendix 3** was downloaded from https://catalog.lindahall.org/discovery/sourceRecord?vid=01LINDAHALL_INST:LHL&docId=alma992679233405961&recordOwner=01LINDAHALL_INST on May 2, 2025. **Appendix 3** confirms the fixed data elements of MARC field tag 008 as “760405c19769999nyumr1p 0 a0eng d”. As discussed above, the first six characters “760405” are in typical “YYMMDD” format and indicate, *SMPTE journal* was cataloged by Linda Hall Library on April 5, 1976. The “c” refers to the publication as a continuing resource currently published since 1976.

30. **MARC** field tag 022 sets forth the unique ISSN as 0036-1682. **MARC** field tag 245 sets forth the title of the publication as \$a*SMPTE journal* : \$bpublication of the Society of Motion Picture and Television Engineers. **MARC**

field tag 260 sets forth information relating to the place of publication and the publisher of *SMPTE journal* as \$a[Scarsdale, NY] : \$bThe Society.

31. Based on the information in **Appendices 2** and **3** attached to this Declaration, standard library practices, and my experience as a librarian *SMPTE journal* containing *Bove* was received by the Linda Hall Library NOV 17 1999 as set forth in the date stamp.

32. Accordingly, based on the information set forth in **Appendices 2** and **3** attached to this Declaration, standard library practices and my experience as a librarian, *SMPTE journal* containing *Bove*, would have been made available to the public within a few days of being checked-in and cataloged on NOV 17 1999. The interested public could have accessed *SMPTE journal* containing *Bove* by searching the Linda Hall Library public catalog, visiting the Library and requesting the volume within a few days of NOV 17 1999.

XI. Public Availability of *Oami*

33. This Declaration relates to the date of receipt, accessibility, and public availability of the following reference: Ryoma Oami, Eiji Kasutani, Akio Yamada, “Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones,” Proc. SPIE 5909, Applications of Digital Image Processing XXVIII, 59090N (16 September 2005). I understand *Oami* is being submitted as **Exhibit 1005**.

34. As detailed below, I have reviewed the print reference of *Applications of Digital Image Processing XXVIII* containing *Oami* which was retained by Wisconsin TechSearch (WTS) from the supplier Linda Hall Library- University of Missouri on May 2, 2025; the U.S. Copyright Record for the *Applications of Digital Image Processing XXVIII* containing *Oami*; and the SPIE Digital Library website record for *Applications of Digital Image Processing XXVIII* containing *Oami*, to determine the date of public accessibility and availability for this reference.

35. **Appendix 4** to this Declaration is a true and accurate copy of the cover page, title page, verso and table of contents for the print issue of Proceedings of SPIE, Applications of Digital Image Processing XXVIII 2-4 August 2005 San Diego, California held by the Linda Hall Library-University of Missouri which includes *Oami*, Ryoma, Eiji Kasutani, and Akio Yamada. “Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones.” **Appendix 4** also sets forth the reference was received by Linda Hall Library 10-31-05, as indicated with a handwritten entry on page iii. In addition, the verso states the “e-First Publication for Proceedings: As of July 2005, Papers in the Proceedings of SPIE are first published electronically in the SPIE Digital Library (www.spiedl.org) and subsequently in print and CD-ROM”.

36. **Appendix 5** to this Declaration is a true and accurate copy of the Linda Hall Library public catalog record for its copy of *Applications of Digital Image Processing XXVIII* containing *Oami*, which was downloaded from https://catalog.lindahall.org/discovery/fulldisplay?docid=alma994591033405961&context=L&vid=01LINDAHALL_INST:LHL&lang=en&search_scope=MyInstitution&adaptor=Local%20Search%20Engine&tab=LibraryCatalog&query=title,contents,Applications%20of%20Digital%20Image%20Processing%20XXVIII,AND&mode=advanced&offset=0 on May 2, 2025.

37. The Linda Hall Library public catalog record, **Appendix 5** sets forth the publication, contents, and holdings information for its copy of *Applications of Digital Image Processing XXVIII* containing *Oami*. **Appendix 5** also denotes *Applications of Digital Image Processing XXVIII* containing *Oami* is located at TA1632 .A66 2005 in the Linda Hall Library.

38. **Appendix 6** contains the MARC record for its holdings of *Applications of Digital Image Processing XXVIII* containing *Oami* which was downloaded from https://catalog.lindahall.org/discovery/sourceRecord?vid=01LINDAHALL_INST:LHL&docId=alma994591033405961&recordOwner=01LINDAHALL_INST on May 2, 2025. **Appendix 6** confirms the fixed date elements of MARC field tag 008 as “051031s2005 wauab b 101 0 eng d”. As discussed above, the first six

characters “051031” are in typical “YYMMDD” format and indicate, *Applications of Digital Image Processing XXVIII* was cataloged by Linda Hall Library on October 31, 2005. The “s” refers to a single known date of publication of 2005.

39. MARC Field tag 020 sets forth the unique ISBN 0819459143. MARC field tag 245 sets forth the title of the publication as \$aApplications of digital image processing XXVIII : \$b2-4 August 2005, San Diego, California, USA / \$cAndrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union. MARC field tag 260 sets forth information relating to the place of publication, the publisher and the date of publication as \$aBellingham, Washington : \$bSPIE, \$cc2005.

40. Based on the information in **Appendices 5** and **6** attached to this Declaration, standard library practices and my experience as a librarian, *Applications of Digital Image Processing XXVIII* containing *Oami* was received by Linda Hall Library on 10-31-05, as indicated with a handwritten entry page iii.

41. Accordingly, based on the information set forth in **Appendices 5** and **6** attached to this Declaration, standard library practices and my experience as a librarian, *Applications of Digital Image Processing XXVIII* containing *Oami*, would have been made available to the public within a few days of being checked-in and cataloged on 10-31-05. The interested public could have been able to access

Applications of Digital Image Processing XXVIII containing *Oami* by (i) searching the Linda Hall Library public catalog, visiting the Library and searching the stacks at Call Number TA1632 .A66 2005; or (ii) by asking a Library staff member and being directed to the print resource located in the Library at Call Number TA1632 .A66 2005 within a few days of 10-31-2005.

XII. United States Copyright Office Record for *Oami*

42. Section XII of this Declaration relates to the date of publication for the United States Copyright Office record for *Applications of Digital Image Processing XXVIII: vol. 5909, 2-4 August 2005*.

43. **Appendix 7** to this Declaration is a true and accurate copy of the record for *Applications of Digital Image Processing XXVIII: vol. 5909, 2-4 August 2005* which was downloaded from https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?Search_Arg=Applications+of+Digital+Image+Processing+XXVIII+&Search_Code=TALL&PID=m02t1xrq0eIEmypEh73li8GT&SEQ=20250506121700&CNT=25&HIST=1 on May 2, 2025.

44. **Appendix 7** denotes a date of creation “2005” and a date of publication “2005-08-02” for the title *Applications of digital image processing XXVIII : vol. 5909, 2-4 August 2005*.

45. **Appendix 8** is a true and accurate Certificate of Registration from the United States Copyright Office displaying the Date of Publication of this Particular Issue as August 2, 2005.

46. In my professional opinion, **Appendices 7 and 8** provide the official United States Copyright Office record and Certificate of Registration for *Applications of Digital Image Processing XXVIII : vol. 5909, 2-4 August 2005* with a date of publication of 2005-08-02 to be accurate, true and correct.

XIII. SPIE Digital Library Record for Oami

47. Section XIII of this Declaration relates to the date of public accessibility for the following reference Oami, Ryoma, Eiji Kasutani, and Akio Yamada. “Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones” in *Applications of Digital Image Processing XXVIII*, edited by Andrew G. Tescher, Proc. SPIE Vol. 59090N (16 September 2005) retrieved from the SPIE Digital Library online platform.

48. The print version indicates e-First Publication for Proceedings as of July 2005 where papers in the Proceedings of SPIE are first published electronically in the SPIE Digital Library, <https://www.spiedl.org/> and subsequently in print.

49. On May 6, 2025, I searched Google Chrome to locate <https://www.spiedl.org/> where I was taken to the current website

<https://www.spiedigitallibrary.org/> attached here as **Appendix 9**. From **Appendix 9**, I selected “Papers”, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie> attached here as **Appendix 10**. From **Appendix 10** I selected “Browse by Volume Number”, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/browse/volume-number> attached here as **Appendix 11**, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/browse/volume-number> and continued my selection of “5000-5999”, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/browse/volume-number/5000-5999> attached here as **Appendix 12**. I further selected the range “5000- 5999”, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/browse/volume-number/5000-5999> attached here as **Appendix 13** further selecting the Volume 5909, Applications of Digital Image Processing XXVIII, <https://www.spiedigitallibrary.org/conference-proceedings-of-SPIE/5909.toc> attached here as **Appendix 14** and navigating to *Oami* attached here as **Appendix 15**, <https://www.spiedigitallibrary.org/conference-proceedings-of-spie/5909/59090N/Mobile-multimedia-library--an-MPEG-7-application-with-camera/10.1117/12.623829.short>.

50. **Appendix 15** sets forth *Oami* was added to the SPIE Digital Library on 16 September 2005. On May 6, 2025 I selected “Purchase Single Article” and was taken to a payment page to purchase a copy of *Oami*. I entered my details

including payment information for the cost of the reference, \$22.26. I next clicked “download pdf” which is attached to my Declaration as **Appendix 16**. I compared **Appendix 4**, a dated print version of *Oami* and **Appendix 16** and conclude they are the same text and figures except for formatting differences, they are the same document.

51. It is my professional opinion that members of the public could have accessed the conference proceeding titled *Applications of Digital Image Processing XXVIII* containing *Oami* by attending the conference in person August 2-4, 2005 or by visiting the SPIE Digital Library website using the steps above on or after 16 September 2005.

XIV. Public Availability of SPIE Digital Library Records

52. As detailed in the steps below, I searched the Internet Archive’s Wayback Machine and located the details where the e-First publication for Proceedings of SPIE are discussed.

53. I first navigated to the Internet Archive website at <https://archive.org/>. I next navigated to The Wayback Machine (which is prominently displayed on the homepage of the Internet Archive website) and clicked on its URL: <https://web.archive.org/>. In the search box, I entered the search term “spie.org” which took me to a page displaying the 32,239 URL’s captured for the URL prefix https://web.archive.org/web/20050715000000*/spie.org. I then navigated to the

year 2005 and began reviewing the July dates for details related to their publications. I then chose (clicked on) July 27, 2005 which was captured 3 times on the date. The first snapshot listed under July 27, 2005 detailed as 03:15:15. This took me to the following extended URL:

<https://web.archive.org/web/20050727031515/http://spie.org/> where I selected

“Publications” then “Proceedings” which details the e-First publication of the Proceedings of SPIE attached here as **Appendix 17**,

<https://web.archive.org/web/20050715090230/http://www.spie.org/app/Publications/index.cfm?fuseaction=proceedings>.

54. The Publications-Proceedings details provided at the archived website on July 15, 2005 states the following procedures: “Editor-reviewed papers from post-meeting Proceedings of SPIE will be available within two to four weeks after each symposium as PDF documents downloadable from the SPIE Digital Library (SDL). The new e-First publication model begins with meetings held after June 2005. Print volumes will be ready to ship by six weeks after the meeting, and collections of proceedings on CD-ROM by eight weeks after the meeting. Papers from on-site Proceedings will be available in the SDL at the start of the symposium. In the e-First model, the date the paper appears in the SDL is the official publication date for that paper.”

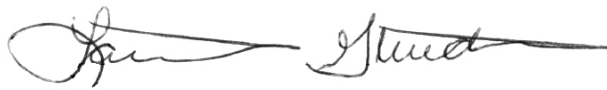
55. Based on the information in **Appendix 17** and my review of the details, members of the interested public could have accessed *Oami* in the SPIE Digital Library as detailed in Section XIII of this Declaration.

XV. Conclusion

56. In signing this Declaration, I understand it will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I understand I may be subject to cross-examination in this case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

57. I declare that all statements made herein of my knowledge are true, that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on June 6, 2025 in State College, Pennsylvania.

A handwritten signature in black ink, appearing to read 'Lauren Gluckman', with a long horizontal flourish extending to the right.

Lauren Gluckman

APPENDIX 1

Priority:**ASAP****Order #:**

1104908

Request Date:

05/02/2025 2:27pm

Contact:

Lauren Gluckman

Delivery Email:

lauren.gluckman@finnegan.com

WTS Staff Contact:

Christy Henrikson

Special Instructions:

Please provide a date stamped reference to include the cover page, title page and reference.

Citation:

2) Adding hyperlinks to digital television BOVE, V. M1 ; DAKSS, J1 ; AGAMANOLIS, S1 ; CHALOM, E1 SMPTE Technical Conference (140 ; Pasadena, CA 1998-10-28) SMPTE journal (1976). 1999, Vol 108, Num 11, pp 795-801 ; ref : 5 ref ISSN 0036-1682
<https://journal.smpete.org/periodicals/SMPTE%20Journal/108/11/>

Library/Supplier:

OCLC 230899865 (1 of 2) OCLC 230899866 (2 of 2) - considering (HOLD FOR CHARGES)



NOVEMBER 1999

SMPTE Journal

**SMPTE Conference & Exhibition
Nov. 19-22, New York City**

SHOW DIRECTORY:





FOUNDED 1916

EXECUTIVE

Robert B. Kisor
Editorial Vice-President

LeRoy E. DeMarsh
Chairman/Board of Editors

Frederick C. Motts
Executive Director

JOURNAL STAFF

Mathew V. Kuriakose
Production/Graphics Supervisor

MaryAnn Fruscianti
Production Editor

Dianne Ross Purrier
Assistant Editor

Marilyn Waldman
Advertising/Editorial/Program Coordinator

ENGINEERING

Carlos V. Girod, Jr.
Director of Engineering

Mark S. Hyman
Staff Engineer

MARKETING

Linda Alexander
Marketing Coordinator

Bryan Nella
Marketing Associate

MEETINGS

Ilisha Stier
Meeting Coordinator

Printed in the USA

SMPTE Journal®

VOL. 108, NO. 11

NOVEMBER 1999

MESSAGE FROM THE CONFERENCE VICE-PRESIDENT

Thomas M. Jordan 745

141ST SMPTE TECHNICAL CONFERENCE AND EXHIBITION

New York City, November 19-22, 1999 748

Conference Sponsors 748

Welcome Letter from Conference Vice-President 748

At Your Service 748

Conference Schedule Overview 750

General Information 754

Hotel Floor Plan 758

Papers Program 760

Honors and Awards 767

Exhibit Floor Plan 776

List of Exhibitors 777

34th Advanced Motion Imaging Conference
San Francisco, CA, February 3-5, 2000 789

NEWS & NOTES 793

TECHNICAL PAPERS

Adding Hyperlinks to Digital Television
V. M. Bove, Jr., J. Dakss, S. Agamanolis, and E. Chalom 795

Time-varying Image Quality: Modeling the Relation Between
Instantaneous and Overall Quality
Roelof Hamberg and Huib de Ridder 802

APPLICATIONS/PRACTICES PAPER

Optimization of DTV Signals Within a Transmission Plant
Brett Jenkins and Ray Kiesel 812

SMPTE NEWS AND DEPARTMENTS

Section Meetings 818

News 820

Obituaries 821

Calendar 821

New Products 822

Professional Services 823

Engineering News 824

Classified 825

Application for Membership 836

Membership Renewal 837

Sustaining Member List 838

STANDARDS AND RECOMMENDED PRACTICES

SMPTE Standards 826

SMPTE Journal (ISSN: 0036-1682) published monthly by the Society of Motion Picture and Television Engineers, 595 West Hartsdale Ave., White Plains, NY 10607. Periodicals postage paid at White Plains, N.Y., and at additional mailing office. Copyright © 1999 by the Society of Motion Picture and Television Engineers, Inc. All rights reserved. Printed in the U.S.A. No part of this issue may be reproduced without written permission of the publisher. Statements and opinions expressed in articles or editorials in the *SMPTE Journal* are expressions of contributors and do not necessarily represent the policies or opinions of the Society. Advertisements appearing in this publication are the sole responsibility of the advertiser. Subscription rates: SMPTE members, \$20.00 included in membership dues; nonmembers, U.S., \$125.00; elsewhere, \$140.00; airmail charge in addition to nonmember rates is \$65.00. Single copy, \$10.00 plus shipping. MEMBERSHIP in SMPTE is available according to qualifications at the following annual dues: Active Member, \$95.00; Student Member, \$25.00. All members receive the *SMPTE Journal* as part of membership.

POSTMASTER: Send change of address to *SMPTE Journal*, 595 West Hartsdale Ave., White Plains, NY 10607-1824.

Adding Hyperlinks to Digital Television

By V. M. Bove, Jr., J. Dakss, S. Agamanolis, and E. Chalom

Hyperlinked video is where objects are made selectable by some form of user interface, and the user's interactions with these objects modify the presentation of the video. Identifying and tracking the objects remains one of the chief difficulties in authoring hyperlinked video; we solve this problem through the use of a video tracking and segmentation algorithm that uses color, texture, motion, and position parameters. An author uses a computer mouse to scribble roughly on each desired object in a frame of video, and the system generates a segmentation mask for that frame and following frames. This technique was applied in the production of a soap opera program, with the result that the user can inquire about purchasing clothing and furnishings used in the show. This paper will discuss this and other uses of the technology, describe our experiences in using the segmentation algorithm for hyperlinked video purposes in both broadcast and on-demand modes, and present several different user-interface methods appropriate for hyperlinked video.

Users of the World Wide Web (WWW) have become accustomed to hyperlinks, whereby clicking on a word or graphic selects another page, or perhaps modifies the current one. The idea of hyperlinked video, for selecting objects, has often been discussed as a desirable possibility. For instance, consider a fashion program in which clicking on an article of apparel provides information about it, or a nature program that lets children go on safari and collect specimens. Playback of such material is well within the capabilities of typical digital television decoders with graphical overlay capability. However, creating it has posed a challenge because of the tediousness of identifying the clickable regions in every frame manually and the difficulty of segmenting and tracking them automatically.

A method, which simplifies the authoring problem, has been developed for tracking and segmenting objects. During editing, the author simply uses the mouse to scribble

roughly on each desired object in a frame of video. The system generates a filled-in segmentation mask for that frame and for preceding and following frames, until there is a scene change or the entrance of new objects. Regions can then have an action (e.g. graphical overlay, switching to a different video data stream, transmission of data on a back channel) associated with them.

The viewer can select objects by a variety of means, including ordinary remote controls or remote controls with inertial sensors. In our demonstrations, a video projector that identifies and locates a laser pointer aimed at its screen.

Our hyperlinked video system has two main technical underpinnings. First, a novel method of using color, texture, motion, and position for object segmentation and tracking of video. While the author is indicating regions of importance by scribbling with the mouse, our system uses a combination of these features to develop multimodal statistical models for each region. The system then creates a segmentation mask by finding entire regions that are statistically similar and tracking them throughout a video scene. The second piece is a scripting language for object-based media called Isis, which enables the

association of the mask with the video frames, and the association of actions with selecting regions at particular times.

This system was utilized to author a prototype hyperlinked video program that resembles an episode of a television serial drama. It enables the viewer to select props, clothing, and scenery and shows purchasing information for the item, including the price and retailer. The content of a program, named "HyperSoap," was produced believing that existing television material could not be adapted effectively to hyperlinked video. This entailed changes in the way "HyperSoap" was scripted, shot, and edited when compared to a traditional television program. A great deal was learned about how people interact with hyperlinked video, and the design of several modes of user interaction was based on this information.

Authoring Hyperlinked Video

Several companies have announced products for hyperlinked video authoring. VisualSHOCK MOVIE from the New Business Development Group of Mitsubishi Electric America, Inc., concentrates on web and local playback (e.g. CD-ROM, DVD-ROM). The authoring tool, the Movie MapEditor, requires specifying a rectangular bounding box around the location of a desired object in the starting and ending frames of a video sequence. A tracking algorithm estimates the position of the hot button in intermediate frames; manual tracking by the operator is also supported. HotVideo from IBM's Internet Media Group has a similar approach, but supports a larger set of geometrical object shapes. Intermediate locations of an object may be specified by multiple keyframes or located by a tracking algorithm. Another current authoring tool for incorporating an object-tracking algorithm is Veon's V-Active.

It was found that this method for

Presented at the 140th SMPTE Technical Conference (paper no. 140-51), in Pasadena, CA, October 28-31, 1998. V. Michael Bove, Jr., Jonathan Dakss, Stefan Agamanolis, and Edmond Chalom are with the MIT Media Lab, Cambridge, MA 02139. An unedited version of this paper appears in *DTV Reality Check—The First Year*, SMPTE, 1998. Copyright © 1999 by SMPTE.

associating links or actions to objects is too coarse for the objects found in many varieties of video sequences. To enable the association of links with arbitrarily shaped regions, a system is needed which can differentiate between objects with higher precision.

Segmenting Objects

Some automated segmentation method is necessary to enable the author to specify clickable regions that follow object outlines rather than simple geometrical shapes, and if we do not want to require manual outlining in every frame. For example, the VisualSEEK system¹ automatically identifies regions based on color and texture distribution. This and other automatic systems are limited by the ability of the system to generate an internal model that corresponds to actual objects as required by the author. Good quality is important for situations in which the labels the system generates for pixels (known as a segmentation mask) are rendered for the viewer, for example to indicate the clickable regions. It must be noted, of course, that playing back video with a detailed segmentation will require much more information to be transmitted or at least retained on the server than would be the case with geometrical regions that can be described in terms of their coordinates.

A novel segmentation system that classifies every pixel in every frame of a video sequence as a member of an object, or group of objects, as defined by an author, has been developed.^{2,3} The system enables the author to define objects by labeling representative clusters of pixels in a single frame. The pixels, which the author has classified, are then tracked through the remaining frames of the sequence. The system computes multiple features for all of the pixels in the sequence, including color, texture, motion, and position. For each frame, statistical models for each object are created based on the features of the classified pixels. These models are then used to classify each unlabeled pixel to the object with highest statistical similarity.

We have created a simple drawing tool that enables the author to indi-

cate regions of importance within a video sequence. The user selects a single representative frame from the sequence (typically one in which the important regions are fully visible) and uses the mouse to highlight representative pixels for each desired object. When the author has finished labeling pixels in the frame, the system estimates the location of these pixels within each of the remaining frames in the sequence using block matching. When this stage is complete, there are pixels in each frame, classified to correspond to objects. The next stage of processing will assign the rest of the pixels in the image to one of these objects.

Starting from a small subset of pixels that the author has indicated in one frame and classifying every pixel in every frame is done by calculating a multimodal statistical model based on a variety of different features including the color, motion, and position of a pixel and the texture and motion properties of a small neighborhood centered upon it. This yields a more robust segmentation, where the quality is minimally affected by misleading data from a single feature. Because a region's model is multimodal, it is possible to specify a region that has red and yellow pixels, and the system will not assign intermediate shades of orange to that object.

After calculating a multidimensional feature vector for every pixel in the video sequence, we use the feature information for the tracked pixels to form a statistical model for each user-identified region. We then compare the feature information of each unclassified pixel to these models and label the pixel as a member of the region to which it bears the greatest statistical similarity. Basically, feature information is used to express each region as a mathematical function so that we can input feature values of unclassified pixels and receive as output the probability that the pixel corresponds to that region.

Because only statistical methods are employed to classify pixels, it is likely that there will be small aberrations in otherwise consistent classification, such as a small group of mislabeled pixels surrounded by properly classified pixels. We use two algo-

rithms during a post-processing stage to help rectify these aberrations. To help remove small groups of misclassified pixels, a process is used that unclassifies each pixel that was originally classified with a certainty below a set threshold. A second algorithm, known as K-nearest neighbor (KNN), reassigns the unclassified pixels to the region that is most popular among its neighboring pixels.

Because objects in video are likely to reoccur in several shots and the system is run on shots individually, at this stage the author may wish to adjust region labels, which have been set locally to values representing a region or object globally. This could also mean adjusting mask values to a special number (i.e., 0), which corresponds to regions where no action or link is desired.

At this stage the system has generated a second video sequence containing the segmentation information for every frame of the input sequence.

Associating Actions with Objects

If pixel-level identification of objects is possible (as it is with our system), a person can interact with the video during playback by indicating one or more pixels comprising a particular object; the playback system can then cause a corresponding event or action to occur. This event could be the display of additional information about the object, such as the rendering of text or an image, or perhaps retrieving information from a URL on the web. An event could also be a change in the way the video is presented, such as showing scenes from alternate camera angles or playing back a different part of the video, which is relevant to the selected object. For example, the hyperlinked video HyperCafe⁴ allows a person to navigate through various video clips of conversations occurring in a cafe, by selecting actors in the video or icons appearing at certain times during the presentation.

One possible implementation for enabling this association is to provide a database in which object labels can be cross-referenced with descriptions of the objects (if one of the desired actions is to render object information) or instructions for executing

actions. Many of the existing hyperlinked video systems utilize this infrastructure. For example, VisualSHOCK maintains an anchor list, accessible by a drop-down menu system, where data structures containing link information are stored and referenced. The playback system for "HyperSoap," built in a scripting environment called Isis,⁵ maintains an internal database consisting of a list of lists, in which each corresponds to a particular object and contains text strings, including the name of the object, as well as actions associated with it.

Interacting with Hyperlinked Video

In this section, we describe various issues confronting us when we designed our playback system. Important user interface issues, including methods in which the presence of links are indicated in video, mechanisms used by the viewer to select objects, and approaches to initiating link events are described in detail. In many cases, our solutions to these issues vary greatly from existing implementations.

One reason for this difference is that our attention focused on user interaction mechanisms apart from the desktop PC. Therefore, we concerned ourselves with the television paradigm. In some ways, using a mouse on a PC may be natural for hyperlinked video, considering that people are used to using a mouse to activate hyperlinks on the web. However, the desktop is not an ideal place to view video; additionally, many of the genres of content suitable for hyperlinked video are programs which people are accustomed to watching in their living rooms.

We designed our system with the intent of simulating a future television-viewing scenario, one in which a device with significant computational ability (e.g. a set-top box, enhanced television receiver, or DVD player) could render video content while receiving directions about playback from both the viewer and the content itself. For our implementation, we used a projection screen and a video projector, which displayed video data transmitted by a workstation running Isis, a scripting environment used to

create our playback software.

When displaying hyperlinked video, it is important to indicate the presence of hyperlinks to viewers in a manner which does not distract the viewer from the content of the video. This also enables a scenario in which the viewer chooses to watch the presentation in a passive manner, as one might read a document on the web without choosing to click on hyperlinked words. Ideally, viewers might be able to adjust the manner in which hyperlink opportunities are presented to suit their own preferences.

Many of the existing hyperlinked video playback tools utilize a common and minimally distracting method for presenting hyperlink opportunities to the viewer. Since the viewer will be interacting with the video using a pointing device (a mouse or its equivalent), when the viewer moves the cursor over a hyperlinked region, the cursor changes shape. At the same time, a name or description for the link is displayed in an area of the screen. IBM's HotVideo software goes a step further and changes the cursor to an icon, which corresponds to the type of information contained in a link, such as a small document if the link is text or a film reel if the object links to another video stream. HotVideo also displays an indicator at all times, which changes color to indicate when there are hyperlink opportunities within the frames of video that are being shown to the viewer. This method is also used by WebTV's WebPIP, as there is always only one hyperlink opportunity in a video segment.

Other approaches can be more distracting. Whenever a frame contains hyperlinked objects, HotVideo can also render the wireframe shapes that the author used to indicate the regions. This method is suitable when there are only a small number of objects present in each frame of video, but can cause the video to become quite cluttered for higher numbers of objects. It can also be confusing to the viewer if the linked object occupies a large portion of the video frame and may surround other objects, such as a wall in an interior shot of a house. Another potentially distracting display method, which a

HotVideo author can choose, is to render the pixels within the wireframe shape in a different manner than the ones outside the shape, such as adjusting the brightness or color tint of those pixels. In favor of having the author determine this method per link, many of these systems do not allow the viewer to choose the method in which all links are indicated.

Since our method allows for identifying objects at a pixel level, the presence of hyperlinks is shown by highlighting all of the pixels that comprise a selected object. This highlighting effect is achieved by using Isis' ability to overlay an image buffer with a video stream in real-time. During playback, there is an image buffer whose pixels are all set to a solid color. When the viewer selects a pixel corresponding to a hyperlinked object, the playback script retrieves the segmentation mask for the next frame and sets the alpha channel of the pixels in the image buffer to correspond to the selected object's pixels in the segmentation mask. In this manner, the only pixels in the image buffer that will be rendered are the ones corresponding to the selected object. Then, when the next frame of video is ready to be shown, Isis composites this frame of video with the image buffer (keeping the image buffer slightly transparent) thus indicating the presence of a hyperlink. If the viewer continues to select pixels from the same object, the playback script updates the alpha channel of the pixels in the image buffer. When the user stops selecting pixels, the script continues to update and render the image buffer but adjusts the transparency of the buffer over time. This gives the viewer the impression that the link is "fading" out of the picture. We have also implemented two of these buffers simultaneously, allowing the system to show the presence of another object while the previously selected object may still be fading.

The means by which the viewer of hyperlinked video selects links should be simple to use and feel as natural to the viewer as possible. Since people are comfortable using a remote control device to browse television programming, it makes sense

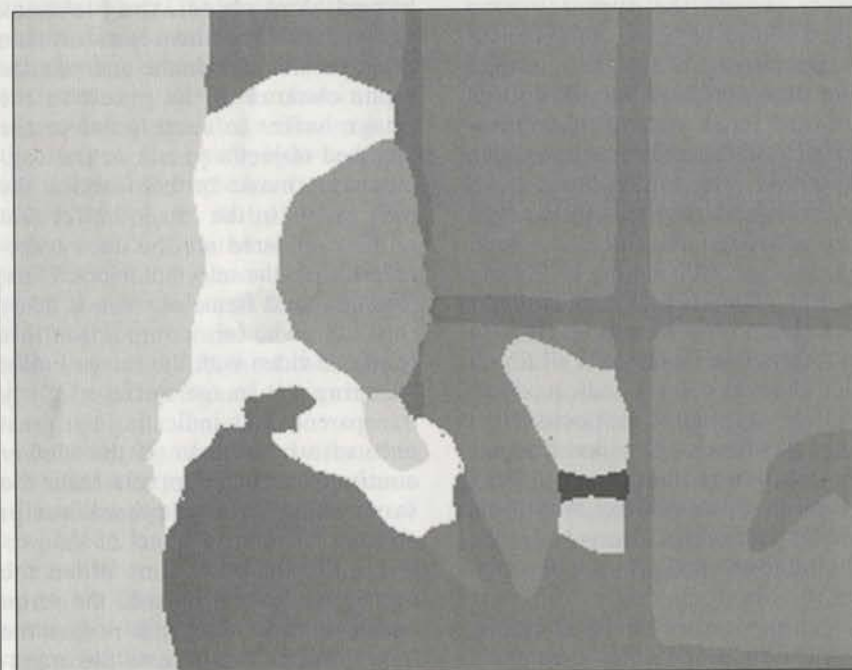


Figure 1. (Top) A video frame with training points overlaid on it; (bottom) the resulting segmentation.

for them to use this device to browse through hyperlinks in that content. WebTV provides buttons on the remote that can be pressed when the icon appears, indicating the presence of a link for the present segment. The viewer can then choose to view the web page content referenced by the link or archive the link for later viewing. When several links are present, we envision a button on the remote

control, which would allow the viewer to cycle through the hyperlinked objects on the screen while a graphical overlay would be provided to indicate the location of each object as it is selected during the cycle. The remote might also incorporate a position touch pad, a joystick, or an inertial sensor.

For our demonstrations, we used a handheld laser pointer as the interac-

tion mechanism and set up a projection device to project frames of video and track the position of the laser on the screen. To do this, we added a small digital camera to a regular video projector and built the functionality into the Isis playback script to use the camera's output. Since the laser pointer forms a red dot with significantly greater intensity than any of the red pixels of the projected video, the system looks at only the red channel of the camera's output and applies a bounding-box algorithm to determine the center of the laser spot. We then apply a linear transformation to correct for possible Keystone distortion and compute a calibrated (x,y) position for the laser in units of screen pixels.

Since the viewer may have no prior knowledge of which objects in the video are hyperlinked, there must be some method to allow the user to browse through objects shown on the screen before selecting one. To achieve this with our laser pointer interface, the playback script tracks the number of frames in which the viewer has pointed to the same object. If the viewer selects the object for more than a set number of frames, the script assumes the viewer intended to select that object. To show the user that the object has been selected, the pixel values in the image buffer for highlighting are set to a solid color, thus modifying the appearance of the highlighting.

The temporal nature of video raises an interesting issue regarding the display and timing of events associated with embedded hyperlinks. Whereas hyperlinked documents on the web are stationary and can be read in any desired fashion, there is a flow to video content generally not controlled by the viewer. While all of the existing systems direct web browsers to video links immediately, we have found that displaying lengthy or detailed information while the video continues to run may cause the viewer to feel distracted from the video content, reading or examining an image simultaneously. Likewise, a delay in showing the link information may blur the context of the information so that it is no longer relevant to the video presentation. Thus, events triggered by selecting hyperlinks

must be presented in a way which maintains the continuity of the video while still presenting the desired link information to the viewer in a timely and meaningful fashion.

Additionally, the system architecture and the capabilities of the viewer's equipment may affect the degrees of freedom available. If the video is playing on-demand from a server, or locally from a DVD, random access and pausing are possible, and it is not necessary to do much caching of content. The segmentation mask does not necessarily need to be transmitted, or could be transmitted only when the user is actually pointing at something on the screen. In a broadcast environment, different content design might be needed since pausing is not possible; also the display device needs the segmentation mask at all times. We've experimented with strategies appropriate for each case. In an early experiment, selecting a link caused the video to be interrupted immediately to show the viewer the link content. After a set period of time, the link content would disappear and the video would resume from the last frame shown before the interruption. While some users enjoyed the instant response from the system, others found it somewhat troubling as their interactions would interrupt important moments within the content, such as cutting an actor off while they were delivering a line. Also, people tended to forget the context of the video content before the interruption and found themselves somewhat lost in the plot of the video when it resumed.

Our next implementation yielded a more pleasant manner of interaction. Once a viewer selected a hyperlink, the script would wait until the video reached a natural breakpoint (for instance, after an actor finishes speaking a set of lines) before interrupting and showing the link's information. This yielded a much more favorable reaction from users of the system, although the initial reaction of many first time users waiting for information was that the playback mechanism was not working properly. Others who felt more interested in the link information than in the video content expressed that they would have preferred instant feedback.

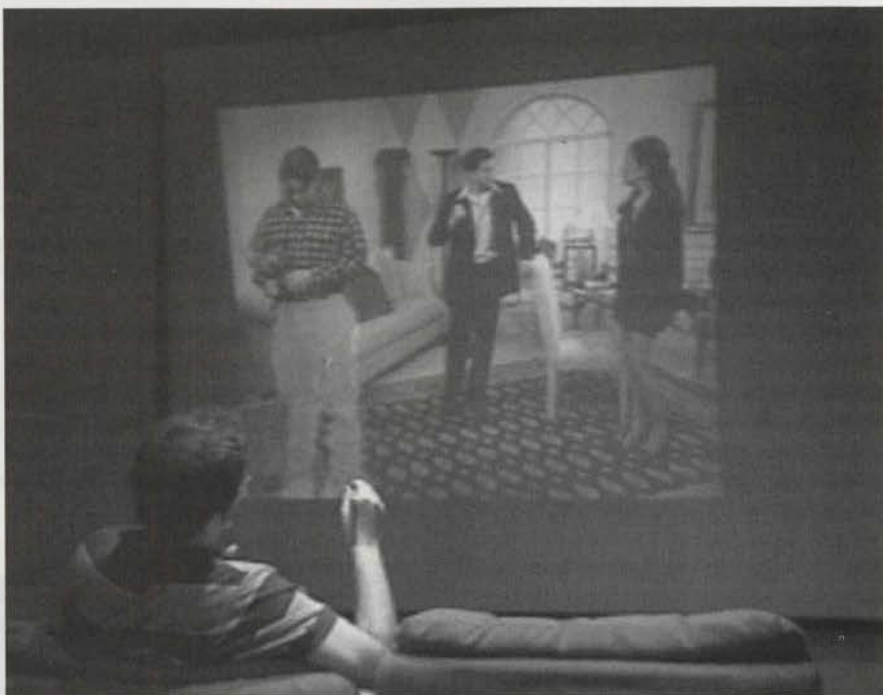


Figure 2. A viewer indicates a desired object by pointing at it with a special remote control.

In addition to this mode of interaction, we decided to implement two additional modes for our system and gave the user the capability to toggle between all three modes during playback. The first additional mode overlaid an abridged version of the link information while the video continued to run. Based on the area of the screen where the user selected an object, the system would try to render the information so that it would be less likely to occlude other hyperlinked objects. This could be combined with a feature similar to WebTV's, where each link was archived to enable the viewer to see the complete information after the end of the program. (Fig. 1.)

"HyperSoap"

We used our authoring and playback systems to create "HyperSoap," a four-minute prototype hyperlinked video drama in which the content was closely patterned after daytime soap operas. We oversaw every aspect of the production of "HyperSoap" content, including scriptwriting, storyboarding and directing the shoot, and post-processing the video and audio. Every item on the set, including clothes, props, and furnishings, was taken from the JCPenney catalog and was for sale. Pointing the laser

remote at the screen highlighted selectable objects and keeping the pointer fixed on one object for an extended period selected it. In instant-feedback mode (simulating a broadcast), a small text window would be displayed on the screen containing the item's description, brand name, price, and a logo for the distributor. In breakpoint mode, more descriptive text boxes would be shown, as well as a still image resembling a photo of the item from a catalog. This same information could also be rendered at the end of the scene. (Figs. 2 and 3.)

In breakpoint mode, we used music to help maintain continuity between the video and the hyperlink information. A soundtrack was composed for "HyperSoap" with the unique quality that it could continue despite interruptions and still help to underscore the action being played out in the video. This was achieved by having the composer write pieces of music for specific parts of the scene, which could be seamlessly looped by the playback script. When the scene changed from one part to the next, the music for the previous part would fade out while the new one began. As a result, when the user selected a link, the background music for the scene would continue through the interrup-

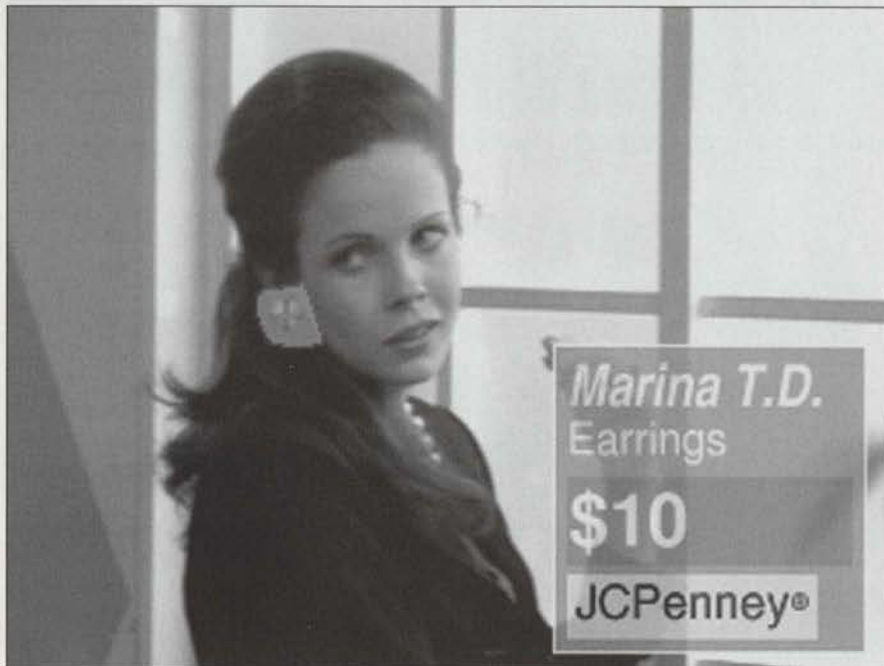


Figure 3. Information called up by selecting an object.

tion. Many users commented that the music helped to bridge the interruptions of the program content when objects were selected.

Our authoring system produced enough pixel-level segmentation of the 40 objects appearing in "HyperSoap" to permit the use of the segmentation mask for highlighting. A total of 45 shots, where the number of linkable objects per shot ranged from 10 to 20, were needed to be run through the system individually. We found that to maintain consistently high-quality segmentation, we needed to retrain the system after approximately 30 frames, or approximately one second of video. This was largely done to prevent the propagation of errors during the tracking of the author-specified pixels; however, it still provided an exceptional level of automation of the segmentation process. Typically, an average of 5,000 pixels were classified in each training frame by the author, so that for each 30-frame sequence the authoring system was required to classify 99.8% of the total pixels. Additionally, we found we were able to reduce the amount of data in the segmentation mask by subsampling the data vertically and horizontally by a factor of two. This had a minimal effect on the quality of the upsampled

mask when it was rendered.

"HyperSoap" raised several interesting issues regarding the differences between traditional television programming and hyperlinked video content. A traditional television soap opera is designed in such a way that simply adding hyperlinks and allowing interruptions would likely detract from its effectiveness because they would obstruct the flow of one of the most important components of this format, the story.

In designing the content for "HyperSoap," we had to move away from a form where the story is central and any product placement or viewer interaction are decorations to a new form where all of these components are equally important and dependent on each other for the survival of the presentation. The knowledge that everything would be for sale and viewers would be interacting often was part of the program's design from the very beginning. This entailed several changes in the way the scene was scripted, shot, and edited when compared to a traditional television program.

For example, shots were designed to emphasize the products for sale while not drawing too much attention away from the flow of the story. This often required close-ups of certain

objects that would have otherwise been difficult to see or select with our interaction device; actions needed to be written in the script so these close-ups would not seem unnatural. Similarly, the scene was edited so that shots were either long enough or returned often enough for viewers to spot products and decide to select them. Also, the story itself involved the products in ways that might increase their value to the consumer. Integrating the story with the products made interruptions to show product information less jarring overall.

Conclusion

We found knowledge of all the objects in "HyperSoap" were for sale and that the program was designed for the dual purpose of entertainment and shopping motivated viewers to stay tuned and to interact with the products. As an additional way to motivate viewers to select products, we also implemented another version of "HyperSoap" in which selecting objects revealed details about the plot and characters in the drama which pertained to those items. We feel that both of these types of programming would create an intriguing form of interaction, were they to be implemented in a broadcast model. Other potential types of content likely to benefit from "hyper" capabilities include how-to videos, particularly those involving complicated components, such as repairing an automobile or learning to fly an airplane. Likewise, training videos that demonstrate complex processes where several people perform tasks simultaneously (for example, a surgical procedure or a football play) could allow viewers with different backgrounds to learn more about the tasks or people whose purpose particularly interests them. This same idea could be applied to documentaries and educational programming, for example enabling children to learn more about particular aspects of nature by selecting animals and plants observed in safari footage, as if they were collecting specimens on a real safari.

Although we used the Isis language for our production, it is quite possible to use standard tools (such as Java). While the computational demands of

segmentation are quite high, display is well within the capabilities of a typical PC or a set-top box with graphical overlay capabilities.

Acknowledgments

The authors would like to thank Michael Ponder, JCPenney, and the JCPenney Headquarters Production Staff for their invaluable assistance with the "HyperSoap" shoot. We would also like to thank Kevin Brooks, Paul Nemirovsky, and Alex Westner for providing their script-

writing, scoring, and Foley sound expertise. Also, a special thank-you to Julia Ann Lewis, Matthew Rupe, and Andrew T. Chandler for their tremendous acting talents. This work has been supported by the Digital Life Consortium at the MIT Media Laboratory.

References

1. J. R. Smith and S. F. Chang, "Local Color and Texture Extraction in Spatial Query," *IEEE Proc. Int. Conf. Image Processing*, III-1011-III-1016, 1996.
2. E. Chalom and V. M. Bove, Jr., "Segmentation of an Image Sequence Using Multi-Dimensional Image Attributes," *IEEE Proc. Int. Conf. Image Processing*, II-525-II-528, 1996.
3. E. Chalom, "Statistical Image Sequence Segmentation Using Multidimensional Attributes," Ph.D. thesis, MIT, 1998.
4. N. Sawhney, D. Balcom, and I. Smith, "Authoring and Navigating Video in Space and Time," *IEEE Multimedia*, 4:30-39, Oct. 1997.
5. S. Agamanolis and V. M. Bove, Jr., "Multilevel Scripting for Responsive Multimedia," *IEEE Multimedia*, 4:40-49, Oct. 1997.

THE AUTHORS



V. Michael Bove



Jonathan Dakss



Stefan Agamanolis



Edmond Chalom

V. Michael Bove, Jr. holds an S.B.E.E., an S.M. in visual studies, and a Ph.D. in mediatechnology, all from the Massachusetts Institute of Technology, where he is currently head of the Media Laboratory's Object-Based Media group. He is the author or co-author of over 40 journal or conference papers on digital television systems, video processing hardware/software design, multimedia, scene modeling, and optics.

Bove holds patents on inventions relating to video recording, hard-copy, and medical imaging, and is technical advisor to WatchPoint Media, Inc. In December 1995, Boston Magazine named him one of the "People Shaping Boston's High-Tech Future."

Bove is on the Board of Editors of the *SMPTE Journal*, served as general chair of the 1996 ACM multimedia conference, and is listed in the current edition of *Who's Who in Entertainment*.

Jonathan Dakss received a B.A. degree in computer science from Columbia University in 1997. He worked at Columbia's Center for Telecommunications Research and aided in the development of algorithms for a wavelet-based image and video compression and segmentation system now in use by the U.S. Air Force, Rome Laboratories.

Dakss received a master's degree in media technology at the MIT Media Laboratory in 1999, where he developed authoring tools for hyperlinked video and produced "Hyper-Soap" and an interactive cooking program that featured Julia Child.

Dakss co-founded WatchPoint Media, Inc., in 1999, which is commercializing the hyperlinking technology he developed at the Media Lab and focusing on its applications for e-commerce.

Stefan Agamanolis is a doctoral candidate at the MIT Media Laboratory where he has worked on a wide vari-

ety of research projects concerning telepresence and telecollaboration, hyperlinked video, interactive storytelling, responsive environments, ambient media, and intelligent authoring tools for multimedia applications. His experimental Isis programming language serves as the foundation for the hyperlinked video prototypes described in this article.

Before coming to the lab, Agamanolis studied computer science, philosophy, and film at Oberlin College.

Edmond Chalom received B.S., M.S., E.E., and Ph.D. degrees from the Massachusetts Institute of Technology. His Ph.D. research was performed at the MIT Media Laboratory. He is currently an MTS at Sarnoff Corp. in the Multimedia Technology Laboratory, specializing in image compression, image analysis, and pattern recognition.

Agamanolis has been a member of Eta Kappa Nu since 1987.

APPENDIX 2

APPENDIX 3

leader 02995cas a2200697 a 4500
001 992679233405961
005 20230906180739.0
008 760405c1976999nyumrlp 0 a0eng d
010 ##\$a 82645600 //r08 \$zsc 79005068 \$zsn 78004288
012 ##\$a3 \$b3 \$es \$jo \$ki \$li \$ml
019 ##\$a2384603
022 0##\$a0036-1682
030 ##\$a\$MPJDF
032 ##\$a285200 \$bUSPS
035 ##\$a(OCOLC)2093452
035 ##\$a(Horizon)370080
035 ##\$a(MoKl)267923-lhalldb
035 ##\$a(lhalldb)267923-lhalldb
035 ##\$a(OCOLC)2384603
037 ##\$a\$MPTE, 862 Scarsdale Ave., Scarsdale, NY 10583
037 ##\$a\$45.00 (US) \$c\$17.00 (students, US) \$c\$53.00 (foreign) \$c\$25.00 (students, foreign)
040 ##\$aNSD \$cNSD \$dOCL \$diUL \$dOCL \$dDLC \$dNST \$dDLC \$dSER \$dNST \$dRCS \$dAIP \$dDLC \$dAIP \$dAGL \$dAIP \$dNST \$diUL \$dNST \$dNSD \$dNST \$dDLC \$dNST \$dCOO \$dNST \$beng
042 ##\$ansdp \$alc
049 ##\$aLHLA
050 00\$aTR845 \$b.56
069 1##\$a527900000
070 ##\$a332.9 \$bSol
072 ##\$aX600
082 0##\$a778.5/05 \$219
210 0##\$a\$MPTE j. \$b(1976)
222 00\$a\$MPTE journal \$b(1976)
245 00\$a\$MPTE journal : \$bpublication of the Society of Motion Picture and Television Engineers.
246 10\$aS.M.P.T.E. journal
246 23\$aSociety of Motion Picture and Television Engineers journal
260 ##\$a[Scarsdale, NY] : \$bThe Society,
310 ##\$aMonthly
362 1##\$aBegan with Jan. 1976.
510 1##\$aapplied science & technology index, \$x0003-6986
510 1##\$aExcerpta medica,
510 1##\$aFilm literature index, \$x0093-6758
510 2##\$aChemical abstracts, \$x0009-2258 \$b1976-1980
510 2##\$aComputer & control abstracts, \$x0036-8113 \$bJan. 1976-
510 2##\$aElectrical & electronics abstracts, \$x0036-8105 \$bJan. 1976-
510 2##\$aEngineering index annual (1968), \$x0360-8557
510 2##\$aEngineering index bioengineering abstracts, \$x0736-6213
510 2##\$aEngineering index energy abstracts, \$x0093-8408
510 2##\$aEngineering index monthly (1984), \$x0742-1974
510 2##\$aInternational aerospace abstracts, \$x0020-5842
510 2##\$aPhysics abstracts. Science abstracts. Series A, \$x0036-8091 \$bJan. 1976-
510 2##\$aReference sources, \$x0163-3546
525 ##\$av91(1982)(?)- incl. a Directory for members.
555 ##\$av90-94(1981-85) publ. as v95 n1, pt. 2; v95-99(1986-90) publ. as v100 n1, pt 2; v100-104(1991-95) publ. as v105 n1, pt 2.1; 105-109(1996-2000) in 110, no. 12.
650 00\$aCinematography \$vPeriodicals
650 00\$aTelevision \$vPeriodicals
710 2##\$aSociety of Motion Picture and Television Engineers.
780 00\$tJournal of the SMPTE \$x0361-4573 \$w(DLC) 90642424 \$w(OCOLC)7000262
785 00\$a\$MPTE motion imaging journal \$w(OCOLC)50872583
850 ##\$aUl \$aCU-I \$aCU-Riv \$aCU-S \$aCU-SB \$aCaAEU \$aCaBVIPA \$aCaOLU \$aCoU \$aDGN \$aDLC \$aDNAL \$aFU \$aINS \$aIaAS \$aInLP \$aInLu \$aMMeT \$aMID \$aMIU \$aNIC \$aNN \$aNRIC \$aNSyU \$aNcD \$aNcRS \$aNCISW \$aPU \$aViB1bv \$aViu \$aWau \$aWn
890 ##\$a\$MPTE journal.
901 ##\$cSer
948 ##\$aLTI 09/01/2008

APPENDIX 4

Priority:**ASAP****Order #:**

1104908

Request Date:

05/02/2025 2:27pm

Contact:

Lauren Gluckman

Delivery Email:

lauren.gluckman@finnegan.com

WTS Staff Contact:

Christy Henrikson

Special Instructions:

Please provide a date stamped reference to include the cover page, title page and reference.

Citation:

1) Ryoma Oami, Eiji Kasutani, and Akio Yamada "Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones", Proc. SPIE 5909, Applications of Digital Image Processing XXVIII, 59090N (16 September 2005);
<https://doi.org/10.1117/12.623829>
<https://www.spiedigitallibrary.org/conference-proceedings-of-spie/5909/59090N/Mobile-multimedia-library--an-MPEG-7-application-with-camera/10.1117/12.623829.short>

Library/Supplier:

OCLC 230898443 (1 of 2) OCLC 230898444 (2 of 2) TXA - NOS/Missing

Applications of Digital Image Processing XXVIII

Andrew G. Tescher
Chair/Editor

2-4 August 2005
San Diego, California, USA



Volume 5909



The International Society
for Optical Engineering

PROCEEDINGS OF SPIE

Applications of Digital Image Processing XXVIII

Andrew G. Tescher
Chair/Editor

2-4 August 2005
San Diego, California, USA

Sponsored and Published by
SPIE—The International Society for Optical Engineering

Organized with the Support of
The International Telecommunications Union

Volume 5909



**The International Society
for Optical Engineering**

SPIE is an international technical society dedicated to advancing engineering and scientific applications of optical, photonic, imaging, electronic, and optoelectronic technologies.

TA1632

.A66

2005

The papers included in this volume were part of the technical conference cited on the cover and title page. Papers were selected and subject to review by the editors and conference program committee. Some conference presentations may not be available for publication. The papers published in these proceedings reflect the work and thoughts of the authors and are published herein as submitted. The publisher is not responsible for the validity of the information or for any outcomes resulting from reliance thereon.

e-First Publication for Proceedings: As of July 2005, papers in the Proceedings of SPIE are first published electronically in the SPIE Digital Library (www.spiedl.org), and subsequently in print and CD-ROM.

Pagination system: The new citation format uses six-digit article identifier numbers (CIDs). Utilization of article CIDs allows proceedings articles to be fully citable as soon as they are published online, using the same identifier for both online and print versions.

The structure of the six-digit article CID number for Proceedings of SPIE is:

- The first four digits indicate the SPIE volume number.
- The last two digits indicate publication order within the volume using a Base 36 numbering system employing both numerals and letters. The CID number appears on each page, followed by the page number within the actual article.

Please use the following format to cite material from this book:

Author(s). "Title of Paper," in *Applications of Digital Image Processing XXVIII*, edited by Andrew G. Tescher, Proc. SPIE Vol. 5909, Article CID Number (2005).

ISSN 0277-786X

ISBN 0-8194-5914-3

Published by

SPIE—The International Society for Optical Engineering

P.O. Box 10, Bellingham, Washington 98227-0010 USA

Telephone 1 360/676-3290 (Pacific Time) · Fax 1 360/647-1445

<http://www.spie.org>

Copyright © 2005, The Society of Photo-Optical Instrumentation Engineers

Copying of material in this book for internal or personal use, or for the internal or personal use of specific clients, beyond the fair use provisions granted by the U.S. Copyright Law is authorized by SPIE subject to payment of copying fees. The Transactional Reporting Service base fee for this volume is \$15.00 per article (or portion thereof), which should be paid directly to the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Payment may also be made electronically through CCC Online at <http://www.copyright.com>. Other copying for republication, resale, advertising or promotion, or any form of systematic or multiple reproduction of any material in this book is prohibited except with permission in writing from the publisher. The CCC fee code is 0277-786X/05/\$15.00.

Printed in the United States of America.

Contents

ix Conference Committee

SESSION 1 IMAGE MODELS AND PROCESSING

- 590901 **Automatic inspection of pavement cracking distress** [5909-01]
B. Xu, Y. Huang, Univ. of Texas at Austin (USA)
- 590902 **Real-time automatic target recognition and identification of ground vehicles for airborne optronic systems** [5909-02]
O. Ruch, J. Dufour, Thales Optronique SA (France)
- 590903 **Hausdorff probabilistic feature analysis in SAR image recognition** [5909-03]
J. A. Saghi, C. Guilas, California Polytechnic State Univ. (USA)
- 590904 **Adaptive noise filtering of white-light confocal microscope images using Karhunen-Loève expansion** [5909-04]
M. Balasubramanian, S. S. Iyengar, P. Wolenski, Louisiana State Univ. (USA); J. Reynaud, R. W. Beuerman, Louisiana State Univ. Health Sciences Ctr./New Orleans (USA)
- 590905 **A detection method of mura on a coated layer using interference light** [5909-05]
K. Taniguchi, K. Ueta, Dainippon Screen Mfg. Co., Ltd. (Japan); S. Tatsumi, Osaka City Univ. (Japan)
- 590906 **Quadratic correlation filter-based target tracking in FLIR image sequences** [5909-06]
A. Bal, M. S. Alam, Univ. of South Alabama (USA)
- 590907 **Minimax distance transform correlation filter-based target detection in FLIR imagery** [5909-07]
J. F. Khan, Univ. of Alabama in Huntsville (USA); M. S. Alam, Univ. of South Alabama (USA); R. R. Adhami, S. M. A. Bhuiyan, Univ. of Alabama in Huntsville (USA)
- 590908 **A new two-level efficient technique for 2D image patches registration via optimized cross correlation, wavelet transform, and moments of inertia with application to ladar imaging** [5909-08]
C. Bejar, D. B. Megherbi, Univ. of Massachusetts Lowell (USA)
- 590909 **Image fusion using DFT, based digital filter banks** [5909-09]
M. Sezgin, I. Erer, Istanbul Technical Univ. (Turkey); O. K. Ersoy, Purdue Univ. (USA)

SESSION 2 NEXT-GENERATION MULTIMEDIA COMPRESSION, TRANSMISSION, AND ARCHITECTURES

- 59090A **ISMuS: interactive, scalable, multimedia streaming platform** [5909-10]
J. Cha, H.-C. Kim, S. Jeong, K. Kim, Electronics and Telecommunications Research Institute (South Korea); C. Patrikakis, National Technical Univ. of Athens (Greece); M. van der Schaar, Univ. of California, Davis (USA)

- 59090B **Joint wavelet packets for groups of frames in MCTF** [5909-11]
M. Trocan, C. Tillier, B. Pesquet-Popescu, GET-ENST (France)
- 59090C **MPEG-21 DIA-based video adaptation framework and its application to rate adaptation** [5909-12]
J. W. Kang, J.-G. Kim, D.-S. Jun, J.-W. Hong, Electronics and Telecommunications Research Institute (South Korea)
- 59090E **An effective MP4 streaming method using the schedule information of image objects** [5909-14]
S. Jeong, J. Cha, H.-C. Kim, K. Kim, Electronics and Telecommunications Research Institute (South Korea)
- 59090F **Performance of scalable video coding with respect to differentiated service requirements** [5909-15]
D. Y. Suh, G. H. Park, S. G. Han, Kyunghee Univ. (South Korea); H.-C. Kim, Electronics and Telecommunications Research Institute (South Korea)
- 59090G **Network support mechanisms for scalable media streaming** [5909-16]
C. Z. Patrikakis, P. Karamolegkos, I. Mihailaris, A. Lampiris, National Technical Univ. of Athens (Greece); H.-C. Kim, Electronics and Telecommunications Research Institute (South Korea)

SESSION 3 IMAGING SECURITY

- 59090H **Study and validation of tools interoperability in the JPSEC framework** [5909-17]
V. Conan, Y. Sadourny, K. Jean-Marie, Thales Communications (France); C. Chan, S. Wee, J. Apostolopoulos, Hewlett-Packard Labs. (USA)
- 59090I **Image replica detection based on support vector classifier** [5909-18]
Y. Maret, F. Dufaux, T. Ebrahimi, Ecole Polytechnique Fédérale de Lausanne (Switzerland)
- 59090J **Supporting secure transcoding in JPSEC** [5909-19]
J. G. Apostolopoulos, S. J. Wee, Hewlett-Packard Labs. (USA)
- 59090K **Scrambling for anonymous visual communications** [5909-20]
F. Dufaux, T. Ebrahimi, Emitall S.A. (Switzerland) and Ecole Polytechnique Fédérale de Lausanne (Switzerland)

SESSION 4 MPEG-7: THE POWER OF STANDARD MULTIMEDIA METADATA AND CONTENT REPRESENTATION

- 59090L **MPEG-7: standard metadata for multimedia content** [5909-21]
W. Chang, National Institute of Standards and Technology (USA)
- 59090M **Easy and effective indexing and browsing schemes for digital photo album application** [5909-22]
S.-K. Kim, Samsung Advanced Institute of Technology (South Korea); S. Yang, K. Seo, Y. M. Ro, Information and Communications Univ. (South Korea); J.-Y. Kim, Samsung Advanced Institute of Technology (South Korea)

- 59090N **Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones** [5909-23]
R. Oami, E. Kasutani, A. Yamada, NEC Corp. (Japan)
- 59090O **Characterizing region of interest in image using MPEG-7 visual descriptors** [5909-24]
M.-S. Ryu, Dongguk Univ. (South Korea); S.-J. Park, Electronics and Telecommunications Research Institute (South Korea); C. S. Won, Dongguk Univ. (South Korea)
- 59090P **Overall framework on digital image searching** [5909-25]
M.-K. Leong, J.-H. Lim, Institute for Infocomm Research (Singapore)
- 59090Q **Embedding MPEG-7 metadata within a media file format** [5909-26]
W. Chang, National Institute of Standards and Technology (USA)

SESSION 5 H.264/AVC APPLICATIONS

- 59090S **Applications of the H.264/AVC fidelity range extensions** [5909-28]
P. Topiwala, Fast VDO LLC (USA)
- 59090U **Rate-adaptive H.264 video for information for global reach (IFGR)** [5909-30]
P. Kota, K. Kannan, Z. Xiong, Texas A&M Univ. (USA); P. Topiwala, FastVDO LLC (USA); D. Hench, Air Force Research Lab. (USA)
- 59090V **Comparative study of JPEG2000 and H.264/AVC FRExt I-frame coding on high-definition video sequences** [5909-31]
P. Topiwala, Fast VDO LLC (USA)
- 59090W **Origins of the performance of H.264/AVC: an account of the development of H.263 and H.264 coding tools** [5909-97]
C. Reader, Reader LLC (USA)
- 59090Y **AVC/H.264 patent portfolio license** [5909-33]
L. A. Horn, MPEG LA, LLC (USA)
- 59090Z **Efficient temporal search range prediction for motion estimation in H.264** [5909-35]
C. Kim, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590910 **An operational rate control scheme for H.264 with two-stage encoding** [5909-36]
D. Kwon, M. Shen, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590911 **Efficient bit stream switching of H.264 coded video** [5909-37]
X. Zhou, C.-C. J. Kuo, Univ. of Southern California (USA)

SESSION 6 IMAGE CODING AND STRUCTURES

- 590912 **Progressive geometry coding of 3D meshes using hierarchical vertex set split** [5909-38]
J. Peng, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590914 **A stereo video coding scheme based on H.264** [5909-41]
J. D. Oh, C.-C. J. Kuo, Univ. of Southern California (USA)

- 590915 **Adaptive cubic convolution interpolation and sequential filtering for color demosaicing of Bayer pattern image sensors** [5909-42]
W. Yu, Electronics and Telecommunications Research Institute (South Korea)
- 590916 **A point-based hybrid cloud modeling technique** [5909-43]
B. Wang, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590917 **Edge directed image interpolation with Bamberger pyramids** [5909-44]
J. G. Rosiles, The Univ. of Texas at El Paso (USA)
- 590919 **Applications based on restored satellite images** [5909-46]
D. Arbel, S. Levin, M. Nir, I. Bhasteker, Sami Shamoon College of Engineering (Israel)
- 59091A **Thermogram analysis and classification in breast cancer diagnostics** [5909-47]
M. Zaviscek, A. Drastich, Brno Univ. of Technology (Czech Republic); K. Dvorak, Faculty Hospital of Brno (Czech Republic)
- 59091B **Multifocus fusion with multisize windows** [5909-49]
R. Redondo, Instituto de Óptica, CSIC (Spain); F. Sroubek, Instituto de Óptica, CSIC (Spain) and Academy of Sciences (Czech Republic); S. Fischer, G. Cristóbal, Instituto de Óptica, CSIC (Spain)
- 59091C **Relaxed color reproduction under different illuminant conditions** [5909-48]
L. Liu, Illinois Institute of Technology (USA) and Panavision SVI (USA); Y. Yang, H. Stark, Illinois Institute of Technology (USA)

SESSION 7 CONTENT MANAGEMENT AND REPRESENTATION

- 59091D **Dynamic fluid animation using scalable grids** [5909-50]
Y. Kwak, Univ. of Southern California (USA); C. Kim, Chinese Univ. of Hong Kong (Hong Kong China); C.-C. J. Kuo, Univ. of Southern California (USA)
- 59091E **Video rate color region segmentation for mobile robotic applications** [5909-51]
A. de Cabrol, L2TI-Lab. de Traitement et Transport de l'Information (France); P. J. Bonnin, V. Hugel, P. Blazevic, LRV-Lab. de Robotique de Versailles (France); M. Chetto, LINA-Lab. d'Informatique de Nantes Atlantique (France)
- 59091F **Temporally optimized edge segmentation for mobile robotics applications** [5909-52]
A. de Cabrol, L2TI-Lab. de Traitement et Transport de l'Information (France); P. J. Bonnin, V. Hugel, K. Bouchefra, P. Blazevic, LRV-Lab. de Robotique de Versailles (France)
- 59091G **Separating patterns and finding the independent components of mixed signals based on non-Gaussian distribution properties** [5909-53]
R. S. Rand, U.S. Army Engineer Research and Development Ctr. (USA); H. Chen, P. K. Varshney, Syracuse Univ. (USA)
- 59091H **Model-based recognition using 3D line sets and multidimensional Hausdorff distance** [5909-54]
M. T. Rahman, M. S. Alam, Univ. of South Alabama (USA)

- 59091I **Data image fusion using combinatorial maps** [5909-55]
K. Boucheфра, P. Bonnin, Lab. de Robotique Versailles (France); A. De Cabrol, Lab. de Traitement et Transport de l'Information (France)
- 59091J **Towards parameter-free classification of sound effects in movies** [5909-56]
S. Chu, S. Narayanan, C.-C. J. Kuo, Univ. of Southern California (USA)
- 59091K **Object detection and segmentation in camouflaged environments** [5909-57]
S. Makrogiannis, M. Trujillo San-Martin, Wright State Univ. (USA)
- 59091L **A feature-based, image registration algorithm using the multi-resolution approach combined with moments of inertia, with application to ladar imaging** [5909-58]
C. Bejar, D. B. Megherbi, Univ. of Massachusetts Lowell (USA)
- 59091M **Automatic image-driven segmentation of cardiac ventricles in cine anatomical MRI** [5909-59]
C. A. Cocosco, Philips Research Labs. (Germany); W. J. Niessen, Univ. Medical Ctr. Rotterdam (Netherlands); T. Netsch, Philips Research Labs. (Germany); E. P. A. Vonken, M. A. Viergever, Univ. Medical Ctr. Utrecht (Netherlands)

POSTER SESSION

- 59091O **A fuzzy approach to supervised segmentation parameter selection for object-based classification** [5909-64]
T. L. Maxwell, Y. Zhang, Univ. of New Brunswick (Canada)
- 59091Q **Investigation of crop nitrogen content based on image processing technologies** [5909-66]
Y. Zhang, M. Li, Z. Xu, X. Zhang, M. Wang, China Agricultural Univ. (China)
- 59091R **Estimating soil moisture based on image processing technologies** [5909-67]
L. Zheng, M. Li, J. Sun, X. Zhang, P. Zhao, China Agricultural Univ. (China)
- 59091U **Use of a local regularity analysis by a wavelet analysis for glitch detection** [5909-72]
C. Ordénovic, C. Surace, Lab. d'Astrophysique de Marseille (France); B. Torrèsani, Lab. d'Analyse, de Topologie et de Probabilités (France); A. Llebaria, J. P. Baluteau, Lab. d'Astrophysique de Marseille (France)
- 59091V **Estimating growth status of winter wheat based on aerial images and hyperspectral data** [5909-73]
Y. Han, M. Li, L. Jia, X. Zhang, F. Zhang, China Agricultural Univ. (China)
- 59091W **Impulsive noise detection and removal with the use of iterative filter based on gradient adaptive neighborhoods** [5909-74]
M. Mozerov, Institute for Information Transmission Problems (Russia); V. Kober, CICESE (Mexico)
- 59091X **Recognition of partially occluded objects using correlation filters with training** [5909-75]
J. A. González-Fraga, V. Kober, J. Álvarez-Borrego, CICESE (Mexico)
- 59091Y **A fast, predictive, block matching motion estimation algorithm based on spatio temporal correlation information** [5909-76]
H. Nisar, T. Choi, Gwangju Institute of Science and Technology (South Korea)

- 59091Z **A reduced color approach to high quality cartoon coding** [5909-81]
Y.-C. Tsai, M. Lee, M.-S. Shen, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590920 **Tempo and tension analysis for MTV-style home video authoring** [5909-82]
S.-Z. Wang, National Chiao Tung Univ. (Taiwan); S.H. Lee, C.-C. J. Kuo, Univ. of Southern California (USA)
- 590921 **Real-time pattern recognition with adaptative correlation filters** [5909-83]
V. H. Diaz-Ramirez, V. Kober, J. Álvarez-Borrego, CICESE (Mexico)
- 590922 **Estimation of the point response function for undersampled data for the infrared camera of the Spitzer Space Telescope** [5909-84]
D. Makovoz, P. Lowrance, California Institute of Technology (USA)
- 590923 **Detecting and measuring human bodies over wide work fields by using stereo cameras** [5909-85]
J. Lu, K. Hamajima, National Institute of Industrial Safety (Japan); S. L. Tuladhar, Japan Science and Technology Corp. (Japan)
- 590924 **Target detection and tracking in airborne video imagery using statistical snake and mean shift** [5909-86]
S. Zhang, CUNY, College of Staten Island (USA); M. Chen, State Univ. of New York at Binghamton (USA)
- 590925 **Preprocessing and compression of digital holographic images** [5909-87]
S. Zhang, CUNY, College of Staten Island (USA); M. Chen, State Univ. of New York at Binghamton (USA)
- 590928 **3D watermarking scheme in stereo vision system** [5909-90]
D.-C. Hwang, K-h Bae, J.-H. Ko, E.-S. Kim, Kwangwoon Univ. (South Korea)
- 590929 **A bi-directional stereo matching algorithm based on adaptive matching window** [5909-91]
K-h Bae, D.-S. Yi, 3R Corp. (South Korea); S. C. Kim, E.-S. Kim, Kwangwoon Univ. (South Korea)
- 59092A **An adaptive image enhancement algorithm and real-time implementation for an infrared imaging system** [5909-92]
A. Kumar, S. Sarkar, R. P. Agarwal, Indian Institute of Technology Roorkee (India)
- 59092B **An adaptive wavelet-based deblocking algorithm for MPEG-4 codec** [5909-93]
T.-K. Truong, I-Shou Univ. (Taiwan); S.-H. Chen, R.-Y. Jhang, Shu-Te Univ. (Taiwan)
- 59092C **A novel face recognition system using the binary phase-only filter via optimal correlation thresholding** [5909-94]
D. B. Megherbi, S. Rastogi, Univ. of Massachusetts Lowell (USA)
- 59092D **Effect of illumination correction on face recognition rates for the binary phase-only filter (BPOF) and the extended normalized correlation methods** [5909-95]
D. Megherbi, S. Rastogi, Univ. of Massachusetts Lowell (USA)
- 59092E **A generic and high performance FPGA based signal processing architecture for staring thermal imaging system** [5909-96]
A. Kumar, S. Sarkar, R. P. Agarwal, Indian Institute of Technology Roorkee (India)

Mobile Multimedia Library: an MPEG-7 application with camera-equipped mobile phones

Ryoma Oami*, Eiji Kasutani, Akio Yamada

Media and Information Research Laboratories, NEC Corporation
1753, Shimonumabe, Nakahara-Ku, Kawasaki, Kanagawa 211-8666, Japan

ABSTRACT

This paper proposes a Mobile Multimedia Library (MML), a new application of information retrieval using camera-equipped mobile phones. MML allows users to get information of an unknown object anywhere and anytime. A user simply takes a picture of the object by a mobile phone and sends it directly to an MML server on which the picture is analyzed to identify the object. The MML server returns the identification result to the mobile phone and the user can browse information of the object on the display of the mobile phone. This application employs the k -Nearest Neighbor approach using eight MPEG-7 visual features to identify objects. A prototype system for animal identification has been developed to demonstrate effectiveness of the MML framework. It takes about ten seconds on average for the identification in which 63% and 79% queries were correctly identified within first four and first ten candidates, respectively, out of 229 categories.

Keywords: MPEG-7, CBIR, content based image retrieval, camera-equipped mobile phone, multimedia library, k -Nearest Neighbor

1. INTRODUCTION

Mobile phones have evolved into ubiquitous communication and information retrieval tools from mere handsets just for telephoning. E-mail exchange and Internet connection functionalities are outstanding features that have spurred this evolution. Many people nowadays exchange e-mails with mobile phones instead of computers thanks to the mobility. Information retrieval services through the Internet are also frequently utilized with mobile phones. On-demand train timetable/transfer information search and restaurant guides are examples of services frequently used in Japan. These services are available anywhere and anytime just only using mobile phones without any preparation in advance. Mobile phones are hence considered to be very valuable terminals despite their inconvenience of inputting characters with a limited number of small buttons.

A camera equipped on recent mobile phones is one of the solutions to improve this inconvenience. The camera is originally incorporated just to simply enjoy taking pictures and sharing them with other people through e-mails, but the camera has become to draw attention as an input means. An OCR (Optical Character Reader) functionality is, for example, installed to automatically input such text strings as URLs and phone numbers to mobile phones without typing them. Barcode reading functionality is another example developed to input text embedded in two dimensional barcodes. Current usage of the camera is just to provide an alternative way to input text strings, but the camera has much more promising potential as an information input means.

Object appearance is substantial information acquired by the camera. This inspires us with an idea that the camera can also be exploited to take unknown objects for acquiring their information. The camera functions as an input means of non-verbal, visual information triggering information retrieval. When a person happens to encounter an unknown flower during trekking, for example, s/he can immediately obtain the flower information just by taking a picture of it. CBIR (Content-Based Image Retrieval) technologies help us to achieve this retrieval.

The research area of CBIR is now very active and many applied systems have been proposed so far. QBIC¹ (Query By Image Content), VisualSEEK², and MARS⁴ are examples of such CBIR systems. In these systems, a query image is provided for searching similar images from an image database. Similarity is measured based on low-level image features representing color, shape and texture. Some of them accept sketch images as a query instead of natural images^{1,9}, and

* r-oami@az.jp.nec.com; phone +81-44-431-7710; fax +81-44-431-7590

keywords as well. CBIR technologies are extended to web-oriented systems, such as WebSEEK³ and RetrievalWare⁵. CBIR applications are now being investigated in various fields, and their application area is still growing.

A new application with camera-equipped mobile phones, Mobile Multimedia Library (MML), is proposed in this paper. This is an information retrieval application that enables us to get information from a picture of an object taken by a camera on mobile phones. The picture is analyzed with a CBIR technology using MPEG-7 visual features^{11,13}, including the color layout description proposed by us⁹. This application inherently solves a general problem of query-by-example for CBIR, that is, how to obtain a proper query to feed CBIR systems. Two major issues are also solved for its practical use.

One of them is restriction on computational resources of mobile phones. Object identification is computationally intensive unlike OCR or barcode reading functionality, so that it is quite difficult to perform it only using the resources on mobile phones. The identification is, therefore, performed on a server in place of resource-restricted mobile phones. E-mail exchange and internet connection functionalities of mobile phones are exploited to exchange information between mobile phones and the server. The other issue is semantic gaps between capability of the current CBIR systems and ability of human recognition⁷. Human interaction is hence incorporated to bypass the semantic gaps.

In Section 2, the concept of MML and the system architecture are depicted. The algorithm of image analysis and MPEG-7 visual features used for this application is then described in Section 3. A prototype developed for identifying animals and its evaluation results are presented in Section 4 and 5. Other possible MML applications are shown in Section 6, which illustrates the usefulness of the MML framework.

2. CONCEPT OF MOBILE MULTIMEDIA LIBRARY AND ITS ARCHTECTURE

2.1. Concept of Mobile Multimedia Library

Mobile Multimedia Library (MML) is an information retrieval application that provides information from a picture of an object taken by a camera on mobile phones, anywhere and anytime. The object is identified using the picture, and the results are presented on a screen of mobile phones.

The advantages of this application are as follows:

(1) No necessity of carrying encyclopedia

This is an overall advantage the application has. A single camera-equipped mobile phone is enough to acquire information of target objects anywhere and anytime. No other preparation is necessary.

(2) Easy input of a visual query

This is an advantage of using camera-equipped mobile phones. Taking a picture of a target object is the only required operation, which is quite easy so that many people including children have already managed a camera on mobile phones. The picture is directly used as a visual query.

(3) No necessity of background knowledge of objects

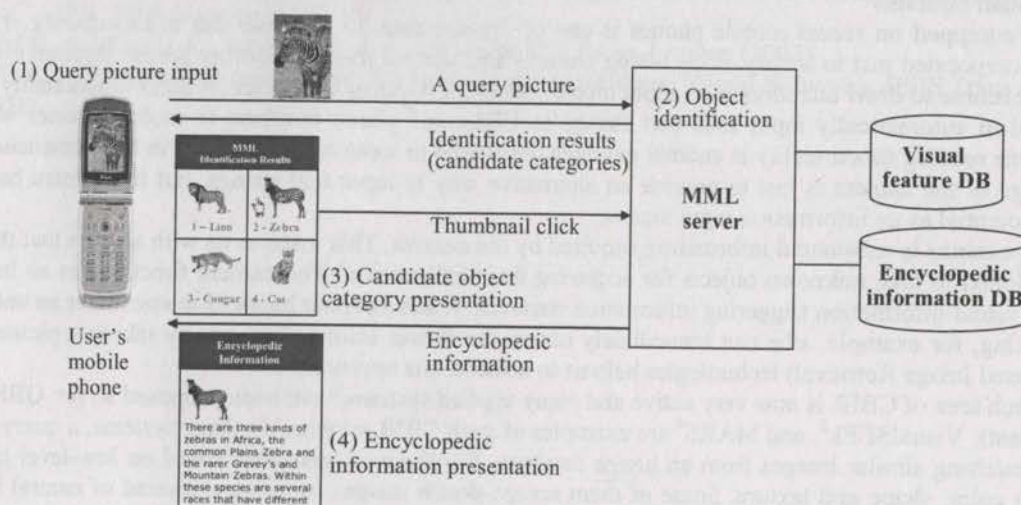


Figure 1: System overview of MML.

This is an advantage CBIR has. The MML application allows us to search objects just from their appearance. No knowledge about taxonomy, which is often needed to consult encyclopedias, is required for getting information of target objects.

(4) Customized search and presentation of retrieved information

This comes from the characteristics of a mobile phone as a private terminal. User preference information, which is stored implicitly in private data such as e-mails or explicitly as a user profile in mobile phones, can be used for this customization. For example, presented information can be customized according to the level of his/her knowledge or interest.

2.2. System overview

The MML application is implemented in a form of a client-server type application that consists of camera-equipped mobile phones and a remote server, which is called MML server, to retrieve information. Object identification process is performed on the MML server instead of resource-restricted mobile phones for its practical use, since the identification process usually needs much computational power.

The system architecture of the MML is shown in Figure 1. Information of an object is retrieved by the MML in the following manner.

1. query picture input,
2. object identification based on query analysis,
3. interactive selection of a proper candidate, and
4. information presentation for selected object category

In Step 1, the user takes a picture of the unknown object by a camera-equipped mobile phone. This picture is uploaded to a remote MML server as a query. In Step 2, the query picture is analyzed on the MML server to identify the object in the picture. MPEG-7 visual features are extracted from the query and compared to the features of training images stored in a visual feature DB. As a result, several object categories are derived as candidates based on the identification results. In the following of the paper, the term "candidate object category" is used to refer object categories selected as candidates that the unknown object can belong to. In Step 3, a web page containing the identification results is presented on the screen of the mobile phone. This page contains representative images corresponding to each candidate object category. Each representative image is linked to encyclopedic information of the corresponding object category. In Step 4, one of the candidates is selected by clicking one of the representative images, and the encyclopedic information about the clicked object category is displayed on the user's mobile phone.

The MML server has two types of databases: a visual feature DB and an encyclopedic information DB, as shown in Figure 1. The former DB contains visual features of training images, which are samples included in each object category and used for query analysis to identify the object. The latter DB has encyclopedic information to be provided to the user.

3. ALGORITHM FOR OBJECT IDENTIFICATION

3.1. Overview

The object identification process on the MML server in Figure 1 is depicted in Figure 2. Visual features are first extracted from a query picture. The distance values between the query features and the training image features in the feature DB are next computed for each elementary feature description standardized as MPEG-7, such as the color layout and the edge histogram. Total distance between two images is then computed by summing the individual distance values that are computed and normalized for each feature description. The obtained total distance values are finally fed to object identification process to determine candidate object categories for the query. In the following subsections, database configuration on the MML server is described. Each step is then described in more detail.

3.2. Visual feature database on the MML server

In the visual feature DB in Figure 2, visual features for plenty of training images are stored for each object category together with its representative image. This is for coping with different appearance and variation. Appearance of objects varies with their viewing angle and capturing conditions, which results in different feature values. Some object categories have several variations, that is, entities that have inherently different appearance even within the same object category (e.g. male lions have manes, but female ones do not). Whereas increasing the number of training images for each

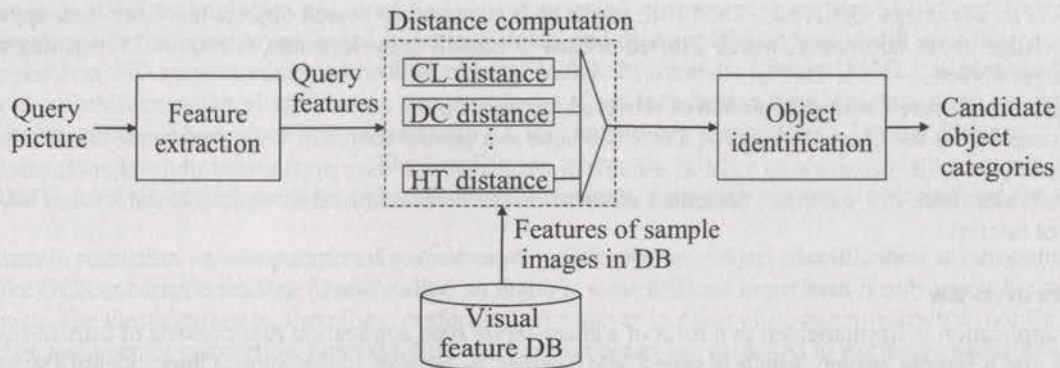


Figure 2: Diagram for query analyzing process to find candidate object categories.

object category improves completeness, this makes processing time longer and this does not always result in improving identification results. A proper number of the training images should be determined in consideration of this trade-off.

3.3. Feature extraction

The feature extraction process in Figure 2 extracts various visual features from the query picture. Several elementary visual feature descriptions defined in MPEG-7¹³ are employed as listed in Table 1 because of their discriminating potential. The interoperability of the features is also their advantage in building an extensible retrieval system. This makes it possible to exchange features and matching results among different systems.

For the object information retrieval application, we employed the following eight elementary feature descriptions: four color features (the color layout, the color structure, the dominant color, and the scalable color), two texture features (the edge histogram, and the homogeneous texture), and two shape features (the contour shape, and the region shape). These features can be extracted automatically, except that object segmentation masks necessary for computing shape features are manually generated. A future work is to incorporate one of conventional automatic segmentation methods into the feature extraction process. The dimensionality of each elementary feature is listed in Table 1. Please see the Appendix for more detailed specification.

3.4. Distance computation between the query image and training images

Distance values are computed between each pair of the query and the training image stored in the visual feature DB. The computed distance values are fed to object identification process.

In the distance computation for each pair, distance values for the eight elementary feature described in Section 3.3 are first individually computed. Distance functions as described in MPEG-7 Part 8¹⁴ are employed for this computation.

The obtained eight individual distance values are then normalized to equalize the range of distance values among different elementary features. The normalization is performed with the mean and the standard deviation of distance distribution in the visual feature DB. Distance values within $[-3\sigma, 3\sigma]$ are linearly mapped onto the interval $[0, 1]$ and the values outside the range are represented with the closest end point, that is, zero or one.

The normalized distance values for individual elementary features are then summed up with a set of certain weighting coefficients to calculate the total distance. The optimal weighting coefficients were derived in a learning process with a set of training images. The combination of eight coefficients was exhaustively searched in the learning process.

3.5. Object identification by k -Nearest Neighbor approach

Object identification is performed using the total distance values obtained in the previous subsection. k -Nearest Neighbor approach¹², which gave the best accuracy among several methods we tried, is taken for this identification. The total distance values are first sorted in the ascendant order, and the number of samples for each object category included in the first k rank is counted. The

Table 1: MPEG-7 elementary visual features used for object identification

Visual Features	Dimensionality
Color Layout	18 dims
Color Structure	256 bins
Dominant Color	≤ 8 colors
Scalable Color	256 bins
Edge Histogram	80 bins
Homogeneous Texture	$2 + 60$ dims
Contour Shape	$\leq (3 + 62)$ dims
Region Shape	35 dims

object categories with high frequency are finally selected as candidates after sorted by the frequency.

4. PROTOTYPE SYSTEM

A prototype system for identifying animals was developed as an example to demonstrate the proposed MML framework. This section describes this animal identification system.

4.1. Prototype system for animal identification

A third-generation mobile phone was used in this prototype system. This has a camera whose resolution is about 300 thousand pixels and its screen size is 33 mm x 45 mm. A computer with a Pentium 3GHz CPU, which has an ADSL connection to the Internet, was utilized for the MML server.

For uploading a query picture from the mobile phone to the MML server, e-mail exchange function of the mobile phone was exploited, that is, the picture was enclosed in an e-mail as an attachment, and sent to the server. On the server side, a program module periodically checking a designated mail account was implemented, and e-mails were immediately downloaded and pictures were separated from the e-mails when they were found. The pictures are then provided to the identification process in Figure 2.

A process for animal identification is launched by the e-mail checking module, and a web page showing identification results are generated. Details about the web page design are described in Section 4.2. Terminal adaptation related to screen size was performed in generating the web page. The size of the representative images is adjusted to the screen size so that all the candidates can be seen without scrolling according to the types of mobile terminals. An environment variable that notifies a platform and a browser type was checked to distinguish different terminals.

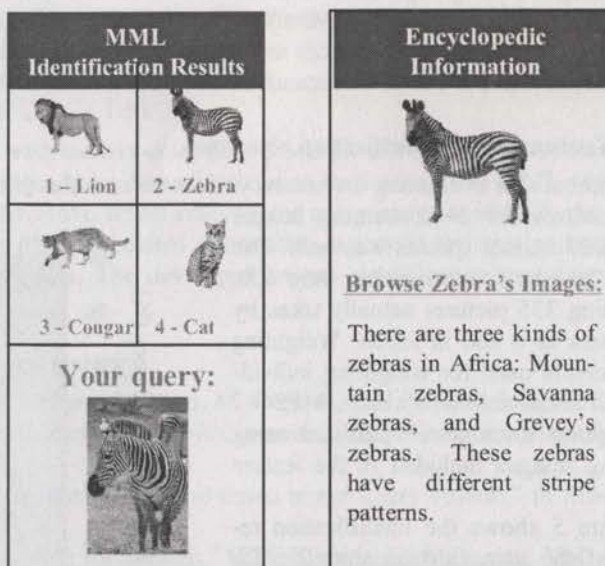
The DB for encyclopedic information for each animal was generated from some sources available on the Internet. The information such as in-depth descriptions of animals and their living environment was included.

4.2. Web page design for showing animal candidates

The web page showing identification results is designed to include representative images for first four candidate categories as shown in Figure 3 (a). This is because "four" is considered to be the maximum number to show the representative images simultaneously without losing details necessary for distinguishing the animals. It is certain that more than four candidates can be displayed with smaller images, but this sometimes makes it hard to differentiate the candidate objects at a glance on a small screen of a mobile phone. In the case of PDAs or smart phones that have a larger screen, more candidates like first ten can be displayed. The query image is also displayed at the bottom just for confirmation.

The web page is designed so as to minimize user operation. The first candidate animal category is focused when the web page is presented on the mobile phones. Hence if the first candidate is the animal category a user wants to know, just one click action is enough to show the information. The focus is easily moved by pressing buttons corresponding to moving focus, so that the user can select one of the representative images.

Encyclopedic information as shown in Figure 3 (b) is displayed on the screen of the mobile phone by clicking one of the four images. The second image is assumed to be clicked in this example. This web page contains the representative image of this category, a link to a page including various images belonging to the same category, and description regarding this animal category. The link is included for checking other images of this category as well to confirm that this is the correct animal category. Thus users can easily get informa-



(a) Web page showing identification results.

(b) Encyclopedic information for the animal selected in (a).

Figure 3: Web pages for identification results and encyclopedic information that is shown after one of the candidate animal categories is selected.



(a) The web Page that shows identification results.

(b) Retrieved information by clicking the picture of the first candidate "lion" in (a).

(c) Detailed descriptions about "lion". These descriptions are shown by scrolling the screen.

Figure 4: Screen shots presented to a user by the prototype system.

tion of animals using the MML application.

5. PROTOTYPE DEMONSTRATION AND EXPERIMENTAL RESULTS

5.1. Demonstration of the prototype system

Figure 4 shows screen shots presented to a user by the prototype system. The left most image shows the web page displaying the first four candidates. If the user clicks the top-left lion image, the information about lion is obtained as shown in the middle picture. When scrolling the screen, detailed descriptions about lions appear as shown in the right-most image. The highlighted description of "Browse Lions Images" is for browsing other lion images.

It was confirmed using this prototype that all the operations required to users, from taking a picture of an animal to browse information displayed, can smoothly be done with only a single hand as designed. The processing time for animal identification is about ten seconds on average, which would be acceptable in ordinary use cases.

5.2. Evaluation of identification algorithm

Experiments for evaluating the accuracy of the animal identification algorithm were carried out. A feature database storing features for 2681 training images from 229 animal species was built. The query images for evaluation were 600, including 335 pictures actually taken by a camera at a zoo in Japan. Weighting coefficients used for weighting individual distance values of each MPEG-7 description tools were optimized using training images included in the feature DB.

Figure 5 shows the identification results of the identification algorithm for several different number k of nearest neighbors. Each curve corresponds to a different k , ranging from 20 to 100. The horizontal axis indicates the rank to which the correct result is included. Hence the accuracy increases monotonically.

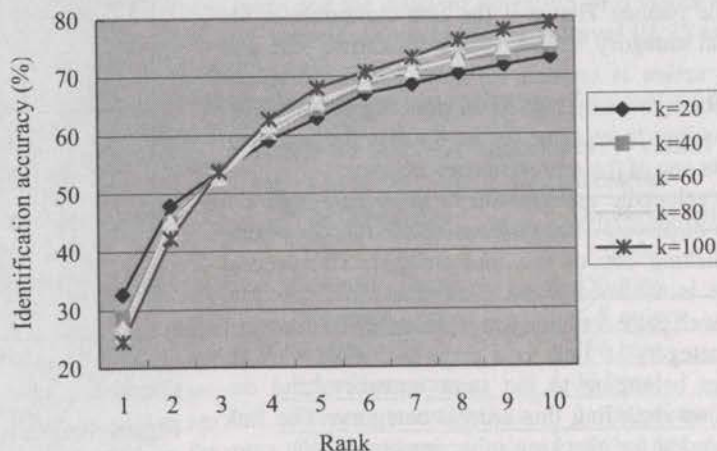


Figure 5: Identification accuracy for several k values.

cally with the rank value.

The accuracy is 63% when first four candidate animal categories are provided, which occurs at $k = 100$. If ten candidates can be displayed, which is the case for smart phones with a larger screen, the accuracy increases up to 79%. Currently, more than 80% of the total species included in the feature DB belong to the dog category, which means many animals with similar appearance are included. If only one candidate is shown to a user, the accuracy is about 33% at most and the best performance is achieved when $k = 20$. This implies the importance of human interaction.

The accuracy is expected to increase if more different types of animals are included with the DB. According to this result, the parameter k was set to 100 in the prototype system mentioned in Section 4.1.

6. OTHER POSSIBLE APPLICATION SCENARIOS FOR MML

The framework of the MML is widely applicable, not limited to animal identification system described above. This framework is quite general, so that many application examples are considered in the same framework. The followings are possible examples for which the MML framework can effectively be applied.

- Flower identifying systems
- Building/monument guide systems
- Museum guide systems
- Product information retrieval systems with logo identifier
- Navigation systems to a nearest store with logo identifier

The flower identifying systems are very similar to the animal identification mentioned in this paper. In this case, location and time information, which can be obtained by GPS (Global Positioning System) or the clock of mobile phones, is also effectively utilized as well as visual features to identify flowers. The building/monument guide systems are also similar to the animal identification, but location information becomes much more relevant. The museum guide systems are used in museums or art galleries to retrieve information of paintings and statues.

The product information retrieval systems are a little different from the previous three examples, which are just for retrieving information of the objects in themselves. In this application, taking a logo of a store makes it possible to access its product information such as price for each product sold there or electric coupons can be retrieved. On the other hand, the nearest store corresponding to the identified logo is searched in the navigation systems, and navigation service to the store is provided with a map and location information obtained by GPS.

The MML framework can, furthermore, be applied to conventional CBIR systems in themselves. One of the problems of the conventional CBIR systems is how to acquire an appropriate query. Camera-equipped mobile phones can be employed to solve this problem.

7. CONCLUSION

A Mobile Multimedia Library, an MPEG-7 application with camera-equipped mobile phones has been proposed in this paper. Mobile phones are exploited as means for inputting a query image, uploading the image to an MML server, and browsing information returned from the MML server. The MML server analyzes the query image to identify the object a user wants to know, and then the identification result with their related information is sent to the user and presented. MPEG-7 visual features are used for identifying the objects. The developed animal identification prototype system demonstrated the usefulness of the MML.

REFERENCES

1. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: the QBIC system," IEEE Computer Magazine, Vol. 28, No. 9, Sep. 1995.
2. J. R. Smith and S-F Chang, "VisualSEEK: A fully automated content-based image query system," in Proc. ACM Multimedia 96, 1996.
3. J. R. Smith and S-F. Chang, "Visually searching the web for content," IEEE Multimedia Magazine, Vol. 4 No. 3, pp. 12-20, July-Sept. 1997.
4. Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in Proc. of International Conference on Image Processing (ICIP), pp. 26-29, Oct. 1997.
5. J. Dowe, "Content-based retrieval in multimedia imaging," in Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993.

6. Y. Rui, T. S. Huang, and S-F. Chang, "Image retrieval: current techniques, promising directions and open issues," *Journal of Visual Communication and Image Representation*, Vol. 10, pp. 39-62, March 1999.
7. A. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 22, No. 12, Dec. 2000.
8. C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 24, No. 8, Aug. 2002.
9. E. Kasutani and A. Yamada, "The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval," in *Proc. of International Conference on Image Processing (ICIP)*, Vol. 1, pp. 674-677, Oct. 2001.
10. V. Castelli and L. D. Bergman, "Image Databases –Search and Retrieval of Digital Imagery—," John Wiley & Sons, Inc., 2002.
11. B. S. Manjunath, P. Salembier and T. Sikora, "Introduction to MPEG-7," John Wiley & Sons, Ltd., 2002.
12. R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Classification," 2nd Ed, John Wiley & Sons, Inc., 2001.
13. "Information Technology –Multimedia content description interface—Part 3: Visual," ISO/IEC 15938-3:2002, 2002.
14. "Information Technology –Multimedia content description interface—Part 8: Extraction and use of MPEG-7 descriptors," ISO/IEC TR 15938-8:2002, 2002.

APPENDIX: MPEG-7 VISUAL FEATURES FOR MML

The following eight MPEG-7 features are utilized in the prototype system of MML for animal identification.

Color Layout

This feature expresses brief color layout information for rectangular regions. This is extracted by computing DCT for 8x8 small image generated by shrinking an entire rectangular region to this size. Low frequency components are selected and quantized to yield the elements of the color layout. The feature data size is usually very small, typically 8 bytes for an rectangular region, but it has a very high discriminative capability. 18 coefficients (six coefficients for all Y, Cb and Cr components) are used in our system.

Color Structure

This feature expresses both color distribution and a local color structure together. An $N \times N$ (typically 4×4) structuring element is sled over a target rectangular region for capturing the color structure. The frequency of each color appearing within the element is counted in the HMMD color space. The total number of bins is 256 in the current system.

Dominant Color

This feature describes dominant colors (up to eight colors) for a rectangular region. Region segmentation is first applied in order to find dominant colors by grouping pixels with similar colors. The portion of each segmented region is also considered when computing distances between two features.

Scalable Color

This is basically color histogram in the HSV color space, which is encoded by a Haar Transform. The main feature of this is its scalable characteristics in terms of the number of bins. This scalability is achieved by Haar Transform for the histogram. The number of bins is 256, which means hue, saturation and value are quantized into 16, 4, and 4 levels, respectively.

Edge Histogram

This feature describes edge histograms for 16 blocks after a rectangular region is divided into the 16 blocks. The histogram is calculated for four directions (horizontal, vertical and two diagonal directions) and for non-directional edges. The number of bins is, therefore, 80 in total.

Homogeneous Texture

This feature expresses characteristics of texture information using mean energy and energy deviation, which are computed for each frequency channel obtained by Gabor functions. The number of channels is 30 designated by six levels in angular direction and 5 levels in radial direction. The number of dimensions is 62, including the mean and the standard deviation for an entire rectangular region.

Contour Shape

This expresses shape information by describing contour of a region based on CSS (Curvature Scale-Shape) representation. The CSS representation decomposes the contour into convex and concave sections by the points with zero curvature. The circularity and eccentricity are derived from the CSS representation, and the magnitude of the largest peak, and the position of the remaining peaks are included in the feature.

Region Shape

This represents a binary shape mask by decomposing it using orthogonal 2-D basis functions defined by the Angular Radial Transform. This can express shapes consisting of multiple disconnect regions and shapes with some holes. Total number of dimensions is 35, which is determined by 12 levels in angular direction and 3 levels in radial direction.

APPENDIX 5

[Library Catalog](#) 

[Advanced Search](#)

[All items](#) [that contain my query words](#) [in the title](#)

For optimum display, please use Chrome or Firefox. Safari users may not be able to see all the relevant information.

[Back to results list](#)



CONFERENCE PROCEEDING
Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union
Tescher, Andrew G.; Society of Photo-optical Instrumentation Engineers; International Telecommunication Union
c2005

[Available at Linda Hall Library Books - 2nd Floor \(TA1632 .A66 2005\) >](#)

- [Top](#)
- [Send to](#)
- [Get It](#)
- [Details](#)
- [Virtual Browse](#)
- [Links](#)

Send to

-  Email
-  QR
-  Permalink
-  Citation
-  EasyBib
-  Export RIS
-  Export to Excel

Get It

[Back to Locations](#)

LOCATION ITEMS

Linda Hall Library
Available, Books - 2nd Floor; TA1632 .A66 2005
(1 copy, 1 available, 1 request)

Item in place	Material Type: Book
Loanable	Location: Linda Hall Library Books - 2nd Floor TA1632 .A66 2005
	Barcode: 33690006100480

Details

Title	Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union
Other title	Digital image processing Proceedings of SPIE--the International Society for Optical Engineering. v. 5909.
Creator/Contributor	Tescher, Andrew G. > Society of Photo-optical Instrumentation Engineers > International Telecommunication Union >
Subjects	Image processing -- Digital techniques > Image compression -- Standards >
Identifier	ISBN : 9780819459145 ISBN : 0819459143 OCLC : (OCoLC)62189828
Series	Proceedings of SPIE ; v. 5909 > Proceedings of SPIE--the International Society for Optical Engineering. v. 5909. >
Publisher	Bellingham, Washington : SPIE
Creation date	c2005
Format	1 v. (various pagings) : ill. (some color), maps ; 28 cm.
General notes	Includes bibliographic references and author index.

Virtual Browse



[Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#) [Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#) [Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#) [Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#) [Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#) [Applications of digital image processing XXVIII : 2-4 August 2005, San Diego, California, USA / Andrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union](#)

Links

[Display Source Record](#)  [>](#)

APPENDIX 6

```
leader 01346nam a22003111a 4500
001 994591033405961
005 20051031111200.0
008 013011c2005 wauab b 101 0 eng d
020 ##$a9780810459145
020 ##$a0813459141
035 ##$a(OCoLC)62189828
035 ##$637546
035 ##$a(WoK)459103-1hal1b
035 ##$a1hal1b0459103-1hal1b
040 ##$aHC $cLH
049 ##$aHLA
090 ##$aTA1632 $b.466 2005
245 00$aApplications of digital image processing XXVIII : $b2-4 August 2005, San Diego, California, USA / $cAndrew G. Tescher, chair/editor ; sponsored and published by SPIE--the International Society for Optical Engineering ; organized with the support of the International Telecommunications Union.
260 ##$aBellingham, Washington : $bSPIE, $c2005.
300 ##$a1 v. (various pagings) : $b111. (some color), maps ; $c28 cm.
490 1$a$uProceedings of SPIE ; $v. 5909
584 ##$aIncludes bibliographic references and author index.
650 00$aImage processing $aDigital techniques $cCongresses
650 00$aImage compression $aStandards $cCongresses.
700 1$aTescher, Andrew G.
710 2$aSociety of Photo-optical Instrumentation Engineers
710 2$aInternational Telecommunication Union
830 00$aProceedings of SPIE--the International Society for Optical Engineering. $v. 5909.
848 ##$aTI 09/01/2006
994 ##$aCB $bLH
```

APPENDIX 7

Try the [Copyright Public Records System \(CPRS\)](#) pilot with enhanced search features and filters.

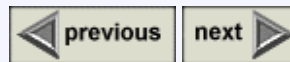
[Help](#)[Search](#)[History](#)[Titles](#)[Start Over](#)

Public Catalog

Copyright Catalog (1978 to present) at DC4

Search Request: Simple Search = Applications of Digital Image Processing XXVIII

Search Results: Displaying 1 of 1 entries

[Labeled View](#)

Applications of digital image processing XXVIII :

Type of Work: Entry not found.

Registration Number / Date: TX0006266473 / 2005-11-18

Title: Applications of digital image processing XXVIII : vol. 5909, 2-4 August 2005.

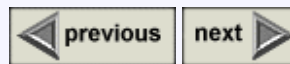
Copyright Claimant: Society of Photo Optical Instrumentation Engineers

Date of Creation: 2005

Date of Publication: 2005-08-02

Copyright Note: Cataloged from appl. only.

Names: [Society of Photo Optical Instrumentation Engineers](#)



Save, Print and Email ([Help Page](#))

Select Download Format



Format for Print/Save

Enter your email address:

Email

[Help](#) [Search](#) [History](#) [Titles](#) [Start Over](#)

[Contact Us](#) | [Request Copies](#) | [Get a Search Estimate](#) | [Frequently Asked Questions \(FAQs\) about Copyright](#) | [Copyright Office Home Page](#) | [Library of Congress Home Page](#)

APPENDIX 8

Certificate of Registration



This Certificate issued under the seal of the Copyright Office in accordance with title 17, *United States Code*, attests that registration has been made for the work identified below. The information on this certificate has been made a part of the Copyright Office records.

Shirley Perlman

United States Register of Copyrights and Director



Short Form SE

For a Serial

UNITED STATES COPYRIGHT OFFICE

REGISTRATION NUMBER

TX 6-266-473



#TX0006266473#

Effective Date of Registration

11-18-05

Application Received
NOV 18 2005

Deposit Received

One

Two

NOV 18 2005

Fee Received

Examined By

TPK

Correspondence

☐

DO NOT WRITE ABOVE THIS LINE.

1

TITLE OF THIS SERIAL AS IT APPEARS ON THE COPY

Applications of Digital Image Processing XXVIII

Volume▼
5909

Number▼

Date on copies▼

2-4 August 2005

ISSN▼

2

NAME AND ADDRESS OF THE AUTHOR AND COPYRIGHT CLAIMANT IN THIS COLLECTIVE WORK MADE FOR HIRE

Society of Photo Optical Instrumentation Engineers
PO Box 10
Bellingham WA 98227-0010

3

DATE OF PUBLICATION OF THIS PARTICULAR ISSUE

Month▼

Day▼

Year▼

August

2

2005

YEAR IN WHICH CREATION OF
THIS SERIAL WAS COMPLETED (IF
EARLIER THAN THE YEAR OF
PUBLICATION)

Year▶

CERTIFICATION*: I, the undersigned, hereby certify that I am the copyright claimant or the authorized agent of the copyright claimant of the work identified in this application, that all the conditions specified in the instructions on the back of this form are met, that the statements made by me in this application are correct to the best of my knowledge.

Handwritten signature (X)

Jacki Sande

Typed or printed name of signer

Jacki Sande

PERSON TO CONTACT FOR CORRESPONDENCE ABOUT THIS CLAIM

Name▶ Jacki Sande

Address (if other than given below)▶

Daytime phone▶ (360) 676-3290 x210

Fax▶ (360) 647-1445

Email▶ jackis@spie.org

DEPOSIT ACCOUNT

Account number▶

Name of account▶

Certificate
will be
mailed
in window
envelope
to this
address:

Name▼

SPIE c/o Jacki Sande

Number/Street/Apt▼

PO Box 10

City/State/ZIP▼

Bellingham WA 98227-0010

YOU MUST:

- Complete all necessary spaces
- Sign your application

**SEND ALL 3 ELEMENTS
IN THE SAME PACKAGE:**

1. Application form
2. Nonrefundable \$30 filing fee in check or money order payable to Register of Copyrights
3. Deposit material

MAIL TO:

Library of Congress
Copyright Office
101 Independence Avenue, S.E.
Washington, D.C. 20559-6000

Fees are subject to change. For current fees, check the Copyright Office website at www.copyright.gov, write the Copyright Office, or call (202) 707-5900.

*17 U.S.C. §506(e). Any person who knowingly makes a false representation of a material fact in the application for copyright registration provided for by section 409, or in any written statement filed in connection with the application, shall be fined not more than \$2,500.

Rev. June 2002—20,000 Web Rev. June 2002 Printed on recycled paper

U.S. Government Printing Office 2000-461-113/20,021

APPENDIX 9

This website uses cookies to provide you with a variety of services and to improve the usability of our website. By using the website, you agree to the use of cookies in accordance with our [Privacy Policy](#).

[Close](#)**SPIE.** **SPIE.** DIGITAL LIBRARY **SPIE.** CAREER CENTER **optics.org**[Sign In](#) | [View Cart](#) | [Help](#)[ADVANCED SEARCH >](#)**SPIE.** DIGITAL LIBRARY

CONFERENCE PROCEEDINGS

[PAPERS](#)[PRESENTATIONS](#)[JOURNALS >](#)[EBOOKS](#)

Search Digital Library



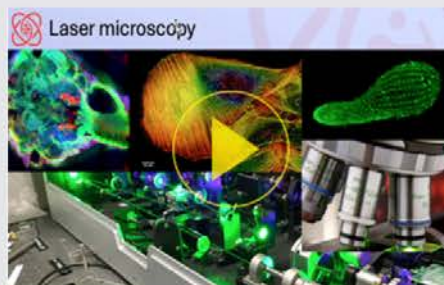
Search the world's largest collection of optics and photonics applied research.

SEE NEW
CONFERENCE
CONTENT
→

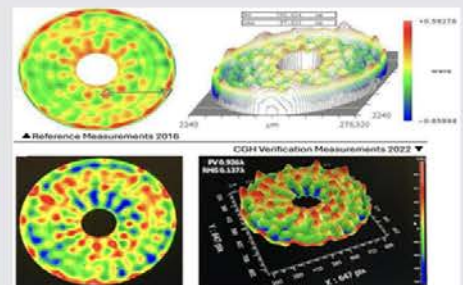
Enter Search Term

[SEARCH >](#)[ADVANCED SEARCH >](#)**SPIE.** MEDICAL IMAGING[GO >](#)**SPIE.** AR|VR|MR[GO >](#)**SPIE.** PHOTONICS WEST[GO >](#)**SPIE.** QUANTUM WEST[GO >](#)

New Proceedings

[Read Recent Conference Proceedings](#)

Featured Presentation

[NONLINEAR QUANTUM BIOIMAGING WITH BRIGHT SQUEEZED LIGHT](#)[Watch the Presentation](#)

Featured Article

[ALIGNMENT PERFORMANCE VERIFICATION OF TWO-MIRROR FOCAL CORRECTOR OPTICS USING COMPUTER-GENERATED HOLOGRAMS FOR ULTRAVIOLET TELESCOPE](#)[Read the Article](#)

ANNOUNCEMENTS

Journals reviewer training program

Page 69 of 112

TECHNOLOGIES

The most viewed articles by technology, from SPIE Journals and Proceedings

ASTRONOMY

[Origins Space Telescope: baseline mission concept](#)

J. Astronomical Telescopes, Instruments, and Systems (2021) Open Access

- Enhance your skills
- Support your community

Learn more >

Reviewer Training Program

Become a great peer reviewer! This essential peer review training course offers a blend of online learning and hands-on peer review experience.

[Get Started with the Reviewer Training Program](#)



LENS.ORG
Solving The Problem Of Problem Solving™

SPIE has partnered with Lens.org, a patent and scholarly research knowledge base, to improve user experience by providing more complete citation information and streamlined citation discovery. Look for the Lens logo on article pages to begin using this new service.

SPIE | editage

Professional English Language Editing Services

Submit a well-written and perfectly formatted manuscript that meets global publication standards.

Use code **SPIE15** to get a 15% discount.

LEARN MORE

ENGLISH LANGUAGE EDITING SERVICES

SPIE partners with Editage to offer English language editing services to researchers worldwide.

[15% discount applied at checkout](#)

The Copernicus CO2M mission for monitoring anthropogenic carbon dioxide emissions from space

Proceedings of SPIE (2021)  Open Access

BIOMEDICAL OPTICS & MEDICAL IMAGING

Identification of different types of tumors based on photoacoustic spectral analysis: preclinical feasibility studies on skin tumors

J. of Biomedical Optics (2023)  Open Access

Kernel Flux: a whole-head 432-magnetometer optically-pumped magnetoencephalography (OP-MEG) system for brain activity imaging during natural human experiences

Proceedings of SPIE (2021)  Open Access

COMMUNICATIONS & INFORMATION TECHNOLOGY

High capacity and access rate, data storage using laser communications

Optical Engineering (2021)  Open Access

Modular FSO Optical System Design for Classical and Quantum Optical Communication Systems

Proceedings of SPIE (2021)  Open Access

DEFENSE & SECURITY

Survey on deep learning applications in digital image security

Optical Engineering (2021)  Open Access

Tactical agility for AI-enabled multi-domain operations

Proceedings of SPIE (2023)  Open Access

ELECTRONIC IMAGING & SIGNAL PROCESSING

Efficient detection of crossing pedestrians from a moving vehicle with an array of cameras

Optical Engineering (2022)  Open Access

Research on bridge defect detection technology based on image processing technology

Proceedings of SPIE (2023)  Open Access

ENERGY

Holographic low concentration optical system increasing light collection efficiency of regular solar panels

J. Photonics for Energy (2021)  Open Access

GALACTIC: high performance alexandrite crystals and coatings for high power space applications

Proceedings of SPIE (2021)  Open Access

LASERS

Structured light for ultrafast laser micro- and nanoprocessing

Optical Engineering (2021)  Open Access

Development of LISA laser system at NASA

Proceedings of SPIE (2023)  Open Access

LIGHT SOURCES & ILLUMINATION

Speckle control for electro-holographic display using high-brightness yellow phosphor light source in projector

Optical Engineering (2023)  Open Access

Modular FSO Optical System Design for Classical and Quantum Optical Communication Systems

Proceedings of SPIE (2021)  Open Access

LITHOGRAPHY & MICROELECTRONICS

Potential of E-beam lithography for micro- and nano-optics fabrication on large areas

J. Micro/Nanolithography, MEMS, and MOEMS (2023)  Open Access

High-NA EUV Lithography exposure tool: advantages and program progress

Proceedings of SPIE (2021)

METROLOGY

Enabling precision coordinate metrology for universal optical testing and alignment applications

Optical Engineering (2021)

Novel diffraction-based overlay metrology utilizing phase-based overlay for improved robustness

Proceedings of SPIE (2021)

NANOTECHNOLOGY

Graphene/molybdenum disulfide nanocomposites: characterization and optoelectronic application

J. Nanophotonics (2023)  Open Access

Carbon nanotube EUV pellicle tunability and lifetime

Proceedings of SPIE (2021)  Open Access

OPTICS

Opto-thermo-mechanical phenomena in satellite free-space optical communications: survey and challenges

Optical Engineering (2023)  Open Access

Optical communication requirements for scientific missions and the Deep Space Gateway

Proceedings of SPIE (2023)  Open Access

REMOTE SENSING

Satellite remote sensing of river discharge: a framework for assessing the accuracy of discharge estimates made from satellite remote sensing observations

J. Applied Remote Sensing (2023)  Open Access

A synchronized remote sensing monitoring approach in the Livingstone Island region of Antarctica

Proceedings of SPIE (2023)  Open Access

SENSORS

Technical concepts of automotive LiDAR sensors: a review

Optical Engineering (2023)  Open Access

New developments of multilinear and multispectral infrared sensors for space applications at LYNRED

Proceedings of SPIE (2023)  Open Access

Site Map

[Home](#)
[Conference Papers](#)
[Conference Presentations](#)
[Journals](#)
[eBooks](#)
[About](#)
[Subscriptions](#)

Information for Authors

[Proceedings Authors](#)
[Journal Authors](#)
[eBook Authors](#)

Information for Reviewers

[Reviewer Guidelines](#)
[Reviewer Training Program](#)

Information for Librarians

[Resources](#)
[Subscriptions](#)

Contact & Support

TECHNICAL SUPPORT
spiedlsupport@spie.org
CUSTOMER SERVICE
+1 360 676 3290
Hours:
8:00 am to 5:00 pm PST
[Help Center](#) | [Contact Us](#)

Connect



SPIE.

[SPIE Privacy Policy](#) | [Terms of Use](#)

© 2025 SPIE

APPENDIX 10

**SPIE. PHOTONICS
WEST**
17–22 January 2026

Share your research

SPIE.**SPIE. DIGITAL
LIBRARY****SPIE. CAREER
CENTER****optics.org**

Sign In

View Cart

Help

ADVANCED SEARCH >

**SPIE. DIGITAL
LIBRARY**

CONFERENCE PROCEEDINGS

PAPERS

PRESENTATIONS

JOURNALS >

EBOOKS

Search Digital Library



Proceedings

SPIE conferences bring together engineers and scientists to present their latest research and to network with peers. Each year SPIE conferences result in approximately 350 proceedings volumes comprising 16,000+ papers and presentation recordings reporting on photonics-driven advancements in areas such as biomedicine, astronomy, defense and security, renewable energy, and more.

[Recent SPIE Conferences](#) | [Browse by Year](#) | [Browse by Volume Number](#)☒ [Subscribe to Digital Library](#)☒ [Receive Email Alerts on Newly Released Proceedings Volumes](#)

NEW PROCEEDINGS

**SPIE. MEDICAL
IMAGING**

GO >

**SPIE. QUANTUM
WEST**

GO >

SPIE. AR | VR | MR

GO >

**SPIE. cjs | PHOTONICS
ASIA**

GO >

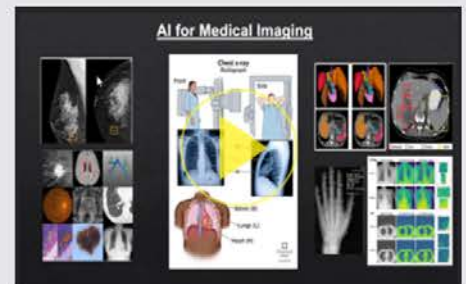
**SPIE. PHOTONICS
WEST**

GO >

**SPIE. LASER
DAMAGE**

GO >

PRESENTATIONS



New Proceedings

RECENTLY PUBLISHED SPIE CONFERENCE PROCEEDINGS

[View the Proceedings](#)

Featured Presentation

CXR-DINO: PAVING THE WAY FOR A
MEDICAL VISION FOUNDATION MODEL
THROUGH SELF-SUPERVISED LEARNING IN
CHEST X-RAY ANALYSIS

[Watch the Presentation](#)

Most Viewed Articles

From the Proceedings of SPIE

Research on cracking WIFI wireless network using Kali-Linux penetration testing software

Lin Wang, Chin Ta Chen, Chih Ming Tsai (2023)

EUV optics: status, outlook and future

Paul Graeupner, Peter Kuerz, Thomas Stammer, Jan van Schoot, Judon Stoeldraijer (2022)

Enhancing SVT-AV1 with LCEVC to improve quality-cycles trade-offs and enhance sustainability of VOD transcoding

Guendalina Coblanchi et al. (2022)  Open Access

Metal oxide resist (MOR) EUV lithography processes for DRAM application

Shinichiro Kawakami et al. (2022)

Memory technology: process and cell architecture

Jeongdong Choe (2023)

Overview of stitching for high NA: Imaging and overlay experimental and simulation results

Natalia Davydova et al. (2023)

Hyper-NA EUV lithography: An imaging perspective

Inhwan Lee, Joern-Holger Franke, Vicky Phillipsen, Kurt Ronse, Stefan De Gendt, Eric Hendrickx (2023)

Quantum imaging overview

Miles Padgett (2023)  Open Access

Soft x-ray: novel metrology for 3D profilometry and device pitch overlay

Christina Porter et al. (2023)

EUV mask technologies: Evolution and ecosystem for devices

Jin Choi (2024)  Open Access

INFORMATION

Proceedings ISSN Identification

Proceedings of SPIE: 0277-786X

Proceedings of SPIE, Progress in Biomedical

Optics and Imaging: 1605-7422

Proceedings Policies, Aim, and Scope

[See all proceedings policies](#)

TECHNOLOGIES

[Astronomy](#)

[Biomedical Optics & Medical Imaging](#)

[Communication & Information Technologies](#)

[Defense & Security](#)

[Electronic Imaging & Signal Processing](#)

[Energy](#)

[Lasers](#)

[Light Sources & Illumination](#)

[Lithography & Microelectronics](#)

[Metrology](#)

[Nanotechnology](#)

[Optics](#)

[Remote Sensing](#)

[Sensors](#)

Call for Papers

SPIE. PHOTONICS
WEST

SPIE. LASER
DAMAGE

SPIE. OPTIFAB

SPIE. FUTURE SENSING
TECHNOLOGIES

SPIE. PHOTONICS
ASIA

SPIE. PHOTOMASK
TECHNOLOGY +
EUV LITHOGRAPHY



Calls for Papers

PRESENT YOUR RESEARCH AT
AN SPIE CONFERENCE



[Open Calls for Papers](#)

Access & Subscriptions

SUBSCRIPTION AND
MEMBERSHIP INFORMATION



[Subscribe Today](#)

Write a Paper

PRESENT AND PUBLISH YOUR
LATEST RESEARCH



[Information for Proceedings Authors](#)

Site Map

[Home](#)
[Conference Papers](#)
[Conference Presentations](#)
[Journals](#)
[eBooks](#)
[About](#)
[Subscriptions](#)

Information for Authors

[Proceedings Authors](#)
[Journal Authors](#)
[eBook Authors](#)

Information for Reviewers

[Reviewer Guidelines](#)
[Reviewer Training Program](#)

Information for Librarians

[Resources](#)
[Subscriptions](#)

Contact & Support

TECHNICAL SUPPORT
spiedsupport@spie.org
CUSTOMER SERVICE
+1 360 676 3290

Hours:
8:00 am to 5:00 pm PST
[Help Center](#) | [Contact Us](#)

Connect



SPIE.

[SPIE Privacy Policy](#) | [Terms of Use](#)

© 2025 SPIE

APPENDIX 11



Browse Proceedings

Browse our growing collection of more than 590,000 conference proceedings papers.

RECENT SPIE CONFERENCES

BY YEAR

BY VOLUME NUMBER

13000-NOW

12000-12999

11000-11999

10000-10999

9000-9999

8000-8999

7000-7999

6000-6999

5000-5999

4000-4999

3000-3999

2000-2999

1000-1999

1-999

+ 13600-13699

+ 13500-13599

+ 13400-13499

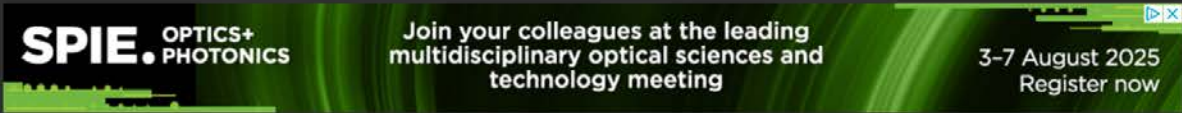
+ 13300-13399

+ 13200-13299

+ 13100-13199

+ 13000-13099

APPENDIX 12



ADVANCED SEARCH >

SPIE DIGITAL LIBRARY

CONFERENCE PROCEEDINGS

PAPERS

PRESENTATIONS

JOURNALS

EBOOKS

Search Digital Library



[Home](#) > [BrowseVolume](#)

Browse Proceedings

Browse our growing collection of more than 590,000 conference proceedings papers.

RECENT SPIE CONFERENCES

BY YEAR

BY VOLUME NUMBER

13000-NOW

12000-12999

11000-11999

10000-10999

9000-9999

8000-8999

7000-7999

6000-6999

5000-5999

4000-4999

3000-3999

2000-2999

1000-1999

1-999

5900-5999

Volume	Volume Title
5900	Optics for EUV, X-Ray, and Gamma-Ray Astronomy II
5901	Solar Physics and Space Weather Instrumentation
5902	Focal Plane Arrays for Space Telescopes II
5903	Astronomical Adaptive Optics Systems and Applications II
5904	Cryogenic Optical Systems and Instruments XI
5905	Techniques and Instrumentation for Detection of Exoplanets II
5906	Astrobiology and Planetary Missions
5907	Photonic Devices and Algorithms for Computing VII
5908	Optical Information Systems III
5909	Applications of Digital Image Processing XXVIII
5910	Advanced Signal Processing Algorithms, Architectures, and Implementations XV
5911	Photorefractive Fiber and Crystal Devices: Materials, Optical Properties, and Applications XI
5912	Operational Characteristics and Crystal Growth of Nonlinear Optical Materials II
5913	Signal and Data Processing of Small Targets 2005
5914	Wavelets XI
5915	Mathematics of Data/Image Coding, Compression, and Encryption VIII, with Applications
5916	Mathematical Methods in Pattern and Image Analysis
5917	Fourth Generation X-Ray Sources and Optics III
5918	Laser-Generated, Synchrotron, and Other Laboratory X-Ray and EUV Sources, Optics, and Applications II
5919	Soft X-Ray Lasers and Applications VI
5920	Ultrafast X-Ray Detectors, High-Speed Imaging, and Applications
5921	Advances in Metrology for X-Ray and EUV Optics
5922	Hard X-Ray and Gamma-Ray Detector Physics VII
5923	Penetrating Radiation Systems and Applications VII

5924	<u>Complex Mediums VI: Light and Complexity</u>
5925	<u>Nanophotonic Materials and Systems II</u>
5926	<u>Tuning the Optical Response of Photonic Bandgap Structures II</u>
5927	<u>Plasmonics: Metallic Nanostructures and Their Optical Properties III</u>
5928	<u>Plasmonic Nano-imaging and Nanofabrication</u>
5929	<u>Physical Chemistry of Interfaces and Nanomaterials IV</u>
5930	<u>Optical Trapping and Optical Micromanipulation II</u>
5931	<u>Nanoengineering: Fabrication, Properties, Optics, and Devices II</u>
5932	<u>Strongly Correlated Electron Materials: Physics and Nanoengineering</u>
5933	<u>Buildings for Nanoscale Research and Beyond</u>
5934	<u>Nonlinear Optical Transmission and Multiphoton Processes in Organics III</u>
5935	<u>Linear and Nonlinear Optics of Organic Materials V</u>
5936	<u>Liquid Crystals IX</u>
5937	<u>Organic Light-Emitting Materials and Devices IX</u>
5938	<u>Organic Photovoltaics VI</u>
5939	<u>Organic Holographic Materials and Applications III</u>
5940	<u>Organic Field-Effect Transistors IV</u>
5941	<u>Fifth International Conference on Solid State Lighting</u>
5942	<u>Nonimaging Optics and Efficient Illumination Systems II</u>
5943	<u>X-ray and Neutron Capillary Optics II</u>
5944	<u>Smart Imagers and Their Application</u>
5945	<u>14th Slovak-Czech-Polish Optical Conference on Wave and Quantum Aspects of Contemporary Optics</u>
5946	<u>Optical Materials and Applications</u>
5947	<u>Liquid Crystals: Optics and Applications</u>
5948	<u>Photonics Applications In Industry and Research IV</u>
5949	<u>Nonlinear Optics Applications</u>
5950	<u>Photonic Crystals and Fibers</u>
5951	<u>Optical Fibers: Technology</u>
5952	<u>Optical Fibers: Applications</u>
5953	<u>Acousto-optics and Photoacoustics</u>
5954	<u>Optical Security Systems</u>
5955	<u>Metamaterials</u>
5956	<u>Integrated Optics: Theory and Applications</u>
5957	<u>Infrared Photoelectronics</u>
5958	<u>Lasers and Applications</u>
5959	<u>Medical Imaging</u>
5960	<u>Visual Communications and Image Processing 2005</u>
5961	<u>Building European OLED Infrastructure</u>
5962	<u>Optical Design and Engineering II</u>
5963	<u>Advances in Optical Thin Films II</u>
5964	<u>Detectors and Associated Signal Processing II</u>
5965	<u>Optical Fabrication, Testing, and Metrology II</u>
5966	<u>Seventh International Symposium on Optical Storage (ISOS 2005)</u>
5967	<u>2004 Shanghai International Conference on Laser Medicine and Surgery</u>
5968	<u>Laser Florence 2004: A Window on the Laser Medicine World</u>
5969	<u>Photonic Applications in Biosensing and Imaging</u>

- [5970 Photonic Applications in Devices and Communication Systems](#)
- [5971 Photonic Applications in Nonlinear Optics, Nanophotonics, and Microwave Photonics](#)
- [5972 Advanced Topics in Optoelectronics, Microelectronics, and Nanotechnologies II](#)
- [5973 Current Research on Laser Use in Oncology: 2000-2004](#)
- [5974 International Conference on Charged and Neutral Particles Channeling Phenomena](#)
- [5975 Topical Problems of Nonlinear Wave Physics](#)
- [5976 Remote Sensing for Agriculture, Ecosystems, and Hydrology VII](#)
- [5977 Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2005](#)
- [5978 Sensors, Systems, and Next-Generation Satellites IX](#)
- [5979 Remote Sensing of Clouds and the Atmosphere X](#)
- [5980 SAR Image Analysis, Modeling, and Techniques VII](#)
- [5981 Optics in Atmospheric Propagation and Adaptive Systems VIII](#)
- [5982 Image and Signal Processing for Remote Sensing XI](#)
- [5983 Remote Sensing for Environmental Monitoring, GIS Applications, and Geology V](#)
- [5984 Lidar Technologies, Techniques, and Measurements for Atmospheric Remote Sensing](#)
- [5985 International Conference on Space Information Technology](#)
- [5986 Unmanned/Unattended Sensors and Sensor Networks II](#)
- [5987 Electro-Optical and Infrared Systems: Technology and Applications II](#)
- [5988 Electro-Optical Remote Sensing](#)
- [5989 Technologies for Optical Countermeasures II; Femtosecond Phenomena II; and Passive Millimetre-Wave and Terahertz Imaging II](#)
- [5990 Optically Based Materials and Optically Based Biological and Chemical Sensing for Defence II](#)
- [5991 Laser-Induced Damage in Optical Materials: 2005](#)
- [5992 25th Annual BACUS Symposium on Photomask Technology](#)
- [5993 Advanced Environmental, Chemical, and Biological Sensing Technologies III](#)
- [5994 Chemical and Biological Sensors for Industrial and Environmental Security](#)
- [5995 Chemical and Biological Standoff Detection III](#)
- [5996 Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality](#)
- [5997 Environmentally Conscious Manufacturing V](#)
- [5998 Sensors for Harsh Environments II](#)
- [5999 Intelligent Systems in Design and Manufacturing VI](#)

- [+ 5800-5899](#)
- [+ 5700-5799](#)
- [+ 5600-5699](#)
- [+ 5500-5599](#)
- [+ 5400-5499](#)
- [+ 5300-5399](#)
- [+ 5200-5299](#)
- [+ 5100-5199](#)
- [+ 5000-5099](#)

Conference Papers
Conference Presentations
Journals
eBooks
About
Subscriptions

Proceedings Authors
Journal Authors
eBook Authors

Reviewer Guidelines
Reviewer Training Program

Information for Librarians

Resources
Subscriptions

TECHNICAL SUPPORT
spiedsupport@spie.org

CUSTOMER SERVICE
+1 360 676 3290

Hours:
8:00 am to 5:00 pm PST
[Help Center](#) | [Contact Us](#)



SPIE.

[SPIE Privacy Policy](#) | [Terms of Use](#)

© 2025 SPIE

APPENDIX 13



[Home](#) > [BrowseVolume](#)

Browse Proceedings

Browse our growing collection of more than 590,000 conference proceedings papers.

RECENT SPIE CONFERENCES

BY YEAR

BY VOLUME NUMBER

13000-NOW

8000-8999

3000-3999

12000-12999

7000-7999

2000-2999

11000-11999

6000-6999

1000-1999

10000-10999

5000-5999

1-999

9000-9999

4000-4999

- +

5900-5999
- +

5800-5899
- +

5700-5799
- +

5600-5699
- +

5500-5599
- +

5400-5499
- +

5300-5399
- +

5200-5299
- +

5100-5199
- +

5000-5099

Site Map

- Home
- Conference Papers
- Conference Presentations
- Journals
- eBooks
- About
- Subscriptions

Information for Authors

- Proceedings Authors
- Journal Authors
- eBook Authors

Information for Reviewers

- Reviewer Guidelines
- Reviewer Training Program

Information for Librarians

- Resources
- Subscriptions

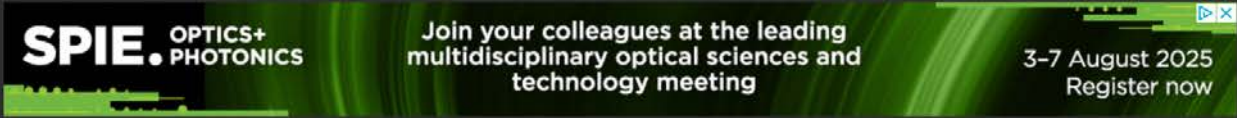
Contact & Support

TECHNICAL SUPPORT
spiedlsupport@spie.org
CUSTOMER SERVICE
+1 360 676 3290
Hours:
8:00 am to 5:00 pm PST
[Help Center](#) | [Contact Us](#)

Connect



APPENDIX 14



Home > Proceedings > Volume 5909

PROCEEDINGS VOLUME 5909

OPTICS AND PHOTONICS 2005 | 31 JULY - 4 AUGUST 2005

Applications of Digital Image Processing XXVIII

Editor(s): [Andrew G. Tescher](#)

[Editor Affiliations +](#)

IN THIS VOLUME

8 Sessions, 74 Papers, 0 Presentations, 0 Posters

- [Image Models and Processing](#) (9)
- [Next-Generation Multimedia Compression, Transmission, and Architectures](#) (6)
- [Imaging Security](#) (4)
- [MPEG-7: The Power of Standard Multimedia Metadata and Content Representation](#) (6)
- [H.264/AVC Applications](#) (8)
- [Image Coding and Structures](#) (9)
- [Content Management and Representation](#) (10)
- [Posters - Wednesday](#) (22)

☒ [Subscribe to Digital Library](#)



[VIEW ALL ABSTRACTS +](#)

IMAGE MODELS AND PROCESSING

Automatic inspection of pavement cracking distress Paper

[B. Xu, Y. Huang D.D.S.](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590901 (2005) <https://doi.org/10.1117/12.613770>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY



Paper

Olivier Ruch, Jean-Yves Dufour

Proceedings Volume Applications of Digital Image Processing XXVIII, 590902 (2005) <https://doi.org/10.1117/12.615053>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Hausdorff probabilistic feature analysis in SAR image recognition



Paper

John A. Saghir, Chessa Guillas

Proceedings Volume Applications of Digital Image Processing XXVIII, 590903 (2005) <https://doi.org/10.1117/12.623699>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Adaptive noise filtering of white-light confocal microscope images using Karhunen-Loeve expansion



Paper

M. Balasubramanian, S. S. Iyengar, P. Wolenski, J. Reynaud, R. W. Beuerman

Proceedings Volume Applications of Digital Image Processing XXVIII, 590904 (2005) <https://doi.org/10.1117/12.615252>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A detection method of mura on a coated layer using interference light



Paper

Kazutaka Taniguchi, Kunio Ueta, Shoji Tatsumi

Proceedings Volume Applications of Digital Image Processing XXVIII, 590905 (2005) <https://doi.org/10.1117/12.616504>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Quadratic correlation filter-based target tracking in FLIR image sequences



Paper

A. Bal, M. S. Alam

Proceedings Volume Applications of Digital Image Processing XXVIII, 590906 (2005) <https://doi.org/10.1117/12.617869>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Minimax distance transform correlation filter-based target detection in FLIR imagery



Paper

J. F. Khan, M. S. Alam, R. R. Adhami, S. M. A. Bhuiyan

Proceedings Volume Applications of Digital Image Processing XXVIII, 590907 (2005) <https://doi.org/10.1117/12.618578>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A new two-level efficient technique for 2D image patches registration via optimized cross correlation, wavelet transform, and moments of inertia with application to ladar imaging



Paper

Carlos Bejar, Dalila B. Megherbi

Proceedings Volume Applications of Digital Image Processing XXVIII, 590908 (2005) <https://doi.org/10.1117/12.620187>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Image fusion using DFT-based digital filter banks



Paper

Murat Sezgin, Isin Erer, Okan K. Ersoy

Proceedings Volume Applications of Digital Image Processing XXVIII, 590909 (2005) <https://doi.org/10.1117/12.616660>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

NEXT-GENERATION MULTIMEDIA COMPRESSION, TRANSMISSION, AND ARCHITECTURES

ISMuS: interactive, scalable, multimedia streaming platform



Paper

Jihun Cha, Hyun-Cheol Kim, Seyoon Jeong, Kyuheon Kim, Charalampos Patrikakis, Mihaela van der Schaar

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090A (2005) <https://doi.org/10.1117/12.623709>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Joint wavelet packets for groups of frames in MCTF



Paper

Maria Trocan, Christophe Tillier, Beatrice Pesquet-Popescu

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090B (2005) <https://doi.org/10.1117/12.623716>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

MPEG-21 DIA-based video adaptation framework and its application to rate adaptation



Paper

Jung Won Kang, Jae-Gon Kim, Dong-San Jun, Jin-Woo Hong

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090C (2005) <https://doi.org/10.1117/12.623716>

[Read Abstract +](#)

DOWNLOAD PAPER

An effective MP4 streaming method using the schedule information of image objects Paper

Seyoon Jeong, Jihun Cha, Hyun-Cheol Kim, Kyuheon Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090E (2005) <https://doi.org/10.1117/12.623815>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Performance of scalable video coding with respect to differentiated service requirements Paper

Doug Young Suh, Gwang Hoon Park, Seung Gyun Han, Hyun-Cheol Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090F (2005) <https://doi.org/10.1117/12.623819>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Network support mechanisms for scalable media streaming Paper

Charalampos Z. Patrikakis, Pantelis Karamolegkos, Ignatios Mihailaris, Alexis Lampiris, Hyun-Cheol Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090G (2005) <https://doi.org/10.1117/12.624537>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

IMAGING SECURITY

Study and validation of tools interoperability in JPSEC Paper

V. Conan, Y. Sadourny, K. Jean-Marie, C. Chan, S. Wee, J. Apostolopoulos

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090H (2005) <https://doi.org/10.1117/12.623822>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Image replica detection based on support vector classifier Paper

Y. Maret, F. Dufaux, T. Ebrahimi

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090I (2005) <https://doi.org/10.1117/12.623823>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Supporting secure transcoding in JPSEC Paper

John G. Apostolopoulos, Susie J. Wee

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090J (2005) <https://doi.org/10.1117/12.623825>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Scrambling for anonymous visual communications Paper

Frederic Dufaux, Touradj Ebrahimi

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090K (2005) <https://doi.org/10.1117/12.623826>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

MPEG-7: THE POWER OF STANDARD MULTIMEDIA METADATA AND CONTENT REPRESENTATION

MPEG-7: standard metadata for multimedia content Paper

Wo Chang

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090L (2005) <https://doi.org/10.1117/12.623824>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Easy and effective indexing and browsing schemes for digital photo album application Paper

Sang-Kyun Kim, Seungji Yang, Kyongsok Seo, Yong Man Ro, Ji-Yeun Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090M (2005) <https://doi.org/10.1117/12.623828>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones Paper

Ryoma Oami, Eiji Kasutani, Akio Yamada

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090N (2005) <https://doi.org/10.1117/12.623829>

[Read Abstract +](#)

[DOWNLOAD PAPER](#)
[SAVE TO MY LIBRARY](#)

Characterizing region of interest in image using MPEG-7 visual descriptors Paper

[Min-Sung Ryu](#), [Soo-Jun Park](#), [Chee Sun Won](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090O (2005) <https://doi.org/10.1117/12.623830>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Overall framework on digital image searching Paper

[Mun-Kew Leong](#), [Joo-Hwee Lim](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090P (2005) <https://doi.org/10.1117/12.623832>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Emedding MPEG-7 metadata within a media file format Paper

[Wo Chang](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090Q (2005) <https://doi.org/10.1117/12.623834>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

H.264/AVC APPLICATIONS

Some applications of H.264/AVC fidelity range extensions Paper

[Pankaj Topiwala](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090S (2006) <https://doi.org/10.1117/12.623852>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Rate-adaptive H.264 video for information for global reach (IFGR) Paper

[Praveen Kota](#), [Karthik Kannan](#), [Zixiang Xiong](#), [Pankaj Topiwala](#), [David Hench](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090U (2005) <https://doi.org/10.1117/12.623855>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Comparative study of JPEG2000 and H.264/AVC FExt I-frame coding on high-definition video sequences

 Paper

[Pankaj Topiwala](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090V (2005) <https://doi.org/10.1117/12.623856>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Origins of the performance of H.264/AVC: an account of the development of H.263 and H.264 coding tools

 Paper

[Cliff Reader](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090W (2005) <https://doi.org/10.1117/12.637515>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

AVC/H.264 patent portfolio license Paper

[Lawrence A. Horn](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090Y (2005) <https://doi.org/10.1117/12.623859>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Efficient temporal search range prediction for motion estimation in H.264 Paper

[Changsung Kim](#), [C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59090Z (2005) <https://doi.org/10.1117/12.613084>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

An operational rate control scheme for H.264 with two-stage encoding Paper

[Do-Kyoung Kwon](#), [Mei-Yin Shen](#), [C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590910 (2005) <https://doi.org/10.1117/12.613833>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Efficient bit stream switching of H.264 coded video Paper

[Xiaosong Zhang](#), [C.-C. Jay Kuo](#)

DOWNLOAD PAPER

[Read Abstract +](#)

SAVE TO MY LIBRARY

IMAGE CODING AND STRUCTURES

Progressive geometry coding of 3D meshes using hierarchical vertex set split Paper

[Jingliang Peng, C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590912 (2005) <https://doi.org/10.1117/12.613076>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A stereo video coding scheme based on H.264 Paper

[Jong Dae Oh, C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590914 (2005) <https://doi.org/10.1117/12.616116>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Adaptive cubic convolution interpolation and sequential filtering for color demosaicing of Bayer pattern image sensors Paper

[Wonpil Yu](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590915 (2005) <https://doi.org/10.1117/12.615852>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A point-based hybrid cloud modeling technique Paper

[Bei Wang, C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590916 (2005) <https://doi.org/10.1117/12.616328>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Edge directed image interpolation with Bamberger pyramids Paper

[Jose Gerardo Rosiles](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590917 (2005) <https://doi.org/10.1117/12.617828>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Applications based on restored satellite images Paper

[D. Arbel, S. Levin, M. Nir, I. Bhasteker](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590919 (2005) <https://doi.org/10.1117/12.616643>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Thermogram analysis and classification in breast cancer diagnostics Paper

[M. Zaviscek, A. Drastich, K. Dvorak M.D.](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091A (2005) <https://doi.org/10.1117/12.617051>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Multifocus fusion with multisize windows Paper

[R. Redondo, F. Sroubek, S. Fischer, G. Cristóbal](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091B (2005) <https://doi.org/10.1117/12.619263>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Relaxed color reproduction under different illuminant conditions Paper

[Li Liu, Yongyi Yang, Henry Stark](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091C (2005) <https://doi.org/10.1117/12.621331>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

CONTENT MANAGEMENT AND REPRESENTATION

Dynamic fluid animation using scalable grids Paper

[Youngmin Kwak, Chang-Suk Kim, C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091D (2005) <https://doi.org/10.1117/12.614280>

DOWNLOAD PAPER

[Read Abstract +](#)

SAVE TO MY LIBRARY

Video rate color region segmentation for mobile robotic applications Paper

[Aymeric de Cabrol](#), [Patrick J. Bonnin](#), [Vincent Hugel](#), [Pierre Blazevic](#), [Maryline Chetto](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091E (2005) <https://doi.org/10.1117/12.614543>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Temporally optimized edge segmentation for mobile robotics applications Paper

[Aymeric de Cabrol](#), [Patrick J. Bonnin](#), [Vincent Hugel](#), [Kamel Boucheffa](#), [Pierre Blazevic](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091F (2005) <https://doi.org/10.1117/12.614548>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Separating patterns and finding the independent components of mixed signals based on non-Gaussian distribution properties Paper

[Robert S. Rand](#), [Hao Chen](#), [Pramod K Varshney](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091G (2005) <https://doi.org/10.1117/12.615427>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Model based recognition using 3D line sets and multidimensional Hausdorff distance Paper

[M. T. Rahman](#), [M. S. Alam](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091H (2005) <https://doi.org/10.1117/12.618424>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Data image fusion using combinatorial maps Paper

[Kamel Boucheffa](#), [Patrick Bonnin](#), [Aymeric De Cabrol](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091I (2005) <https://doi.org/10.1117/12.615712>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Towards parameter-free classification of sound effects in movies Paper

[Sellina Chu](#), [Shrikanth Narayanan](#), [C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091J (2005) <https://doi.org/10.1117/12.616217>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Object detection and segmentation in camouflaged environments Paper

[S. Makrogiannis](#), [M. Trujillo San-Martin](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091K (2005) <https://doi.org/10.1117/12.618357>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A feature-based image registration algorithm using the multi-resolution approach combined with moments of inertia, with application to ladar imaging Paper

[Carlos Bejar](#), [D. B. Megherbi](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091L (2005) <https://doi.org/10.1117/12.624113>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Automatic image-driven segmentation of cardiac ventricles in cine anatomical MRI Paper

[Chris A. Cocosco](#), [Wiro J. Niessen](#), [Thomas Netsch](#), [Evert-jan P. A. Vonken](#), [Max A. Viergever](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091M (2005) <https://doi.org/10.1117/12.625575>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

POSTERS - WEDNESDAY

A fuzzy approach to supervised segmentation parameter selection for object-based classification Paper

[Travis L. Maxwell](#), [Yun Zhang](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091O (2005) <https://doi.org/10.1117/12.614435>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Investigation of crop nitrogen content based on image processing technologies Paper

[Yane Zhang](#), [Minzan Li](#), [Zenghui Xu](#), [Xijie Zhang](#), [Maohua Wang](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091Q (2005) <https://doi.org/10.1117/12.614682>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Estimating soil moisture based on image processing technologies Paper

[Lihua Zheng](#), [Minzan Li](#), [Jianying Sun](#), [Xijie Zhang](#), [Peng Zhao](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091R (2005) <https://doi.org/10.1117/12.614684>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Use of a local regularity analysis by a wavelet analysis for glitch detection Paper

[C. Ordénovic](#), [C. Surace](#), [B. Torrèsanj](#), [A. Liebaria](#), [J. P. Baluteau](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091U (2005) <https://doi.org/10.1117/12.614966>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Estimating growth status of winter wheat based on aerial images and hyperspectral data Paper

[Yunxia Han](#), [Minzan Li](#), [Liangliang Jia](#), [Xijie Zhang](#), [Fusuo Zhang](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091V (2005) <https://doi.org/10.1117/12.615717>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Impulsive noise detection and removal with the use of iterative filter based on gradient adaptive neighborhoods Paper

[Mikhail Mozerov](#), [Vitaly Kober](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091W (2005) <https://doi.org/10.1117/12.615727>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Recognition of partially occluded objects using correlation filters with training Paper

[J. Angel González-Fraga](#), [Vitaly Kober](#), [Josue Álvarez-Borrego](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091X (2005) <https://doi.org/10.1117/12.615796>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A fast predictive block matching motion estimation algorithm based on spatio temporal correlation information

 Paper

[Humaira Nisar](#), [Tae-Sun Choi](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091Y (2005) <https://doi.org/10.1117/12.615803>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A reduced color approach to high quality cartoon coding Paper

[Yi-Chen Tsai](#), [Ming-Sui Lee](#), [Meiyin Shen](#), [C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 59091Z (2005) <https://doi.org/10.1117/12.617424>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Tempo and tension analysis for MTV-style home video authoring Paper

[Shen-Zheng Wang](#), [Shih-Hung Lee](#), [C.-C. Jay Kuo](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590920 (2005) <https://doi.org/10.1117/12.617542>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Real-time pattern recognition with adaptative correlation filters Paper

[Victor H. Díaz-Ramírez](#), [Vitaly Kober](#), [Josue Álvarez-Borrego](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590921 (2005) <https://doi.org/10.1117/12.617815>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Estimation of the point response function for undersampled data for the infrared camera of the Spitzer Space

Telescope  Paper

[David Makovoz](#), [Patrick Lowrance](#)

Proceedings Volume Applications of Digital Image Processing XXVIII, 590922 (2005) <https://doi.org/10.1117/12.617953>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Detecting and measuring human bodies over wide work fields by using stereo cameras Paper

Jian Lu, Kyoko Hamajima, Subarna Lata Tuladhar

Proceedings Volume Applications of Digital Image Processing XXVIII, 590923 (2005) <https://doi.org/10.1117/12.618156>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Target detection and tracking in airborne video imagery using statistical snake and mean shift Paper

Shuqun Zhang, Mo Chen

Proceedings Volume Applications of Digital Image Processing XXVIII, 590924 (2005) <https://doi.org/10.1117/12.618171>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Preprocessing and compression of digital holographic images Paper

Shuqun Zhang, Mo Chen

Proceedings Volume Applications of Digital Image Processing XXVIII, 590925 (2005) <https://doi.org/10.1117/12.618599>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

3D watermarking scheme in stereo vision system Paper

Dong-Choon Hwang, Kyung-hoon Bae, Jung-Hwan Ko, Eun-Soo Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 590928 (2005) <https://doi.org/10.1117/12.619602>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A bi-directional stereo matching algorithm based on adaptive matching window Paper

Kyung-Hoon Bae, Dong-Sik Yi, Seung-Cheol Kim, Eun-Soo Kim

Proceedings Volume Applications of Digital Image Processing XXVIII, 590929 (2005) <https://doi.org/10.1117/12.611560>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

An adaptive image enhancement algorithm and real time implementation for an Infrared Imaging system

 Paper

Ajay Kumar, S. Sarker, R. P. Agarwal

Proceedings Volume Applications of Digital Image Processing XXVIII, 59092A (2005) <https://doi.org/10.1117/12.613571>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

An adaptive wavelet-based deblocking algorithm for MPEG-4 codec Paper

Trieu-Kien Truong, Shi-Huang Chen, Rong-Yi Jhang

Proceedings Volume Applications of Digital Image Processing XXVIII, 59092B (2005) <https://doi.org/10.1117/12.624367>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A novel face recognition system using the binary phase-only filter via optimal correlation thresholding Paper

D. B. Megherbi, S. Rastogi

Proceedings Volume Applications of Digital Image Processing XXVIII, 59092C (2005) <https://doi.org/10.1117/12.624467>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Effect of illumination correction on face recognition rates for the binary phase-only filter (BPOF) and the extended normalized correlation methods Paper

Dalila Megherbi, Soumitra Rastogi

Proceedings Volume Applications of Digital Image Processing XXVIII, 59092D (2005) <https://doi.org/10.1117/12.624468>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

A generic and high performance FPGA based signal processing architecture for staring thermal imaging system

 Paper

Ajay Kumar, S. Sarker, R. P. Agarwal

Proceedings Volume Applications of Digital Image Processing XXVIII, 59092E (2005) <https://doi.org/10.1117/12.613572>

[Read Abstract +](#)

DOWNLOAD PAPER

SAVE TO MY LIBRARY

Site Map

[Home](#)
[Conference Papers](#)
[Conference Presentations](#)
[Journals](#)
[eBooks](#)
[About](#)
[Subscriptions](#)

Information for Authors

[Proceedings Authors](#)
[Journal Authors](#)
[eBook Authors](#)

Information for Reviewers

[Reviewer Guidelines](#)
[Reviewer Training Program](#)

Information for Librarians

[Resources](#)
[Subscriptions](#)

Contact & Support

TECHNICAL SUPPORT
spiedsupport@spie.org
CUSTOMER SERVICE
+1 360 676 3290

Hours:
8:00 am to 5:00 pm PST
[Help Center](#) | [Contact Us](#)

Connect



SPIE.

[SPIE Privacy Policy](#) | [Terms of Use](#)

© 2025 SPIE

APPENDIX 15

SPIE. PHOTONICS WEST
17-22 January 2026

Share your research

SPIE.

SPIE. DIGITAL LIBRARY

SPIE. CAREER CENTER

ophcs.org

Sign In

View Cart

Help

ADVANCED SEARCH >

SPIE. DIGITAL LIBRARY

CONFERENCE PROCEEDINGS

PAPERS

PRESENTATIONS

JOURNALS >

EBOOKS

Search Digital Library



Paper

16 September 2005

Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones

Ryoma Oami, Eiji Kasutani, Akio Yamada

Author Affiliations +

Proceedings Volume 5909, Applications of Digital Image Processing XXVIII; 59090N (2005)

<https://doi.org/10.1117/12.623829>

Event: Optics and Photonics 2005, 2005, San Diego, California, United States

ARTICLE

CITED BY ▾

Abstract

This paper proposes a Mobile Multimedia Library (MML), a new application of information retrieval using camera-equipped mobile phones. MML allows users to get information of an unknown object anywhere and anytime. A user simply takes a picture of the object by a mobile phone and sends it directly to an MML server on which the picture is analyzed to identify the object. The MML server returns the identification result to the mobile phone and the user can browse information of the object on the display of the mobile phone. This application employs the k -Nearest Neighbor approach using eight MPEG-7 visual features to identify objects. A prototype system for animal identification has been developed to demonstrate effectiveness of the MML framework. It takes about ten seconds on average for the identification in which 63% and 79% queries were correctly identified within first four and first ten candidates, respectively, out of 229 categories.

© (2005) COPYRIGHT Society of Photo-Optical Instrumentation Engineers (SPIE). Downloading of the abstract is permitted for personal use only.

Citation [Download Citation ▾](#)

Ryoma Oami, Eiji Kasutani, and Akio Yamada "Mobile multimedia library: an MPEG-7 application with camera-equipped mobile phones", Proc. SPIE 5909, Applications of Digital Image Processing XXVIII, 59090N (16 September 2005); <https://doi.org/10.1117/12.623829>

PROCEEDINGS
9 PAGES

DOWNLOAD PAPER

SAVE TO MY LIBRARY

GET CITATION

CITATIONS

Cited by 10 patents.

[Explore citations on Lens.org](#)

Advertisement

A Tougher Compact Hexapod

- Fast delivery
- Many models in stock

Advertisement

SPIE. CAREER CENTER

Specialized, streamlined, and seamless job searching

Find work that works for you >

ACCESS THE FULL ARTICLE

ORGANIZATIONAL

Sign in with credentials provided by your organization.

Organizational Username

Organizational Password

Page 97 of 112

PURCHASE THIS CONTENT

SUBSCRIBE TO DIGITAL LIBRARY

50 downloads per 1-year subscription

Members: \$195

Non-members: \$335

ADD TO CART

25 downloads per 1-year subscription

Members: \$145

Non-members: \$250

ADD TO CART

ORGANIZATIONAL SIGN IN

INSTITUTIONAL

Select your institution to access the SPIE Digital Library.

SELECT YOUR INSTITUTION

PERSONAL

Sign in with your SPIE account to access your personal subscriptions or to use specific features such as save to my library, sign up for alerts, save searches, etc.

PERSONAL SIGN IN

No SPIE Account? [Create one](#)

PURCHASE SINGLE ARTICLE

Includes PDF, HTML & Video, when available

Members: \$17.00

Non-members: \$21.00

ADD TO CART

RIGHTS & PERMISSIONS

[Get copyright permission](#) 

KEYWORDS

[Cell phones](#)

[Cameras](#)

[Visualization](#)

[Prototyping](#)

[Feature extraction](#)

[Multimedia](#)

[System identification](#)

[Show All Keywords](#)

RELATED CONTENT

Mobile-based biology edutainment application for secondary schools

Proceedings of SPIE (March 14 2013)

Multi-user smartphone-based interaction with large high-resolution displays

Proceedings of SPIE (February 04 2013)

Mapping low-level image features to semantic concepts

Proceedings of SPIE (January 01 2001)

A novel methodology for querying web images

Proceedings of SPIE (January 17 2005)

Browsing through rapid-fire imaging: requirements and industry initiatives

Proceedings of SPIE (December 20 1999)

Agent-like support for content-based retrieval and navigation

Proceedings of SPIE (December 23 1997)

 [Subscribe to Digital Library](#)

 [Receive Erratum Email Alert](#)

Site Map

[Home](#)
[Conference Papers](#)
[Conference Presentations](#)
[Journals](#)
[eBooks](#)
[About](#)
[Subscriptions](#)

Information for Authors

[Proceedings Authors](#)
[Journal Authors](#)
[eBook Authors](#)

Information for Reviewers

[Reviewer Guidelines](#)
[Reviewer Training Program](#)

Information for Librarians

[Resources](#)
[Subscriptions](#)

Contact & Support

TECHNICAL SUPPORT
spiedlsupport@spie.org

CUSTOMER SERVICE
+1 360 676 3290

Hours:
8:00 am to 5:00 pm PST

[Help Center](#) | [Contact Us](#)

Connect



SPIE.

[SPIE Privacy Policy](#) | [Terms of Use](#)

© 2025 SPIE

APPENDIX 16

Mobile Multimedia Library: an MPEG-7 application with camera-equipped mobile phones

Ryoma Oami*, Eiji Kasutani, Akio Yamada

Media and Information Research Laboratories, NEC Corporation

1753, Shimonumabe, Nakahara-Ku, Kawasaki, Kanagawa 211-8666, Japan

ABSTRACT

This paper proposes a Mobile Multimedia Library (MML), a new application of information retrieval using camera-equipped mobile phones. MML allows users to get information of an unknown object anywhere and anytime. A user simply takes a picture of the object by a mobile phone and sends it directly to an MML server on which the picture is analyzed to identify the object. The MML server returns the identification result to the mobile phone and the user can browse information of the object on the display of the mobile phone. This application employs the *k*-Nearest Neighbor approach using eight MPEG-7 visual features to identify objects. A prototype system for animal identification has been developed to demonstrate effectiveness of the MML framework. It takes about ten seconds on average for the identification in which 63% and 79% queries were correctly identified within first four and first ten candidates, respectively, out of 229 categories.

Keywords: MPEG-7, CBIR, content based image retrieval, camera-equipped mobile phone, multimedia library, *k*-Nearest Neighbor

1. INTRODUCTION

Mobile phones have evolved into ubiquitous communication and information retrieval tools from mere handsets just for telephoning. E-mail exchange and Internet connection functionalities are outstanding features that have spurred this evolution. Many people nowadays exchange e-mails with mobile phones instead of computers thanks to the mobility. Information retrieval services through the Internet are also frequently utilized with mobile phones. On-demand train timetable/transfer information search and restaurant guides are examples of services frequently used in Japan. These services are available anywhere and anytime just only using mobile phones without any preparation in advance. Mobile phones are hence considered to be very valuable terminals despite their inconvenience of inputting characters with a limited number of small buttons.

A camera equipped on recent mobile phones is one of the solutions to improve this inconvenience. The camera is originally incorporated just to simply enjoy taking pictures and sharing them with other people through e-mails, but the camera has become to draw attention as an input means. An OCR (Optical Character Reader) functionality is, for example, installed to automatically input such text strings as URLs and phone numbers to mobile phones without typing them. Barcode reading functionality is another example developed to input text embedded in two dimensional barcodes. Current usage of the camera is just to provide an alternative way to input text strings, but the camera has much more promising potential as an information input means.

Object appearance is substantial information acquired by the camera. This inspires us with an idea that the camera can also be exploited to take unknown objects for acquiring their information. The camera functions as an input means of non-verbal, visual information triggering information retrieval. When a person happens to encounter an unknown flower during trekking, for example, s/he can immediately obtain the flower information just by taking a picture of it. CBIR (Content-Based Image Retrieval) technologies help us to achieve this retrieval.

The research area of CBIR is now very active and many applied systems have been proposed so far. QBIC¹ (Query By Image Content), VisualSEEK², and MARS⁴ are examples of such CBIR systems. In these systems, a query image is provided for searching similar images from an image database. Similarity is measured based on low-level image features representing color, shape and texture. Some of them accept sketch images as a query instead of natural images^{1,9}, and

* r-oami@az.jp.nec.com; phone +81-44-431-7710; fax +81-44-431-7590

keywords as well. CBIR technologies are extended to web-oriented systems, such as WebSEEk³ and RetrievalWare⁵. CBIR applications are now being investigated in various fields, and their application area is still growing.

A new application with camera-equipped mobile phones, Mobile Multimedia Library (MML), is proposed in this paper. This is an information retrieval application that enables us to get information from a picture of an object taken by a camera on mobile phones. The picture is analyzed with a CBIR technology using MPEG-7 visual features^{11,13}, including the color layout description proposed by us⁹. This application inherently solves a general problem of query-by-example for CBIR, that is, how to obtain a proper query to feed CBIR systems. Two major issues are also solved for its practical use.

One of them is restriction on computational resources of mobile phones. Object identification is computationally intensive unlike OCR or barcode reading functionality, so that it is quite difficult to perform it only using the resources on mobile phones. The identification is, therefore, performed on a server in place of resource-restricted mobile phones. E-mail exchange and internet connection functionalities of mobile phones are exploited to exchange information between mobile phones and the server. The other issue is semantic gaps between capability of the current CBIR systems and ability of human recognition⁷. Human interaction is hence incorporated to bypass the semantic gaps.

In Section 2, the concept of MML and the system architecture are depicted. The algorithm of image analysis and MPEG-7 visual features used for this application is then described in Section 3. A prototype developed for identifying animals and its evaluation results are presented in Section 4 and 5. Other possible MML applications are shown in Section 6, which illustrates the usefulness of the MML framework.

2. CONCEPT OF MOBILE MULTIMEDIA LIBRARY AND ITS ARCHTECTURE

2.1. Concept of Mobile Multimedia Library

Mobile Multimedia Library (MML) is an information retrieval application that provides information from a picture of an object taken by a camera on mobile phones, anywhere and anytime. The object is identified using the picture, and the results are presented on a screen of mobile phones.

The advantages of this application are as follows:

- (1) No necessity of carrying encyclopedia

This is an overall advantage the application has. A single camera-equipped mobile phone is enough to acquire information of target objects anywhere and anytime. No other preparation is necessary.

- (2) Easy input of a visual query

This is an advantage of using camera-equipped mobile phones. Taking a picture of a target object is the only required operation, which is quite easy so that many people including children have already managed a camera on mobile phones. The picture is directly used as a visual query.

- (3) No necessity of background knowledge of objects

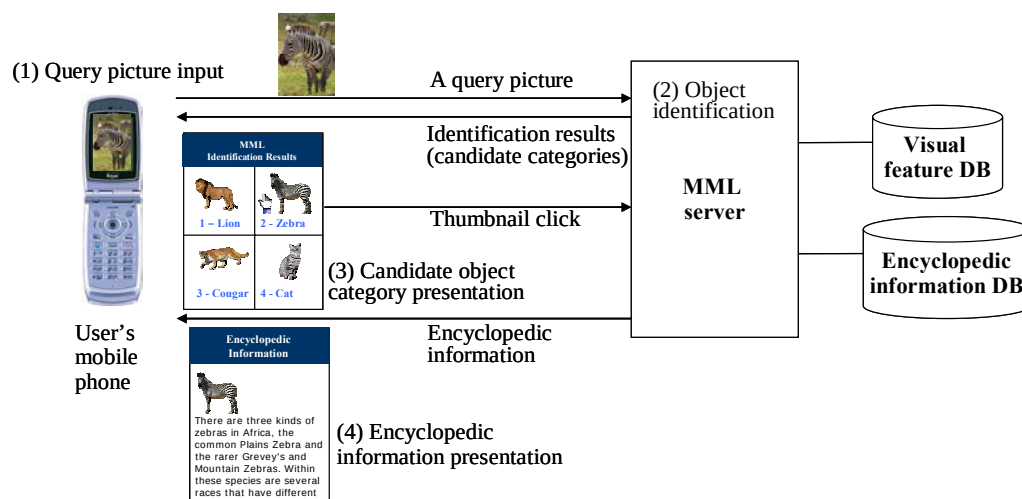


Figure 1: System overview of MML.

This is an advantage CBIR has. The MML application allows us to search objects just from their appearance. No knowledge about taxonomy, which is often needed to consult encyclopedias, is required for getting information of target objects.

(4) Customized search and presentation of retrieved information

This comes from the characteristics of a mobile phone as a private terminal. User preference information, which is stored implicitly in private data such as e-mails or explicitly as a user profile in mobile phones, can be used for this customization. For example, presented information can be customized according to the level of his/her knowledge or interest.

2.2. System overview

The MML application is implemented in a form of a client-server type application that consists of camera-equipped mobile phones and a remote server, which is called MML server, to retrieve information. Object identification process is performed on the MML server instead of resource-restricted mobile phones for its practical use, since the identification process usually needs much computational power.

The system architecture of the MML is shown in Figure 1. Information of an object is retrieved by the MML in the following manner.

1. query picture input,
2. object identification based on query analysis,
3. interactive selection of a proper candidate, and
4. information presentation for selected object category

In Step 1, the user takes a picture of the unknown object by a camera-equipped mobile phone. This picture is uploaded to a remote MML server as a query. In Step 2, the query picture is analyzed on the MML server to identify the object in the picture. MPEG-7 visual features are extracted from the query and compared to the features of training images stored in a visual feature DB. As a result, several object categories are derived as candidates based on the identification results. In the following of the paper, the term “candidate object category” is used to refer object categories selected as candidates that the unknown object can belong to. In Step 3, a web page containing the identification results is presented on the screen of the mobile phone. This page contains representative images corresponding to each candidate object category. Each representative image is linked to encyclopedic information of the corresponding object category. In Step 4, one of the candidates is selected by clicking one of the representative images, and the encyclopedic information about the clicked object category is displayed on the user’s mobile phone.

The MML server has two types of databases: a visual feature DB and an encyclopedic information DB, as shown in Figure 1. The former DB contains visual features of training images, which are samples included in each object category and used for query analysis to identify the object. The latter DB has encyclopedic information to be provided to the user.

3. ALGORITHM FOR OBJECT IDENTIFICATION

3.1. Overview

The object identification process on the MML server in Figure 1 is depicted in Figure 2. Visual features are first extracted from a query picture. The distance values between the query features and the training image features in the feature DB are next computed for each elementary feature description standardized as MPEG-7, such as the color layout and the edge histogram. Total distance between two images is then computed by summing the individual distance values that are computed and normalized for each feature description. The obtained total distance values are finally fed to object identification process to determine candidate object categories for the query. In the following subsections, database configuration on the MML server is described. Each step is then described in more detail.

3.2. Visual feature database on the MML server

In the visual feature DB in Figure 2, visual features for plenty of training images are stored for each object category together with its representative image. This is for coping with different appearance and variation. Appearance of objects varies with their viewing angle and capturing conditions, which results in different feature values. Some object categories have several variations, that is, entities that have inherently different appearance even within the same object category (e.g. male lions have manes, but female ones do not). Whereas increasing the number of training images for each

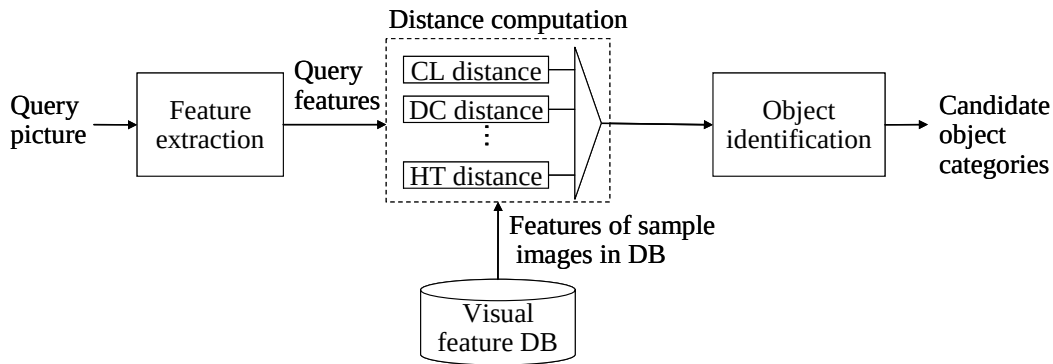


Figure 2: Diagram for query analyzing process to find candidate object categories.

object category improves completeness, this makes processing time longer and this does not always result in improving identification results. A proper number of the training images should be determined in consideration of this trade-off.

3.3. Feature extraction

The feature extraction process in Figure 2 extracts various visual features from the query picture. Several elementary visual feature descriptions defined in MPEG-7¹³ are employed as listed in Table 1 because of their discriminating potential. The interoperability of the features is also their advantage in building an extensible retrieval system. This makes it possible to exchange features and matching results among different systems.

For the object information retrieval application, we employed the following eight elementary feature descriptions: four color features (the color layout, the color structure, the dominant color, and the scalable color), two texture features (the edge histogram, and the homogeneous texture), and two shape features (the contour shape, and the region shape). These features can be extracted automatically, except that object segmentation masks necessary for computing shape features are manually generated. A future work is to incorporate one of conventional automatic segmentation methods into the feature extraction process. The dimensionality of each elementary feature is listed in Table 1. Please see the Appendix for more detailed specification.

3.4. Distance computation between the query image and training images

Distance values are computed between each pair of the query and the training image stored in the visual feature DB. The computed distance values are fed to object identification process.

In the distance computation for each pair, distance values for the eight elementary feature described in Section 3.3 are first individually computed. Distance functions as described in MPEG-7 Part 8¹⁴ are employed for this computation.

The obtained eight individual distance values are then normalized to equalize the range of distance values among different elementary features. The normalization is performed with the mean and the standard deviation of distance distribution in the visual feature DB. Distance values within $[-3\sigma, 3\sigma]$ are linearly mapped onto the interval $[0, 1]$ and the values outside the range are represented with the closest end point, that is, zero or one.

The normalized distance values for individual elementary features are then summed up with a set of certain weighting coefficients to calculate the total distance. The optimal weighting coefficients were derived in a learning process with a set of training images. The combination of eight coefficients was exhaustively searched in the learning process.

3.5. Object identification by k -Nearest Neighbor approach

Object identification is performed using the total distance values obtained in the previous subsection. k -Nearest Neighbor approach¹², which gave the best accuracy among several methods we tried, is taken for this identification. The total distance values are first sorted in the ascendant order, and the number of samples for each object category included in the first k rank is counted. The

Table 1: MPEG-7 elementary visual features used for object identification

Visual Features	Dimensionality
Color Layout	18 dims
Color Structure	256 bins
Dominant Color	≤ 8 colors
Scalable Color	256 bins
Edge Histogram	80 bins
Homogeneous Texture	$2 + 60$ dims
Contour Shape	$\leq (3 + 62)$ dims
Region Shape	35 dims

object categories with high frequency are finally selected as candidates after sorted by the frequency.

4. PROTOTYPE SYSTEM

A prototype system for identifying animals was developed as an example to demonstrate the proposed MML framework. This section describes this animal identification system.

4.1. Prototype system for animal identification

A third-generation mobile phone was used in this prototype system. This has a camera whose resolution is about 300 thousand pixels and its screen size is 33 mm x 45 mm. A computer with a Pentium 3GHz CPU, which has an ADSL connection to the Internet, was utilized for the MML server.

For uploading a query picture from the mobile phone to the MML server, e-mail exchange function of the mobile phone was exploited, that is, the picture was enclosed in an e-mail as an attachment, and sent to the server. On the server side, a program module periodically checking a designated mail account was implemented, and e-mails were immediately downloaded and pictures were separated from the e-mails when they were found. The pictures are then provided to the identification process in Figure 2.

A process for animal identification is launched by the e-mail checking module, and a web page showing identification results are generated. Details about the web page design are described in Section 4.2. Terminal adaptation related to screen size was performed in generating the web page. The size of the representative images is adjusted to the screen size so that all the candidates can be seen without scrolling according to the types of mobile terminals. An environment variable that notifies a platform and a browser type was checked to distinguish different terminals.

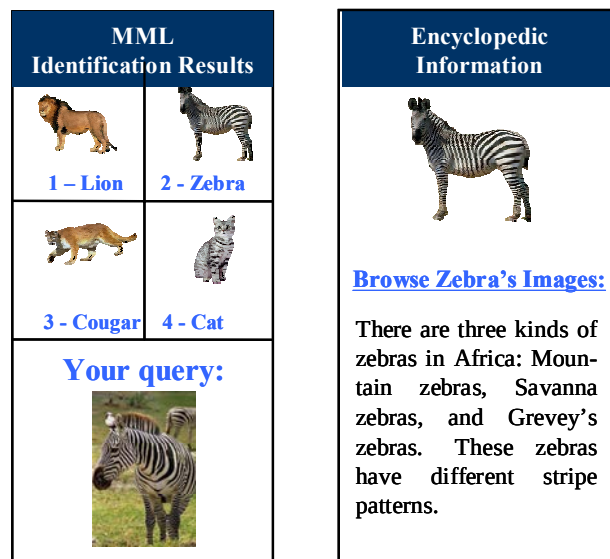
The DB for encyclopedic information for each animal was generated from some sources available on the Internet. The information such as in-depth descriptions of animals and their living environment was included.

4.2. Web page design for showing animal candidates

The web page showing identification results is designed to include representative images for first four candidate categories as shown in Figure 3 (a). This is because “four” is considered to be the maximum number to show the representative images simultaneously without losing details necessary for distinguishing the animals. It is certain that more than four candidates can be displayed with smaller images, but this sometimes makes it hard to differentiate the candidate objects at a glance on a small screen of a mobile phone. In the case of PDAs or smart phones that have a larger screen, more candidates like first ten can be displayed. The query image is also displayed at the bottom just for confirmation.

The web page is designed so as to minimize user operation. The first candidate animal category is focused when the web page is presented on the mobile phones. Hence if the first candidate is the animal category a user wants to know, just one click action is enough to show the information. The focus is easily moved by pressing buttons corresponding to moving focus, so that the user can select one of the representative images.

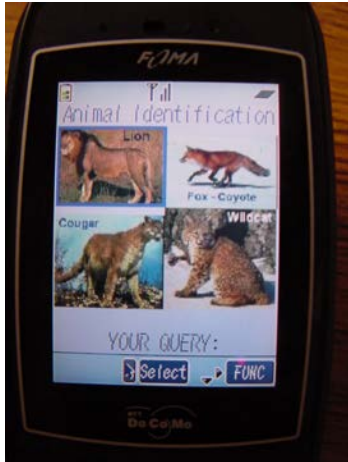
Encyclopedic information as shown in Figure 3 (b) is displayed on the screen of the mobile phone by clicking one of the four images. The second image is assumed to be clicked in this example. This web page contains the representative image of this category, a link to a page including various images belonging to the same category, and description regarding this animal category. The link is included for checking other images of this category as well to confirm that this is the correct animal category. Thus users can easily get informa-



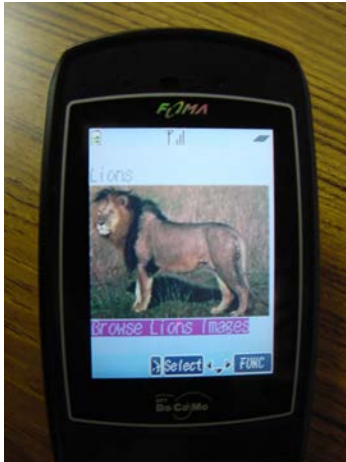
(a) Web page showing identification results.

(b) Encyclopedic information for the animal selected in (a).

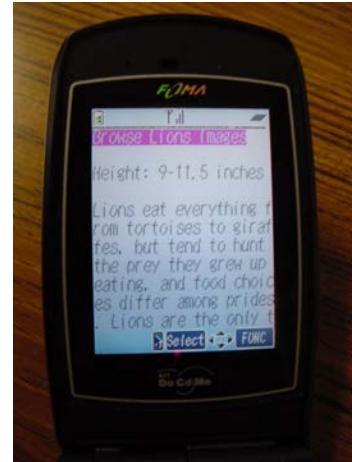
Figure 3: Web pages for identification results and encyclopedic information that is shown after one of the candidate animal categories is selected.



(a) The web Page that shows identification results.



(b) Retrieved information by clicking the picture of the first candidate "lion" in (a).



(c) Detailed descriptions about "lion". These descriptions are shown by scrolling the screen.

Figure 4: Screen shots presented to a user by the prototype system.

tion of animals using the MML application.

5. PROTOTYPE DEMONSTRATION AND EXPERIMENTAL RESULTS

5.1. Demonstration of the prototype system

Figure 4 shows screen shots presented to a user by the prototype system. The left most image shows the web page displaying the first four candidates. If the user clicks the top-left lion image, the information about lion is obtained as shown in the middle picture. When scrolling the screen, detailed descriptions about lions appear as shown in the right-most image. The highlighted description of "Browse Lions Images" is for browsing other lion images.

It was confirmed using this prototype that all the operations required to users, from taking a picture of an animal to browse information displayed, can smoothly be done with only a single hand as designed. The processing time for animal identification is about ten seconds on average, which would be acceptable in ordinary use cases.

5.2. Evaluation of identification algorithm

Experiments for evaluating the accuracy of the animal identification algorithm were carried out. A feature database storing features for 2681 training images from 229 animal species was built. The query images for evaluation were 600, including 335 pictures actually taken by a camera at a zoo in Japan. Weighting coefficients used for weighting individual distance values of each MPEG-7 description tools were optimized using training images included in the feature DB.

Figure 5 shows the identification results of the identification algorithm for several different number k of nearest neighbors. Each curve corresponds to a different k , ranging from 20 to 100. The horizontal axis indicates the rank to which the correct result is included. Hence the accuracy increases monotonically.

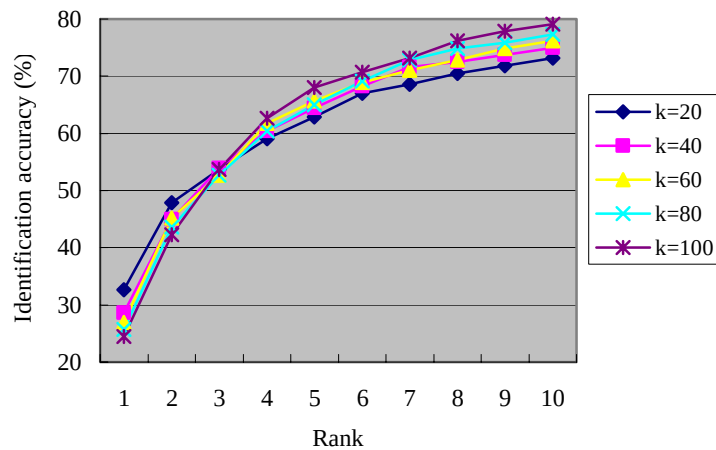


Figure 5: Identification accuracy for several k values.

cally with the rank value.

The accuracy is 63% when first four candidate animal categories are provided, which occurs at $k = 100$. If ten candidates can be displayed, which is the case for smart phones with a larger screen, the accuracy increases up to 79%. Currently, more than 80% of the total species included in the feature DB belong to the dog category, which means many animals with similar appearance are included. If only one candidate is shown to a user, the accuracy is about 33% at most and the best performance is achieved when $k = 20$. This implies the importance of human interaction.

The accuracy is expected to increase if more different types of animals are included with the DB. According to this result, the parameter k was set to 100 in the prototype system mentioned in Section 4.1.

6. OTHER POSSIBLE APPLICATION SCENARIOS FOR MML

The framework of the MML is widely applicable, not limited to animal identification system described above. This framework is quite general, so that many application examples are considered in the same framework. The followings are possible examples for which the MML framework can effectively be applied.

- Flower identifying systems
- Building/monument guide systems
- Museum guide systems
- Product information retrieval systems with logo identifier
- Navigation systems to a nearest store with logo identifier

The flower identifying systems are very similar to the animal identification mentioned in this paper. In this case, location and time information, which can be obtained by GPS (Global Positioning System) or the clock of mobile phones, is also effectively utilized as well as visual features to identify flowers. The building/monument guide systems are also similar to the animal identification, but location information becomes much more relevant. The museum guide systems are used in museums or art galleries to retrieve information of paintings and statues.

The product information retrieval systems are a little different from the previous three examples, which are just for retrieving information of the objects in themselves. In this application, taking a logo of a store makes it possible to access its product information such as price for each product sold there or electric coupons can be retrieved. On the other hand, the nearest store corresponding to the identified logo is searched in the navigation systems, and navigation service to the store is provided with a map and location information obtained by GPS.

The MML framework can, furthermore, be applied to conventional CBIR systems in themselves. One of the problems of the conventional CBIR systems is how to acquire an appropriate query. Camera-equipped mobile phones can be employed to solve this problem.

7. CONCLUSION

A Mobile Multimedia Library, an MPEG-7 application with camera-equipped mobile phones has been proposed in this paper. Mobile phones are exploited as means for inputting a query image, uploading the image to an MML server, and browsing information returned from the MML server. The MML server analyzes the query image to identify the object a user wants to know, and then the identification result with their related information is sent to the user and presented. MPEG-7 visual features are used for identifying the objects. The developed animal identification prototype system demonstrated the usefulness of the MML.

REFERENCES

1. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: the QBIC system," IEEE Computer Magazine, Vol. 28, No. 9, Sep. 1995.
2. J. R. Smith and S-F Chang, "VisualSEEk: A fully automated content-based image query system," in Proc. ACM Multimedia 96, 1996.
3. J. R. Smith and S-F. Chang, "Visually searching the web for content," IEEE Multimedia Magazine, Vol. 4 No. 3, pp. 12-20, July-Sept. 1997.
4. Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in Proc. of International Conference on Image Processing (ICIP), pp. 26-29, Oct. 1997.
5. J. Dowe, "Content-based retrieval in multimedia imaging," in Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993.

6. Y. Rui, T. S. Huang, and S-F. Chang, "Image retrieval: current techniques, promising directions and open issues," *Journal of Visual Communication and Image Representation*, Vol. 10, pp. 39-62, March 1999.
7. A. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 22, No. 12, Dec. 2000.
8. C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, Vol. 24, No. 8, Aug. 2002.
9. E. Kasutani and A. Yamada, "The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval," in *Proc. of International Conference on Image Processing (ICIP)*, Vol. 1, pp. 674-677, Oct. 2001.
10. V. Castelli and L. D. Bergman, "Image Databases –Search and Retrieval of Digital Imagery—," John Wiley & Sons, Inc., 2002.
11. B. S. Manjunath, P. Salembier and T. Sikora, "Introduction to MPEG-7," John Wiley & Sons, Ltd., 2002.
12. R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Classification," 2nd Ed, John Wiley & Sons, Inc., 2001.
13. "Information Technology –Multimedia content description interface—Part 3: Visual," ISO/IEC 15938-3:2002, 2002.
14. "Information Technology –Multimedia content description interface—Part 8: Extraction and use of MPEG-7 descriptors," ISO/IEC TR 15938-8:2002, 2002.

APPENDIX: MPEG-7 VISUAL FEATURES FOR MML

The following eight MPEG-7 features are utilized in the prototype system of MML for animal identification.

Color Layout

This feature expresses brief color layout information for rectangular regions. This is extracted by computing DCT for 8x8 small image generated by shrinking an entire rectangular region to this size. Low frequency components are selected and quantized to yield the elements of the color layout. The feature data size is usually very small, typically 8 bytes for an rectangular region, but it has a very high discriminative capability. 18 coefficients (six coefficients for all Y, Cb and Cr components) are used in our system.

Color Structure

This feature expresses both color distribution and a local color structure together. An $N \times N$ (typically 4×4) structuring element is slid over a target rectangular region for capturing the color structure. The frequency of each color appearing within the element is counted in the HMMD color space. The total number of bins is 256 in the current system.

Dominant Color

This feature describes dominant colors (up to eight colors) for a rectangular region. Region segmentation is first applied in order to find dominant colors by grouping pixels with similar colors. The portion of each segmented region is also considered when computing distances between two features.

Scalable Color

This is basically color histogram in the HSV color space, which is encoded by a Haar Transform. The main feature of this is its scalable characteristics in terms of the number of bins. This scalability is achieved by Haar Transform for the histogram. The number of bins is 256, which means hue, saturation and value are quantized into 16, 4, and 4 levels, respectively.

Edge Histogram

This feature describes edge histograms for 16 blocks after a rectangular region is divided into the 16 blocks. The histogram is calculated for four directions (horizontal, vertical and two diagonal directions) and for non-directional edges. The number of bins is, therefore, 80 in total.

Homogeneous Texture

This feature expresses characteristics of texture information using mean energy and energy deviation, which are computed for each frequency channel obtained by Gabor functions. The number of channels is 30 designated by six levels in angular direction and 5 levels in radial direction. The number of dimensions is 62, including the mean and the standard deviation for an entire rectangular region.

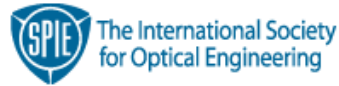
Contour Shape

This expresses shape information by describing contour of a region based on CSS (Curvature Scale-Shape) representation. The CSS representation decomposes the contour into convex and concave sections by the points with zero curvature. The circularity and eccentricity are derived from the CSS representation, and the magnitude of the largest peak, and the position of the remaining peaks are included in the feature.

Region Shape

This represents a binary shape mask by decomposing it using orthogonal 2-D basis functions defined by the Angular Radial Transform. This can express shapes consisting of multiple disconnect regions and shapes with some holes. Total number of dimensions is 35, which is determined by 12 levels in angular direction and 3 levels in radial direction.

APPENDIX 17



Furthering Innovation in Optics and Photonics

SPIE HOME	PUBLICATIONS	CONFERENCES	EXHIBITIONS	MEMBERSHIP	EDUCATION
---------------------------	------------------------------	-----------------------------	-----------------------------	----------------------------	---------------------------

SPIE BOOKSTORE	DIGITAL LIBRARY	JOURNALS	PROCEEDINGS	SPIE PRESS	MAGAZINES	AUTHOR INFORMATION
--------------------------------	---------------------------------	--------------------------	-----------------------------	----------------------------	---------------------------	------------------------------------

BY YEAR	BY VOLUME	BY CONFERENCE
-------------------------	---------------------------	-------------------------------

 SEARCH PUBLICATIONS » »
 USING InCITE®

☒ **Volumes**
☐ **Papers**

 ADVANCED
SEARCH

 VIEW
CART

 BROWSE
PUBLICATIONS

- [Nanotechnology](#)
- [Defense & Security](#)
- [Aerospace, Remote Sensing, & Astronomy](#)
- [Automation, Inspection, & Product Engineering](#)
- [Biomedical Optics](#)
- [Communications & Fiber Optics](#)
- [Electronic Imaging, Displays, & Medical Imaging](#)
- [Lasers & Applications](#)
- [Microelectronics, Optoelectronics, & Micromachining](#)
- [Optical Physics, Chemistry, & Biology](#)
- [Optical Science & Engineering](#)
- [Signal & Image Processing](#)

Publications

PROCEEDINGS

Proceedings of SPIE begins e-First publication

Editor-reviewed papers from post-meeting Proceedings of SPIE will be available within two to four weeks after each symposium as PDF documents downloadable from the [SPIE Digital Library](#) (SDL). The new e-First publication model begins with meetings held after June 2005.

Print volumes will be ready to ship by six weeks after the meeting, and collections of proceedings on CD-ROM by eight weeks after the meeting. Papers from on-site Proceedings will be available in the SDL at the start of the symposium.

In the e-First model, the date the paper appears in the SDL is the official publication date for that paper.

SPIE Digital Library: for fastest access to editor-reviewed papers

The SPIE Digital Library is the largest and most extensive resource available on optics and photonics in the world. Researchers get unprecedented access to SPIE Proceedings and Journals from 1990 to the present—approximately 200,000 articles with 15,000 new articles added each year.

- [More information about the SPIE Digital Library.](#)

Printed Proceedings Volumes

If you are only interested in editor-reviewed papers from a single conference or want an archive of the conference that includes your paper, you can get the Yellow book faster than ever before; *within six weeks of the conference!*

Featured Proceedings:

- [View the Proceedings from Optics & Photonics and order online.](#)
- [View the Proceedings from Fluctuations and Noise, and order online.](#)
- [View the Proceedings from Optical Metrology and order online.](#)
- [View the Proceedings from Microtechnologies for the New Millennium and order online.](#)

Proceedings of SPIE on CD-ROM

If you are interested in searching editor-reviewed papers from multiple conferences and a broad topical area. You can search for specific areas of interest. CD-ROMs are now available *within 8 weeks of the meeting!* You no longer have to wait a long, long time to have the added value of the CD-ROM.

- [View the list of CDS volumes.](#)



RESOURCES

- [Conference Calendar](#)
- [Proceedings Author Information](#)
- [Due Dates & Editor Contacts](#)
- [Electronic Manuscript Submission](#)
- [Critical Reviews](#)
- [Contract Publishing Services \(pdf\).](#)

Selected SPIE Papers on CD-ROM

This series features edited collections of papers from the Proceedings of SPIE and SPIE scholarly journals on core subjects in optics, imaging, and photonics. All CD-ROMs are in Adobe PDF. Viewing, searching, and printing software is included.

- [View the list of CDP volumes.](#)

Critical Reviews

Invited review papers by leading researchers providing in-depth analysis and review in key optical technologies.

- [View the list of CR volumes.](#)

About Proceedings of SPIE

This widely-known, used, and referenced body of research contains original, previously unpublished work reflecting presentations at SPIE-sponsored events.

- Timely information about new research and development
- Fast conference to publication timeline
- Synergy of cross-disciplinary participation at SPIE Events
- Focus on applied technologies
- Snapshots of ongoing development
- Strong international participation provides a global view of current technologies
- Access to new information as it develops
- The latest innovations before they are published in a journal
- Content created by engineers and scientists from academia, industry, and research labs

Papers of SPIE Proceedings are indexed in leading scientific databases such as Physics Abstracts, Ei Compendex, Chemical Abstracts, INSPEC, International Aerospace Abstracts, Index to Scientific and Technical Proceedings, and published in the [SPIE Digital Library](#).

SPIE also publishes technical conference proceedings for other organizations that sponsor optics and optics-related conferences around the world, as part of its [Contract Publishing Services \(PDF file\)](#) program.

| [SPIE Home](#) | [Publications](#) | [Conferences](#) | [Exhibitions](#) | [Membership](#) | [Education](#) |

Telephone: +1 360/676-3290 | Fax +1 360/647-1445 | Email: spie@spie.org

© 1994– 2005 SPIE—The International Society for Optical Engineering

| [Privacy Policy](#) |

SPIE is a not-for-profit international society dedicated to advancing optics and photonics.