Links for ACM DL'97 Tutorial

Museums and Images

- The Museum Educational Site Licensing Project and local cache
- MESL Content Selection Cumulative Content Summary and local cache
- New Models for Distributing Digital Content, A paper for the Data Processing Clinic, University of Illinois, March 25, 1996, J. Trant, Getty Art History Information Program and local cache
- Enabling Educational Use of Museum Digital Materials: The Museum Educational Site Licensing (MESL) Project, by J. Trant, Getty Art History Information Program and local cache
- Jennifer Trant, esp. Recent Papers and Publications and local cache
- Introduction to Imaging, Issues in Constructing an Image Database, Howard Besser and Jennifer Trant and local cache
- Introduction to Imaging, Issues in Constructing an Image Database: Contents, Howard Besser and

Jennifer Trant and local cache
- MESL Related Resources and local cache
- Standards Framework for the Computer Interchange of Museum Information and local cache
- Preparing Quality Images for Computer Networks and local cache
- Image and Multimedia Database Resources (Besser) and local cache

Singapore

- Institute of Systems Science, NUS, Singapore and local cache
- ISS Visual Information Processing Group and local cache
- ISS VIPG projects: Robust Face Recognition System, Robust Fingerprint Recognition System, Web-based Multimedia Spatial Information System with query language, FACEit - a content-based face image archival and retrieval system, STAR - System for Trademark Archival and Retrieval, Video Processing facilitates manipulation, indexing and retrieval of MPEG encoded video, Personalized Information Network delivers personalized multimedia news and local cache

NSF/ARPA/NASA DLI

- DIGITAL LIBRARIES INITIATIVE, Funded through a joint initiative of NSF/DARPA/NASA and local cache
- The Alexandria Project and local cache
- Alexandria Project Tutorial Table of Contents and local cache
- Alexandria Digital Library Annual Report Table of Contents and local cache
- The Second Alexandria Design Review, Report on the Workshop, 19-21 February, 1997, Santa Barbara, CA and local cache
- Image Browsing in the Alexandria Digital Library (ADL) Project and local cache
- Universe: Spatial Data Available Over the Web and local cache
- Informedia Digital Video Library Technology Outreach and local cache
- Informedia Goes to School: Early Findings from the Digital Video Library Project and local cache
- CMU's Informedia Digital Video Library and local cache
- Edge Detector Demo and local cache
- CMU's Face Detector Demo and local cache
- Computer Vision Demos - worldwide and local cache
- CMU Image Database - data sets and local cache
- Informedia: News-on-Demand and local cache

- U. Illinois Chicago Database and Information Systems Laboratory and local cache
- U. Illinois Chicago Database and Information Systems Laboratory and local cache

Search

- QBIC(TM) -- IBM's Query By Image Content and local cache
- Virage Inc. VIR Technology Demo - find similar images and local cache
- Visual Information Retrieval: A Virage Perspective and PDF version and local cache and local PDF
- Finding Images/Video in Large Archives: Columbia's Content-Based Visual Query Project and local cache
- Webseek: A Content-Based Image and Video Search and Catalog Tool for the Web and local cache
- Content-based Multimedia Data Management and Efficient Remote Access, James Griffioen, Brent Seales, Raj Yavatkar and local cache
- Storage and Retrieval Techniques for Multimedia Data by P. Zezula and C. Galindo-Legaria (IEI-CNR) and PDF format and local cache and local PDF

Compression

- BIBLIOGRAPHY OF WAVELET AND WAVELET-RELATED DOCUMENTS; see also: Subject: Wavelet Digest, Vol. 4, Nr. 1. From: wavelet@math.scarolina.edu (Wavelet Digest) and local cache
- WAVELET PROJECT figures and local cache
- Fractal Image Encoding Announcements and Questions and local cache
- Fractal Image Encoding - resources and local cache
- A Survey of Compressed Domain Processing Techniques, Brian C. Smith and local cache

Misc.

- Image Description on the Internet: A Summary of the CNI/OCLC Image Metadata Workshop, September 24 - 25, 1996, Dublin, Ohio and local cache
- Reconnecting Science and Humanities in Digital Libraries, A Symposium Sponsored by The University of Kentucky and The British Library, 19-21 October 1995, Marriott's Resort at Griffin Gate, Lexington, KY and local cache
- ERCIM Computer Graphics Network Report and Recommendations from the VRML Workshop, 29/30 January 1997,The Coseners House, Abingdon, UK and local cache

- Specifying the PREMO Synchronization Objects, D.J. Duke, D.A.Duce, I. Herman, G.Faconti and PDF version and local cache and local PDF
- The Nordic Council for Scientific Information and local cache
- Macromedia Director and local cache

Education

- Multimedia Educational Materials and local cache
- Coalition on Reinventing Information Science, Technology and Library Education and local cache
- CRISTAL-ED Core Curriculum and local cache
- CRISTAL-ED Specializations and local cache
- U Mich. courses on Digital Tools and local cache
- Impact of New Information Resources: Multimedia and Networks, Winter 1995 Homepage (Besser) and local cache
- Evaluation: Besser Course - Student Essays and local cache

MESL CONTENT SELECTION working group

MESL Content Selection Cumulative Content Summary

The content selection period for the 1996 MESL distribution is now complete. This table provides access to the content summaries and lists for all museums for both distributions, as well as to their institutional descriptions. The lists currently linked in as text files are the definitive distribution lists provided by John Weise at Michigan.

cds, 8/25/96

MESL Particpating Museums

Museum Descriptions	1995 Distribution		1996 Distribution	
Fowler Museum of Cultural History	summary	list	summary	list
George Eastman House	summary	list	summary	list
Harvard University Art Museums	summary	list	summary	list
Library of Congress	summary	list	summary	list
Museum of Fine Arts, Houston	summary	list	summary	list
National Gallery of Art	summary	list	summary	list
National Museum of American Art	summary	list	summary	list

To view the content selection web archive, CLICK HERE.

George Eastman House

Description of the Museum

GEORGE EASTMAN HOUSE
PHOTOGRAPHY COLLECTIONS

Before 1949 several major photography collections had been assembled at the Eastman Kodak Research Laboratory, where their initial purpose was to support research in photographic science. The Eder Collection purchased by Kodak in 1922 from the prolific Viennese photo-historian, scientist, and teacher, Josef Maria Eder, and the Cromer Collection, purchased in 1939 from the widow of Parisian lawyer and collector Gabirel Cromer, formed the nucleus of the still photography collection that was later transferred to the home of George Eastman, founder of Eastman Kodak Company. (At his death in 1932 George Eastman had willed his home to the University of Rochester, which subsequently deeded it to the newly incorporated photography museum.)

Photography collections contain more than 400,000 prints and negatives; substantial holdings of 19th-century French and American photographs, including the world's most extensive daguerreotype collections; major 20th-century holdings, including substantial collections of work by Ansel Adams, Margaret Bourke- White, Alvin Langdon Coburn, Lewis W. Hine, Gertrude Kasebier, and Edward Steichen; and representative collections of commercial, scientific, and contemporary photography.

[MESL Content Selection]

Last Modified: Tuesday, 13-Aug-96 18:39:53 EDT

National Museum of American Art

Description of the Museum

The National Museum of American Art (NMAA) is the nation's museum dedicated to the arts and artists of the United States from the earliest colonial times to the present. NMAA's collection of more than 37,500 works of art--paintings, sculptures, prints, drawings, photographs, crafts, and folk art--is the largest collection of works by American artists in the world With its Renwick Gallery, devoted to the study and exhibition of American crafts and decorative arts, the museum addresses the full spectrum of American artistic pursuits from colonial times to the present.

A visit to NMAA's galleries offers an opportunity to experience over two hundred years of the nation's visual heritage from the eighteenth century to the present. Early highlights include masterworks of the colonial era and early republic, and exceptional portrait miniatures. Western art includes more than

480 paintings of Indian life by George Catlin, varied western subjects, and important works by Native American artists. Among nineteenth- century landscapes are views of the Hudson River Valley and Niagara Falls, epic-scale western vistas, and more. Nineteenth-century marble sculptures, Paul Manship's twentieth-century bronzes and major contemporary works highlight the museum's sculpture collection. Impressionism and the American Renaissance are represented by in-depth holdings of Childe Hassam, John Twachtman, Albert Pinkham Ryder, and Thomas Wilmer Dewing and masterworks by John Singer Sargent, Winslow Homer, James McNeill Whistler, Mary Cassatt, and others. The museum has the finest collection of works by African-American artists of any general museum, representing more than 100 artists.
Hispanic-American art includes Luis Jimnez's "Vaquero" sculpture at the museum's entrance, religious painting and sculpture, carvings, folk art, and contemporary works. The museum's folk art collection is anchored by the Herbert Waide Hemphill, Jr., collection of historic and contemporary folk art--more than 400 objects. Photography and graphic arts include documentary photographs and prints from the 1930s and 1940s, extensive holdings of photographs by Aaron Siskind and Irving Penn, contemporary photos, and graphics from the nineteenth and twentieth-centuries. With a rich collection of figurative

and scene paintings, including the nation's largest collection of art produced for New Deal-era projects, and the Patricia and Phillip Frost Collection of American abstraction, NMAA offers an excellent survey of twentieth century American art.

[MESL Content Selection]

Last Modified: Tuesday, 13-Aug-96 18:41:30 EDT

New Models for Distributing Digital Content

A paper for the Data Processing Clinic, University of Illinois
March 25, 1996.

J. Trant, Getty Art History Information Program

jtrant@io.org
http://www.io.org/~jtrant

Contents

I. Background: Intellectual Property Rights

- Rights in digital images complex
- Existing rights administration systems inefficient
- Museums must rethink rights administration process
- Current rights administration systems may

- Current rights administration systems may impede compliance
- Quality information at risk

II. MESL a response

- Rights holders require
- Rights users require

III. Models for IP Administration

- 'Pay per Bit' problematic
- Site License model for MESL
- Collective Administration of Rights

IV. Why test with museums and universities?

- Uses for museum information
- MESL a multi-institution collaboration
- Participating Museums
- Participating Universities

V. Progress to Date

- Cooperative Agreement
- Selection of test images
- Data Distribution
- Text Formats
- Image Formats
- Objectives

- Use on campus
- Evaluation
- Economics
- Goals

VI. Further Information

I. Background: Intellectual Property Rights

Rights in digital images complex

- legal framework slow to respond to technology
- major barrier to creation of image database systems
- management systems require balance between concerns of rights holders and rights users
- Rights may exist in...
  - original work of art
  - photographic reproduction
  - digital image
  - manipulated digital image
- Specific rights may be associated with each form/version
- Rights in different media are often disassociated
- Having rights in one media does not mean you

have rights in another

Existing rights administration systems inefficient

- based on single item negotiations
- unable to handle volume generated by digital use
- true costs for locating and using images unreasonably high

Museums must rethink rights administration process

- ease of digital copying removes gatekeeper role
- no longer control access to quality
- no longer able to enforce compliance
- system expensive to maintain

Current rights administration systems may impede compliance

- difficult to locate suitable images
- cost of using individual images high
- terms and conditions vary
- museums unable to supply high volume at reasonable cost
- high overhead throughout system

Quality information at risk

- top 10 syndrome
- content defined based on ease of access
- "ersatz reproductions" easier to use than "real thing"
- museum documentation sacrificed

II. MESL a response

- establish the terms and conditions for the educational use of museum images and associated information
- facilitate distribution of high quality information
- enable collaboration between rights holders and rights users

Rights holders require

- respect of original object
- integrity of information
- protection and security
- acknowledgement
- remuneration

Rights users require

- easy access to large body of material
- central source to locate content
- common terms and conditions
- reasonable fee structure
- effective and efficient administration

III. Models for IP Administration

- Pay per use/bit
- Site license

'Pay per Bit' problematic

- focus on individual image
- high monitoring requirements
- per-use charge may inhibit access
- discourages exploration of resource
- costs unpredictable; budgeting difficult
- shifts cost to individual

Site License model for MESL

- focus on collections of images
- reasonable monitoring requirements
- information free at point of use

- encourages exploration, serendipitous use
- predictable costs; no hidden charges

Collective Administration of Rights

- one-stop shopping for image users
- standard terms and conditions
- standard data delivery format
- efficient administration
- enables communication between rights holders and rights users

---

IV. Why test with museums and universities?

- museums have images and information
- universities have delivery systems
- both have educational mandate
- share common culture and interests

Uses for museum information

- teaching and research in the Humanities
    - art history, history, anthropology, cultural and religious studies

- multimedia development
  - distance learning, visual literacy, independent study, life-long learning
- information and computer science
  - image database research, access, description, search and retrieval, image processing

MESL a multi-institution collaboration

- selection based upon competative call for participation
- interdisciplinary project teams on each campus
- central coordination at AHIP
- management committee to advise
- January 1995 - June 1997

Participating Museums

- Fowler Museum of Cultural History at UCLA
- The George Eastman House, Rochester, NY
- The Harvard University Art Museums, Cambridge, MA
- The Library of Congress, Washington, DC
- The Museum of Fine Arts, Houston, TX
- The National Gallery of Art, Washington, DC
- The National Museum of American Art, Washington, DC

Participating Universities

- American University, Washington, DC
- Columbia University, New York, NY
- Cornell University, Ithaca, NY
- University of Illinois, Urbana-Champaign, IL
- University of Maryland, College Park, MD
- University of Michigan, Ann Arbor, Dearborn and Flint, MI
- University of Virginia, Charlottesville, VA

V. Progress to Date

- Cooperative Agreement signed
- image selection
- image distribution
- evaluation

Cooperative Agreement

- terms and conditions of experiment
  - educational use on campus network
  - includes research, teaching, student projects
  - no redistribution
  - no commercial use

- will form basis for model site license

Selection of test images

- over 4000 images distributed in spring 95
- additional set to be distributed spring 96
- challenge to create coherent data sets
    - support teaching
    - work with existing/planned digitization
    - negotiation to meet curricular needs

Data Distribution

- University of Michigan as "distribution central"
- each campus mounted own set of images
- delivery decisions based on local infrastructures

Text Formats

- Data Dictionary developed
    - object description as structured text
    - unstructured text as linked documents
        - conservation reports
        - exhibition history
        - bibliography
- Challenge to create consistency in merged data set

Image Formats

- quality as high as museums comfortable releasing
  - 758 x 512 through 1536 x 1024
  - 24 bit colour
- file formats
  - JFIF with JPEG
  - PhotoCD
  - TIFF
- sites determined image delivery formats

Objectives

- separate content from delivery systems
- learn about issues by studying choices
- compare delivery system designs with common data set
- acknowledge heterogeneous nature of data creation and delivery infrastructure

Use on campus

- critical to success of project
- required of participants
- broad range of activities
  - UMD, joint studio/art history project
  - UVA, Religious Studies
  - UMI, SILS

Evaluation

- statistics about use
- profiles of distribution systems
- assessment of interface choices
- examination of search capabilities

Economics

- costs and benefits of new technology
- costs of rights administration
- model self-sustaining system

Goals

- enable the educational use of museum digital assets
- propose a famework for the collective administration of museum intellectual property rights
- suggest a scalable system for implementation

VI. Further Information

- http://www.ahip.getty.edu/mesl

- Demonstration Tonight
  - Charles Bauer, University of Illinois
  - Beth Sandore, University of Illinois

Last Modified: Wed, Jun 19, 1996 12:36 PM

Return to my Home Page.

MESL ABOUT MESL

Enabling Educational Use of Museum Digital Materials:

The Museum Educational Site Licensing (MESL) Project

by J. Trant
Getty Art History Information Program

A paper for the Electronic Imaging and the Visual Arts Conference, Florence, Italy, February 8-9, 1996.

Contents

- Introduction
- Why Site Licenses?
- MESL Progress to Date
  - Cooperative Agreement
  - Image Selection
  - Data Distribution
  - Text

    - Images
    - Academic Use
- Evaluation and Measurement
- Economics of Information
- Conclusion
- Further Information
- Footnotes

Introduction

The use of digital imaging and network communications technologies offers the promise of bringing cultural heritage collections out of museums and into the public eye. When captured and stored in digital form, images of works of art and artefacts - along with their accompanying textual descriptions - can be used in new, exciting ways, placing works in context rather than isolating them. Networked information resources remove some of the physical barriers to the enjoyment of cultural heritage collections, making them available to wider audiences, including those who might never normally enter a museum building.

Few markets are more primed to make use of museum collections in digital form than the educational community, where a strong tradition in creating and exploiting networked information

resources exists (at least in the United States). College and university campuses are developing sophisticated networks, and enthusiastically embracing networked multi-media technologies. In the process, the nature of teaching and research is changing. For this transformation to be completed, however, a critical mass of digital information must exist, and it must be available in standard forms.

The creative development of interactive educational multi-media programs has been hampered, however, by difficulties in obtaining of quality content. Intellectual property rights agreement are needed which balance the interests of rights holders and the desires of those who use images for study, teaching and research. Such agreements must reflect a common understanding of the rights, permissions and restrictions associated with digital museum materials. A common framework for administering rights, which reflected widely accepted terms and conditions for the use of materials, would further the educational use of digital materials.

The Museum Educational Site Licensing Project (MESL) brings representative museums, colleges, and universities from the United States together to define the terms and conditions for educational use of museums' digital images and information on campus-wide networks. During this two-year

experiment, launched in 1995, a select group of educational and collecting institutions are collaborating in good faith to study the capture, distribution, and educational use of digital images and their associated texts.

The MESL project is a collaboration of seven collecting institutions and seven universities,1 which is defining the terms and conditions governing the educational use of digitized museum images and related information.1 MESL participants are developing a model educational site license, testing and evaluating procedures for the collection and distribution of museums' digital images and information, and assessing the impact of this distribution, in both technical and economic terms. At the end of this experiment, which will run for two-academic years (until June 1997) the participants will propose a broadly-based system that could support ongoing distribution and educational use of museum images and text.

Why Site Licenses?

Many schemes have been suggested to manage intellectual property on the "information highway". The majority of these have focused on the development of metering systems, that operate on a 'pay per bit' model. For several reasons, this premise

does not translate well into the educational or museum community. First, monitoring at this level of detail requires an investment in infrastructure and a commitment of systems resources which may be beyond the capability of many cultural or educational institutions. Secondly, charging for the use of information seems antithetical to intellectual exploration and learning, goals shared by both educational institutions and museums.

"Pay-per-bit" models may actually inhibit access to information. Who would be most penalized by limits on access? Both the student struggling to come to terms with the subject, and the scholar striving to develop an in-depth understanding of a particular area of research would generate disproportionately high charges. And who would be responsible for paying them? The individual themselves? The department? The institution? Would each user have a particular allowance? How would such an allocation be made? How can an institution budget for the acquisition of electronic resources, when there are neither set charges nor a ceiling on payments in place? How can finite resources be allocated when costs are unpredictable?

The site licensing model adopted by the MESL project addresses these concerns, and reflects the following principles:

- Information should be free at its point of use; hidden usage charges should be avoided. (Charging specific access fees would be similar to charging individuals each time they check-out library books from an institutional collection.)
- Costs for assembling collections of electronic resources should be borne as institutional expenditures. Passing fees on to individuals inhibits access to information and disregards the heritage of libraries, assembled as an information resource for the public good.
- Costs need to be predictable; libraries must be able to budget for the acquisition or use of electronic resources.
- Costs should be reasonable, and based on the costs of generating electronic resources; fees should be structured so that the widest possible use of resources is encouraged. While educational institutions may be able to pay for the use of digital information sources, they do not have the resources to bear the costs of conversion.
- Monitoring and security requirements should be reasonable.

The site-license model offers a means of satisfying these concerns. Through an annual subscription fee, an educational institution could gain access to a wide

body of quality electronic information about many museum collections. This virtual archive (or selections from it) could be distributed on a campus network and made available to the full campus community. No specific charges would be incurred by the users of the information. Security requirements would be similar to those of maintaining the campus network as a whole, and would not require monitoring at the individual-access level.

Site licenses could be administered by a not-for-profit entity which acts as an intermediary between museums and educational institutions. A shared administrative framework would ease both the costs of distribution and the administration of rights. A common agreement would remove the institutional burden of negotiating licenses on a case-by-case basis, a very labor intensive proposition. The licensing of material would provide a constant revenue stream for the museum community, which could be used to add to the available stock of digital images. Educational institutions would have access to a predictable supply of quality images and information about works in museum collections, and would also have the opportunity to act as information providers, making works from their own special collections and archives available for distribution. A self-sustaining distribution system could be established, which would operate for the benefit of

both participating museums and educational institutions.

MESL Progress to Date

The Museum Educational Site Licensing Project was structured to lay the ground-work for such a cooperative administration of the rights to use digital museum information for educational purposes. It is engaged in an experimental interchange of museum information, which enables the participants to identify and appreciate the issues involved, and propose a future direction based on collective experience.

Project participants were selected through a competitive Call for Participation issued in the fall of 1994. Since that time, over 4000 images and accompanying text records have been distributed. This interchange is governed by a Cooperative Agreement, signed by all participating institutions, which outlines the terms of the project and preliminary conditions of use.2 Programs for evaluating the impact of networked digital images on teaching, and assessing the economics of digital distribution of museum images are under development.

Inter-disciplinary project teams in each participating institution have contributed to the project's

achievements. At the universities, staff from departments including computer services, art history, instructional technology and the library have collaborated in the creation of local delivery systems. On the museum side, the preparation and delivery of images and data has involved teams from information technology, registration, photo services, education, and curatorial departments.

The majority of the work in MESL is conducted in task forces, centered around specific issues. Seven working groups have been formed:

- Documentation and Distribution;
- Monitoring and Security;
- Base Measurement;
- Evaluation;
- Faculty Training and Support;
- Content Selection; and the
- World Wide Web.

Each group is chaired by a MESL Project Team member, and has a Management Committee liaison.

The majority of MESL work takes place electronically.2 Each working group has its own listserv, for the discussion of specific issues. A general project listserv is used for updates on activities. WWW pages which document the work of

the project are available at http://www.ahip.getty.edu/mesl/home.html. This site contains both completed project documentation and work-in-progress. As a result, some sections are restricted to project participants. In recognition of the interest in the project, a listserv MESL-OBSERVE@AHIP.GETTY.EDU has been established to enable those not actively involved in the project to offer their comments and insights.

Cooperative Agreement

A Cooperative Agreement, outlining the terms of collaboration in the project, has been signed by all participating institutions.3 It offers the first articulation of 'acceptable use' under a possible site license. Images are made available for "educational use, including faculty research, teaching and student projects." Participants have agreed that images may be incorporated into student hypermedia projects, or faculty courseware. However, these products and the images themselves, are not to be redistributed beyond the participating educational institution's site, without the express written permission of the contributing museum. As the project progresses, a set of guidelines which further defines "educational use" will be developed. The MESL agreement has also provided a model for the participants' agreement signed in CIMI's Project CHIO.

Image Selection

The first distribution of MESL museum data was made to participating universities for the fall term of the 1995 academic year. Over 4,000 images and accompanying textual documentation have been made available on each of 7 university campuses. The selection of this test set of images involved a complex negotiation between participating museums and faculty who desired to use the images in teaching and research. The challenge was to minimize the demands placed on the museums for content, while satisfying the desires of the faculty for a coherent image set.

Initially, museums stated (at a general level) the digital images they had available, while universities stated (at a subject-level) their areas of interest. Where these overlapped, item level lists were provided to the universities for detailed review and selection. As museums were unable to commit significant resources to the creation of new digital image content specifically for the MESL project, universities had to chose from the slate of existing images, or from works for which photographic images existed that could be easily scanned.

Several other factors limited the selection of works

from MESL museums. Initially, images which posed intellectual property issues, such as works made by living artists, works on loan to museum collections and works which entered the collection with specific donor restrictions were ruled out. (These issues will be addressed once the mechanics of distribution itself are worked out). Museums also selected images which had existing textual documentation of sufficient quality to be distributed to the universities.

These restraints made creating coherent image sets, that supported the teaching goals of the participating faculty, a challenge. In some cases, MESL museums were able to reorganize digitization schedules to make specific works available. In other cases, documentation was reviewed and revised. Direct contact between museum curators and university faculty produced some exciting results; for example, through negotiation, the Fowler Museum of Cultural History was able to assemble an image set in African art which met the pedagogical needs of faculty at Columbia University.

In the second data distribution, planned for the spring of 1996, we are experimenting with the use of WWW-based tools to facilitate this negotiation. This work is being facilitated by the Digital Image Center at the University of Virginia.

Data Distribution

Once the images were selected, museums provided images and text in electronic form to the University of Michigan, which agreed to act as MESL's 'Distribution Central'. Michigan assembled the image and text sets from all seven museums, duplicated the merged data set, and distributed it on CD-ROM to the participating universities, and to several museums who were also interested in mounting the data. FTP was later used to distribute updates to the text files. Each university then devised a strategy for distributing the data on its campus network, which responds to local needs and infrastructure.

Text

To facilitate the creation of the merged data set, a MESL Data Dictionary was developed. This format, which looked closely at both the AITF Categories for the Description of Works of Art, and the evolving CIMI SGML DTD for museum objects, includes both structured and unstructured text.5

Structured museum text (i.e. text in database fields) is grouped into the following broad areas: Object Identification, Object Description; Credit Information; Pointers to Image Files; Pointers to Text Files; Image Capture Information; and Version Identification. This

information is exported from museums' collections management systems, into an agreed upon tagged ASCII text format.

Unstructured text files (i.e. free text or prose which exists in electronic form) includes such types of information as Exhibition History; Publication History; Curatorial Notes; Conservation History; Bibliography; Published Texts and Unpublished Texts. This information was included in the data transferred to the museums as a strong interest was expressed by the participating faculty for in-depth analysis of the objects. These files are sent in text format, along with the structured data files for the objects.

A revised version of the MESL Data Dictionary will include further examples, and include guidelines for the presentation of data, to begin to address issues of consistency in the merged data, including questions of character sets and data. In addition, the University of Michigan has agreed to play a data validation role in the second distribution, to ensure the structural integrity of all database records distributed.

Images

Images were distributed in the highest quality that the participating museums were willing to release. Color images were distributed in 24-bit (16.7 million colors)

and grayscale images in 8-bit (256 shades of gray) files, at resolutions ranging from 758 x 512 pixels to 1536 x 1024 pixels. Image files were in one of four file formats: JFIF, with minimum JPEG compression; PhotoCD, Lossless Compressed TIFF; and GIF (for line art).

High-quality images were distributed to each of the campuses. Each university then made its own decisions about the creation of derivative images, such as thumbnails or screen-size images, based on the requirements of their local delivery system. This has resulted in some redundant processing, and the next distribution of images may reassess this methodology.

The experimental nature of MESL, however, may mitigate against a centralized approach. One of the goals of the project is to examine the individual choices made by each of the participating universities. Each has been encouraged to developed delivery systems which reflected their specific network infrastructures and instructional needs. As a part of the evaluation exercise, we will begin to identify what commonalties exist between these implementations,.

Academic Use

One of the key objectives of MESL is introduce museum digital images and information into the teaching and research activities of participating universities. In the short time since the images have been available on campuses, a number of faculty have integrated MESL images into their teaching plans. Each participating university has committed to using the image set in at least one course in each academic year of the project.

As an example, at the University of Maryland, Dr. Sally Promey used images in an American Landscapes: Art and Technology course offered in the fall of 1995. The digital images were projected during classes, which are held in a specially equipped teaching theater. The digital images also provided a vehicle for the development of collaborative projects between students in Art History and those enrolled in the Art Studio course. An art history student provides a verbal description of a MESL image. This is passed to studio student, who then creates a new work based on the description. The two students then collaborate on an analysis of the two works of art.

At the University of Michigan, Dr. Howard Besser has students testing content-based retrieval systems and examining rights-holder protection schemes in an Image Databases course. Dr. Olivia Frost will be using MESL images for team-based learning in a

course on the organization of non-textual information resources.

Evaluation and Measurement

The availability of the MESL data on seven campuses provides a great opportunity to assess the impact of electronic resources on the teaching process. The MESL Evaluation Working Group is defining an evaluation and measurement program which will document and evaluate the distribution of museum images and information over campus networks. Baseline data, documenting teaching using photographic images, will be collected and compared with that gathered from a number of studies.

The impact of the use of images in classrooms and across campus networks for research and individual study will be measured and evaluated using a combination of qualitative and quantitative methods. Information gathered from observation, interviews and other interactions with users and potential users of the systems will be supplemented by statistics derived from system transaction monitoring, survey questionnaires and other statistical methods, to gain a full picture of how system use, instruction and learning are affected.

Economics of Information

MESL provides a testbed for the analysis of the economics of a future museum Rights and Reproductions Organization (RRO). Such a not-for-profit organization would be administered by rights holders to distribute museum content for educational uses, and could potentially become a vehicle for collective licensing for commercial uses as well. In order to be positioned to carry their activities forward after the end of their two-year project, MESL participants are undertaking research on the costs and benefits of such an organization. An inter-disciplinary team of academically-based researchers drawn from economics, business, arts management and information telecommunications systems will be assembled to identify and analyze the costs associated with the creation and distribution of digitized museum information. (David Bearman is reporting on the development of this study, elsewhere on the EVA program.)

Conclusion

The electronic distribution of information enables cultural institutions to reach a broad audience. This adoption of communications technologies could enable museums to enhance their educational

mission and re-establish their relevance, by delivering quality content which can be readily integrated into established and developing curriculum. To realize the potential of electronic access to cultural heritage, however, museums need to explore the nature and exploitation of their digital information resources. The MESL project is building our understanding of the educational use of museums' digital knowledge, and establishing the preconditions for the efficient and effective interchange of information between these two communities.

Further Information

The MESL WWW site can be found at http://www.ahip.getty.edu/mesl/home.html

Footnotes

1 Participating museums: Fowler Museum of Cultural History at the University of California, Los Angeles; George Eastman House, Rochester, New York; Harvard University Art Museums, Cambridge, Massachusetts; Library of Congress, Washington, DC; Museum of Fine Arts, Houston, Texas; National Gallery of Art, Washington, DC; and National Museum of American Art, Washington, DC. Participating universities: American University,

Washington, D.C.; Columbia University, New York; Cornell University, Ithaca; University of Illinois at Urbana-Champaign; University of Maryland at College Park; University of Michigan at Ann Arbor, Dearborn, and Flint; and University of Virginia, Charlottesville. The MESL project was launched by the Getty Art History Information Program, in association with MUSE Educational Media. MESL is administered by AHIP and advised by a Management Committee comprised of Maxwell Anderson, David Bearman, Howard Besser, Karl Katz, Patti McClung (MESL Project Manager) and Jennifer Trant.

2 Three project meetings, involving the Project Coordinators and Working Group Chairs, have also been held (in February, June and December of 1995).

3 Copies of the Cooperative Agreement are available upon request. It can also be found on the MESL WWW site.

4 Copies of the MESL Data Dictionary are available upon request. It can also be found on the MESL WWW site.

EVA - Florence
February 1996

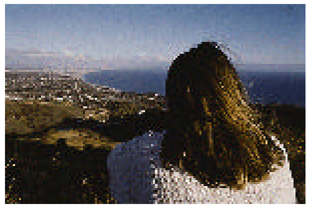
---

http://www.io.org/~jtrant/papers/jt.eva.florence.9602.ht

Last updated: April 14, 1995

Return to J. Trant's Home Page.

Jennifer Trant

Photographs: Ben Davis, Santa Monica, CA, 1995

Partner and Principal Consultant, Archives & Museum Informatics

Pittsburgh, Pennsylvania

I am a consultant regarding the management of information about cultural heritage. My areas of expertise include strategic information management, use of network technologies for access to cultural heritage information, computerized museum collection documentation (in both text and image form), international standards development initiatives, project design and management.

Director, Arts Information Management

Toronto, Ontario

I also consult regarding the application of technology to the mission of art galleries and museums. Until June of 1996, I was under contract to manage the Imaging Initiative, for the Getty Art History Information Program. This included directing the activities of the Museum Educational Site Licensing Project (MESL), participating in the development of imaging standards, and developing educational materials and programs. During 1996, I also worked with the Arts and Humanities Data Service, King's College, London, England, on issues related to Collections and Standards Development.

A specialist in arts information management, I've worked with automated documentation systems in major Canadian museums, including the National Gallery of Canada and the Canadian Centre for Architecture, where I developed and implemented common cataloguing standards for the Prints and Drawings, Photographs, and Archives Collections. My past work as an independent consultant includes preparation of the report of the Art Information Task Force, a collaborative initiative of the College Art Association and the Getty Art History Information Program, entitled *Categories for the Description of Works of Art*. I also consulted on Luna Imaging's

CD-ROM publication Frank Lloyd Wright: Presentation Drawings and Conceptual Sketches, distributed by Oxford University Press.

Throughout my career, I have been actively involved in the definition of museum data standards, participating in numerous committees and regularly publishing articles and presenting papers about issues of access and intellectual integration. My current interests center around the use of information technology and communications networks to improve access to cultural heritage information.

I am Chair of the Multi-Media Working Group of the International Documentation Committee (CIDOC) of the Internatonal Council of Museums (ICOM), and on the Boards of CIDOC, the Media and Technology Committee of the American Association of Museums(AAM), and the International Visual Arts Information Network (IVAIN).

Further details can be found in my Curriculum Vitae.

Recent Papers and Publications

Daniel Greenstein and Jennifer Trant, "The Arts and Humanities Data Service. Collecting digital research data; building a framework for digital resource

preservation and interchange" *Ariadne*, (July/August, 1996).

New Models for Distributing Digital Content, A presentation for the Data Processing Clinic, University of Illinois, March 25, 1996, and paper to be published in the conference procedings.

Enabling Educational Use of Museum Digital Materials: The Museum Educational Site Licensing (MESL) Project, A paper for the Electronic Imaging and the Visual Arts Conference, Florence, Italy, February 8-9, 1996.

"The Museum Educational Site Licensing (MESL) Project: An Update," *Spectra*, Vol. 23, No. 3, Spring 1996.

"Framing the Picture: Standards for Imaging Systems". *International Conference on Hypermedia and Interactivity in Museums*, San Diego, California, October, 1995.

Introduction to Imaging: Issues in Constructing an Image Database with Howard Besser, Getty Art History Information Program, 1995.

The Getty AHIP Imaging Initiative: A Status Report, *Electronic Imaging and the Visual Arts* (EVA), The

National Gallery, London, July, 1995. Also appearing in *Archives and Museums Informatics*, Cultural Heritage Information Quarterly, Vol. 9, no. 3, 1995, 262-278.

"The Museum Educational Site Licensing Project", *Spectra,* Winter 1994-95, 19-21.

> "A woman must have money and a room of her own if she is to write ... "
> Virginia Woolf, *A Room of One's Own*, 1929.
>
> What Virginia had lots of, it seems, is time...

Send comments to J. Trant jtrant@archimuse.com
See who's been here lately.

Created: September 26, 1995
Last Modified: Feb. 1, 1997 URL: http://www.io.org/~jtrant

GETTY INFORMATION INSTITUTE

Introduction to Imaging

Issues in Constructing an Image Database

Howard Besser and Jennifer Trant

CONTENTS GLOSSARY REFERENCE

 Cover image created from Jan Breughel the Elder, *The Entry of the Animals into Noah's Ark.* 1613, oil on panel, 21 1/2 x 33 in. (54 x 83 cm) J. Paul Getty Museum, Malibu, California.

PREVIOUS GLOSSARY REFERENCE NEXT

Introduction to Imaging: Issues in Constructing an Image Database

---

Contents

Authors

Acknowledgements

Foreword

1. **Introduction**
2. **What is a Digital Image?**
3. **Monitor vs. Print Display**
4. **Dynamic Range**
5. **Resolution**

6. **Image Capture**
7. **Compression**
8. **Various Users, Various Uses**
9. **Documentation**
10. **Access**
11. **Selecting Scanners**
12. **Image Quality**
13. **Quality Control**
14. **Display Devices**
15. **The Working Environment**
16. **Displaying Images and Text**
17. **Networks**
18. **Security Policy and Procedures**

19. **Technical Standards**

20. **Research Agenda**

21. **Other Resources**

22. **Conclusion**

- Terms highlighted throughout the text are linked to the Glossary.

PREVIOUS GLOSSARY REFERENCE NEXT

Comments and suggestions are welcomed. Adapting the print publication --*Introduction to Imaging: Issues in Constructing an Image Database* -- raised significant issues regarding the development and repurposing of materials for print and on-line delivery, especially in relation to image quality, file formats and network speed.

Howard Besser and Jennifer Trant, *Introduction to*

Imaging: Issues in Constructing an Image Database
The Getty Information Institute Imaging Initiative

Related Resources

This is a growing collection of web links to material related to the major issues in the MESL project.

Participants who have suggestions about items to add to the list should send them to the WWW Working Group.

Intellectual Property and Copyright

- United States Copyright Office: http://www.loc.gov/copyright

  One of the links from the Library of Congress Homepage, this site has a wealth of information and U.S. Copyright Office documents. Copyright Basics talks about what copyright is, who can claim it, what works are protected, what aren¢t protected, a notice of copyright, and more.

- IFLA Copyright and Intellectual Property Resources: http://www.nlc-bnc.ca/ifla/II/cpyright.htm

An extensive bibliography of primary and secondary information sources on copyright and intellectual property.

- Copyright Clearance Center: http://www.copyright.com

  Contains a listing of their services, online registration to their services, and a catalog of registered titles to preview. Under *Some New/Some Newer* is an article about the startup of a collective licensing agency for photographs. Media Photographers Copyright Agency has signed a partnership agreement with CCC to maintain a searchable database of photographic images that will be sent to clients via CD-ROM or online. It also has links to other web copyright information.

- Stanford University Copyright & Fair Use site: http://fairuse.stanford.edu/

  This site is sponsored by the Council on Library Resources, FindLaw Internet Legal Resources and the Stanford University Libraries and Academic Information Resources.

- Coalition for Networked Information (cni.org)

- CNI Copyright Archive
- CNICop yright Documents (includes FAQs)

Protection of Digital Information

- Howard Besser's Fall 1996 Course on Protecting Digital Information: http://www.sims.berkeley.edu/courses/is290-1/f96/

  Includes information and links to resources on digital signatures, encryption, watermarking, containers and encapsulation, authentication and permission controls

- LITA/LAMA Presentation on the Encryption: http://129.79.33.62./jm docs/lita96now.html

  Slides from a presentation at the LITA/LAMA conference, October 1996, presenting the basic concepts surrounding encryption and authentication of digital documents.

- IBM Digital Library: http://www.software.ibm.co m/is/dig-lib/

  One solution to protection and authentication, as an example of emerging efforts from the commercial sector.

Organizations

- American Association of Museums: http://www.netready.com/AAM/
- Association of Art Museum Directors' Art Museum Network: http://www.ago.on.ca/AAMDO/
- Consortium for the Computer Interchange of Museum Information: http://www.cimi.org/cimi/
- Getty Information Institute: http://www.gii.getty.edu
- Museum Computer Network: http://world.std.com/~mcn/MCN.h tml

Imaging and Image Databases

- Framing the Picture: Standards for Imaging Systems / by Jennifer Trant, International Conference on Hypermedia and Interactivity in Museums / Museum Computer Network Joint Conference, San Diego, California, October 1995
- Introduction to Imaging: Issues in Constructing an Image Database / by Howard Besser and Jennifer Trant, Getty Art History Information Program, 1995
- Prepari ng Quality Images for Computer Networks / by John P. Weise, The University of Michigan School of Information and Library Studies
- The Digital Work of Art in an Age of Immaculate

and Promiscuous Reproduction / by Steve Dietz, National Museum of American Art
- Image and Multimedia Database Resources / links gathered by Howard Besser
- Image and Multimedia Retrieval, Research Agenda for Cultural Heritage on Information Networks, The Getty Art History Information Program / by Donna M. Romer, Eastman Kodak
- The Clearinghouse of Image Databases and IMAGELIB listserv archives, the University of Arizona

Standards for Image Description and Data Interchange

- Data Standards Committee of the Visual Resources Association
- SGML for Cultural Heritage Information / by Joseph Busch
- Standards Framework for the Computer Interchange of Museum Information / by David Bearman and John Perkins
- Describing Image Files from the Fall1994 Meeting of the Coalition for Networked Information, Howard Besser and Jennifer Trant
- CNI/OCLC Metadata WorkshopWorkshop on Metadata for Networked Images

---

Home	About MESL	Working Groups	Participants

Last modified: Wednesday, 13-Nov-96 10:31:24

URL: *http://www.ahip.getty.edu/mesl/resources.html*

Standards Framework for the Computer Interchange of Museum Information

David Bearman and John Perkins

First Edition May 1993

Museum Computer Network

This document is also being published as a double issue of *SPECTRA* Vol 20 No 2 and 3. *SPECTRA* is the quarterly publication of the Museum Computer Network.

Questions or comments for either author or MCN may be addressed to


```
jperkins@fox.nstn.ns.ca

  or

MCN
8720 Georgia Ave. Suite 502
Silver Spring  MD 20910-3602
USA
301-495-0900
```

Table of Contents

[EXECUTIVE SUMMARY]

[FOREWORD]

[READING THIS REPORT]

ACKNOWLEDGEMENTS

GETTING MORE INFORMATION

I. INTRODUCTION TO CIMI

Origins of CIMI
Assumptions
History
Second CIMI Committee Meeting
Third CIMI Committee Meeting
Final CIMI Committee Meeting
CIMI Task Groups

II. INFORMATION STANDARDS

Standards organizations
Standards Frameworks
OSI
OSE
CALS
Interchange

III. REQUIREMENTS FOR COMPUTER INTERCHANGE OF MUSEUM INFORMATION

General Functional Requirement

Museum Applications and their Interchange Requirements
Database Building
Information Retrieval
Business Transactions
Message Handling
Document Handling
File Transfer
Realtime Links

IV. STANDARDS OPTIONS FOR MUSEUM DATA INTERCHANGE

Options for Telecommunications and Media Standards
Options for Museum Interchange Applications
Collections Catalogues and Reference Databases
Information Retrieval
Business Transactions
Messaging and Directory Services
File Transfer
Document Handling
Human/Machine Interaction

V. RECOMMENDATIONS

Museum recommendations
Vendor and Network recommendations
Museum profession recommendations

VI. IMPLEMENTING THE STANDARDS FRAMEWORK FOR CIMI

Implementation strategy
Tasks for present and future CIMI task groups
The future of the CIMI initiative

APPENDIX A: GLOSSARY OF CIMI TERMS

APPENDIX B: REFERENCES

Preparing Quality Images for Computer Networks

- by John P. Weise
- The University of Michigan
- School of Information and Library Studies

Table of Contents

Series Introduction

Emphasizing Image Quality

- **Limitations of Different Image Forms**
- **Components of Digital Image Quality**
- **Perspectives on Digital Image Quality**
  - **The Copyright Holder and/or the Image Creator**
    - Questions To Ask
  - **The Subject Matter Expert**
    - Questions To Ask
  - **The End User**
    - Questions To Ask
  - **The Archivist**

        - Questions To Ask
    - **Summary: Emphasizing Image Quality**

Data Transfer Time

- **Common Factors That Affect Data Transfer Time**
    - **The Transfer Rate**
        - Table: Typical Data Transfer Rates
    - **The Quantity of Data**
    - **Network Traffic**
    - **Throughput Speed of Storage Devices**
- **Summary: Data Transfer Time**

Bitmapped Image Technology Primer

- Purpose
- **Objectives**
    - Define the Essential Terms
    - Provide a Framework...

- **Bitmapped vs. Object Oriented Images**
- **More About Bitmapped Images**
    - **Resolution**
    - **Image Size**
    - **Color Depth**
        - Table: The Number of Colors Available at Common Bit Depths

- **Calculating Image Data**
- **Summary: Bitmapped Images**
- **Displaying the Image on a Computer Screen**
  - Scaling the Size and Resolution
  - Color Reduction (Quantization)
  - Dithering
  - Advantages of Color Reduction and Dithering
  - Disadvantages of Color Reduction and Dithering

Balancing Image Quality and Transfer Time

- Introduction
- Goal
- Objectives

- **Techniques for Reducing Bitmapped Image Data**
  - **Purpose**
  - **Approach**
  - **Reducing Resolution**
    - Overview
    - Advantages
      - Graph: Affects of Resolution on Data
    - Disadvantages

- **Image Cropping**
  - Overview
  - Advantages
  - Disadvantages
- **Reducing Color (Bit) Depth**
  - Overview
  - Advantages
  - Disadvantages
  - Is Color Really Necessary?
- **Image Data Compression**
  - Overview
  - Advantages of Lossy Compression
  - Disadvantages of Lossy Compression
  - Advantages of LossLESS Compression
  - Disadvantages of LossLESS Compression

JPEG Compressed Still-Images

- **Introduction**
- **Image Compression**
  - Lossless versus Lossy
  - JPEG Lossy Compression
- **JPEG File Formats**
  - PICT & JPEG
  - TIFF & JPEG
  - Encapsulated PostScript & JPEG
  - GIF & JPEG
- **Other Sources of JPEG Information**

Bibliography

Image and Multimedia Database Resources

Paul Peters dies11/18/96

Index to this set of links

General Overviews of Digital Imaging
Information about Metadata and Standards
Ethical & Legal Issues
Image Capture and Compression
Image Quality & Conservation
Technical Protection for Images
Retrieval issues and search engines for WWW Databases
Content-based Retrieval

Image Databases on the Net
Cultural Repositories' Collections
Unique or Special Image Databases
Three Dimensional Images
Stock Photographs, Commercial Image DBs
Large-Scale Projects and Initiatives
Smaller-Scale Projects and Initiatives
Sample implementations with Image Browsers

Additional Resources

Multimedia Resources
Research Papers
Jobs
Misc

Welcome to

- Upcoming Events
- About ISS
- Contact Directory
- News/Job Opportunities
- ISS Training Courses
- Publications
- Facilities
- Staff Directory
- ISS Organisation
- ISS Hosted Sites
- Collaborations

You Are Driver Number
64284
on The ISS Superhighway Since 26th June 1996.

Questions about this service should be directed to:
webmaster@iss.nus.sg

VIP Group

Visual Information Processing Group

The "Visual Information Processing Group" at the Institute of Systems Science of the National University of Singapore is a part of the ISS "New Initiatives".

Projects | People | Publications | What's New

Mission:

Problem driven research and development in the area of representation, analysis, storage and retrieval of visual information such as images, maps and video. Particular areas of concentration are multimedia information systems, content-based retrieval of image and video data, spatial information systems and image and map understanding.

International Collaboration:

RWCP Real World Computing Partnership (Japan) project:
Active Integrated Media System (AIMS) (1996-2002).

CLIPS-IMAG, University of Joseph Fourier, France, on Indexing and Retrieval of Visual Information".

Contact Information:

For more details on any of the activities of the VIPG, please contact:

Dr. WU, Jian Kang
Institute of Systems Science
National University of Singapore
Kent Ridge, Singapore 119597
Ph: (65) 772-2544
Fax: (65) 774-4990
Email: *jiankang@iss.nus.sg*

For comments or suggestion, please email: *bkteo@iss.nus.sg*

This page last updated on: 28 Jan 1997.

VIP Group

Visual Information Processing Group

Projects:

- *Robust Face Recognition System*

- *Robust Fingerprint Recognition System*

- *NeuroVec* - A new generation paper map digitization system.

- *Web-based Multimedia Spatial Information System* with visual spatial query language.

- *FACEit* - a content-based face image archival and retrieval system.

- *STAR* - System for Trademark Archival and Retrieval

- *Video Processing* facilitates manipulation, indexing and retrieval of MPEG encoded video.

- *Agent-based Meeting organizer* manages the calendar and schedules meetings.

- *Personalized Information Network* delivers personalized multimedia news.

VIPG Home | People | Publications | What's New

Robust Face Recognition System

Aim

To recognize a person's face regardless of changes in facial expression and image capturing conditions.

Approach

A *sequence of images* of the area in front of the terminal is captured and submitted to the following processes:-

- ***Detection* of the presence of face images, and *segmentation* of facial features. The methods are invariant to changes of distance and angle between camera and the person, as well as to image capturing conditions.**
- **Feature extraction methods, and**

methods for fusion of multiple feature measures. The feature measures extracted are invariant to facial feature changes, such as growing fatter or thinner, with or without moustache.

- **Novel recognition paradigm *``recognition by recall''* and *incremental learning*, aiming at a breakthrough on recognition problems having large amount of classes and small amount of samples in each class.**

Contact:
LAM, Chian Prong
Institute of Systems Science
National University of Singapore
Heng Mui Keng Terrace
Singapore 119597
Fax: +65 774 4990
email: prong@iss.nus.sg

Last modified: Wed Mar 5 13:23:45 SGT

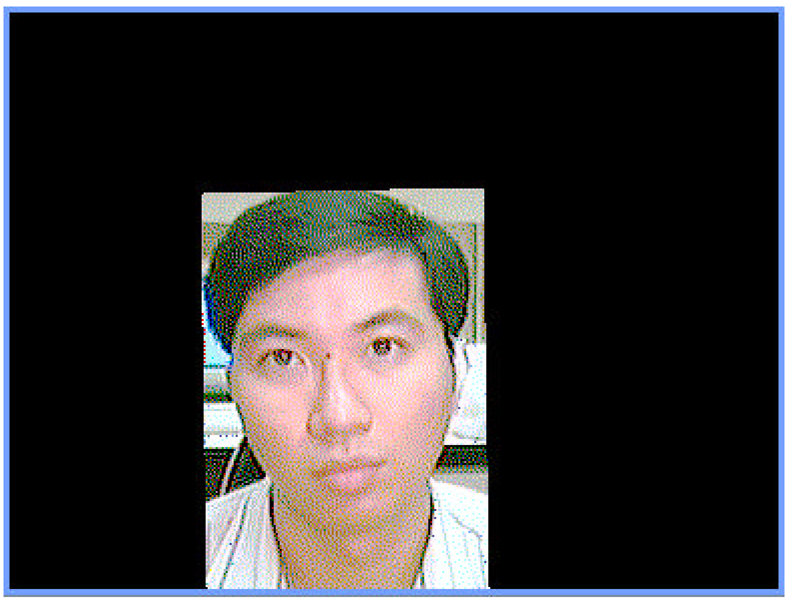
Raw Image Sequence

An image sequence of a few frames is captured and sent to the face segmenter to extract the region where the face is.


MoFFeR
Live
Capture
Load Video
Step Process
Cont Process
Train
Active Image
Capturing live images .....Wait a moment please.....

Face Segmentation

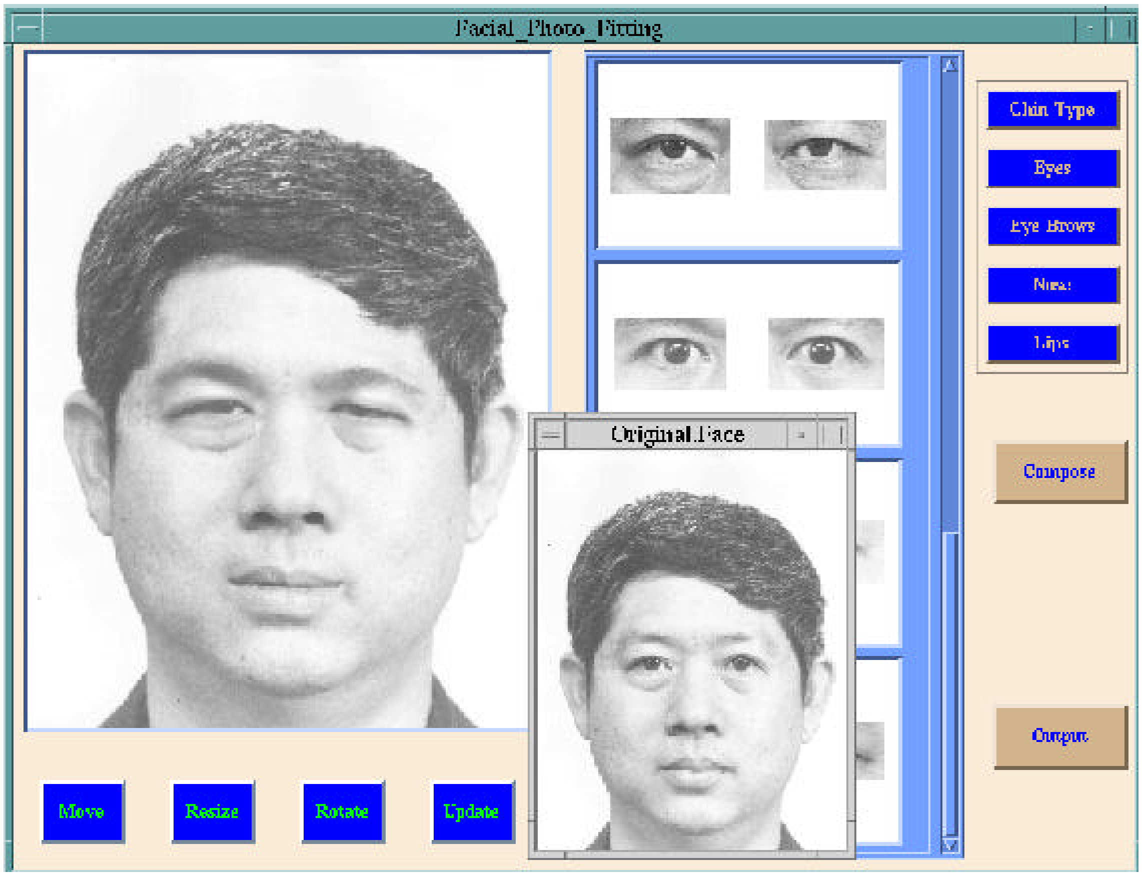
The system automatically detects and segments out the face in an image irregardless of the background.



Facial Image Composition and Retrieval

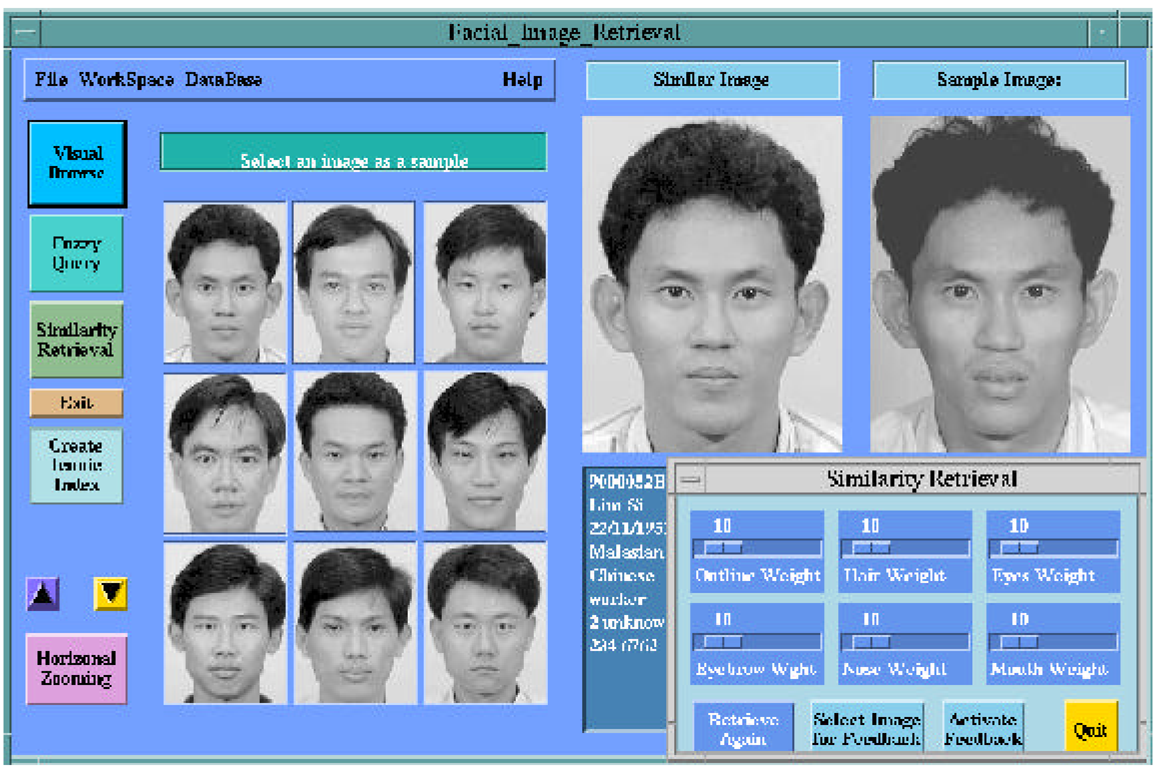
Facial Image Composition

- A good wireframe model;
- A well-designed fitting algorithm;
- Template databases.


Facial_Photo_Fitting
Original Face
Chin Type
Eyes
Eye Brows
Nose
Lips
Compose
Move
Resize
Rotate
Update

Similarity Retrieval

- **Normalization of images;**
- **Measures for 6 facial features;**
- **Use of multiple feature measures.**



Retrieved a similar image of the same individual taken at 6 months earlier.

Content-based Index (ContIndex)

Challenges

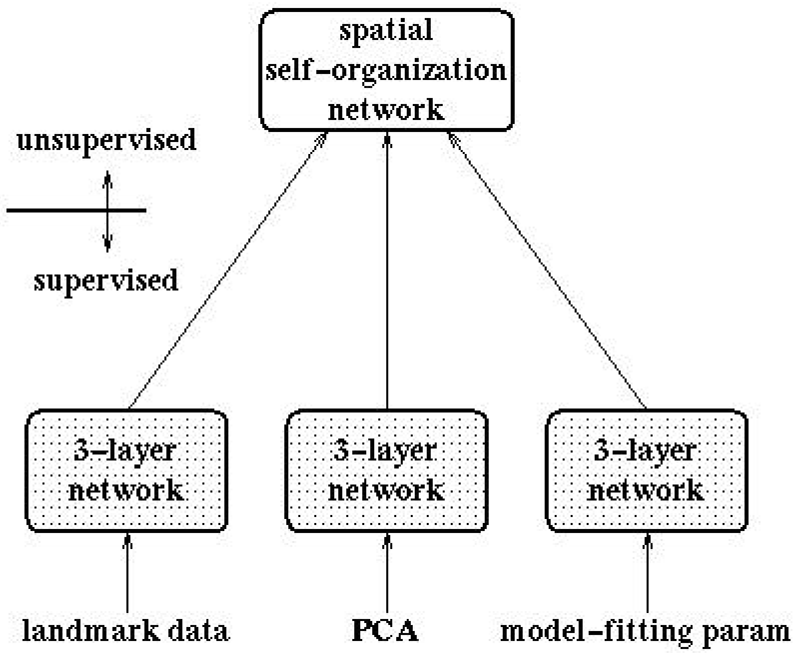
- **Complex image visual features: represented by vectors or structure;**
- **The index should be consistent with categories of visual features;**
- **Icon image for each node to facilitate visual browsing;**
- **For viewing purpose, nodes should be spatially organized.**
- **Facilitate vertical, horizontal, and multi-resolution visual browse.**

Generalization for Content-based Index

- **Allow attribute be abstract data type (vector, array, data structure, fuzzy sets).**
- **LEP neural network is used to fuse multiple features. The nodes are self-organized to balance node population, and to gain spatial arrangement.**
- **Key attributes can be different for different levels.**

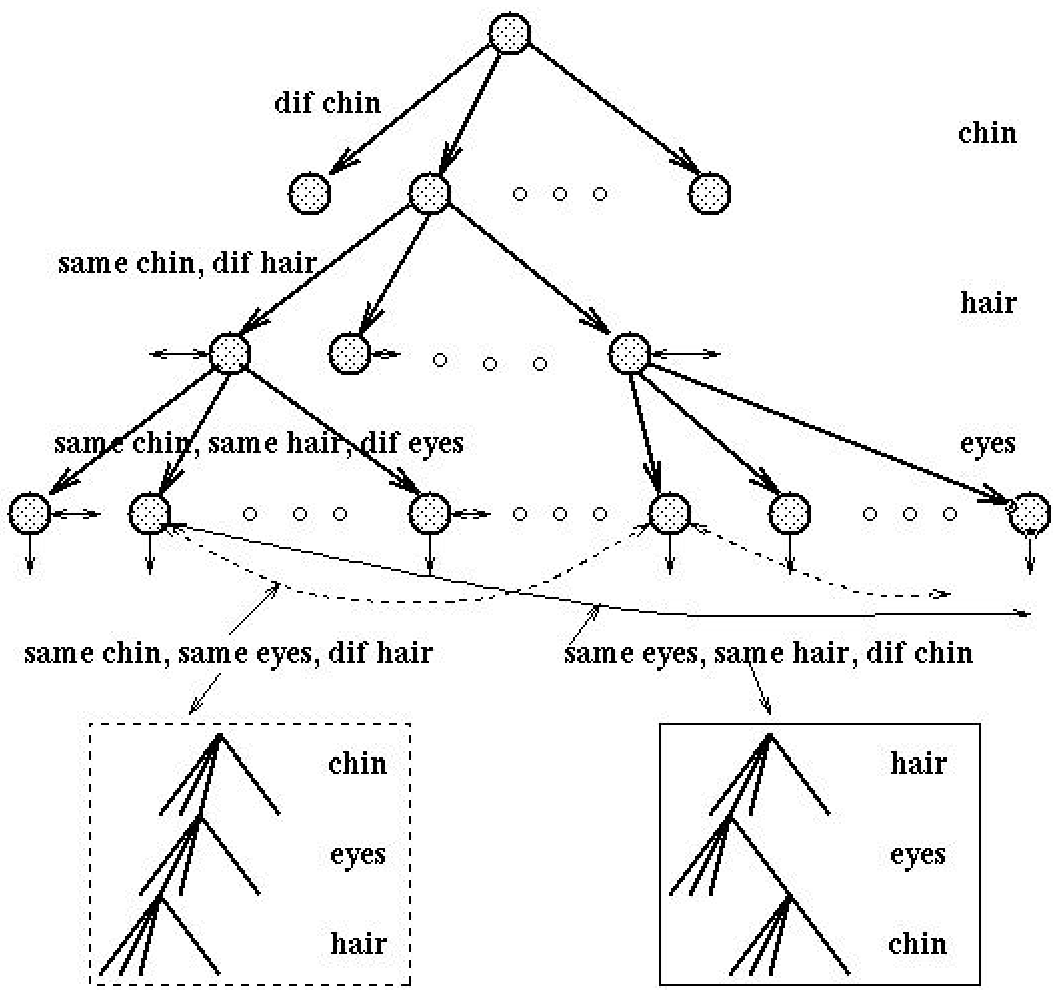
Next Page

The neural network model LEP (Learning based on Experience and Perspectives) is used to create self-organized index nodes using multiple featuer measures.



Next Page

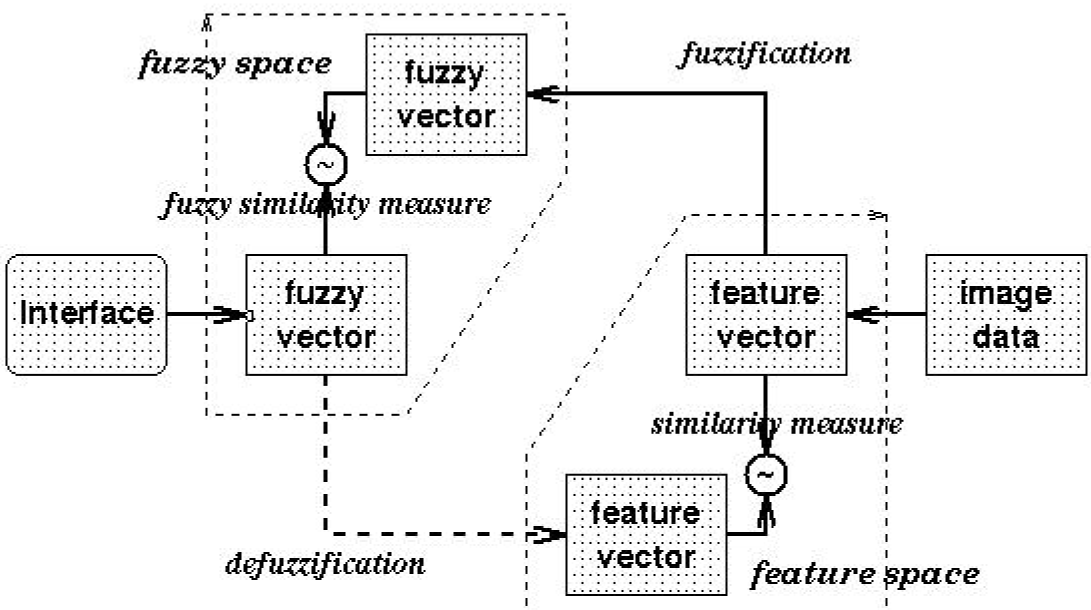
The structure of ContIndex. Horizontal links increase the flexibility of the index.



Next Page

Fuzzy Retrieval

Processing Fuzzy Query in Fuzzy Space or Feature Space?



Because incomplete fuzzy query definition, the fuzzy query can only be processed in fuzzy space.

Next Page

VIP Group

Visual Information Processing Group

Spatial Information System

Description

The work under this theme is a part of our research project for RWC. It is to demonstrate the development of a complex intelligent media agent in a network environment. Several core technologies are being developed. The major ones are as follows:

(a) Processing visual queries with spatial algebra

This is an open and fruitful research area. As most of the real world data has spatial property, and most of the decision making processes need to invoke spatial reasoning, spatial query languages will be a core technique in real world computing. Apart from existing database query techniques, query processing based on spatial algebra is necessary. We have theoretically defined the spatial algebra, and implemented it based on our unique

dual-vector-raster data structure proposed by Wu, etc in 1990.

(b) Presenting the results with multimedia tools

Real world applications need multimedia tools to present complete information to users. Today, there are choices of hardware and software tools available. However, the difficult issues are speed and synchronization, and handling (fast access, 3D visualization, rendering and texture mapping) of large volumes of images and video data. We have proposed a formal approach to semantic specification of objects using a trace calculus. It will be used as the theoretical basis.

(c) Answering questions with reasoning capability

People usually ask for best solutions. Best solutions are a result of reasoning processes. Here, the examples can be: various routes to go to work, the best site for a supermarket, etc. In this case, the reasoning involves not only knowledge represented as logic or semantic networks, but also verification of spatial relationships, computation of spatially distributed entities, etc.

(d) Querying distributed complex databases

It requires a computing environment which enables operational sharing of distributed data resources and transparent processing of those information and data in a distributed computing environment. We have recently developed the Kleisli/CPL technology for integrating heterogeneous distributed complex data sources by simply using CPL query language. We have also successfully applied this technology in the human genome arena. We propose to apply this technology as a tool for combining data from heterogeneous spatial data bases residing in multiple sites.

A preliminary demonstration prototype has been implemented by using the idea of processing visual queries with spatial algebra (see demo description).

The discussion with a few of potential customers are being carried out.

Key Achievements

* Development of Spatial Algebra.

* Proposal of dual-vector-raster data structure.

* Semantic specification of objects using a trace calculus.

* Development of Kleisli/CPL technology.

Sponsors of the project

RWCP The Real World Computing Program is a research program funded by the Japanese Ministry of International Trade and Industry and aims at the realization of flexible information processing systems for diversified information in the real world by introducing intuition-like processing functions. The Program was launched in 1992, and is scheduled to complete in 2002.

Related Publications

Demo

The prototype has been implemented in UNIX platform, it has the following main features:

(a) choose interesting area on the map by simply using mouse for click and zoom-in, zoom-out

(b) submit request for finding a property through a friendly GUI using spatial query language (with spatial operators, such as "near", "cross", "within")

(c) multi-media data (image, audio, video) provided for user to get a good understanding of the specific

property

(d) fuzzy concepts, such as "near", can be defined by user to meet his own need and personalized preference.

A demo of JAVA applet with similar functions has also been implemented to allow user to retrieve information through network.

Pathname for the demo will be available soon!

Names of people involved in the project

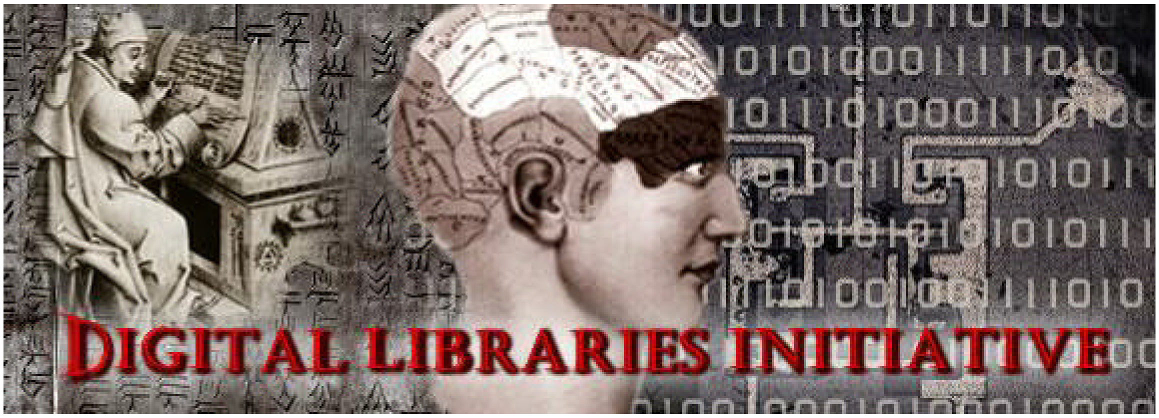
Ding Liya
Mohan S. Kankanhalli
Wang Chuan Chu
Wong Lim Soon
Wu Jian Kang

Contact Information:

For more details, please contact:
Ding Liya
Institute of Systems Science
National University of Singapore
Kent Ridge, Singapore 119597

Ph: (65) 772-3110
Fax: (65) 774-4990
Email: *liya@iss.nus.sg*

VIPG Home | Projects | People | Publications | What's New



DIGITAL LIBRARIES INITIATIVE

Funded through a joint initiative of

NSF/DARPA/NASA

University of California at Berkeley
Environmental Planning and Geographic Information Systems

University of California at

The Initiative's focus is to dramatically advance the means to collect, store, and organize information in digital forms, and make it available for searching, retrieval, and processing via communication networks -- all in user-friendly ways.

Digital Libraries basically store materials in electronic format

Santa Barbara
The Alexandria Project: Spatially-referenced Map Information

Carnegie Mellon University
Informedia Digital Video Library

University of Illinois at Urbana-Champaign
Federating Repositories of Scientific Literature

University of Michigan
Intelligent Agents for Information Location

Stanford

and manipulate large collections of those materials effectively. Research into digita libraries is research into network information systems, concentrating on how to develop the necessary infrastructure to effectively mass-manipulate the information on the Net.. The key technological issues are how to search and display desired selections from and across large collections.

Summaries of the six DLI projects from the May 1996, Special Issue on Digital Libraries in the Institute of Electrical and Electronics Engineers, IEEE Computer Magazine.

The magazine of digital library research, the D-Lib Magazine, including the July/August 1996 issue The DLI Testbeds: Today and Tomorrow.

University Interoperation Mechanisms Among Heterogeneous Services

Digital Library conference information, publications, related projects and resources to the DLI, Digital Library Related Information and Resources.

DLI Project Contacts

NSF Digital Libraries Contact

DLI Workshop Series

DLI Publications

National Synchronization for the Digital Library Initiative is being coordinated by the University of Illinois at Urbana-Champaign, and supported by a supplemental grant by the National Science Foundation.

Comments to Susan Harum at the University of Illinois DLI Project
10/31/96

The Alexandria Project

Introduction

Welcome to the home page of the **Alexandria Project**. We are a consortium of researchers, developers, and educators, spanning the academic, public, and private sectors, exploring a variety of problems related to a **distributed digital library for geographically-referenced information.**

Distributed means the library's components may be spread across the Internet, as well as coexisting on a single desktop. **Geographically-referenced** means that all the objects in the library will be associated with one or more regions (``footprints'') on the surface of the Earth.

The centerpiece of the Alexandria Project is the **Alexandria Digital Library** (ADL), an online information system inspired by the Map and Imagery Laboratory (MIL) in the Davidson Library at the University of California, Santa Barbara. The ADL currently provides access over the World Wide Web to a subset of the MIL's holdings, as well as other geographic datasets.

- **ADL on the Web** NEW
- **Public documents**
  - Annual progress report (best overview document)
  - SNEP - Executive Summary and Volume 1 report NEW
- **Participants**
- **Other digital libraries**
- **Spatial Data Available Over the Web**
- **Personnel**
- **Contact Information**

- **Job Opportunities**

- **Alexandria Design Review Panel**
  This link is for the members of the Design Review Panel only and
  requires a user id and password to proceed.

Please take a moment to let us know what you think of the Alexandria Web. Use our online comment form, or send email to webmaster@alexandria.sdc.ucsb.edu, or contact individual Web Team members directly.

Thanks for stopping by!

Note: clicking on the Alexandria logo at the top or bottom of a page will always bring you back to this (Alexandria home) page.

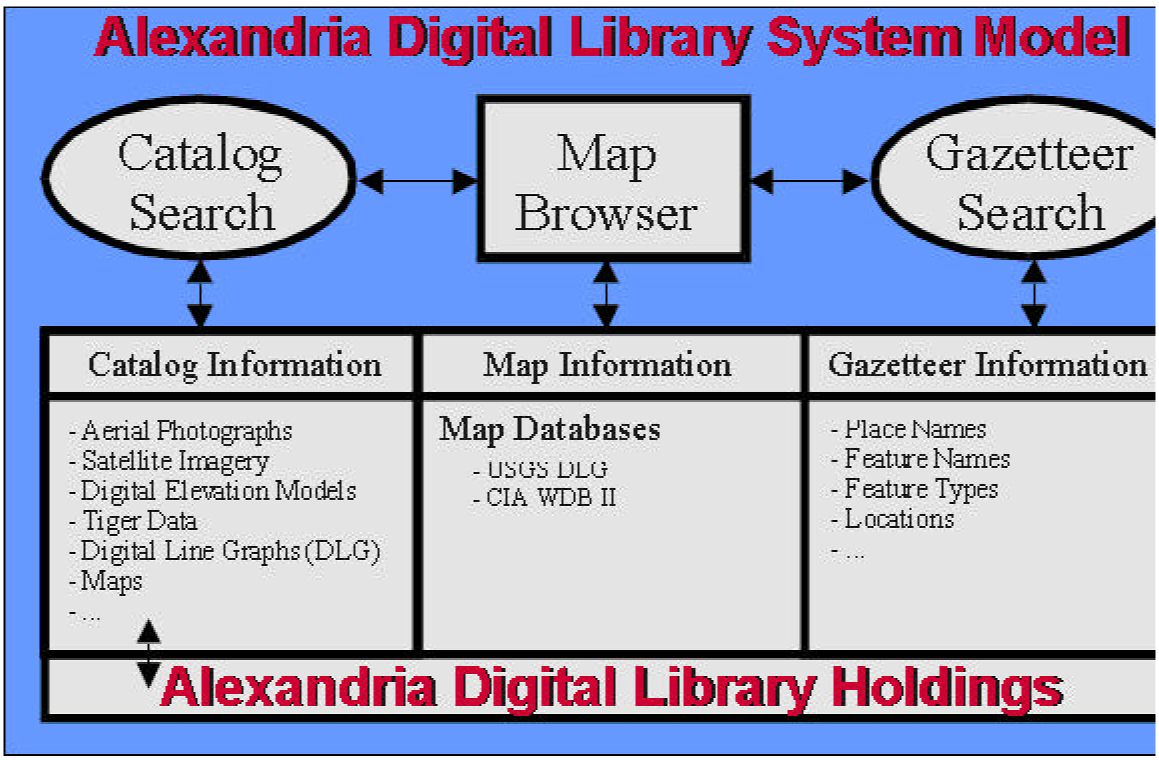
Alexandria Digital Library

Last modified on 1997-01-08 at 16:52 GMT by the

Alexandria Web Team

Tutorial Table of Contents

Prev Next



Conceptual model of the Alexandria Web interface

- Conventions
- Session / System Setup

- Map
- Gazetteer
- Catalog
- Overview of Current Holdings
- Walkthroughs (Example Sessions)
- Feedback
- Technical Reference
- Acknowledgements
- Access ADL

- World Wide Web Help
  (How to use a web browser. That's what you're using right now.)

Prev Next

Next Up Previous
Next: PROJECT SUMMARY

ALEXANDRIA DIGITAL LIBRARY

ANNUAL REPORT

NSF Program: CISE (IRIS), NSF 03-141 Award Number: IRI94-113

PI Name: Terence R. Smith Period Covered By This Re

PI Institution: UCSB Date: 15 February, 1997

PI Address: Department of Computer Science
UCSB
Santa Barbara
CA 93106

ALEXANDRIA DIGITAL LIBRARY

ANNUAL PROGRESS REPORT

- PROJECT SUMMARY
- SIGNIFICANT EVENT

- OVERVIEW OF PROGRESS IN LIBRARY DEVELOPMENT
- CURRENT STATUS OF TESTBED DEVELOPMENT
    - Implementation Team
    - Architecture of the Testbed System
        - Components of Testbed Resulting from R&D Team Research
        - Distributed Object Design Activities and Next Testbed System
    - User Interface Component
    - The Catalog Component
        - Alexandria Gazetteer
        - Loading and Maintenance of Metadata
    - Collections
        - Collection Loading Strategy
        - Loading of Datasets
        - Scientific Datasets and Metadata
    - Computing Support for the Testbed: Hardware and Communications
        - Current Computing Equipment
        - Current Networking Support
        - Current ``Other'' Hardware and Software Support
        - Equipment and Facilities Needs
    - Development of an Operational Library
    - Abstracts of Published Papers
- RESEARCH ACTIVITIES AND PROGRESS
    - LIBRARY TEAM

    - USER REQUIREMENTS SUBTEAM
    - METADATA AND CATALOG INTEROPERABILITY SUBTEAM
    - ALEXANDRIA ATLAS SUBTEAM
- INTERFACE DESIGN AND EVALUATION TEAM
    - UCSB Component of the Interface Evaluation Team
    - Colorado Component of the Interface Evaluation Team
    - Interface Design Subteam
    - Abstracts of Published Papers
- INFORMATION SYSTEMS TEAM)
    - Pharos and Resource Discovery
    - Multidimensional indexing
    - Content Based Retrieval
    - Data Placement
    - Tertiary Storage
    - Extensible Data Store
    - Database Performance Tuning
    - Abstracts of Published Papers
- IMAGE PROCESSING TEAM
    - Multiresolution Browsing: Balanced Rounding (BR)--Transform
    - Faulty Storage and/or Transmission of Corrupted Wavelet Coefficients
    - Texture based retrieval
    - Abstracts of Published Articles
- PERFORMANCE AND PARALLEL

PROCESSING TEAM
        - Fast subregion retrieval and image compression using wavelet transforms
        - A scalable WWW server on multicomputers
        - SWEB++: Distributed scheduling and adaptive client-server computing for improving response times of WWW applications
        - Abstracts of Published Articles
- COMPARISON OF ACTUAL AND PREVIOUSLY PLANNED R&D ACTIVITIES
    - TESTBED TEAM
    - LIBRARY TEAM
    - INTERFACE DESIGN AND EVALUATION TEAM
    - INFORMATION SYSTEMS TEAM
    - IMAGE PROCESSING TEAM
    - PERFORMANCE AND PARALLEL PROCESSING TEAM
- MANAGEMENT REPORT
    - UCSB Personnel
        - Professional Staff
        - Faculty Investigators
        - Graduate Students Associated with the Project
        - Visitors to the Project
    - Colorado Personnel
        - Faculty Investigators

            - Graduate Students Associated with the Project
        - Organizational Structure
            - Research and Development Team Structure
            - Project Retreats
        - Alexandria Advisory Board
            - The Composition of the Advisory Board
            - Report of the Second Board Meeting
            - Report of the Third Board Meeting
        - Alexandria Design Review Panel and Meetings
            - The First Alexandria Design Review
            - The Second Alexandria Design Review
    - EXTERNAL RELATIONS AND INTERACTIONS
        - Interactions with Old and New Partners
            - American Geological Institute (AGI) and The University of Tulsa
            - California Environmental Resources Evaluation System (CERES)
            - Central Imagery Office (CIO)
            - Central Intelligence Agency (CIA)
            - CIESIN
            - Davidson Library, UCSB
            - Defense Mapping Agency
            - Digital Equipment Corporation (DEC)
            - ERDAS
            - Environmental Systems Research Institute (ESRI)

        - Hughes
        - Informix/Illustra
        - Library of Congress
        - Microsoft
        - NASA
    - National Imagery and Mapping Agency (NIMA)
        - Oracle
        - O2
        - San Diego Supercomputer Center (SDSC)
        - Sierra Nevada Ecosystem Project (SNEP)
        - SPOT Image
        - The Analytic Science Corporation (TASC)
        - United States Geological Survey (USGS)
        - United States Navy, Stennis
        - United States Navy, San Diego (NRaD)
        - Earth Data Analysis Center (EDA), University of New Mexico
        - Utah State University/Mojave Database Cooperative
        - Xerox
    - Interoperability Agreements and Activities
        - Stanford DLI Project
        - Berkeley DLI Project
        - Illinois DLI Project
        - CMU DLI Project
    - Visits and Demonstrations
- AUGMENTATION OF THE RESOURCES OF THE PROJECT

    - Funding Request Under Consideration
    - Unsuccessful Funding Requests
    - Professional interactions with DL Community
    - Publicity
- EDUCATIONAL ACTIVITIES
    - Courses Concerning Digital Library at UCSB and U.Colorado
    - Information and Training Programs for Project Personnel
    - Project Seminars
    - Educational Activities in Relation to Schools
    - Completed Dissertations
- PUBLICATIONS
    - References to Published Articles
    - Presentations
- FINANCIAL REPORT
- Certification
- About this document ...

Terence R. Smith
Thu Feb 20 13:50:53 PST 1997

Version: 5/12/97

The Second Alexandria Design Review

Report on the Workshop

19-21 February, 1997

Santa Barbara, California

Workshop Co-Leaders:

Barbara Buttenfield, U. Colorado-Boulder

Linda Hill, U. California-Santa Barbara

Workshop Co-Conveners:

Joseph Boisse, U. California-Santa Barbara

James F. Williams II, U. Colorado-Boulder

The Alexandria Digital Library Project

Terry Smith, Project Director

Michael Goodchild, Associate Director

Table of Contents

Executive Summary *
Acknowledgments *
1. Description and Scope of the Workshop *
2. The Workshop *
2.1 Steering Committee *
2.2 Meeting Objectives and Pre-Meeting Preparation *
2.3 Meeting Participants *
2.4 Meeting Format *
2.4.1 General Format *
2.4.2 Presentations and Software Demonstrations *
2.4.3 Follow-Up Activities *
3. Themes Covered in the Small Groups *
3.1 Search, Browse, and Retrieval Session *
3.1.1 Setting the Issues *

3.1.2 Discussion in Small Groups and Plenary Session: Search, Browse, and Retrieval *
3.1.3 ADL Summary of Comments and Response: Search, Browse, and Retrieval *
3.2 Content and Processing Session *
3.2.1 Setting the Issues *
3.2.2 Discussion in Small Groups and Plenary Session: Content and Processing *
3.2.3 ADL Summary of Comments and Response: Content and Processing *
3.3 Interface and Navigation Session *
3.3.1 Setting the Issues *
3.3.2 Discussion in Small Groups and Plenary Session: Interface and Navigation *
3.3.3 ADL Summary of Comments and Response: Interface and Navigation *
4. Initial Summary by ADL Staff (Summing Up Session at the Workshop) *
4.1 Discussion *
5. Summary of Evaluation Comments on the Workshop *
6. Impact of the Design Review on ADL *
6.1 ADL Collection Development Strategy *
6.2 ADL User Profile Strategy *
6.3 ADL Human Support Strategy for Digital Library Users *
6.4 ADL Interface Design *
7. Future Plans the Design Review Panel *
Appendix A. Attendees *
Appendix B. Meeting Agenda *

Executive Summary

This report is a summary of the Second Alexandria Design Review Panel. The workshop was held in Santa Barbara, California, 19-21 February 1997 to review user requirements for the Alexandria Digital Library and to discuss plans for testbed development and research progress in the coming eighteen months. The challenge to build an operational digital library testbed is improved by the research accomplishments of the ADL team. It is important to recognize that some research questions cannot be addressed in the absence of an operational testbed. This balance of emphasis will continue to be a concern in the coming eighteen months.

Thirty-two representatives from the public sector, private sector, and academia came from U.S. (one from Canada) for two-and-a-half days to iterate between plenary and small group discussion, software demonstrations, and informal discourse. The intention is that panel members will continue to be proactive in the coming year, through online activities

(working with the Library testbed), through personal contact with ADL project staff, and by face-to-face meetings at national conferences in the coming year. The next Alexandria Design Review Workshop will be held approximately one year hence.

In preparation for the meeting, the Steering Committee agreed upon a set of objectives to guide the content of the agenda. A three-part focus was established to deal in turn with issues of search, browse, and retrieval; content and processing; and interface and navigation. These three areas encompass a number of the more difficult problems to be resolved immediately. Their selection was based primarily upon feedback gained in focus groups, survey questionnaires, and videotaped or recorded interactions with actual and potential digital library patrons over the first two years of the Alexandria Project. Many specific topics were discussed including ways for users to view the contents of the library (an overview of the holdings), giving the user a "road map" of the services of the library, visualization (presentation) of search results, iterative search processes, and customization of the user interface.

From the wide range of comments and recommendations contributed by the Panelists that are detailed in this report, some strong themes can be identified. First, the Panel is supportive of the

efforts of the research teams working at Santa Barbara and at Colorado. Panelists and funding agency representatives agreed that keeping a careful balance between developing an operational digital library and contributing to theoretical and applied research is necessary. The importance of the content of the library was strongly supported, as were the enabling technologies for distributed ingesting of data with attention to the representation of data provenance and verification. Collection development criteria were discussed at length, with emphasis on focusing on particular geographic areas and collecting all forms of georeferenced information about that place. Panelists recommended that ADL develop enabling technologies and specifications for connecting local collections to ADL, with distributed searching capability.

In connection with ADL user interface design, the implementation of user profiles was discussed with ideas for multiple profiles per person and for the way in which the profiles could influence the interface presentation and conduct of the search. Support for a method of providing "something with a heartbeat" to be available to assist in the use of ADL was expressed in several ways, with the related recommendation that ADL aim to create a professional tool for information specialists as well as for other user communities. A discussion of *measures*

Image Browsing in the Alexandria Digital Library (ADL) Project

B. S. Manjunath
Electrical and Computer Engineering Department
University of California at Santa Barbara
manj@ece.ucsb.edu

D-Lib Magazine, August 1995

The management of images, video, and in general, multimedia data, is an important issue in the design of digital libraries. In particular, two problems stand out: efficient storage and fast retrieval. We outline below the general approach taken to address these two problems in the University of California at Santa Barbara (UCSB) Alexandria Digital Library project whose goal is to create a database of spatially indexed data. Maps and satellite images are among the main data sets in this project.

- The Problem of Terr(or)a Bytes

- Multiresolution Browsing and Wavelets
- Texture Features for Retrieval
- Acknowledgements and References

d-lib forum | d-lib magazine | d-lib next

The Problem of Ter(ror)a Bytes

In the Alexandria Digital Library (ADL) project, the size of the image files tends to be extremely large -- from few mega bytes to hundreds of megabytes of data. Even with access to very high speed networks, it is often impractical to transmit a large image as a single item, particularly if the user is in a browsing mode and trying to find items of interest. A simple solution is to maintain low resolution "thumbnail" images (e.g., subsampled image) for each of the large images. The thumbnails may then be used to support such browsing. While the storage of thumbnails consumes storage space, this overhead is typically insignificant compared to the advantages from their use.

Multiresolution Browsing and Wavelets

Storing thumbnails and the original data does appear to be useful. Low resolution images are currently being used in many existing geographic information systems (GIS) as well as in the rapid prototype of the ADL project. However, this strategy addresses only one specific issue -- that of fast browsing through large number of images. But in an interactive system, users are likely to do much more than make binary decisions based on simple thumbnail images. They may want to select a certain region within the image and zoom-in on it. Or, perhaps the thumbnail does not offer enough information to make such a binary decision but getting a slightly better resolution image might help.

Such operations are not possible using just these low resolution images as thumbnails. Further, different groups of users may have different requirements. For example, a school teacher using a LANDSAT image for a certain demonstration may not need the same high resolution image as a scientist trying to classify the image data. Clearly, what is needed is a means of storing images at different intermediate resolutions -- that is, a hierarchical multiscale representation of these images. An obvious solution to this problem is the use of wavelet transforms.

Figures 1a and 1b show an image and its wavelet transform.

Figure 1a: The original image

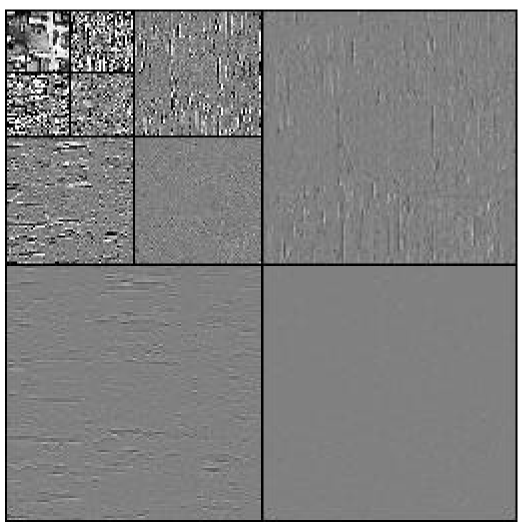


Figure 1b: The wavelet decomposition

Wavelets have been widely used in many image processing applications including compression, enhancement, reconstruction, and image analysis, and a wavelet transformation provides a multiscale

decomposition of the image data. The lowest resolution image (top left hand corner of Figure 1b) is now the thumbnail that can be used for browsing. Notice that the number of transform coefficients is exactly the same as the number of pixels in the original image. Fast algorithms exist for computing the forward and inverse wavelet transforms. Desired intermediate levels can be easily reconstructed as illustrated in Figure 2.

Figure 2a: The reconstruction at level 1

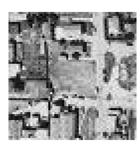


Figure 2b: The reconstruction at level 2

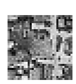


Figure 2c: The reconstruction at level 3
(Note that this corresponds to the low resolution image at the top left hand corner of Figure 1b)

Storing the transformed images (wavelet coefficients)

facilitates the design of hierarchical storage structures. Coarse resolution data are accessed more frequently than the higher resolution information, and hence can be stored in faster devices for efficient browsing.

Important issues related to wavelet based storage include the choice of decomposition (i.e. choice of filters) appropriate for the different image databases. Image compression is also important in storing large amount of data. Many GIS and medical imaging applications often require lossless compression. Although the total number of wavelet coefficients equals the number of pixels in the images, their storage requirements differ. The original intensity data, in most cases, consists of only integer numbers. Wavelet coefficients, in general, are real numbers, thus requiring more memory. Even for the case of no compression, these coefficients need to be quantized and encoded appropriately to ensure that they do not take more space than the original image data. How these coefficients can be quantized while maintaining a near perfect reconstruction is an important research problem.

Texture Features for Retrieval

Content based retrieval is about developing tools for intelligent browsing of the data. In traditional

alpha-numeric databases, such as an on-line library catalog, we search using keywords, author names, or book titles. Similarly, generic image attributes useful for search include color, histogram, texture, and shape. However, research on content-based image retrieval is still in its very early stages. We now briefly describe our recent work on using texture information for image retrieval.

Examples of texture images include photographs of water, sand, a brick wall, a wire fence, or aerial photographs of agricultural regions. Textured images, in general, are hard to describe (i.e., they do not have good structure). Often, the resolution and distribution of objects in the scene determines if it is " textured&quote; or not. For example, consider a bunch of coffee beans spread on the ground. While each bean is clearly an identifiable object, the random distribution of the beans as a whole is more like a texture pattern. Natural textures tend to be more irregular than man made ones. During the past six months, we have made considerable progress in developing algorithms for texture based search. The basic idea is to pre-process the images at the time of storage and extract the texture information. This is done using Gabor filters, which are modulated Gaussians. Processing through a bank of these Gabor filters is (approximately) equivalent to extracting line edges and bars in the images, at

different scales and orientations.

Simple statistical moments such as the mean and standard deviation of the filtered outputs can now be used as indices to search the database. Figure 3 (586 Kbytes) shows an application to browsing large air photos.

Instead of Gabor filters, one may also use the same orthogonal wavelet transform that was used for storing the image data. But extensive experiments on a large set of textured images show that retrieval performance is better using Gabor filters than when using conventional orthogonal wavelets. Why not use Gabor transforms for storage? Because Gabor functions do not form an orthogonal basis set, and hence the representation will not be compact. Further, no efficient algorithms exist for computing the forward and inverse transformations, which is important in a digital library context. While data ingest is off-line and can be computationally intensive, data retrieval should be fast and be performed in real time using existing hardware. Orthogonal wavelets are good for such implementations whereas non-orthogonal Gabor wavelets are promising for image analysis.

Efforts are currently underway to incorporate wavelet based storage and texture feature based search

using Gabor filters into the main testbed of the ADL. Many of the issues related to multiresolution browsing appear to be design problems. These include the choice of wavelet filters and storage of the different subbands on the disk. The parallel processing group is investigating efficient parallel algorithms for computing the wavelet transforms. Our initial results on texture based search are very encouraging, and in collaboration with the database researchers, we are investigating methods for indexing using these features.

We have given here only a brief outline of image processing issues related to the ADL. Many excellent books and journal articles are available on topic of wavelet transforms (see, for example, [1]) More details on the content based search using texture features can be found in [2].

Acknowledgments:

Thanks to Norbert Strobel and Wei-Ying Ma for generating the results, to Christoph Fischer for the air photo image, and to Sanjit Mitra and Terry Smith for their help in writing this article.

[1] M. Vitterli and C. Herley, ``Wavelets and filter banks: theory and design,'' IEEE Trans. Signal Process. vol. 40, Sept. 1992, 2207-2232.

[2] B. S. Manjunath and W. Y. Ma, ``Texture Features for browsing and retrieval of image data,'' Technical Report CIPR-TR-95-06, July 1995.

---

d-lib forum | d-lib magazine | d-lib next

---

hdl:cnri.dlib/august95-manjunath

Universe

Alexandria Digital Library: ADL

[comment] [suggestions] [information] [add a URL]

Universe

**[UNIVERSE] [EARTH] [AFRICA] [AMERICAS] [ANTARCTICA]
[ASIA] [EUROPE] [OCEANIA]
[By Subject] [By Title]**

Earth	**Jupiter**	**Mars**	**Moon**
Neptune	**Saturn**	**Uranus**	**Venus**

Misc

Remote-sensing images

MAGAZINE

Informedia Digital Video Library

Technology Outreach

Howard D. Wactlar
Carnegie Mellon University
Howard.Wactlar@cs.cmu.edu

D-Lib Magazine, July/August 1996

ISSN 1082-9873

Background

The Informedia Digital Video Library at Carnegie Mellon University is one of the NSF/DARPA/NASA jointly funded Digital Library Initiative projects, established in 1995. This particular effort focuses on search and discovery in the video medium. The Informedia project will establish a large, on-line digital video library by developing intelligent, automatic

mechanisms to populate the library and allow for full-content and knowledge-based search and retrieval via desktop computer and metropolitan area networks. Initially, the library will be populated with several thousand hours of raw and edited video drawn from licensed public television documentaries and broadcast news and special events. The library is being deployed in testbeds at local area K-12 schools, at Carnegie Mellon University, and as demonstration systems at government sponsors.

The distinguishing feature of our technical approach is the integrated application of speech, language and image understanding technologies for efficient creation and exploration of the library. Using a high-quality speech recognizer, the sound track of each videotape or broadcast, combined and aligned with closed-captioning information when available, is converted to a textual transcript. A language understanding system then analyzes and organizes the transcript and stores it in a full-text information retrieval system. Likewise, image understanding techniques are used for segmenting video sequences by automatically locating boundaries of shots, scenes, and conversations. The system thus partitions video into small-sized segments and provides alternate representations and abstractions of video content to better support information retrieval and manipulation. Exploration of the library is based

on these same techniques.

Component and Content Availability

Present

The highly modular system structure and implementation of the Informedia Digital Video Library system is itself a fertile testbed for researchers in many disciplines. Any of the component systems (e.g., speech recognition, image sequence segmentation; user interface display and control tools; text indexing, search and retrieval; video servers; network streaming protocols; dynamic pricing algorithms) can be exported for use in other research projects elsewhere. It is our intent to encourage investigation by DLI researchers who have interests in any of the components as well as the overall system use and application. We can also import components from DLI members to incorporate into the Informedia system (such as natural language processing, speech recognition, or image segmentation systems, etc.), if built to our interfaces and data types. One application, News on Demand, has already been described in this magazine (September 1995) and a discussion of some of the education-related applications will be forthcoming in

the fall.

Future

External research groups will have much the same set of opportunities, with restricted licensing and a different cost structure. Requests for involvement by external researchers will be evaluated by the project's principal investigators. Criteria include anticipated impact on the performance or function of the overall system and costs to integrate and verify their contributions if implementation is involved.

Maturing Informedia into a universally-usable system will enable easier access to researchers. We are currently moving towards an HTML Informedia client interface, utilizing commonly available technology to allow access over the Internet. To date, the interface has been a customized, proprietary, Windows 95 application. Research into Informedia's data and networking architecture will lead ultimately to using emerging commercial servers for data distribution, and satisfying their standards and protocols. Data and derived metadata in the Informedia library are collected under license, and can be licensed by others. We are now pursuing public domain data as well. NetBill, our network billing component, is a separable body of code (both in client and server) that is being made available to other DLI sites for use

as desired.

The Informedia library will continue to exist beyond the end of the current project; we expect that user support and services will be provided by third parties. We anticipate future applications of the technology in the health field, education and training, etc. Work on the various components of the Informedia Digital Video Library system (such as speech, language processing, and image understanding) will continue at Carnegie Mellon for related research efforts. We will maintain the infrastructure for creation and dissemination of digital video content, with network access as appropriate.

An important and explicit goal of this project is to accelerate acceptance of Informedia Library technologies by seeding the network community and priming the providers, both non-profit and commercial. We have assembled the project partners and organized the project structure with this goal in mind. The partnerships we have established for resources, field testing, and productization will enable us to achieve a more pervasive impact and potential commercial realization, and ultimately allow the Informedia Digital Video Library system to survive beyond its research infancy.

Home Magazine Comments

MAGAZINE

Informedia Goes to School

Early Findings from the Digital Video Library Project

Michael G. Christel
Krishna Pendyala
Carnegie Mellon University
Pittsburgh, Pennsylvania
mac@sei.cmu.edu
pendyala@cmu.edu

D-Lib Magazine, September 1996

ISSN 1082-9873

Introduction

The Informedia Digital Video Library Project at Carnegie Mellon University studies how multimedia digital libraries can be established and utilized

efficiently. Based on feedback from a series of prototype demonstrations, focus groups, and teacher workshops, an initial IDVL system was fielded at a K-12 school here in Pittsburgh in the Spring of 1996. We will use the acronym "IDVL" (for Informedia Digital Video Library) to describe both the contents of the video library and the interface to that data.

Three high school science teachers utilized the IDVL system in their Biology and Physics classes, where students made use of six client stations in a science lab to access the video library. In addition, two client stations were placed in the school library for general use. The system subsequently was used by students ages 7 to 17 enrolled in various summer school programs. This paper will report on the lessons learned from these first experiences with the IDVL, describe how that feedback has been used to design the next iteration, and conclude with future plans as the new IDVL system is delivered this month.

The Informedia Project focuses on the video medium. The project began with the premise that full content retrieval in a video library can be a difficult and time-consuming process, as evidenced by anyone trying to find a clip of interest in a videotape, or worse from a stack of videotapes. In order to improve access to relevant information, the initial IDVL logically partitions video into small-sized segments

and provides alternate representations and abstractions of video content. In the tradeoff between recall and precision, we chose to err on the side of recall and provide tools to users supporting their focusing strategies. We return a set of video segments for a query, and provide browsing mechanisms and content abstractions so that the set can be quickly reviewed by a user to find the segments of interest.

The Initial IDVL System

The initial IDVL contained 1600 video segments spanning 45 hours of video from QED Communications in Pittsburgh, the U.K.'s Open University, and other sources. The video material concentrated in the areas of biology, math, and physics. The interface to this library was a Microsoft Windows application with query and browsing capabilities, an early prototype of which was described in a September 1995 D-Lib article. Users included teachers and students in high school biology and physics classes, and students in summer school programs emphasizing science topics. Primary tasks were fact-finding and the creation of multimedia essays [56 KB] using video drawn from the library.

Transaction logs were captured automatically, recording the system's usage by teachers and

students. IDVL users were also invited to record their comments concerning the system. A few informal interviews were conducted. Some students were given a subset of QUIS and other on-line questionnaires similar to those employed by the Alexandria Project as part of a more formal evaluation of a portion of the interface.

Findings

Table 1 details the changes made to the IDVL as a result of the initial user feedback. The presentation style mirrors that used in the D-Lib article by Van House, Butler, Ogle, and Schiff, who stress the importance of iterative, user-centered design. We also adhere to and recognize the importance of that design philosophy. The changes in the IDVL shown in Table 1 were derived from user feedback, will need to be validated with user testing, and most likely will be refined further in the future. The remainder of this section will highlight a few findings and discuss the user feedback and any corrective action taken for each finding.

Automatic transaction logging is a useful feedback tool

Often students are unwilling or unable to be interviewed because of time constraints and

personality traits. However, a log of their actions with the IDVL provides a rich data set for analysis leading to IDVL improvements. For example, queries in the transaction logs which produced no results were analyzed. With the addition of better word stemming and a spell checker, over half of those queries would have became meaningful and would have returned results, and so these features were added to the new IDVL.

The transaction log helps identify how users spend their time with the IDVL. Timing every transaction, including all mouse clicks and media control interface commands, rather than only a subset of user actions, was added to the new IDVL. Future analyses can now determine the percentage of time users spend watching videos, browsing for them, formulating queries, and so on. Ideally such complete transaction data can then be used in support of further research into digital library activities. For example, Paepcke identifies five broad categories of activities in digital libraries: discover, retrieve, interpret, manage, and share. Better logs can be used to support research into how time is spent on these tasks in an educational setting with a video library.

Presentation choices will affect system usage

The manner in which search results are presented

definitely influences which result items are chosen by the user. The first users of the IDVL had the results presented to them in a pyramid layout [32 KB]. The access pattern [22 KB] reveals that the top result was accessed twice as much as those on the second row of the pyramid, while the second row items were accessed twice as much as those on the third pyramid row. Obviously the pyramid layout biased which results were selected for subsequent viewing. Given that accessing video can be expensive in terms of viewing time and perhaps data transfer, the user should be encouraged to browse through a set of results to better select an appropriate video segment, rather than being enticed into selecting the top of a pyramid layout. The new IDVL will present results in a grid of three columns, with better search results browsing tools so that the user can quickly see which words in a query phrase correspond to which search result items.

Users accessed the most visible features of the system (word query, playing a video) much more frequently than advanced features (focusing on words in a query, browser interface, filmstrips). To encourage use of other features believed to be beneficial, on-screen instructions have been added and the new built-in help system describes these features more completely. To test whether such features actually are of benefit in fact-finding and

essay creation tasks, small structured interface evaluation studies have been conducted this summer focusing on a few of these features; the data is currently being analyzed.

During interviews with users they frequently asked for the ability to do conjunctive searches ("and" all of the query words together), rather than the default disjunctive "or" processing. Like most Web search engines, the IDVL allowed conjunctive search but only by making a choice in a "Search Options" dialog which was displayed only after a menu action. Because the option was hidden from view, it was not accessed. The new IDVL has put this option on the displayed query form.

Reduce the cognitive load on the user

Users left the system in its default setting (for color, maximum results to return, style of results presentation, and numerous other interface settings) in over 90% of the sessions. The wealth of options useful for Informedia project demonstrators, e.g., data library locations, dialog synchronization features, and speech recognition settings, overwhelmed users in the school. Even beneficial options were hence ignored. The options were simplified and organized so that only a small but relevant subset is accessible to the school users. Also, context-sensitive help was

extended to better describe available options to the users.

In browsing the search results, users could move the mouse over a search results item and its one line abstract would be displayed in a pop-up text window. However, this worked only if the search results window was the active window which received the mouse events, and it took a mouse click by the user to activate the search results window. Many users were frustrated by forgetting to click the search results window in order to activate it. As a result, the extra click is not necessary in the new IDVL and the search results window becomes the active window automatically when the mouse moves into its area. Similarly, other extra mouse or key clicks are no longer needed in the new IDVL.

One of the primary tasks for the users in the school was embedding relevant clips from the IDVL into their own multimedia essays and presentations. The "embed" operation involved the use of a separate dialog box to specify the start and end times for the video to be copied by reference. (The attribution for the video segment, such as producer and title, was included in the reference to satisfy one of the first requirements we received from the school librarian.) Having the dialog box and the video playback window be separate windows meant that they could overlap

with one another, one could be hidden from view, and so on. The new IDVL fixes this obvious deficiency and puts all controls which act on the video, such as selecting a portion of the video for copying, into the video window.

Plans for Future IDVLs

The teachers are excited by the potential of using the Informedia system in their classrooms. They and the students would like to see more video content in the library, which will more than double in size with the new IDVL. Many teachers in other subjects expressed an interest to incorporate the library in their curricula if the domain would be extended beyond science topics, and so the project is pursuing copyright issues with news corporations and other content providers.

We are currently moving towards a WWW Informedia client interface, utilizing commonly available technology to allow access over the Internet. The inclusion of public domain video in the new IDVL makes it easier to share this data with others.

We will continue to work closely with teachers and students to iteratively refine the IDVL based on their needs. Initially the video itself was compelling enough to motivate both teachers and students to use the

library. We may find the need to provide more scaffolding for educational use of the library in the future. For example, one teacher in particular had great success in motivating the use of the initial IDVL for the creation of multimedia essays on student-selected topics. Other teachers may need help to replicate such experiences, perhaps in the form of workshops, Web pages, and other documentation outlining and supporting the practice of inquiry-based learning via digital resources. The UMDL Teaching and Learning Project is especially relevant to such an effort.

We realize that through better communication with the teachers and students we can improve upon the iterative, user-centered design process used to refine the IDVL. Through incorporation of others' work in the area, including work discussed at the Allerton Institute and UCLA Social Aspects of Digital Libraries Workshop, we can better track the needs and activities of our DL users. Following our first full iteration with the teachers and students, we believe the new IDVL will be used more and with greater satisfaction and benefit than was the initial system, and that this new IDVL is a step closer toward the goal of educationally beneficial digital video libraries.

Acknowledgements

This project is funded by the NSF/NASA/ARPA Digital Libraries Initiative. A complete list of sponsors is available at http://www.informedia.cs.cmu.edu/sponsors.ht ml.

Table 1. Changes to the Informedia Digital Video Library (IDVL)

FEATURE	INITIAL IDVL	FALL '96 IDVL
Library Size	45 hours	100 hours
Platform	Windows 3.1	Windows 95
Query Form	example with misspelling[4 KB]	example with misspelling[3 KB]
Spell Checking	none	potential misspellings flagged quickly for user; suggestions/corrections take place at user request
Results Display	example [33 KB]	example [46 KB]
Browsing Results Display	simple tools	extended tools with better support (less clicking; on-screen instructions)
Video Embedding Tool	example [34 KB]	example [35 KB]
Options	many tailorable features	reduced feature set for school users
Transaction Logs	simple	extended

Home Magazine Comments

NEXT ▶

hdl://cnri.dlib/september96-christel

Edge Detector Demo

Enter the URL of the image to process:

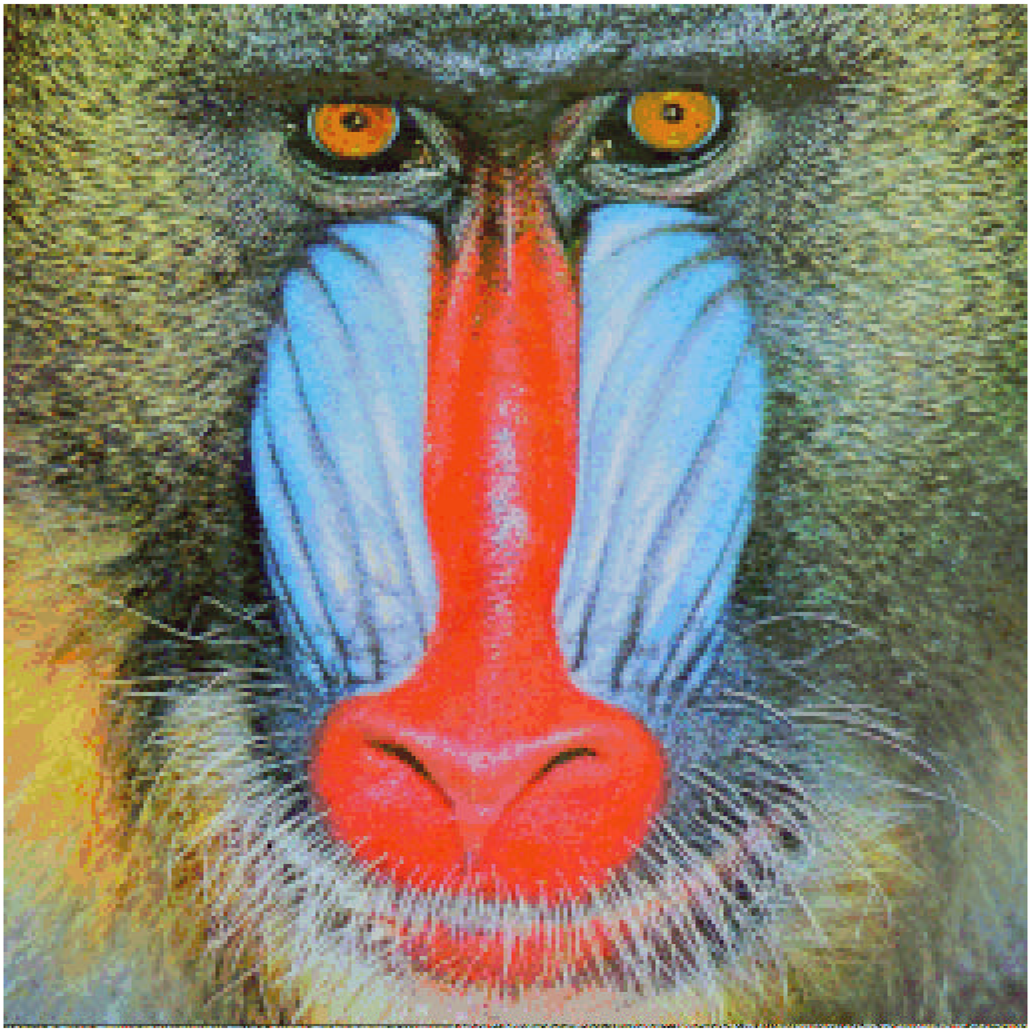
http://www.ius.cs.cmu.edu/usr/users/har/Edge/mandrill.ape.gi

For security reasons, this URL must be fully specified, and thus local file accesses are not allowed. All images are converted to grayscale before edge detection is performed. A variety of image formats are supported.

Run Edge Detector | Reset Form

Warning: The edge detection and file format conversions performed by this system can be rather time consuming... It might take a couple minutes for the result to be returned.

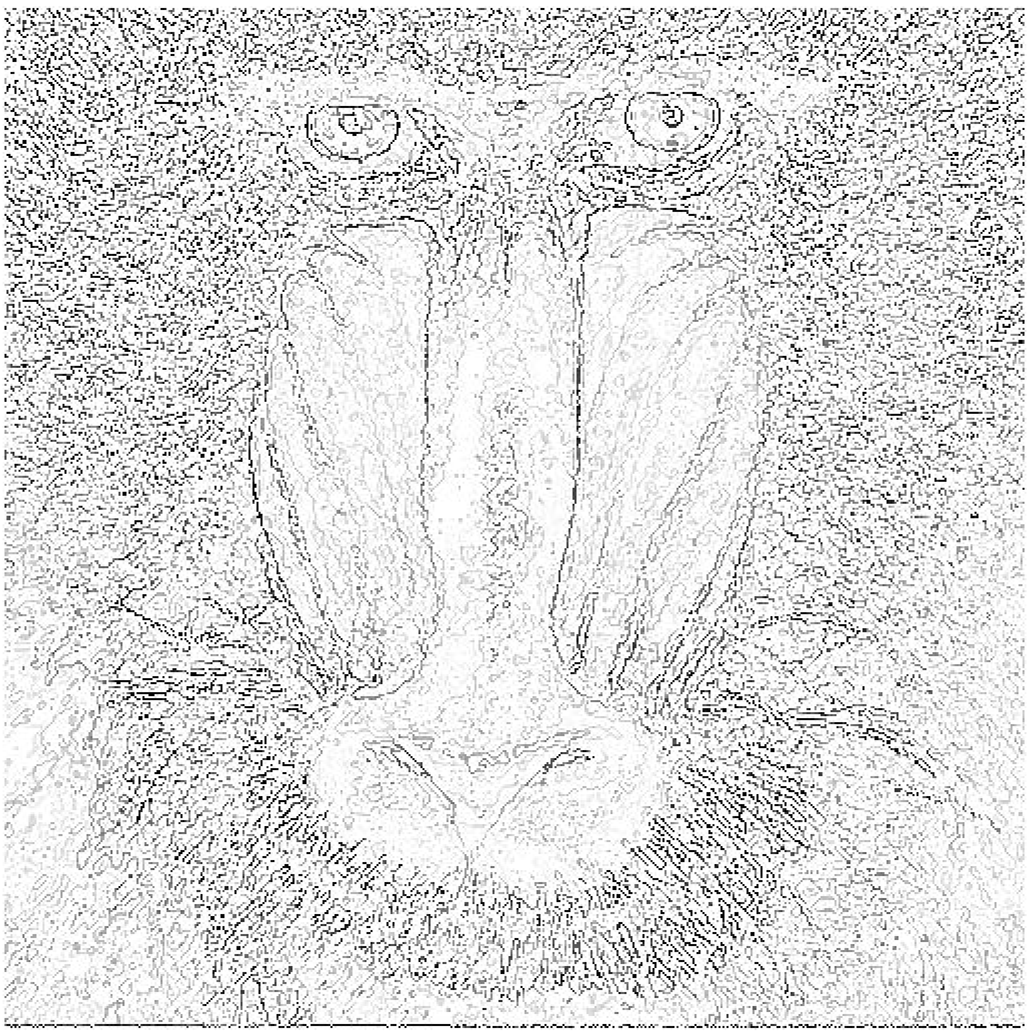
Henry A. Rowley
har@cs.cmu.edu



Edge Detector Output

For URL:

`http://www.ius.cs.cmu.edu/usr/users/har/`



CMU's Face Detector Demo

This is the front page for an interactive WWW demonstration of a face detector developed here at CMU. A detailed description of the system is available. The face detector can handle pictures of people (roughly) facing the camera in an (almost) vertical orientation. The faces can be anywhere inside the image, and range in size from 20 pixels (at the smallest) to covering the whole image.

Since the system does not run in real time, this demonstration is organized as follows. First, you can submit an image to be processed by the system. Your image may be located anywhere on the WWW. After your image is processed, you will be informed via an e-mail message.

After your image is processed, you may view it in the gallery (gallery with inlined images). There, you can see your image, with green outlines around each location that the system thinks contains a face. You can also look at the results of the system on images supplied by other people.

If you work in the area of face detection, you might find our database of test images useful.

Back to the face detector home page.
Back to the computer vision demos page.

Henry A. Rowley (har@cs.cmu.edu)
Shumeet Baluja (baluja@cs.cmu.edu)
Takeo Kanade (tk@cs.cmu.edu)

Computer Vision Demos

The Web is a great way to show off image processing algorithms. You don't even have to release your code to the public, just set up a server to do the processing locally and let folks feed it their own images.

Warning! Some of these can take a long time!

- **Behavioral model of active visual perception and invariant recognition** **(BMV)** *(Rostov State U)*
- **Canny Edge Detector** *(Carnegie Mellon U)*
- **Content-based Image Retrieval** *(Virage, Inc)*
- **Content-based Image Retrieval** - Leiden 19th Century Portrait Database *(Leiden U)*
- **Direct Least Square Fitting of Ellipses**
- **Epipolar Geometry computation** *(INRIA)*
- **Face Detection** *(CMU)*
- **ImageScape** - An AltaVista-like web search engine for images and text *(Leiden Imaging and Multimedia Group / Leiden University)*
- **ImageSearch** - A Java based image-query system which allows the user to draw sketches

and place icons which represent objects such as faces, trees, sky, etc. *(Leiden Imaging and Multimedia Group / Leiden University)*

- **Imaging Machine** *(Visioneering Research Laboratory)*
- **JACOB** - Content-based query system for video databases *(U Palermo)*
- **Kanade/Okutomi Variable Window Stereo** *(CMU)*
- **Multiresolution Contrast Based Thresholding**
- **Online Image Processing** - Including direct stereo, correlation stereo, dominant texture direction, and scale-space segmentation *(U Bremen)*
- **Segmentor** - A colorful image segmentation java applet
- **Shape Queries Using Image Databases (SQUID)**
- **Web OmniCamera** *(Columbia)*
- **WebSeer** - An image search engine for the WWW
- **ZOMAX** - A general image processing and retrieval system for the World Wide Web. Query by image content (e.g. hue) or process images (edge detection, etc). *(Intelligent Sensory Information Systems (ISIS) group / University of Amsterdam)*

Computer Vision Home Page (last updated Thu Jul

10 20:01 EDT 1997)
Text only version of this page.
Please submit new links using our forms interface or send email to vision+@cs.cmu.edu

Welcome to Vision and Autonomous Systems Center's Image Database. We were down for about a year, but we're back up now, with new features!

The goal of the VASC Image Database is to share image data sets with researcheres around the world. To facilitate this, we have created this site, which contains over 5000 images split up over nearly 200 different data sets. We also provide a mechanism by which any one of these 200 sets can be downloaded in either .tgz or .zip format.

We will be adding new data to this site as time permits. Furthermore, we will now accept datasets from other researchers, to add to our archive. If you wish to contribute stereo, motion, or other image data (particularly if you have some form of ground truth available, please contact Parag Batavia.

As a caveat, many of these datasets are quite old (5-10 years, in some cases), and do not have any form of ground truth data. What is on here is made available as-is. If you don't see any additional information with the images, there isn't any available.

Each data set has thumbnail images in jpg format. However, the full sized images themselves are stored in PNG (Portable Network Graphics) format. Because this is a relatively new image format, we have collected a set of PNG resources, available on our PNG Resources page.

Data Set Overview

- Motion Data - A large set of motion series. In most cases, sampling rate is not available.
- Road Sequences - Many road image sequences, taken from our Navlab series of vehicles.
- Stereo Data - A large set of stereo (left/right) images. Baseline information is generally not available.
- JISCT Data - Datasets prodvided by research groups at JPL, INRIA, SRI, CMU, and Teleos.

Additional Resources

- PNG Resources page.
- CMU Computer Vision Home Page - Lots of links to computer vision resources.
- Other Image Databases around the world

Maintained by:
Parag Batavia, The Robotics Institute, Carnegie Mellon University
parag@ri.cmu.edu
Last modified: Fri Jun 27 11:00:54 EDT 1997

Database and Information Systems Laboratory

Under construction

People | Research | Publications

People



Research

Publications

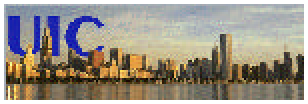
Y. A. Aslandogan, C. Thier, C. Yu, "A System for Effective Content Based Image Retrieval", ACM Multimedia '96 pp.429-30 (Prototype Demonstration)

Y. A. Aslandogan, C. Thier, Clement T. Yu, C. Liu and K. R. Nair "Design, implementation and evaluation of SCORE (a System for COntent based REtrieval of Pictures)", IEEE-ICDE-11, March 1995, pp.280-287.

K. Vadaparty, Y. Aslandogan and G. Ozsoyoglu "Towards a Unified Database Access", ACM-SIGMOD Conference 1993, pp357-366.

Y. A. Aslandogan "VQL: A Visual Query Language for Unified Database Access", M.Sc. Thesis, Case Western Reserve University, June 1994.

C.T.Yu, K.R.Nair, C.Liu, A.Aslandogan "Content based approximate picture retrieval",RIAO 94 (Prototype Demonstration)



EECS Database and Information System Laboratory

Laboratory Directors:	Clement Yu
Associated Faculty:	Jorge Lobo, A. Prasad Sistla, and Ouri Wolfson

About a dozen workstations connected by Ethernet are available for research by faculty and graduate students. The main research projects are (1) query processing in a heterogeneous environment containing relational databases, object-oriented databases, fuzzy relational databases, textual databases, pictorial databases and deductive databases; (2) retrieval of pictures based on approximate contents. This has wide applications in entertainment, education, law enforcement and military industries. (3) Automatic analysis of unstructured (in natural language) data. Consumer product surveys can be analyzed and concisely described without examining individual documents. (4)

Determining relationships between names of attributes/entities. Relationships are in the form of synonyms, homonyms, is-a relationships and aggregate relationships. This information is useful in schema integration. (5) Determine the reliability of complex systems when some of its component systems may fail. This has applications in complex systems such as nuclear reactors and oil plants. (6) Parallelization in database environment. This is important when the amount of data is huge, there is large number of users and stringent time requirements are imposed. (7) Active databases. These types of databases have immense applications in computer systems, traffic management systems, security trading and communication networks. (8) Data Replication in Mobile Computing. Minimize charges by using appropriate data replication schemes. (9) Fuzzy relational systems. This allows imprecise queries to be answered in the presence of imprecise data.

[back to Research Lab Facilities]

Home | Undergraduate Programs | Graduate Programs
EECS Newsletter | Calendar | Seminar
Positions Available | Departmental & Research Labs | EECS People
About EECS Department | About the EECS Web

For additional information, call, write or email:

The University of Illinois at Chicago
Department of **E**lectrical **E**ngineering and **C**omputer **S**cience
Science and Engineering Offices
851 South Morgan Street (M/C 154)
Chicago, Illinois 60607-7053

ph: 312/996-3422, info@eecs.uic.edu

webmaster@eecs.uic.edu
Last modified 06/30/97

The QBIC Project

This site received a 4 star rating from McKinley Group's editorial team.

QBIC(TM) -- IBM's Query By Image Content

On-line collections of images are growing larger and more common, and tools are needed to efficiently manage, organize, and navigate through them. We have developed a system called QBIC which allows queries of large image databases based on visual image content -- properties such as color percentages, color layout, and textures occurring in the images. Such queries use the visual properties of images, so you can match colors, textures and their positions without describing them in words. But, for example, a query using a red car on a green background will match a red bird on a similar green background. So we often combine content based queries with keyword and text predicates to get powerful retrieval methods for image and multimedia databases.

NEW QBIC is available for download with a free 90 day

trial license. The download package includes the image indexing and search engine (for AIX, Linux, Windows NT/Windows95, and OS/2), a Web front end, APIs for imbedding QBIC in other applications or extending QBIC with new query functions, and even a sample image collection. You can download it from IBM software download site.

NEW Examples of sites using QBIC include:

- The Art and Art History QBIC project at UC Davis.
- **Imagebase** at the Fine Arts Museums of SanFrancisco.
- Image collections from the French Ministry of Culture.

For full use licenses, contact drewc@ibm.net. Also, check out QBIC's availability in the DB2 Image Extenders, which are components of IBM's scalable, multimedia, Web-enabled DB2 Universal Database.

To try the World Wide Web QBIC search engine on a set of approximately 1,900 images, called our "demo" catalog, press the "Run Demo" button below. Other catalogs may be available in the future. You need an HTML 3.0 capable browser.

Catalog name: demo Run Demo

If you have any comments on this on-line demo, or would like to contact the QBIC group, write us at: qbicwww@almaden.ibm.com

If you would like to be on our mailing list, please enter your name in the following box and press Enter:

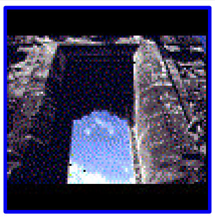
Your e-mail address:

Other links related to QBIC:

- DB2 Extenders, which now include QBIC technology in DB2 through the Image Extender.
- Ultimedia Manager 1.1, a stand-alone product that incorporates QBIC technology.
- IBM Digital Library - Related technologies for information management.
- IBM MediaMiner - Software tools for building multimedia information retrieval and mining applications.
- Technical paper requests on QBIC.

[IBM home page | Order | Search | Contact IBM | Help | (C) | (TM)]

prev **Random Images** next | Random

 view full size	 view full size	 view full size	 view full size
 view full size	 view full size	 view full size	 view full size

Columns: 4 5 6 7 8 **Rows:** 1 2 3 4

Similarity measure: Color Percentages

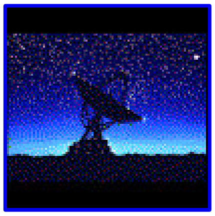
URL, file name or keywords to search on:

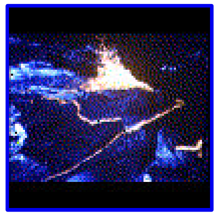
There have been* 87727 *quer

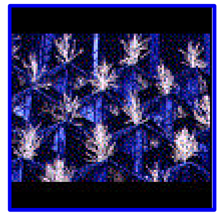
Help

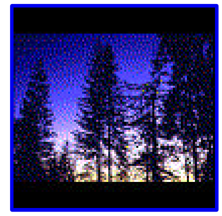
- **Image Similarity Search**: Select method, c
- **Text Only Search**: Type in keywords in the
- **URL Only Search**: Type in the URL or file r
- **Text & Similarity**: Enter keywords, select s
- Click on **Customized Search** for custom qu

prev **Images 1-8 out of 50** next Random


view full size

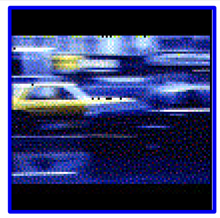

view full size

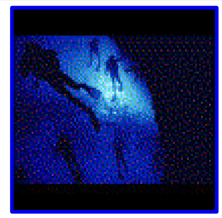

view full size


view full size


view full size


view full size


view full size


view full size

Columns: 4 5 6 7 8 **Rows:** 1 2 3 4

Similarity measure: Color Percentages

URL, file name or keywords to search on:

There have been 87728 ***quer***

- **Image Similarity Search**: Select method, c
- **Text Only Search**: Type in keywords in the
- **URL Only Search**: Type in the URL or file r
- **Text & Similarity**: Enter keywords, select s
- Click on **Customized Search** for custom qu

Help

Virage Inc.

VIR TECHNOLOGY DEMO

CLICK AN IMAGE TO FIND SIMILAR IMAGES

RANDOM

Color 10

Composition 5

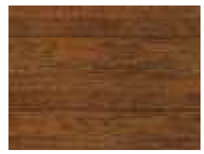
Texture 5

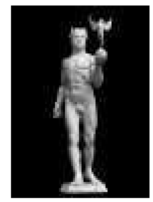
Structure 5

HELP

Show 12 images

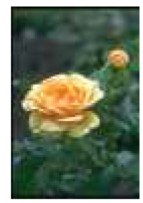
info


info


info


info


info

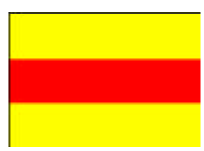

info

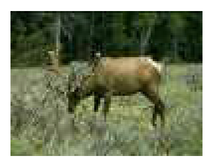

info


info


info


info


info


info

Virage Inc.

VIR TECHNOLOGY DEMO

CLICK AN IMAGE TO FIND SIMILAR IMAGES

RANDOM

Color 10

Composition 5

Texture 5

Structure 5

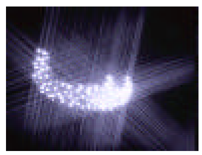
HELP

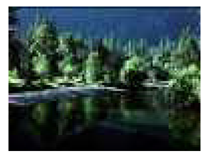
Show 12 images

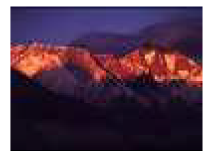

info

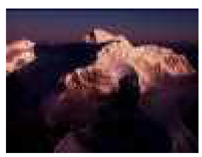

info


info

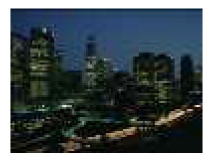

info

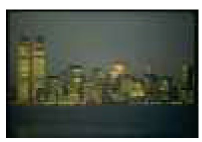

info

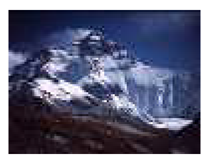

info

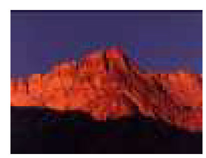

info


info


info


info


info


info

Visual Information Retrieval Technology
A Virage Perspective

Revision 4

Dr. Amarnath Gupta
Amarnath@virage.com

Virage Inc.
177 Bovet Road
Suite 520
San Mateo, CA 94402
(415) 573-3210
Fax: 573-3211

info@virage.com
http://www.virage.com

The authors wish to thank the entire Virage team for creating the vision of the technology and for giving if form.

Overview

Motivation

A significant event in the world of information systems in the past few years is the development of multimedia information systems. A multimedia information system goes beyond traditional database systems to incorporate various modes of non-textual digital data, such as digitized images and videos, in addition to textual information. It allows a user the same (or better) ease of use and flexibility of storage and access as traditional database systems. Today, thanks to an ever-increasing number of application areas like stock photography, medical imaging, digital video production, document imaging and so forth, gigabytes of image and video information are being produced every day. The need to handle this information has resulted in new technological requirements and challenges:

- Image and video data are much more voluminous than text, and need supporting technology for rapid and efficient storage and retrieval.
- There are several different modes in which a user would search for, view, and use images and videos.
- Even if multimedia information resides on different computers or locations, it should easily be available to the user.

Thus, representation, storage, retrieval, visualization and distribution of multimedia information is now a central theme both in the academic community and industry alike. At present, there is no technology available in the market that has the capability to manage this information. This white paper presents technology to meet this urgent need, produced by Virage, Inc., a pioneering company in Visual Information Management.

Virage, Inc. was formed in April of 1994 by Professor Ramesh Jain, Director of the Visual Computing Laboratory at the University of California, San Diego. The company's core technical team has done extensive academic research in multimedia information system technology and has developed a new model for such systems, called the Visual Information Management System (VIMSYS) model. Unlike traditional database systems, this model recognizes that most users prefer to search image and video information by what the image or video actually contains, rather than by keywords or descriptions associated with the visual information. This requires an information system very different from a traditional DBMS. In a traditional DBMS, an image is treated as a file name, or the raw image data exists as a binary large object (BLOB). The limitation is clear: a file name or the raw image data is useful for displaying the image, but not for describing it. In fact, textual descriptors such as a set of keywords are also inadequate to describe an image, simply because the same image might be described in different ways by different people. The only proper method by which the user can get access to the content of the image is by using image-analysis technology to extract the content from an image or video. Once extracted, the content represents most of what the user needs in order to organize, search, and locate necessary visual information.

This breakthrough concept of content extraction alleviates several technological problems. The foremost benefit is that it gives a user the power to retrieve visual information by asking a query like "Show me the pictures that look like this one." The system satisfies the query by comparing the content of the query picture with that of all target pictures in the database. This is called Query By Pictorial Example (QBPE), and is a simple form of content-based retrieval, a new paradigm in database management systems. An important concept in content-based retrieval is to determine how similar two pictures are to one another. The notion of similarity (versus exact matching as in database systems) is appropriate for visual information because multiple pictures of the same scene will not necessarily "match," although they are identical in content. In the paradigm of content-based retrieval, pictures are not simply matched, but are ranked in order of their similarity to the query picture. Another benefit is that content extraction results in very high information compression. The content of an image file may be expressed in as little as 1Kb or 2Kb, regardless of the original image size. As an image is inserted into a Virage database, the system extracts the content in terms of generic image properties such as its color, texture, shape and composition, and uses this information for all subsequent database operations. The original image is not accessed except for display. An additional strength of content extraction is that it allows the use of distributed database techniques for information storage and exchange across different computers without any platform dependence and without overloading the network bandwidth. Naturally, the VIMSYS model also supports textual attributes as would all standard databases.

The Virage Model of Visual Information

Following the aforementioned data model for visual information, Virage technology admits four layers of information abstraction: the raw image (the Image Representation Layer), the processed image (the Image Object Layer), the user's features of interest (called the Domain Object Layer) and the user's events of interest for videos (the Domain Event Layer). The top three layers form the content of the image or video.

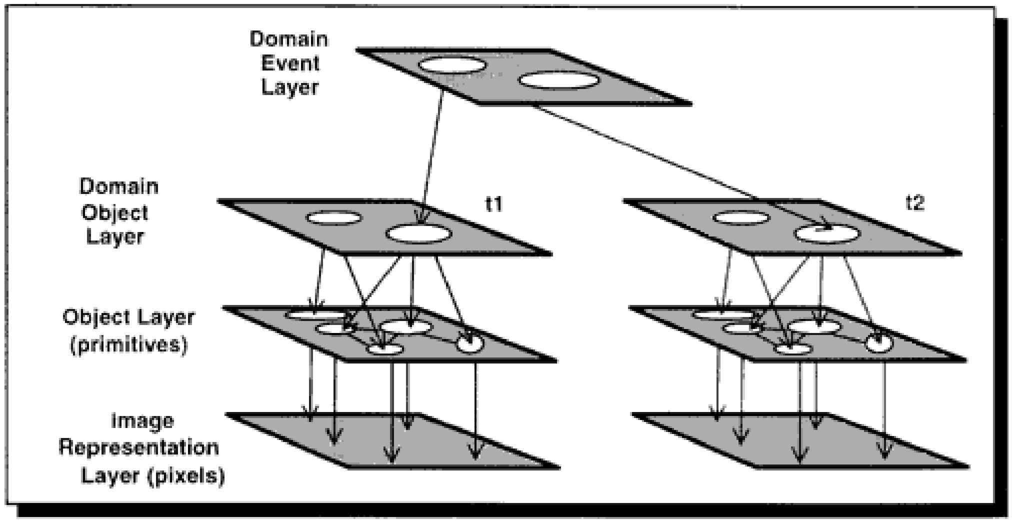


Fig. 1. The VIMSYS Model

Primitives

Image objects are computable image properties that can be localized in the spatial domain (arrangement of color), the frequency domain (sharp edge fragments), or by statistical methods (random texture). Virage calls these computed features primitives. Primitives are either global, computed over an entire image, or local, computed over smaller regions of the image. For each generic image property such as color, texture, and shape, a number of primitives are computed.

Distance Metrics

Since primitives are extracted by different computational processes, they belong to different topological spaces, each having different distance metrics defined for them. Computationally, these metrics are designed to be robust to small perturbations in the input data. Because the abstracted image primitives are defined in topological spaces, searching for similarity in any image property corresponds to finding a (partial) rank order of distances between a query primitive and other primitives in that same space. Also, since the space of image properties is essentially multidimensional, several different primitives are necessary to express the content of an image. This implies that individual distance metrics need to be combined into a composite metric using a method of weighted contributions.

Primitive Weighting

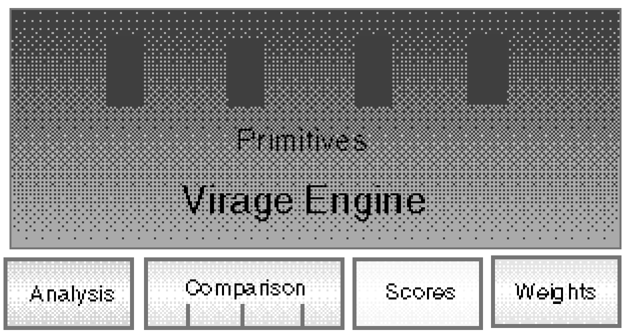
Our experience is that the overall similarity between two images lies literally "in the eye of the beholder." In other words, the perceptual distance between images is not computable in terms of topological metrics. he same user will also change his or her interpretation of similarity depending on the task at hand. To express this subjective element, Virage provides visual tools to the user to control which relative combinations of individual distances satisfies his or her needs. As the user changes the relative importance of primitives by adjusting a set of weighting factors (at query time), the Virage system dynamically recomputes the overall similarity between a single query image and the full set of target images.

Domain Events

The VIMSYS event model relates to time-dependent media such as digital video, as well as image sequences such as mamograms of the same individual over time. Time-dependent features are things such as object motion, object discontinuities, scene breaks or cuts, and (particularly in the case of video) editing features such as dissolves, fades, and wipes.

The information model described above is central to the architecture of the Virage technology. All other aspects such as the keywords associated with images, the exact nature of data management and so forth are somewhat secondary and depend on the application environments in which the technology is used. In the following section, the software aspect of this core technology is explained. This is followed by an explanation of the different environments in which the core model is embedded.

The VIR Image Engine



Virage technology is built around a core module called the Virage Engine and operates at the Image Object Level of the Virage model. There are three main functional parts of the Engine: Image Analysis, Image Comparison, and Management. These are invoked by an application developer. Typically, an application developer accesses them during image insertion, image query, and image requery (a query with the same image but with a different set of weighting factors). The following section describes the function of each unit, and how the application developer uses the Virage Application Programmer's Interface (API) to exchange information with the VIR Image Engine. The full capabilities of the Engine are decomposed into two API sets: the base Engine, and the Extensible VIR Image Engine. The base Engine provides a fixed set of primitives (color, texture, structure, etc.) while the Extensible Engine provides a set of mechanisms for defining and installing new primitives (discussed in detail later).

Image Analysis

The Image Analysis functions perform several preprocessing operations, such as smoothing and contrast enhancement, to make the image ready for different primitive-extraction routines. Each primitive-extraction routine takes a preprocessed image, and, depending on the properties of the image, computes a specific set of data for that primitive. A vector of the computed primitive data is stored in a proprietary data structure. The application simply hands the Engine a raw image buffer, and the Engine returns a pointer to a set of data containing the extracted primitive data.

The application is then responsible for storing and managing the data in a persistent fashion. The Engine operates in a "stateless" fashion (akin to Web Serving), which means it has no knowledge of how the image data is organized and stored, or how the results of queries are managed. There is no transaction management at the Engine API level. This property means that system developers and integrators need not worry about conflicts between the Virage Engine and other application components such as databases, client-server middleware, etc.

Comparisons

There are several ways to compare images using the API. Each method involves computing one or more similarity distances for a pair of primitive vectors. The computation of the similarity distance is performed in two steps. First, for each primitive such as "structure" or "global color," a similarity distance is computed. These are then combined with weights by a judiciously cho-

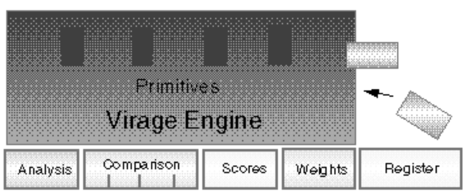
sen distance function that forms a final score that is used to rank results by similarity. Of course, the definition of "similarity" at this point is determined by the set of weights used. Applications may also synthesize a property weighting (such as "composition") by intelligently applying weights during comparisons. If "composition" is weighted low, then global primitives should be emphasized; if it is weighted high, then local primitives should be emphasized.

In order to get the result of an image comparison, the application supplies the precomputed primitive vectors from two images, together with a set of weights. The system fills in the score data structure and returns a pointer to the caller. The score structure from a comparison can be used to efficiently recompute a new score for a different set of weights (but the same query image, in other words, a re-query). An additional API call will take a score structure and a weights structure and recompute a final score (ranking) without needing to recompute the individual primitive similarities. A third API call is an extension of the first, in that the user also supplies a threshold value for the score. Any image having a distance greater than this value is considered non-qualifying, which can result in significant performance gains since it will probably not be necessary to compute similarity for all primitives.

Management

There are several supporting functions that fall in the category of "management." These include initialization, allocation and de-allocation of weights and scores structures, and management of primitive vector data.

The Extensible VIR Image Engine



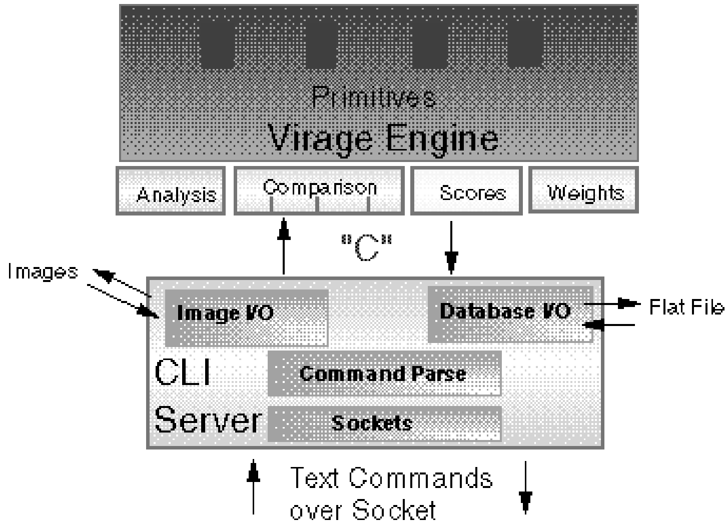
The purpose of the Extensible Engine is to provide to the application developer the flexibility of creating and adding custom-made primitives to the system. There are three important aspects of this type of development. The first is the notion of a schema of primitives, which constitute the overall visual matching mechanism for the application being developed. Applications using the Extended Engine may specify a schema of primitives that is tailored to a specific application; for example, grayscale images don't require color primitives. The second is the definition of custom primitives and incorporating them into a schema. The third is a set of image-processing support tools to assist in easily developing new primitives.

In order to define a new primitive, the developer supplies custom functions to:

- compute the data associated with the extracted features for the primitive
- compute the distance between two sets of feature data previously extracted
- perform a byte swap of the feature data for endian management
- print the values of the primitive for debugging purposes

These functions are then registered with the system and associated with a primitive ID tag. From there, it can be incorporated into any schema definition by referencing the ID tag just like a built-in primitive.

Virage Command Line Interface (VCLI)



In order to help create applications that collaborate with other applications or systems, Virage provides a command-line interface. The VCLI not only allows a user to analyze, adjust weights, and perform searches, it also permits the creation of a database that can be saved and loaded at some later time. It manages fields for alphanumeric information, allows control over the number of results displayed, and is able to display a random set of images from a database. The VCLI is invoked with a port address, and so can be used over a network. Virage has used this facility to create a demo at its web site (www.virage.com) that can be run over the World Wide Web. In this product, the user at any remote site has clickable buttons and icons that trigger VCLI calls to the server running the software.

Image Tools

Image Processing

The Virage Image Processing Toolkit (VIP) provides a set of image-processing functions to perform common operations needed in image-management applications such as contrast normalization, scaling, cropping, and so on. It also provides more advanced features such as convolutions, histograms, geometric transforms, masking operations, and so on. Of course, developers are free to use any available image-processing toolkit in conjunction with the VIR Image Engine, but the VIP toolkit provides a clean integration path as well as a complete and efficient module.

Image Format Support

Virage also provides a library of image file format readers and writers, incorporating format conversions, colorspace conversions, and various compression options. The library currently provides the support for the following formats:

❍ BMP (O/S 2)	❍ PCX	❍ SGI
❍ BMP (Windows)	❍ PICT	❍ Scitex CT
❍ GIF	❍ PSD	❍ TGA
❍ JPEG	❍ RLE (Windows)	❍ TIFF
❍ MAC	❍ RLE (Utah)	❍ PhotoCD

Integration Technologies

In addition to the tools presented so far, Virage has developed a robust set of competencies in various aspects of user interfaces and system integration. These components and techniques can be combined to form custom solutions in specific areas of visual-asset management.

Graphical User Interfaces

In addition to the programmer's interface, Virage also has a fully operational set of GUI tools necessary to develop a complete application. These includes facilities for image insertion, image query, weight-adjustment tools for requery, dialog boxes for creating field names for meta-data, inclusion of keywords, and support for several popular image file formats.

Query Canvas

The Query Canvas is a specific user-interface mechanism that is an enhancement to the query-specification environment. It consists of a bitmap editor in which the user can sketch a picture with drawing tools and color it using a color-selection palette. In addition, the user can drag and drop an image from an existing collection onto the canvas and modify it using the same drawing tools. Once an image has been created, it can be submitted as a query to the system. This tool saves the user significant initial browsing time in those cases where he or she already has an idea of what the target images should look like. Since the query canvas allows modification of images, it encompasses the functionality of the "query-by-sketch" paradigm.

Light Table

As a workgroup support tool for visual information retrieval among desktop publishers, Virage has incorporated a sharable workspace called a Light Table. During a query session, selected output from a query result can be placed in a Light Table. Once in the Light Table, they can be arranged in groups for inspection under various background illuminations, and can be annotated for opinion sharing in the group. Light Tables can be sent around by e-mail (MIME) or put on a floppy disk.

Applications

The VIR Image Engine directly implements the Visual Information Model previously described and acts as the hub around which all specific applications are constructed. The Engine serves as a central visual information retrieval service that fits into a wide range of products and applications. The Engine has been designed to allow easy development of both horizontal and vertical applications.

Vertical Applications

Because the facility of content-based image retrieval is generic, there is a large potential for developing the Virage Technology in several vertical application areas, such as:

- digital studio
- document management for offices
- digital libraries
- electronic publishing
- face matching for law enforcement agencies
- radiological information systems
- environmental image analysis
- on-line shopping
- trademark searching
- Internet publishing and searching
- remotely sensed image management for defense

To explain why the Virage Technology is a central element in these applications, let us consider some application possibilities in detail.

Environmental Imaging

Environmental scientists deal with a very large number of images. Agencies such as NASA produce numerous satellite images containing environmental information. As a specific example, the San Diego Bay Environmental Data Repository is geared towards an . . .

> ". . . understanding of the complex physical, biological and chemical processes at work in the Bay . . . it is possible to correlate these different kinds of data in both space and time and to present the data in a visual form resulting in a more complete picture of what is and what is not known about the Bay. . . . This is the kind of information that is required to assist decision makers in allocating scarce resources in more effective and informative monitoring programs by sharing data, eliminating redundant monitoring and reallocating resources to more useful and effective purposes. Another key component of this work is to provide all of these data and resultant analyses to the public-at-large . . . through the World-Wide-Web of the Internet."[1]

For such applications, the methods are applicable to any geographic area in the world. Many of the datasets for environmental information are in the form of directly captured or computer-rendered images, which depict natural (mostly geological) processes, their spatial distribution, and time progression of measurands. It is a common practice for environmental scientists to search for similar conditions around the globe, which amounts to searching for similar images.

Medical

A significant amount of effort is being spent in nation-wide health care programs for early detection of cancer. Image comparison is one of the fundamental methods for detecting suspicious regions in a medical image. Specifically, consider a cancer-screening center where a large number of fine needle aspiration cytology (FNAC) tests are conducted daily for breast cancer. We can envision a system that uses Virage's image-similarity techniques to provide an intelligent screening aid for the practicing cytologist. After the slide is prepared, it is scanned by a camera-equipped microscope at different levels of magnification. At each magnification level, the slide is compared to a database of other slides (or an existing pre-annotated atlas) at the same magnification, and similarity is computed in terms of cell density, number of nuclei, shapes of nuclei, and number of dividing cells. Suspicious regions of the slide are presented to the cytologist for closer inspection. If nothing suspicious is found, the system might suggest skipping the next higher level of magnification. The cytologist could always override the suggestion, but in general, it would save the cytologist the tedium of scanning through the entire slide, and thus increase his or her productivity.

Multimedia

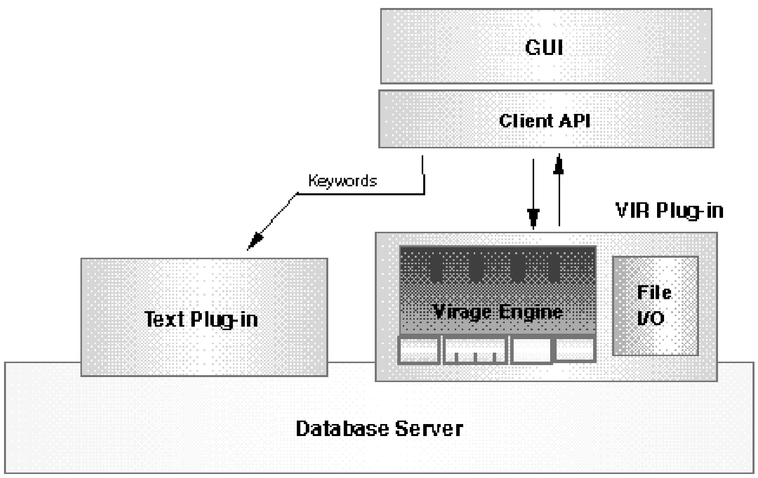
Digital libraries of videos are becoming common due to the large number of sports, news, and entertainment videos produced daily. Searching capabilities for a video library should allow queries such as "show other videos having sequences like this one." If the query sequence has a car chase in it, the system should retrieve all videos with similar scenes and make them available to the user for replay. The basic technology to achieve this relies on detection of edit points (cuts, fade-ins, and dissolve), camera movements (pan and zoom), and characterizing a segmented sub-sequence in terms of its motion properties. Also needed is a smooth integration with a database system containing textual information (such as the cast, director, and shooting locations), and other library facilities for which software products already exist.

---

1. From the San Diego Bay Project home page at
http://www.sdsc.edu/SDSC/Research/Comp_Bio/sdbay/sdbay.html

Application Integration

The Virage Engine naturally accommodates various types of databases and application frameworks. In these scenarios, the Virage Engine can be licensed to other system integrators and application developers to enrich their technology by including Virage's image-searching capabilities. Typical applications could be:



- Inserting the engine into the framework of popular database products such as Oracle, Informix, and other databases built to support multimedia. The engine can be integrated to accommodate both direct and indirect data-access models.
- An image-manipulation and processing tool like Photoshop from Adobe, Image Manager from Microsoft, or CorelDraw from Corel could use the Virage Engine for image search and management.
- An Internet crawler and search engine like WebCrawler, InfoSeek and Lycos can extend their capabilities with image finding in the network and thus help build a searchable image repository distributed over the network.

Each of these Virage-enriched applications have further growth potential in several vertical markets.

Future Directions

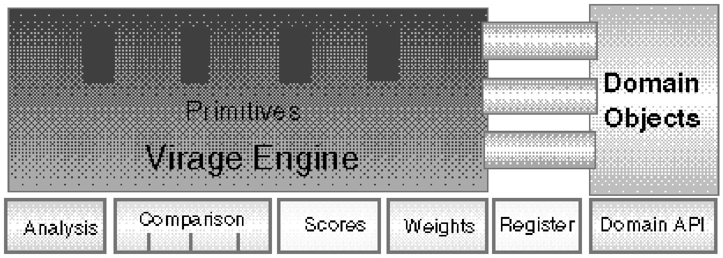
Traversing the path from the current state of visual information retrieval to full visual information management requires active research and development of new technology. The open architecture strategy latent in the design of the VIR Image Engine will allow it to interface with several kinds of software products. In this section, we briefly discuss some of the directions Virage is pursuing in achieving this goal.

Video and Multimedia Information

Virage has all the technological expertise to apply this technology to video information retrieval and expects to develop it fully in the near future. The focus of the video effort will be in providing basic video-manipulation tools for segmentation, storage, and retrieval based on scene breaks. Methods for fast preview, retrieval, timeline editing, and updating will be developed using temporal variation of low-level image properties, including motion characteristics of objects in the video. A key element of the technology is to derive index mechanisms and similarity metrics to capture the temporal change of visual features. We are exploring the market for applications of this technology, such as building video annotation systems, integration with digital video production systems, and providing an iterative query mechanism for video databases.

Domain Definition Mechanism

Recall the Domain Object Layer in the VIMSYS model, which referred to features meaningful to the users. In many vertical applications, such as medical imaging or facial image matching



for surveillance, the user has a well-defined model of the features that are of interest. These features are typically sub-objects present in the image (like an artery in an image of the cardiothoracic region), or they can be features that relate to distributional properties of image features (for example, this shade of green with that texture designates a certain farming practice). In such applications, a visual information retrieval system needs a method to specify the domain objects in terms of features computed as image objects. To incorporate these specifications into the fold of visual information retrieval, Virage is planning to develop a set of parametrically specifiable primitives, a set of domain-specific primitives, and a domain-specific mechanism of constructing domain-specific objects using these primitives.

Query Specification Mechanisms

In cases where the user has a clearer specification of the visual attributes or arrangements to be retrieved, more expressive queries can be formed. Area range restrictions can be imposed on image regions (for example, all green regions should have an area between A1 and A2), or Boolean combination of spatial attributes can be mentioned (for example, the red circles could be here or here, but not there).

Equally important is the extension of our current Query Canvas as a more general re-query mechanism. In this extension we would like the user to edit a query for each feature, such as color, texture, and structure. We are exploring what kind of analysis tools need to be provided to the user to make this re-query mechanism most effective.

Virage also plans to have a more intelligent means of handling image browsing operations by using the user's viewing history. This will increase the user's search efficiency (mean time from random to focus) for very large image collections. We also expect to provide a query- time feature-definition facility for advanced users. With this facility, the user will be able to define a feature on the fly while making a query. This new feature will be computed by the system, used for the current query session and added to the system's built-in set of features for future use.

Conclusion

The visual sense of humans is a powerful system in terms of any quantifiable dimension you might choose: color discrimination, spatial resolution, temporal response, overall bandwidth, etc. It therefore makes sense that computers should interact with humans in a way that takes advantage of this power. Visual techniques for navigation and processing of information will undoubtedly be the foundation for information systems in the future. Virage's visual search technology is a major step in this direction.

Until recently, database and imaging technologies were separate islands serving different needs in separate application areas. Now they are each expanding in scope, merging and overlapping. The intersection of these technologies is creating a new category of applications. The Virage Engine represents a concrete and viable tool to begin constructing these new applications today. It has broad horizontal capabilities to address diverse needs in imaging and visual management. Its interface is flexible and extensible. It is the first of several steps to be taken by Virage as we deliver yet more powerful tools to application developers to enable entirely new application areas in visual information management.

MAGAZINE

Finding Images/Video in Large Archives

Columbia's Content-Based Visual Query Project

Shih-Fu Chang, John R. Smith
Horace J. Meng, Hualu Wang, and Di Zhong
Department of Electrical Engineering and
Center for Telecommunications Research
Columbia University

{sfchang,jrsmith,jmeng,hwang,dzhong}@ctr.columbia.

D-Lib Magazine, February 1997

ISSN 1082-9873

Table of Contents

- An Application Driven Problem
- State of the Art
- Research Strategies

- Prototype Systems
- Testbed Support and User Evaluation
- Open Issues
- References

An Application Driven Problem

How do we find a photograph from a large archive which contains thousands or millions of pictures? How does a CNN video journalist find a specific clip from the myriad of video tapes, ranging from historical to contemporary, from sports to humanities? How do people organize and search the content of personal video tapes of family events, travel scenes, or social gatherings?

The era of "the information explosion" has brought about the wide dissemination and use of visual information, particularly, digital images and video, which we are also seeing in combination with text, audio, and graphics. The development of tools and systems that enhance image functionalities, such as searching and authoring, is critical to the effective use of visual information in the new media applications.

The current research and development of images and video search tools is driven by practical applications.

We are seeing the establishment of large digital image and video archives, such as the Corbis catalog, which includes the Bettman Archive; the Picture Exchange, which is a joint venture between Kodak and Sprint; and many digital video libraries in various domains (e.g., environment, politics, arts), such as the on-line CNN news archives.

The systems for the search and retrieval of images and video from these archives require the development of efficient and effective image query tools.

State of the Art

The use of comprehensive textual annotations provide one method for image and video search and retrieval. Today, text-based search techniques are the most direct, accurate, and efficient methods for finding "unconstrained" images and video. Text annotation is obtained by manual effort, transcripts, captions, embedded text, or hyperlinked documents. In these systems, keyword and full text searching may be enhanced by natural language processing techniques to provide great potential for categorizing and matching images.

The searching of images by their visual content complements the text-based approaches. Very often,

textual information is not sufficient. Visual features of the images and video also provide a description of their content. By exploring the synergy between textual and visual features, these image search systems may be further improved. However, it is a significant challenge to automatically reconcile inconsistence between input from these features.

Many content-based image search systems have been developed for various applications. There has been substantial progress in developing powerful tools which allow users to specify image queries by giving examples, drawing sketches, selecting visual features (e.g., color and texture), and arranging spatial structure of features. Using these approaches, the greatest success is achieved in specific domains, such as remote sensing and medical applications. The reason is that in constrained domains, it is easier to model the users' needs and to restrict the automated analysis of the images, such as to a finite set of objects.

The integration of computer vision and image processing promises a wealth of techniques for solving the image and video search problems. But new challenges remain. In unconstrained images, the set of known object classes is not available. Also, use of the image search systems varies greatly. Users may want to find the most similar images, find an

appropriate class of images, browse the image collection quickly, and so on. One unique aspect of image search systems is the active role played by users. By modeling the users and learning from them in the search process, we can better adapt to the users' subjectivity. In this way, we can adjust the search system to the fact that the perception of the image content varies between individuals, or over time.

The general system architecture for a content-based visual query system is included in Figure 1. The analysis of images and feature extraction plays important roles in both off-line and on-line processes. Other important aspects of the system include the closed interaction loop (including users), the supporting database components for retrieval and indexing, the integration with multimedia features, and the efficient user interfaces for query specification and image browsing.

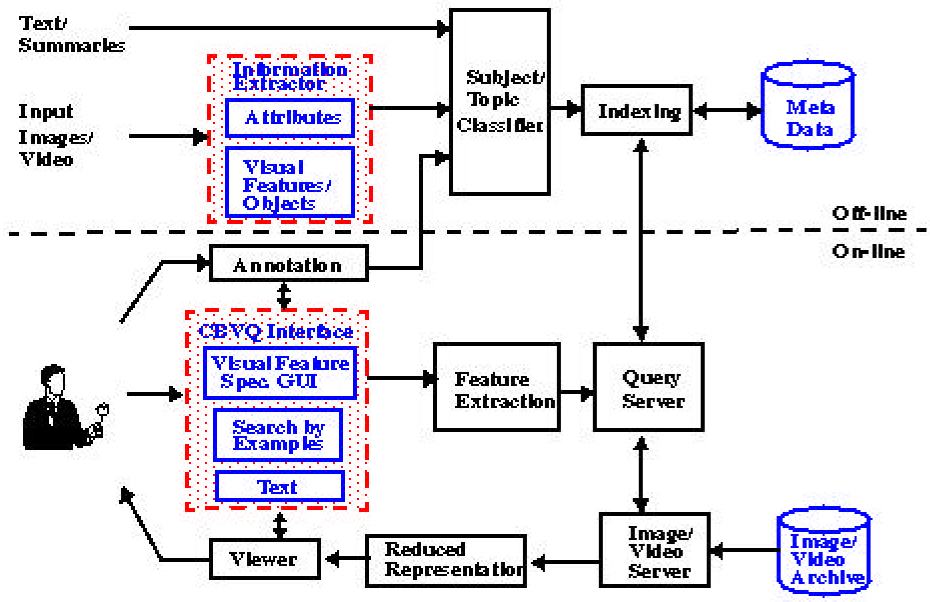


Figure 1. A general CBVQ system architecture.

The search of images is an emerging field with many exciting research challenges. The research tasks are practical, important, but not easy. In the following, we present our research strategies, prototype systems for image/video search, and our views on the important open research issues.

Research Strategies

We present our strategies for tackling the above challenging issues in this section.

Create a visual feature library by automatic image

analysis

Although today's computer vision systems cannot recognize high-level objects in unconstrained images, we are finding that low-level visual features can be used to partially characterize image content. These features also provide a potential basis for abstraction of the image semantic content. The extraction of local region features (such as color, texture, face, contour, motion) and their spatial/temporal relationships is being achieved with success. We argue that the automated segmentation of images/video objects does not need to accurately identify real world objects contained in the imagery. Our goal is to extract the "salient" visual features and index them with efficient data structures for fast and powerful querying. Semi-automated region extraction processes and use of domain knowledge may further improve the extraction process.

In the later sections, we discuss the use of automatically extracted spatial/color regions for image search, and the integration of multiple visual features for video object indexing. We use a hierarchical object based schema for feature indexing and high-level object abstraction [4] (see Figure 2). The fusion of multiple visual features improves the region extraction process. We also show that the aggregation of regions into higher level objects is

influenced by the spatial/temporal relationships of the regions. For example, Figure 3 shows the results of automatic video object segmentation and tracking. The visual features and spatial/temporal attributes of regions generate an index for searching for the video objects stored in the archive.

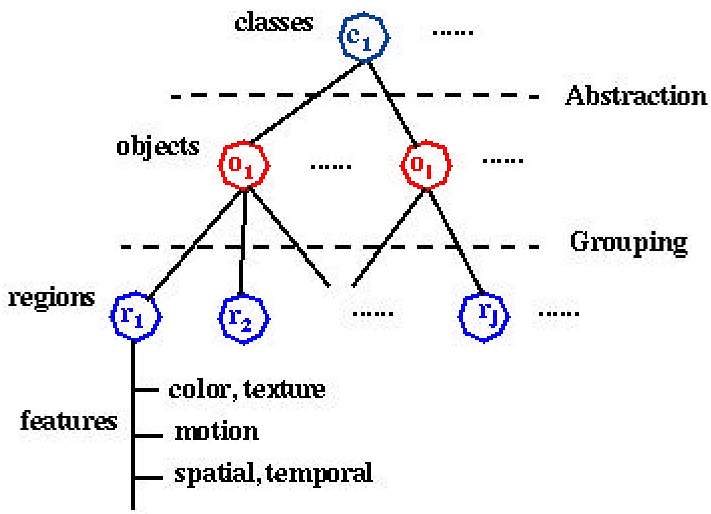


Figure 2. A hierarchical object based schema for images/video.

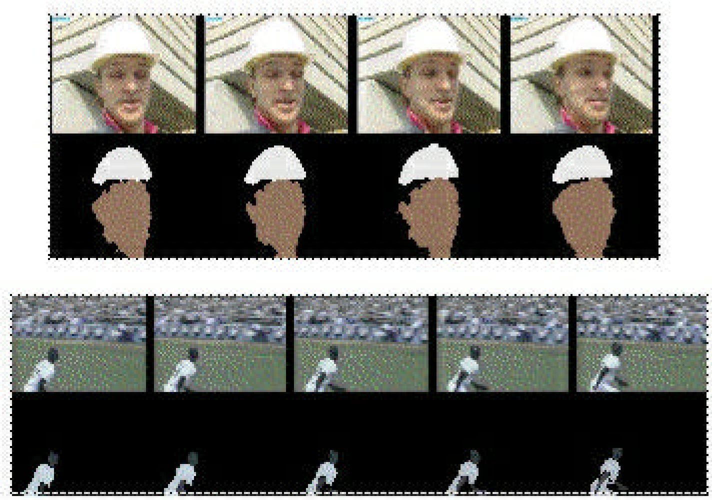


Figure 3. Examples of automatically segmented and tracked video objects.

Explore the synergy between compression and functionalities

It's impossible to anticipate the users' needs completely at the feature extraction and indexing stage. The ideal solution is that images and video are represented (for compression also) in a way that is amenable to dynamic feature extraction. Today's compression standards (such as JPEG, MPEG-1, MPEG-2), are not suited to this need. The objective in the design of these compression standards was to reduce bandwidth and increase subjective quality. Although many interesting analysis and manipulation tasks can still be achieved in today's compression

formats (as described later), the potential functionalities of the images were not considered. However, recent trends in compression, such as MPEG-4 and object-based video, have shown interest and promise in this direction. The goal is to develop a system in which the video objects are extracted, then encoded, transmitted, manipulated, and indexed flexibly with efficient adaptation to users' preference and system conditions.

Learn from users and domain ontologies

To break the barrier of decoding semantic content in images, user-interaction and domain knowledge is needed. These systems learn from the users' input as to how the low-level visual features are to be used in the matching of images at the semantic level. For example, the system may model the cases in which low-level feature search tools are successful in finding the images with the desired semantic content. In this way, the categories can be monitored and better analyzed by the system. Learning and other techniques in artificial intelligence provide great potential for these systems.

If the applications require the definition of specific semantic subjects, the feature models of images in these classes are constructed by hand and then used to match objects in the unknown images/video. This

object recognition and subject classification method provides a system for on-line information filtering. We see great potential for improving image search systems to link the low-level visual features with high-level semantics. However, in unconstrained application domains, we expect only moderate success early on.

Integrate visual and other multimedia features

Exploring the association of visual features with other multimedia features, such as text, speech, and audio, provides another potentially fruitful direction. Our experience indicates that it is more difficult to characterize the visual content of still images compared to video. Video often has text transcripts and audio that may also be analyzed, indexed, and searched. Also, images on the World Wide Web typically have text associated with them. In this domain, the use of all potential multimedia features enhances image retrieval performance.

Prototype Systems

We have developed several content-based visual query prototype systems. WebSEEk and VisualSEEk explore the problem of efficiently searching large image archives. WebClip focuses on browsing, search, and content editing of networked video.

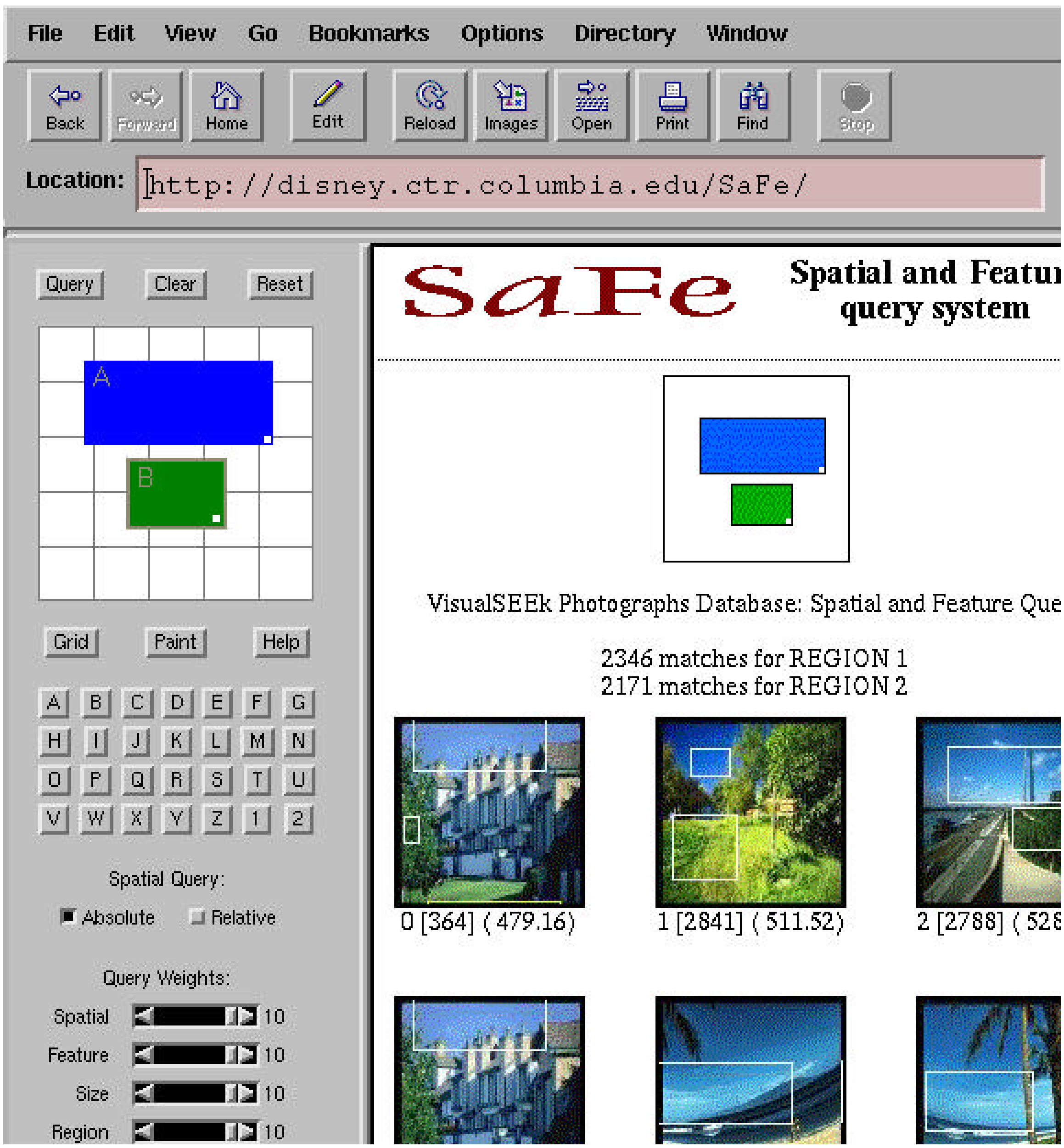
In WebSEEk, the images and video are analyzed in two separate automatic processes:

> (1) visual features (such as color histograms and color regions) are extracted and indexed off-line,
>
> (2) the associated text is parsed, and utilized to classify the images into subject classes in a customized image taxonomy (including more than 2000 classes).

More than 650,000 unconstrained images and video clips from various sources have been indexed in the initial prototype implementation. Users search for images by navigating through subject categories, or by using content-based search tools. The details of the system design and operation are described in [1].

One objective of WebSEEk is to explore the synergy between visual features and text. We also demonstrate the feasibility of image searching in a large scale testbed, the World Wide Web. We are developing more sophisticated content-based image search techniques in the VisualSEEk system [2]. VisualSEEk enhances the search capability by integrating the spatial query (like those used in geographic information systems) and the visual

feature query. Users ask the system to find images/video that include regions of matched features and spatial relationships. Figure 4 shows a query example in which two spatially arranged color patches were issued to find images with blue sky and open grass fields.



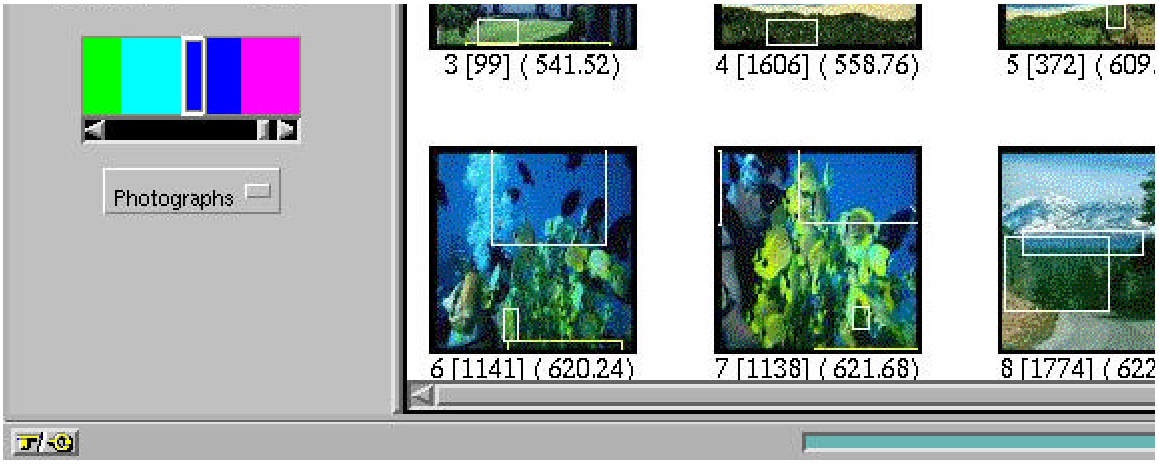


Figure 4. An example query using VisualSEEk.

For video, we have developed a system called WebClip [3], which allows for efficient browsing and editing of compressed video over the Web. One objective is to demonstrate the benefits of using compressed video without full decoding during the content analysis and manipulation stages. Visual features (like scene changes, foreground motion objects, and icon streams) can be extracted directly from the compressed video. Web users do not need high-end video decoding or processing facilities like those used in professional studios. Another objective of WebClip is to integrate the search and editing functionalities in the same environment. Tools developed in image search systems (like the above mentioned WebSEEk and VisualSEEk systems) are being ported to the video system. We are also adding new tools for searching by motion feature and temporal characteristics. After retrieval of matched

video clips, the users use web-based tools to edit the video and compose new presentations with various video special effects.

Figure 5 shows the functionality components of WebClip. The compressed video sequences are parsed to obtain visual features and objects. The browsing and search interface provides a tree-structure hierarchical scene-based interface. This display can be adapted to different browsing modalities:

(1) the time-based model,
(2) the story-based model, and
(3) the feature-based model.

The time-based model hierarchically lays out the icons of key frames from each video scene. This allows for rapid inspection of video content according to a sequential order of time. The story-based model recognizes (automatically or manually) the story structure within the video (e.g., a complete news story) and groups all video scenes belonging to the same story under a single node in the tree. The feature-based model clusters all video scenes to classes within each of which all video scenes have similar visual features. We have also undertaken new efforts to extend the joint spatial/feature query tools of the VisualSEEK system to the video domain. Video is

indexed and searched by spatial/temporal relationships and visual features of video objects contained in the video sequence.

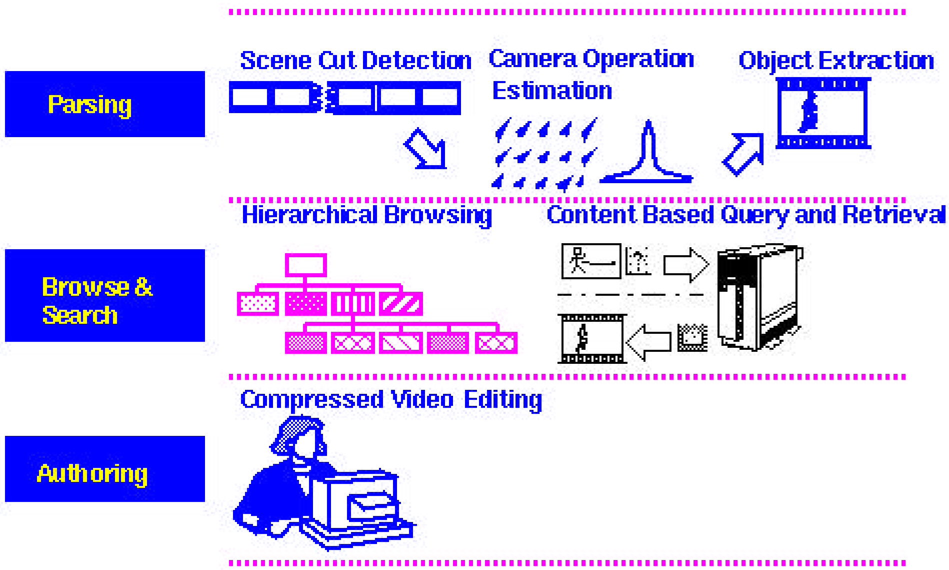


Figure 5. Functionality components of WebClip.

Testbed Support and User Evaluation

Most of the test images and video in our testbed are collected from the public domain, including data on the Web, copyright free photograph stock from commercial CD's, MPEG simulation test video, and proprietary content from local research groups. Features extracted from these images are stored in our SGI ONYX-based server, which has 50GB storage space on disk arrays, and 50GB tertiary

space on a tape archive.

Network facilities include standard Internet connections (via a T-3 line to outside), ATM connections within the campus and with external wide area networks (NYNET), and internal wireless networks running mobile IP. A video-on-demand (VoD) system which supports software-based video servers, MPEG-2 transport, and heterogeneous client terminals has been developed in the Image and Advanced TV lab. We envision the integration of our search systems with the VoD system soon to provide integrated image services.

An important work plan for the near future is the collaboration with faculty and students in the School of Journalism and at Teachers College, Columbia University. User studies and performance evaluation are being conducted in the news and education domains. One example is the Columbia Digital News Systems group [5], which integrates our efforts with others on information tracking, natural language processing, and multimedia briefing.

Open Issues

Image/video searching is a relatively new field, but it has many exciting research issues. It requires close interaction between multiple technical disciplines and

applications users. Researchers have made great progress in recent years, but a few critical issues have still not been addressed adequately. In particular, we believe that further breakthroughs need to be made in the following areas before image search systems can make significant impacts on real applications.

Effective evaluation metrics and testset

Today, there are no satisfactory methods for measuring the effectiveness of image search techniques. Precision/recall types of metrics have been used in some of the literature but are impractical due to the tedious process of measuring image relevancies. There are no standard image corpus or benchmark procedures. We believe that resolution of this issue is of top priority for researchers and users in this field.

Dynamic extraction and matching of visual features

As mentioned earlier, the image indexing and search schemes must adapt to dynamic user needs, resource conditions, and input data. In particular, the user needs and application requirements vary over time. A static set of features and matching schemes is limited. Efficient, if not real-time, methods should

be developed to perform dynamic feature extraction, matching and abstraction. Real-time is defined in three different aspects:

> (1) fast enough to process live information (like live video),
> (2) fast enough to process a large amount of new information on-line (like on-line information filtering), and
> (3) fast enough to re-process existing data in the archive.

The degree of time urgency decreases in the same order. All these aspects demand breakthroughs in image/video representation and dynamic content analysis.

Linking low-level features to high-level semantics

Today's content-based image search systems allow for image queries based on image examples, feature specification, and primitive text-based search. The WebSEEk system uses automatically extracted text in image subject classification. Other researchers have also shown some success in using newspaper photograph captions and video transcripts to assist visual content analysis. Adaptive visual feature organization through user interaction has also been proposed. But the linkage between low-level visual

features and high-level semantics is still very weak. Non-technical, general users tend to expect the same level of functionalities as those seen in today's text search systems. We admit that this is a difficult objective. But, as they are driven by critical application needs, image search systems will benefit from any breakthrough made in this direction.

Acknowledgements

This project is supported in part by the ADVENT industry partnership project at the Image and Advanced TV Lab of CTR, Columbia University, Columbia Digital Library project, and National Science Foundation (IRI-9501266). We appreciate the research collaboration in this area with Dr. Chung-Sheng Li of IBM, Dr. Kenrick Mock of Intel, Dr. Harold Stone of NEC, Dr. HongJiang Zhang of Hewlett-Packard, and Mr. Jan Stanger.

References

1. J. R. Smith and S.-F. Chang, "Searching for Images and Videos on the World-Wide Web," to appear in IEEE Multimedia Magazine, Summer, 1997. (also Columbia University CU/CTR Technical Report #459-96-25). Demo: http://www.ctr.columbia.edu/webseek

ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/

2. J. R. Smith and S.-F. Chang, "VisualSEEk: A Fully Automated Content-Based Image Query System," ACM Multimedia Conference, Boston, MA, Nov. 1996. Demo: http://www.ctr.columbia.edu/VisualSEEk) ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/

3. J. Meng and S.-F. Chang, "CVEPS: A Compressed Video Editing and Parsing System," ACM Multimedia Conference, Boston, MA, Nov. 1996. Demo: http://www.ctr.columbia.edu/WebClip ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/

4. D. Zhong and S.-F. Chang, "Video Object Model and Segmentation for Content-Based Video Indexing," IEEE Intern. Conf. on Circuits and Systems, June, 1997, Hong Kong. (special session on Networked Multimedia Technology & Application) ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/

5. A. Aho, S.-F. Chang, K. McKeown, D. Radev, J. Smith, and K. Zaman, "Columbia Digital News Systems," to appear in Workshop on Advances in Digital Libraries, 1997.

Approved for release, February 14, 1997.

Meng, Hualu Wang, and Di Zhong

hdl:cnri.dlib/february97-chang

at Columbia University

A Content-Based Image and Video Search and Catalog Tool for the Web

(press here to Browse all subjects)

Animals
birds, dinosaurs, monkeys, fishes

Architecture
bridges, lighting, domes heating

Art
painting, illustr, sketching cezanne, monet, vangogh

Ast
nasa eclip

Cats
leopards, lions, kittens, cheetahs

Celebrities
bullock, aniston, monroe, keanu

Dogs
bulldogs, puppies, coyotes, wolves

Foo
apple pizza fruits

Horror
godzilla, aliens, skeletons, monsters

Humour
simpsons, beavis, dilbert, ren/stimpy

Movies
batman, starwars, jurassic, python, blade runner, actresses

Mus
beatl rock, zepp guita

Nature
sunsets, flowers, weather, mountains

Sports
baseball, basketball, swimming, hockey, olympics, surfing

Transportation
cars, planes, titanic, motorcycles, porsches

Tra
asia, newy austr mexi

Image/Video Topic

(single word)

Search

all **videos** **color photos** **gray images** **graphi**

[*WebSEEk*] [browse] [add urls] [postcards]

WebSEEk has catalogued 665115 images and videos

Best experienced with

Microsoft Internet Explorer FREE

Click here to start.

Get the picture,

WEBSEEK

see what I mean?

This site is best viewed with

NETSCAPE Now!

Netscape Navigator.

Cool Site of the Day on Nov. 29, 1996

Cool Site of the Hour on Nov. 28, 1996 (at 10PM)

Internet Digest Best Sites of the Week for 02-Dec-96

WebSEEk is a Site

Next Up Previous
Next: Introduction **Up:** Searching for Images and

Searching for Images and Videos on the World-Wide Web

John R. Smith and Shih-Fu Chang

Department of Electrical Engineering and Center for Image Technology for New Media, Columbia University, New York, N.Y. 10027

{jrsmith, sfchang}@itnm.columbia.edu

Center for Telecommunications Research Technical Report # 459-96-25

Abstract:

We describe a prototype visual information system for searching for images and videos on the World-Wide Web. New visual information in the form of images, graphics, animations and videos is being published

on the Web at an incredible rate. However, cataloging this visual data is beyond the capabilities of current text-based Web search engines. The key to cataloging it is the marriage of text-based processing and content-based visual analysis of the images and videos. In this paper, we describe a complete system by which visual information on the Web is (1) collected by automated agents, (2) processed in both text and visual feature domains, (3) catalogued and (4) indexed for fast search and retrieval. We introduce an image and video search engine which utilizes both text-based navigation and content-based technology for searching visually through the catalogued images and videos. Finally, we provide an initial evaluation based upon the cataloging of over one half million images and videos collected from the Web.

Keywords - content-based visual query, image and video storage and retrieval, World-Wide Web.

- Introduction
- Image and Video Collection Process
- Subject Classification and Indexing
- Search, Browse and Retrieval
- Content-based Techniques

- Evaluation
- Summary and Future Work
- Acknowledgments
- References
- About this document ...

John Smith
Fri Aug 16 11:09:46 EDT 1996

James Griffioen, Brent Seales, Raj Yavatkar are assistant professors in the Department of Computer Science at the University of Kentucky

Content-based Multimedia Data Management and Efficient Remote Access

James Griffioen, Brent Seales, Raj Yavatkar

Abstract:

The recent availability of low-cost powerful workstations and PCs has caused a staggering increase in the amount and quality of digital data, which includes still images, video clips and sound. The problem of managing such a large volume of information has become very important for researchers who need efficient and simple access to the data. In this paper we present a framework for managing the storage, analysis, and access of multimedia data. We address the problems of user access and data management via an extensible graphical user interface, which is the front-end to a

semantic object-oriented multimedia management tool. We present a new method for efficient content-based searching of image and video data, and we examine several proposed solutions to the problem of network access to large multimedia databases.

Introduction

The Internet has seen explosive growth during the last few years and continues to grow at a staggering rate, adding new users from every imaginable profession each day. The primary reason for the enormous popularity of the Internet is the access it provides to vast and up-to-date repositories of information ranging from television footage to research publications. However, the large number of users and amount of information involved has prompted the need for development of utilities and software tools (termed *middleware*) that facilitate organization of repositories for easy and efficient access (searching, browsing, and retrieval).

More recently, middleware applications and services such as the World Wide Web, WAIS, gopher, and many others have attempted to organize the masses of information and provide efficient mechanisms for accessing and searching the information space.

Although these accomplishments represent a significant advance, the rapid proliferation of new information media, particularly multimedia information such as images, video, and sound, will require new or enhanced middleware services to effectively handle such information.

We are exploring new methods for enhancing middleware services to support effective and efficient access to multimedia information. Specifically, we are developing techniques that allow Internet users to access the information embedded within multimedia documents. The system allows users to manipulate, enhance, and annotate existing information and share these interpretations and modification with other users. Our design extends middleware services to support tools for multimedia data modeling and organization, embedded information extraction (semantic tagging), and efficient network transfer policies such as cooperative proxy server caching, automatic data prefetching, and hierarchical data transfers. The next subsections outline our approaches to multimedia data modeling, embedded information extraction, and efficient network transfer policies.

Multimedia Data Modeling Tools

Multimedia information contains an enormous amount of "embedded" information. Moreover, such documents have the potential to convey an infinite amount of "derived" or "associated" information. To provide access to this embedded information, we are developing tools/services to extract and derive the information hidden in the multimedia data. We are building such services (manipulating the data and interactively defining views and annotations) using the semantic modeling methods we have developed in the MOODS system [GMY93,GYM95]. The system allows users to extract automatically and manually the desired information from an image and/or assign user-dependent information to an image. The extracted and assigned information is then entered back into a database which allows users to efficiently search and query the multimedia information.

Entering the derived or generated data into the database and associating it with the original information gives other users automatic access to the conclusions, thoughts, and annotations of previous users of the information. This ability to modify, adjust, enhance, or add to the global set of information and then share that information with others is a powerful and important service. This type of service requires cooperation between the multimedia data manipulation tools described above and the

information repositories scattered across the network. Generated or extracted information must be deposited and linked with existing information so that future users will not only benefit from the original information but also from the careful analysis and insight of previous users.

Embedded Information Extraction Techniques

Conventional semantic information extraction techniques typically involve manual labeling of data.

For example, video footage is frequently annotated by hand for later searches and retrievals. We present a new efficient method for automatic extraction of semantic information from images and video. Our approach works on standard compression schemes such as JPEG, MJPEG and MPEG that are commonly used to capture, store, and transfer (via a network) video and images.

Our embedded information identification method extends the eigenspace approach which was originally proposed in a different context [,]. We show how the eigenspace approach can be modified to work on compressed image and video streams in their compressed format. The result is that we can

search images and video very quickly for visual objects like shapes, faces and textures. The objects that are found can be entered back into the database automatically using tools based on MOODS, coupling this powerful search technique with a highly organized and flexible data management framework.

Enhancing Internet Middleware

To support an evolving global information base such as the Internet, we must integrate multimedia data manipulation services with existing Internet data repositories and their access methods. In particular, we are integrating the multimedia modeling methods of MOODS and the data extraction techniques we have developed for images and video into the World Wide Web hypertext language (HTML) and the web's data access methods provided by HTTP and ftp.

To support efficient on-line interactive access to digitized information, we are working on enhancements to the existing HTML/HTTP middleware services. These enhancements include strategies for faster network retrieval such as parallel retrieval, cooperative data caching, prefetching and hierarchical data transfers.

The remaining sections discuss in more detail our approach to enhanced middleware. Section 2 elaborates on the object-orient ed framework for modeling multimedia data called MOODS. Section 3 describes one of our new techn iques for efficiently identifying objects in images and video clips. Section 4 describes techniques to impro ve network access to large digital databases. Finally, Section 5 concludes with a summary of our contributions.

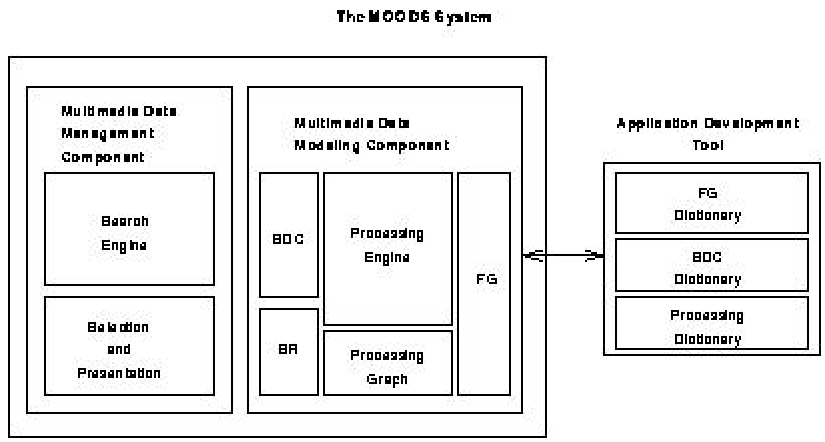
The MOODS Multimedia Management System

The goal of our research is to develop a highly flexible information management system capable of providing full access to the embedded information contained in multimedia data. Such a system must support both automatic and manual extraction of embedded information, support for multicomponent multimedia information, user and application specific views of the embedded information, and support for evolving views of the information.

The MOODS system takes an innovative approach to the design of information management systems. Unlike conventional database systems which only

support alpha-numeric data or limited access to visual data (e.g., simple storage, retrieval, and display), MOODS provides full access to all the embedded information contained in multimedia data and treats multimedia data like a firstclass object. MOODS combines a database management system with an interactive multimedia data processing system that allows users to access and manipulate the embedded semantic information. The resulting semantic information can then be queried and search using conventional database primitives. Not only does the resulting system support interactive queries and retrievals, but it also provides the functionality required to dynamically extract new or additional embedded information, the functionality to reinterpret existing embedded information, and functionality to define new data/interpretation models.

System Components



: The components of the MOODS multimedia information management system.

Figure 1 illustrates the various components of the MOODS information management system. The system consists of two main components: a *database* component and a *data modeling* component. In addition, an *application development tool* aides in the construction of modeling systems and facilitates reuse of functions an modeling systems.

The *data modeling* system distinguishes MOODS from conventional information management systems by providing the capability to extract embedded information from stored data. The ability to process data dynamically is important for several reasons. First, the processing engine provides each user and application domain with the ability to dynamically extract the embedded information that is of

importance to the user or domain. Second, a dynamic processing engine allows view to be changed or to evolve as new processing techniques become available. Third, dynamic processing reduces the data storage requirements without limiting the amount of information that can be access. Recall that the amount of embedded information is limitless. Static processing schemes severely limit the potential information that can be accessed. This is one of the primary problems plaguing conventional multimedia databases that statically link keys to embedded information. Under MOODS, the system extracts the desired information on demand.

The data modeling system consists of five components. A set of *function groups (FG)* defines the logical operations that can be applied to the data. The types of semantic information known to the system are represented by a set of *semantic data classes (SDC)*, and a set of *semantic relationship (SR)* definitions describe the logical relationship between the various semantic data classes. The *processing graph* defines the transformations that a data item may legally undergo over its lifetime. Finally a runtime *processing engine* allows users of the system to interactively apply functions to extract and identify embedded information subject to the constraints imposed by the five components described above.

The *database* component is tightly integrated with the modeling system. After the modeling system has extracted the embedded information, the extracted semantic information is automatically entered into and becomes available to the database for search or browsing via conventional database techniques. The transfer of information from the modeling system to the database means that any new information identified (or reinterpreted) by the modeling system will be made available to the user(s). Using the SR definitions, the database also supports queries on data that has not been processed or has not been processed to the level required by the query. For example, a query for a ``Cheverolet Camaro'' may be satisfied by locating all data items that contain a car and analyzing them further to determine the car's model. In this case, the database dynamically invokes the services of the modeling system.

The *application development* system is an auxiliary tool that aids in the development of new modeling systems. In particular, it maintains a complete database of all known (previously defined) function groups, semantic classes, semantic relationship models, and processing graphs. The tool allows users to construct new data models by defining their own semantic classes, relationships, and graphs, or by incorporating and combining existing structures.

Data model reuse is fostered by allowing newly developed data models to be added to the tool's database for use in future models. The system also allows users to quickly modify or convert an existing model to meet their personal needs or the needs of their application domain.

The Runtime System

To utilize the information present in multimedia data, the data must first be processed (automatically or manually) to obtain its meaningful components and the information it conveys. We refer to these basic components as *structural objects*. Example structural objects in an image might be regions of uniform intensity, an edge boundary, curved segments, or regions of uniform texture. An audio stream might contain regions consisting of a single pitch or frequency or regions containing bursts of sound. These structural objects contain no semantic information other than basic structural features. However, given an application domain (e.g., medical imaging, cartography, or meteorology), additional semantic information can be associated with a structural object or set of structural objects. Each structural object corresponds to some object or semantic category in the current application domain.

We call the domain-dependent categories *domain objects*. Each application domain typically has a reasonably well-defined set of domain objects (e.g., ventrical or tumor in medical imaging) that provide useful domain-dependent information about the structural object. It is precisely this *structural object* ↔ *domain object* mapping that must be established during the embedded information extraction process.

To illustrate the MOODS runtime operation, consider a meteorology system. Initially an (expert) meteorologist would use the application development tool to construct a multimedia data model that can extract weather information from satellite photos, heat maps, etc. This involves defining the image processing transformations and image processing functions that can be applied to extract and assign semantic meaning to the embedded information in the images.

Having defined the set of legal transformations a satellite photograph may undergo to extract the semantic meaning, the meteorologist must then define data-specific transformations that will later be used to enter data into the database. For example, the meteorologist may define one transformation process for satellite images taken at close range of smaller regions (e.g., the state of Kentucky) and another transformation process for satellite images of

larger regions (e.g., the continental United States). Once defined, new images will be automatically processed according to the appropriate transformation model and the resulting semantic information will be entered into the database.

A different meteorologist may then search the database for images containing storm activity. Using the extracted semantic information, the system will return a list of images containing stormy activity. The meteorologist may then decide to invoke the processing engine on one of the returned images to further identify the intensity of a particular storm in the image. This information will then be entered back into the database and will respond to searches asking for storms with a particular intensity.

The Data Modeling Engine

The MOODS data modeling component supports an object-oriented model for manipulating multimedia data. To extract the embedded information from a multimedia object, a user manually or automatically pushes the multimedia object through a series of transformations to identify the important semantic information in the object. At any given point in the transformation process, the object has a semantic

meaning which is made known to the database so that the object and can be selected via conventional database search/query strategies. To support semantic objects that can be manually or automatically manipulated by the user, the MOODS system provides an enhanced object-oriented programming/object-manipulation environment with the following salient features:

Dynamic Data Semantics:

The semantics associated with the data in an object will typically change often over the object's lifetime. Therefore, it is important to dynamically change the set of functions (operations) associated with an object after it is instantiated. That is, the set of functions associated with an object depends on the object's current semantic meaning. As the object is processed its semantic meaning changes resulting in a new set of functions that can be applied. For example, a satellite photograph of urban city may initially have only have the semantic meaning ``an image'' and applicable functions ``enhance'', ``segmentation'' and ``edge-detect''. However, after executing an edge-detection function, the semantic meaning may have changed to ``an image of edges'' and now has available new functions to identify edges that represent roads and functions that identify other edges a

buildings. As the semantics of the image evolve, the set of functions available to the object become more specific and powerful. Also, functions that no longer make sense may be removed. For example, segmentation on an ``edge-image'' makes no sense and would be removed.

Abstract Function Types:

Given an image, one usually has a wide range of functions available that can perform a particular image processing operation. An example of such an operation is edge detection. Several edge detection techniques exist and new methods continue to emerge. Each technique implements a different/new algorithm and may be appropriate under different circumstances. Thus, given an image object, it is advantageous not to bind the data to a particular function, but rather to an abstract function class and dynamically select an appropriate function from the abstract class at runtime.

Abstract functions simply define a logical operation, not the implementation, and postpones the binding of the actual implementation until runtime. In many cases, binding can be done automatically at runtime based on the current semantics of the data. As a result, the abstract

function completely hides the implementation of the function from the user much like an abstract data type hides its implementation.

Inheritance:

Given a raw image, two or more users (or applications) might process the same image and obtain different semantic data to be used for different purposes. For example, a satellite image might be of interest to both a geologist and a military analyst. Initial processing of the raw image (such as noise removal, enhancements) might be common to both applications before domain-specific processing takes over. In such situations, it is important to group common operations and objects together in a class and then let individual applications inherit the image attributes in a subclass avoiding duplication of efforts.

Composition:

Composition refers to the merging of two or more distinct objects into a new object. For example, two independent pictures of the same scene may be merged together to produce additional information about the scene (e.g., the depth of objects in the scene). Alternatively, a single facial image might be processed to extract the individual facial features, modify some of the

features, and then merge the modified facial features back together to create an object that can be easily searched for certain combinations of facial feature characteristics. In other words, the system must provide the ability to combine processed objects together into compound objects with new functionality that did not exist when the objects were independent.

History mechanism:

As discussed earlier, an image typically goes through a series of transformations that extract information from the image or compute new information based on the image. Thus, the state of an object must contain not only the current data and associated semantics, but also the sequence of operations that were applied to the original object to bring it to the current state. Such a history is especially useful in an interactive programming system where a user may wish to test a variety of alternatives for processing the visual information.

Note that a history mechanism also includes the notion of saving the results of past processing so that it can be reused by future users/applications. Alternatively, the original user may wish to backtrack to a previous version of the data and resume processing from that point onward.

A Prototype System

We have implemented a prototype system that allows users to define abstract function groups, semantic data classes, and a semantic processing graphs that can be used to interactively processes images to extract the embedded semantic information. The semantic information, along with the associated images, are then entered into a database (Illustra) for later retrieval. To demonstrate the power of the system, We have use the prototype to construct a processing tool that does simple feature recognition, music recognition, document title/author identification, and paleographic letter form identification in the Beowulf manuscript.

Semantic Searching

Identifying the content of digital data forms can be done most reliably by hand, but the large volume of data that is now available makes autonomous techniques necessary. Researchers in image processing and computer vision have long addressed problems such as pattern classification, object

recognition and motion detection. Very few techniques, however, have been applied to the problem of extracting and storing information from video clips.

Our new approach for tagging semantic information in video clips is based on the Karhunen-Loeve (KL) transform [Fuk90], which has been used by others for face recognition [TP91] and object recognition [MN95]. The key idea to the efficiency of our approach is that we exploit the fact that video clips are stored digitally in a transformed, compressed format. The potential loss of fidelity is completely controllable, ranging from almost no loss to a large amount of distortion. This transformed format is the basis for the video storage standards that are now in place, including the full motion JPEG (MJPEG) and MPEG formats.

The standard MJPEG and MPEG formats are actually algorithms that transform a video through many steps with the goal of reducing the storage size. Classical approaches to the video semantic search problem completely decode (uncompress) each frame of the video and then search the pixels of the frame for objects to classify or recognize. Our approach can classify objects in the compressed stream without fully restoring (uncompressing) the video to the individual pixel level. Avoiding the complete

restoration to the pixel level improves the system's computational efficiency without a loss in classification ability.

Recognition using Eigenspace

The basic idea behind the eigenspace method is to represent a large number of ``training'' images with a lower-dimensional subspace. When a new image is seen, it can be classified as being similar or very different from the training images with just a few operations in the pre-computed subspace. For example, one could compile a large number of images of an object like the White House. Once these training images are distilled using the Karhunen-Loeve (KL) transform, any new image can be classified as ``containing the White House'' or ``not containing the White House''.

Specifically, let the input to the KL process be a set of images $\mathbf{f} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\}$, represented as vectors, where each image in $\mathbf{f}$ has $n = x \times y$ pixels. We treat an image as a vector by placing the pixel values in scanline order. The first step in computing the KL transform is to subtract from each input vector the average value

$$\mathbf{m} = \frac{1}{k}\sum_{i=1}^{k}\mathbf{f}_i \qquad (1)$$

of the input vector set. This generates a new vector set $\hat{\mathbf{f}} = \{\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \ldots, \hat{\mathbf{f}}_k\}$ where $\hat{\mathbf{f}}_i = \mathbf{f}_i - \mathbf{m}$ for $i = 1, 2, \ldots, k$.

Now we compute the covariance matrix $\mathbf{C}$ of the mean-adjusted input vectors $\hat{\mathbf{f}}$:

$$\mathbf{C}_{n \times n} = \mathbf{U}\mathbf{U}^T = \begin{bmatrix}\hat{\mathbf{f}}_1 & \hat{\mathbf{f}}_2 & \ldots & \hat{\mathbf{f}}_k\end{bmatrix}\begin{bmatrix}\hat{\mathbf{f}}_1^T \\ \hat{\mathbf{f}}_2^T \\ \vdots \\ \hat{\mathbf{f}}_k^T\end{bmatrix} \qquad (2)$$

The KL transform is obtained from the eigenvectors and eigenvalues of $\mathbf{C}$ by solving an eigenstructure decomposition of the form:

$$\lambda_i \mathbf{\Lambda}_i = \mathbf{C}\ \mathbf{\Lambda}_i \qquad (3)$$

This decomposition produces **n** eigenvectors $\mathbf{\Lambda} = \{\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \ldots, \mathbf{\Lambda}_n\}$ and their associated eigenvalues $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Because $\mathbf{C}$ is real and symmetric, the eigenvectors are complete and orthogonal, forming a basis that spans the **n**-dimensional space. Thus, using the eigenvectors as a basis, the original input vectors $\mathbf{f}_i$ can be described exactly as

$$\mathbf{f}_i = \sum_{j=1}^{k} p_{i_j} \mathbf{\Lambda}_j + \mathbf{m} \qquad (4)$$

where, as before, $\mathbf{m}$ is the mean of the input vector set.

All **k** eigenvectors are needed in order for the equality to hold. But one attractive property of the eigenspace representation is that the eigenvectors can be ordered according to their associated eigenvalues. The first (largest) eigenvector is the most significant, encoding the largest variation of the input vector set. This ordering allows the original input vectors to be approximated by the first **w** eigenvectors $\{\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \ldots, \mathbf{\Lambda}_w\}$ with $w \ll k$:

$$\mathbf{f}_i \approx \sum_{j=1}^{w} p_{i_j} \mathbf{\Lambda}_j + \mathbf{m} \qquad (5)$$

Given two vectors that are projected into eigenspace, it is a well-known fact that the closer the projections in eigenspace, the more highly correlated the original vectors [MN95]. This gives us the following important fact:

> Distance in eigenspace is an approximation of cross-correlation in the image domain.

The result is that the eigenspace gives a very powerful way to classify an image with respect to a set of known images. One can simply project the new image to eigenspace and measure its Euclidean distance to the other known points in eigenspace. The closest neighbors are projected images which correlate the highest with the input.

The Compressed Domain

The main idea of our work is to formulate the eigenspace approach on input vectors that are *not* image pixel values, but have been transformed from the image domain to another representation. The semi-compressed domain is convenient since it is an intermediate form that compressed video frames and images must pass through during decompression.

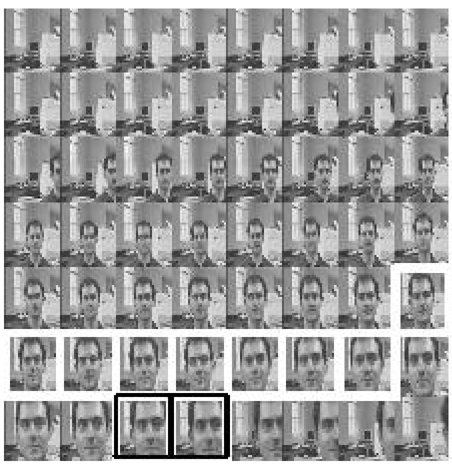
The critical fact we have proven[] is that

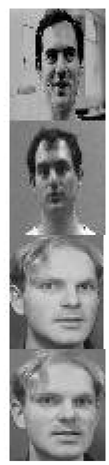
> Distance in the eigenspace constructed from semi-compressed vectors is an approximation of cross-correlation in the image domain.

One implication of this fact is that videos can be searched automatically for content using the eigenspace approach *without first decoding them*.

This gives a big gain in efficiency without much loss of classification performance.

Experimental Results





: The input video sequence appears on the left while the faces identified appear on the right. The white boxed frames are those recognized as containing subject 27. Black boxed frames are false matches.

We demonstrate the results of our method using a short (56 frame) video clip of a human face. The goal is to identify people (faces) that appear in an MPEG video clip without completely decoding the MPEG.

We constructed a semi-compressed-domain eigenspace from two poses of 26 people in the Olivetti Face Database and two poses from a local subject. The local subject, the 27th person, was also the subject pictured in the test video clip. Figure 2 shows the 56 frame video clip on the left. The four pictures on the right show the poses from the database that were found to be present somewhere in the video. The top two images on the right are the two poses of subject 27 who was correctly identified. The two poses of subject 27 were taken on different days, with different clothes and under different lighting conditions than the video clip.

Matching was done in the semi-compressed eigenspace, using only the DCT frames of the MPEG sequence. We used 10 eigenvectors (of the possible 54) for the results shown. This was because of scale, since the scale of subject 18 matches those frames better than subject 27. The highlighted boxes in the video sequence on the left identify the frames that were found to match an image from the database. Frames 40 - 47 were correctly recognized as subject 27. Only frames 51 and 52 were incorrectly identified as matching subject 18 from the database. Despite the efficiency and accuracy of our algorithm, this example shows that false positives are still possible

and a higher-level thresholding policy must be used to eliminate these false matches.

Finally, we measured the computational time necessary for scanning the frames of an MPEG video clip. We instrumented the public domain MPEG software from UC-Berkeley, and compared the execution time for obtaining DCT frames (before the inverse DCT) to the time necessary to obtain the complete image (after the inverse DCT). We found that the dominant cost is the inverse DCT, and that obtaining only the DCT frames takes one half the time of completely decoding the images.

Network Access To Multimedia Data

Although existing Internet middleware tools such as the WorldWideWeb, Gopher, Archie, and others provide access to multimedia information, we believe that they are not capable of dealing with the massive scale and size of typical digitized documents and their associated information i.e., annotations. For instance, a Beowulf manuscript consists of hundreds of pages of high fidelity images where each image is 21 MB or larger [Kie95,Kie94a,Kie94b]. Existing tools are not designed to allow user manipulation of the information on this scale and to efficiently handle

(display/search/browse) the massive digitized documents we envision.

To support efficient on-line interactive access to such digitized information, we are investigating the following enhancements to the existing HTML/HTTP middleware services:

Cooperative Data Caching and Parallel Retrieval

Currently, HTTP proxy servers support a limited form of information caching[LA94] in which a p roxy server caches information accessed on behalf of several nearby clients, typically located on a LAN. However, proxy servers located within the same geographical region do not communicate with each other to share information in their caches. Instead, a proxy server contacts the original information provider (e.g. server at a given URL) whenever the desired information is not found in its cache. This organizational structure can still cause bottlenecks at popular URL servers and does not use network bandwidth efficiently. An alternative is to dynamically discover proxy servers in the immediate vicinity that currently cache the desired information. For scalability reasons, such cooperation among proxy servers should be as stateless as possible for scalability.

We have extended the HTTP proxy server mechanism to exploit the IP/UDP Multicast (group communication) to dynamically locate other proxy servers that have the desired URL document in their caches. To reduce the host processing bottlenecks at proxy servers and allow real-time retrieval and playback of multimedia documents, we have added extensions for parallel retrieval of parts of a document from different proxy servers.

Data Prefetching

To hide the communication latencies across a WAN that spans a huge geographic area as the Internet requires heuristic algorithms that predict future access patterns with a high degree of confidence and moves documents to local or proxie caches in advance of the need for the data. We are examining automated methods for predicting future document accesses based on past access histories [GA94] . Such methods must be careful not to waste valuable (shared) Internet bandwidth on unnecessary data but yet make the maximal amount of data locally available before the data is accessed.

We have developed a heuristic-based prefetching

algorithm that can, with high accuracy, retrieve multimedia documents before they are needed. The algorithm takes a probabilistic approach using a probability graph that limits the amount of information that needs to be maintained about past accesses but yet can select prefetch documents with a high degree of confidence.

Hierarchical Data Transfer

Responsiveness of an on-line data access system depends largely on the quality of the underlying communication link. With the proliferation of various network technologies including wireless communication and slow-speed dial-up links, an on-line data access/retrieval system must be designed to perform well over slow speed links as well as congested network paths. One increasingly popular technique for accommodating heterogeneous networks is hierarchical encoding. Multimedia information is stored at multiple resolutions, and the appropriate level of resolution is selected and transferred automatically based on parameters such as the speed of the link. We are currently investigating techniques for the use of hierarchical data encoding in digital libraries. For example, if a user wishes to view an archived manuscript on-line

when the network is congested, the system automatically compensates for the low transfer rate by fetching a lower-resolution version of the image data. Users who wish to wait can still run the system in a fixed-resolution mode. However, quick access to lower resolution contents is often desirable when browsing a document. Dynamic hierarchical data transfer requires new encoding and storage formats and also adaptable communication protocols. This is the focus of some of our ongoing research.

Conclusion

Digitized multimedia data such as images, video, and audio is rapidly becomming commonplace will soon replace conventional alpha-numeric data as the primary data format. New techniques are need to access, manage, and search these new multimedia data types. We have presented a flexible multimedia data management system called MOODS that treats the embedded semantic information contained in a multimedia document as a first class entity that can be identified and searched for like conventional data. We also presented an efficient algorithm for automatic identification of semantic information in a compressed images and video. Finally, we described extensions to existing middleware languages and

protocols such as HTML/HTTP to improve remote network access to multimedia data, an increasingly important problem as we move to a completely networked world.

References

Fuk90
K. Fukunaga. *Introduction to Statistical Recognition*. Academic Press, 1990.

GA94
James Griffioen and Randy Appleton. Reducing File System Latency using a Predictive Approach. In *Proceedings of the 1994 Summer USENIX Conference*, June 1994.

GMY93
James Griffioen, Rajiv Mehrotra, and Rajendra Yavatkar. An Object-Oriented Model for Image Information Representation. In *Proceedings of the Second International Conference on Information and Knowledge Management*, Nov 1993.

GYM95
J. Griffioen, R. Yavatkar, and R. Mehrotra. A Semantic Data Model for Embedded Image

Information. In *Proceedings of the 1994 IST/SPIE Symposium on Electronic Imaging Science and Technology*, February 1995.

Kie94a
Kevin Kiernan. Electronic Beowulf. *National Geographic*, December 1994.

Kie94b
Kevin Kiernan. Scholarly Publishing on the Electronic Networks: Gateways, Gatekeepers, and Roles in the Information Omniverse. *Digital Preservation, Restoration, and Dissemination of Medieval Manuscripts*, 1994.

Kie95
Kevin Kiernan. The Electronic Beowulf. *Computers in Libraries*, February 1995.

LA94
Ari Luotenen and Kevin Altis. World Wide Web Proxies. In *Proceedings of the CERN WWW Conference*, August 1994.

MN95
H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, 14:5--24, 1995.

TP91
M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991.

About this document ...

Griffioen/Seales

...Access
This work is supported in part by the National Science Foundation and the Center for Computational Sciences and the College of Arts and Sciences at the University of Kentucky

...Database
We gratefully acknowledge the Olivetti Research Laboratory for making their data publicly available at `http://www.cam-orl.co.uk/facedatabase.ht`

James Griffioen
Mon Oct 9 17:54:30 EDT 1995

ERCIM Research Report - 04/95-R035

Storage and Retrieval Techniques for Multimedia Data

by P. Zezula and C. Galindo-Legaria (IEI-CNR)

Abstract

The success of multimedia information systems to adequately meet the needs of accessing and presenting essential information from a large multimedia information space, depends heavily on the existence and proper application of storage techniques suitable for this task. This bibliography contains surveys of some papers which have recently appeared on this subject. The main emphasis is given to the modeling, content retrieval, and management of continuous multimedia objects. However, distributed algorithms and storage

techniques for database mining and mobile wireless computing, i.e. the emerging applications which are closely related to the multimedia domain, are also brie y referenced.

Interested?

load a PostScript file of the report to your disk (use the "load to disc" option)
PDF version of the report (For PDF files, a free viewer (Acrobat Reader) is available from Adobe Systems. You can use Acrobat Readers with the World-Wide Web. Here's How.)

---

return to the research reports content list

Storage and Retrieval Techniques for Multimedia Data

(annotated bibliography)

by
P. Zezula and C. Galindo-Legaria
IEI-CNR, Via S. Maria 46
56126 Pisa, Italy

Abstract

The success of multimedia information systems to adequately meet the needs of accessing and presenting essential information from a large multimedia information space, depends heavily on the existence and proper application of storage techniques suitable for this task. This bibliography contains surveys of some papers which have recently appeared on this subject. The main emphasis is given to the modeling, content retrieval, and management of continuous multimedia objects. However, distributed algorithms and storage techniques for database mining and mobile wireless computing, i.e. the emerging applications which are closely related to the multimedia domain, are also briefly referenced.

1 Introduction

The existence of massive volumes of multimedia data makes the performance aspect of multimedia information management systems the critical implementation issue. Several very important application environments, such as digital video libraries or medical information systems, have claimed requirements for storage of many trillions of bytes of data. Because of the huge

amount of data available, such environments naturally demand the use of a storage hierarchy, and a carefully optimized environment for data placement and retrieval of real-time, delay-sensitive, and synchronized multimedia data is necessary.

But, it is not only the large volume and the real-time characteristics of the multimedia data which make the performance problem complex. Another serious performance related problem concerns efficiency in which the information, a user is looking for, is provided to him. This demands not only the use of powerful, uniform, and extensible models for describing the concepts stored in the information base, but also index structure extensions for effective and efficient content-based retrieval of information.

Even though the area of storage techniques for multimedia data is quite new, the available literature is already extensive. What makes the situation even more complex is that the boundaries of this topic are not very well defined, because the current research in such areas as distributed algorithms, parallel data processing, mobile wireless computing, database mining, and computer networks, does consider multimedia data as well.

The following annotated bibliography should be considered as a summary of literature which the authors of this paper have studied in order to enter this new, interesting, and promising area of research. For this reasons, there are no claims for completeness or optimal structuring of the bibliography. But, since we felt that such text might be useful also for other people interested in this subject, we have decided to share the knowledge we have learned.

2 Multimedia data modeling

Data modeling is essential to the design and maintenance of information systems, as it provides the conceptual structure for the interpretation and use of stored data. Recent work has started to address the issue of modeling of multimedia data, whose *continuous* nature cannot be handled appropriately by traditional database modeling tools. This media continuity is made manifest in two senses. First, time flow is central to the meaning of items, which have a specific duration and whose relationships are heavily time-dependent —e. g. sound is recorded on a CD at 44100 samples per second, and its reproduction must be coordinated in time with a video signal in case the sound consists of an explanation of the video images. Second, items are not always

decomposable into discrete, atomic units, from a conceptual point of view —e. g. a digital video signal may be composed of a sequence of bits, but only subsequences, images, or image fragments may make sense at the conceptual level; there are no natural "primitive components."

A collection of high-level modeling primitives applicable, in principle, to different time-based data is presented in [22]. The theme of the paper is that databases should handle multimedia data not as BLOBs (i. e. large sequences of uninterpreted bits), but as *streams*, with explicit modeling of time, derivation, and composition dependencies between items.

An modeling technique for video/audio data based on film theory is presented in [26]. The paper adapts known concepts from film analysis, emphasizing the meaningful composition of elements in terms of a narrative. It also shows the difficulty of building a unique conceptual structure for a multimedia item, whose human interpretation is open-ended in comparison with the highly schematic tuple of string/integer values traditionally used in databases.

A set of n-ary relation primitives to describe time-dependencies between multimedia items is presented in [32]. Dependencies modeled include simultaneity of sequentiality, for example. A logical interpretation of the primitives is described, and algorithms to create and traverse this representation are given.

There is relatively little theoretical work on multimedia databases. Some initial work on databases for geometrical information is presented in [36]. The paper tries to adapt the basic theoretical concept of "genericity" of relational databases to the case where data does have some intrinsic semantics attached to it (geometrical, in this case).

summary:

- modeling of time-based media [22]
- video information contents and film theory [26]
- interval-based models for time-dependent multimedia data [32]
- theory of spatial database queries [36]

3 Content-based retrieval

Access structures which support efficient content-based retrieval of multimedia objects, form one of the most important parts of multimedia information management systems. One important stream of research effort in developing indexing structures for spatial and temporal data is reviewed in [41]. In the following, we briefly survey some recent techniques suitable for image and text data indexing.

3.1 Image and space retrieval

The problem of multi-dimensional range searching in both main memory and secondary memory has been the subject of much research. Many elegant data structures (such as the priority tree, segment tree, and interval tree) have been proposed for use in main memory for special cases of 2-dimensional range searching. The path caching technique, proposed in [37], is able to convert many of these in-core data structures into secondary storage structures. These new data structures have optimal query response time at the expense of small storage overhead.

When points are stored in a spatial access method, such as an R-tree, an important problem for a query optimizer is the estimation of the search effort for executing range queries. Traditional assumptions for such methods are the uniformity and independence, which make the analysis tractable, but which do not reflect the reality where the data are indeed skewed. A formula that estimates the number of disk accesses for range queries for R-trees, given only the fractal dimension of the point set and its count, is developed in [17]. Experiment on real data sets show that the formula is very accurate: the relative error is usually below 5%, and it rarely exceeds 10%.

The content-based image indexing problem is formulated in [15] as a multi-dimensional nearest-neighbor search problem. An optimistic vantage-point tree algorithm is developed that can dynamically adapt the indexed search process to the characteristics of given queries. The performance study indicates that the system typically needs to touch less than 20% of the index entries for a query.

summary:

- skewed space data distribution [17]

- path caching (external 2-dimensional searching) [37]
- content-based image indexing [15]
- indexing spatial and temporal data [41]

3.2 Text retrieval

The application of database technology is seen essential to the operation of a conventional business enterprise. However, there is a universe of business information, namely text, which is currently stored, accessed, and manipulated in an add hoc fashion with none of the consistency and discipline of the database approach. A tutorial on theory, practice and experience in text dominated databases can be found in [23].

Free text retrieval through signature files is an important indexing technique which can significantly benefit from a parallel architecture. A parallel bit-sliced signature file method is proposed in [35], which can index very large files where the size of the signature file exceeds the available (parallel) main memory. The method uses a partial fetch slice swapping algorithm, thus accesses only the necessary number of bit slices. Arithmetic examples show that the method is able to handle a 128GB database with a 2sec response time on a machine with the characteristics of the Connection Machine.

The fundamental problem in the signature construction is to find the optimum signature weight to minimize the false drop probability. Even though the problem was extensively studied in the past, a new formula for computing the false drop probability is presented in [44]. The formula is easier to analyze so that optimal solutions can be more adequately derived. Performance results show that these solutions are better than the solutions obtained before.

summary:

- text dominated databases [23]
- parallel bit-sliced signature files [35]
- performance analysis of superimposed signature files [44]

4 Management of multimedia data

Recent years have witnessed a significant increase in electronic management of information. In particular, many types of information that traditionally were considered to be analog, such as image and sound, are now processed and utilized in digital form. According to [9], this advancement comes with tremendous opportunities and challenges for information systems to better meet information users' needs to manipulate multimedia information in a natural and effective manner.

To embed various media into divers applications, multimedia information systems (MMIS), which allow for creation, processing, storage, management, retrieval, transfer, and presentation of multimedia information, have been proposed. These systems resulted from the synergism of many different fields, including databases, digital signal processing, image processing, optical communication, mass storage, artificial intelligence, new paradigms for programming languages, and text processing.

SuiteSound, [40], is a tool for manipulating multimedia flows, such as digital audio. It supports development of collaborative applications using multimedia by integrating flexible, easy-to-use digital audio with a flexible, easy-to-use object-based system. SuiteSound applications readily combine live or recorded audio with editable representations of user objects.

A low-cost storage architecture for a movie on demand (MOD) server that relies principally on disks is presented in [34]. The high bandwidths of disks in conjunction with a clever strategy for striping movies on them is utilized in order to enable simultaneous access and transmission of different portions of a movie.

summary:

- multimedia information systems [9]
- a system for distributed collaborative multimedia [40]
- storage server for a movie on demand [34]

4.1 Continuous media management

Though the basic storage support for multimedia data types may indeed be the same as for conventional database types, such as records, the kinds of

data manipulations and application interfaces needed, especially for continuous media types, are vastly different from those found in conventional data processing. According to [33], addressing these requires significant changes in storage system architectures, both for database management systems and for the underlying operating system and network facilities.

For continuous media, the mode of delivery of data is all important. The human eye, and more so the human ear, are sensitive to shifts and skips in the time domain, so rates of presentation are significant. There are also synchronization demands, because most continuous media presentations combine more than one data stream. The data storage volumes demand tertiary storage, which introduces seconds of latency on some accesses. Data rates are high enough, relative to hardware capabilities, that buffering of data in the storage system or application memory is only suitable for overcoming very small latencies or time shifts. In fact, getting data too early becomes as much a problem as getting data too late, so the schedule with which data is delivered to an application becomes critical.

A study of the ways, in which storage system architectures must change in order to provide the constrained-latency storage access on continuous media (taking into account operating system and network support as well as database management) is presented in [33]. The problems of temporal synchronization of various data streams in multimedia information, exchanged between users over a high speed network, are addressed in [39].

Acme, [6], is a network server for digital audio and video I/O. It lets users specify their synchronization requirements through an abstraction called a logical time system.

The Continuous Media File System, CMFS [7], supports real-time storage and retrieval of continuous media data (digital audio and video) on a disk. CMFS clients read or write files in "sessions," each with a guarantied minimum data rate. Multiple sessions, perhaps with different rates, and non-real-time access can proceed concurrently. Issues, such as real-time semantics of sessions, disk layout, an acceptance test for new sessions, and disk scheduling policy are also addressed in [7].

The problems which appear in integrating storage and transmission of multimedia data with computing are discussed in [38]. The design of a high-performance multimedia storage server that addresses the above complex design issues, is the subject of this paper.

Physical storage organizations for time-dependent multimedia data are

also studied in [11]. A model is derived to relate disk characteristics to different media recording and playback rates in order to derive their storage patterns. These storage organizations guarantee that as long as a multimedia delivery process is running, "starvation" never appears.

Optical disks are among the most promising secondary storage devices for multimedia data. The placement of multimedia data on optical disk is of primary importance because this data is being extracted from the optical disk in real-time (e.g. real-time for audio data means that system hiccups greater than 30 msec cannot be tolerated). Methods for an efficient placement of audio data on optical disks for real-time applications are developed in [48].

summary

- audio data placement on optical disks for real-time applications [48]
- I/O server and its synchronization mechanism [6]
- synchronization of multimedia data streams [39]
- a file system for continuous media [7]
- storage techniques for continuous multimedia [38]
- organizations for time-dependent multimedia data [11]
- storage system architectures for continuous media data [33]

4.2 Distribution, replication, and storage hierarchies

The special requirements of multimedia data require novel hardware configurations, and algorithms tuned for such configurations. First, the large storage requirements introduce the need of tertiary storage not as an infrequently used archive, but as an essential component of the normal operation of a system. Second, the strict rate-of-delivery requirements introduce, on one hand, the need of parallel disk access to increase bandwidth; and, on the other hand, the need of scheduling algorithms sensitive enough to deliver multiple streams of data with little variance on a per-stream specified rate-of-delivery, in order not to overflow the buffers of the reproduction device.

Various research results derived from a terrain-visualization project developed at the Lawrence Berkely Laboratory are reported in [45, 12, 13]. An

overview of the system architecture, which declusters data over a high-speed, geographically distributed network is given in [45]. The architecture relies on ATM technology to merge multiple slow streams arriving from disks into a single combined stream of the appropriate bandwidth. Only off-the-shelf, relatively inexpensive hardware is required. Motivated by this project, [12] deals with the problem of how to optimize the time of retrieval of a set of partially replicated items —i. e. which replica to read of each item. Also, [13] deals with the optimization of allocation of object streams on sequential tertiary storage. Optimal allocation must balance the time required to mount, say a tape, with the time required to seek the required data within such tape.

A series of papers by Ghandeharizadeh and his colleagues examines in depth the alternative architectures and policies to deliver multiple streams of data retrieved from tertiary storage. [19] opens this series of articles by presenting the problem, and proposing that objects should be downloaded and replicated from tertiary storage into multiple disks dynamically, depending on the order of requests, free resources, and required bandwidth of each stream. Then, [20, 21] examine, using simulation, several policies for the replacement of data in the disk buffers, which are different from traditional buffer replacement policies due to timing constraints. Finally, [10] refines and improves the earlier work using striping.

A related paper, [14], describes the problems and some solutions introduced by the conceptually simple fast-forward replay, for browsing purposes. Modifying the rate-of-delivery affects architectures and policies designed under different assumptions. In addition, an appropriate solution has to be media-dependent: It has to take into account encoding standard (e. g. differential coding of frames in MPEG), and also a "reasonable" appearance to users.

summary:

- distributed parallel data storage systems [45]
- optimal response time for retrieval of replicated data [12]
- optimizing allocation of objects in tertiary storage [13]
- continuous retrieval of streams using parallelism [19]
- replicas in parallel multimedia information systems [20]

- hierarchical storage structures [21]
- staggered striping [10]
- video server fast-forward functionality [14]

5 New perspectives

New applications, such as the multimedia data processing, and new hardware/software technology are major factors to drive database research to new directions. A survey of the effects of up-to-date storage hardware technology on data structure and algorithms is presented in [30]. Specifically, the following memory technologies are discussed in this paper: high-speed DRAM, flash memory, dielectrical thin film memory, dual-port RAM, and content addressable memory.

5.1 Distributed algorithms

Data-intensive computer applications are posing ever-increasing demands in terms of storage and performance capacity. One cost-effective approach to meeting such requirements is to exploit distributed storage and computing resources in a client-server architecture. The most salient feature of this approach is that its communication overhead is largely independent of the number of servers and clients in the system; it is considered a scalable approach.

More precisely, the objective of scalability means the following. Starting with a system where one server manages a file of a specific size that is accessed by a specific number of clients at a specific rate, a scalable approach can efficiently manage a file that is n times bigger and accessed by n times more clients at the same per-client rate, by adding servers and distributing the file across these servers. Furthermore, the response time of the clients' requests should be as good as in the one-server case.

Distributed dynamic hashing, [16], extends the idea of dynamic hashing by using a novel autonomous location discovery algorithm that learns the bucket locations instead of using a centralized directory. Unlike many other works in this field, [46] considers explicitly the cost/performance ratio of the system by aiming to minimize the number of servers that are acquired to

provide the required performance. The advantage of the distributed search tree method, [31], is that it can perform well also for non-uniform data request distributions, such as the linear order of keys, nearest neighbour, or range queries.
summary:

- distributed dynamic hashing [16]
- scalable distributed file organization [46]
- distributed search tree [31]

5.2 Mobile wireless computing

In the mobile wireless computing environment of the future, a large number of users, equipped with low powered palmtop machines, will query databases over the wireless communication channels. However, the physical requirements of such environment make the problem of organizing wireless broadcast data different from data organization on the disk.

One specific problem is how to improve battery utilization because small palmtop units typically run on small AA batteries. Several indexing schemes are designed in [28, 29]. These papers also demonstrate that the proposed algorithms lead indeed to significant improvement of battery life while still retaining a low access time.

Palmtop based units will often be disconnected for prolonged periods of time due to the battery power saving measures, but palmtops will also frequently relocate between different cells and connect to different data servers at different times. Caching of frequently accessed data items will be an important technique that will reduce connection on the narrow bandwidth wireless channel. However, cache invalidation strategies will be severely affected by the disconnection and mobility of the clients. In fact, the server may no longer know which clients are currently residing under its cell and which of them are currently on. A taxonomy of different cache invalidation strategies is proposed in [8] and the impact of client's disconnection times on their performance analyzed.

Other way to reduce the expensive wireless communication is to apply data replication, which has been proposed and analyzed in [27].
summary:

- caching strategies [8]
- data replication [27]
- indexing on air [28, 29]

5.3 Database mining

Database mining refers to the efficient construction and verification of models of patterns embedded in large, typically multimedia, databases, and is emerging as a major application area for database technology. The problem is that current database systems do not provide the necessary functionality for all kind of data which have already been collected. There are many cases reported in which massive amount of data have been stored on tertiary storage, but are very slowly migrating to database systems. Quest, [3, 1], is the name of a project aiming at enhancing database technology to address this problem. Specifically, Quest focuses on three classes of database mining problems involving classification, association, and sequences.

The importance of sequence query processing is motivated in [42] where also a framework for the optimization of sequence queries based on several novel techniques is presented. An indexing method for processing similarity queries on time sequences is proposed in [2]. Time sequences are mapped to a frequency domain by considering the first few coefficients of the Discrete Fourier Transformation. The transformed sequences are organized as R-tree, and in this way, efficient answer to similarity queries is obtained. The method is generalized in [18] to answer approximate-match queries for subsequences of arbitrary length. Combinatorial pattern discovery or combinatorial data mining method of [47] is able to find structural or topological patterns in data that can lead to important conclusions or prediction of new phenomena.
summary:

- tutorial [1]
- text database discovery [24]
- combinatorial pattern discovery [47]
- subsequence matching [18, 42]

- similarity search in sequence databases [2]
- Quest project [3]

5.4 Databases in networks

Various types of large, complex, and expensive real-time computer systems contain a database engine as a critical component. These systems share some of common database issues with conventional applications, but they also exhibit rather unique characteristics that present challenging database issues. According to [4], major database issues for network management include choosing the right data model, handling two different kinds of data terms of integrity and recovery constraints, supporting temporal queries, satisfying real-time performance and high availability requirements, and several miscellaneous issues. Some of these issues have been investigated in various areas of database research stage. Advances in these areas that result in actual integrated implementations for data-intensive, real-time and temporal applications are eagerly awaited.

Research on distributed databases, multimedia databases, federated databases, will enable integration of databases to a greater or lesser extent, but does not contribute towards problems such as application interoperability and distributed system management [25]. For these functions, the research on distributed objects shows greatest promise.

summary:

- telecommunication networks [4]
- distributed objects [25]

References

[1] Agrawal, R., Tutorial on Database Mining. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 75–76.

[2] Agrawal, R., Faloutsos, C., and Swami, A., Efficient Similarity Search in Sequence Databases. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 69–84.

[3] Agrawal, R., Carey, M., Faloutsos, C., Ghosh, S., Houtsma, M., Imielinski, T., Iyer, B., Mahboob, A., Miranda, H., Srikant, R., and Swami, A., Quest: A Project on Database Mining. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 514-514.

[4] Ahn, I., Database Issues in Telecommunications Network Management. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 37–43.

[5] Ahn, I., Filtered Hashing. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 85–100.

[6] Anderson, D.P. and Homsy, G., A Continuous Media I/O Server and Its Synchronization Mechanism. *IEEE Computer*, October 1991, 24(10):51–57.

[7] Anderson, D.P., Osawa, Y., and Govindan, R., A File System for Continuous Media. *ACM Transactions on Computer Systems*, November 1992, 10(4):311–337.

[8] Barbara, D. and Imielinski, T., Sleepers and Workaholics: Caching Strategies in Mobile Environments. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 1–12.

[9] Berra, P.,B., Golshani, F., Mehrotra, R., and Liu-Sheng, O.R., Guest Editors' Introduction: Multimedia Information Systems. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):545–550.

[10] Berson, S. and Ghandeharizadeh, S., Staggered Striping in Multimedia Information Systems. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 79–90.

[11] Chen, H. and Little, T.D.C., Physical Storage Organizations for Time-Dependent Multimedia Data. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 19–34.

[12] Chen, L.T. and Rotem, D., Optimal Response Time Retrieval of Replicated Data. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 36–44.

[13] Chen, L.T. and Rotem, D., Optimizing Storage of objects on Mass Storage Systems with Robotic Devices. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 273–286.

[14] Chen, M.S., Kandlur, D.D., and Yu, P.S. Support to Fully Interactive Playout in a Disk-Array-Based Video Server. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 391-398.

[15] Chiueh, T., Content-Based Image Indexing. *Proceedings of the 20th VLDB Conference*, Santiago, Chile, 1994, pp. 582–593.

[16] Devine, R., Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 101–114.

[17] Faloutsos, C. and Kamel, I., Beyond Uniformity and Independence: Analysis of R- tree Using the Concept of Fractal Dimension. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 4–13.

[18] Faloutsos, C., Ranganathan, M., and Manolopoulos, Y., Fast Subsequence Matching in Time-Series Databases. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 419-429.

[19] Ghandeharizadeh, S. and Ramos, L., Continuous Retrieval of Multimedia Data Using Parallelism. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):658–669.

[20] Ghandeharizadeh, S. and Shahabi, C., Management of Physical Replicas in Parallel Multimedia Information Systems. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 51–68.

[21] Ghandeharizadeh, S. and Shahabi, C., On Multimedia Repositories, Personal Computers, and Hierarchical Storage Systems. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 407-416.

[22] Gibbs, S., Breiteneder, C., and Tsichritzis, D., Data Modeling of Time-Based Media. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 91–101.

[23] Gonnet, G.H., Tutorial on Text Dominated Databases, Theory Practice and Experience. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 301–302.

[24] Gravano, L., Garcia-Molina, H., and Tomasic, A. The Effectiveness of GlOSS for Text Database Discovery Problem. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 126–137.

[25] Herbert, A., Databases in Distributed Systems: The New Frontier. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 2–6.

[26] Hjelsvold, R., Video Information Contents and Architecture. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 259–271.

[27] Huang, Y., Sistla, P., and Wolfson, O. Data Replication for Mobile Computers. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 13–24.

[28] Imielinski, T., Viswanathan, S., and Badrinath, B.R., Energy Efficient Indexing on Air. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 25–36.

[29] Imielinski, T., Viswanathan, S., and Badrinath, B.R., Power Efficient Filtering of Data on Air. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery (Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 245–258.

[30] Kambayashi, Y., Takakura, H., and Meki, S., Data Structure and Algorithms for New Hardware Technology. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 164–196.

[31] Kroll, B. and Widmayer, P. Distributing a Search Tree Among a Growing Number of Processors. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 265–276.

[32] Little, T.D.C. and Ghafoor, A., Interval-Based Conceptual Models for Time-Dependent Multimedia Data. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):551– 563.

[33] Maier, D., Walpole, J., and Staehli, R., Storage System Architectures for Continuous Media Data. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 1–18.

[34] Ozden, B., Biliris, A., Rastogi, R., and Silberschatz, A., A Low-Cost Storage Server for Movie on Demand Databases. *Proceedings of the 20th VLDB Conference*, Santiago, Chile, 1994, pp. 594–605.

[35] Panagopoulos, G. and Faloutsos, C., Bit-Sliced Signature Files for Very Large Text Databases on a Parallel Machine Architecture. *Lecture Notes in Computer Science*, No. 779, M. Jarke, J. Bubenko, and K. Jeffery

(Eds.). Proceedings of EDBT'94, Cambridge, United Kingdom, March 1994, Springer-Verlag, pp. 379–392.

[36] Paredaens, J., Van den Bussche, J., and Van Gucht, D., Towards a Theory of Spatial Database Queries. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 279–288.

[37] Ramaswamy, S. and Subramanian, S., Path Caching: A Technique for Optimal External Searching. *Proceedings of ACM PODS'94*, ACM Symposium on Principles of Database Systems, Minneapolis, Minnesota, USA, May 1994, pp. 25–35.

[38] Rangan, P.V. and Vin H.M., Efficient Storage Techniques for Digital Continuous Multimedia. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):564–573.

[39] Ravindran, K. and Bansal, V., Delay Compensation Protocols for Synchronization of Multimedia Data Streams. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):574–589.

[40] Riedl, J., Mashayekhi, V., Schnepf, J., Claypool, M., and Frankowski, D., SuiteSound: A System for Distributed Collaborative Multimedia. *IEEE Transactions on Knowledge and Data Engineering*, August 1993, 5(4):600–610.

[41] Salzberg, B., On Indexing Spatial and Temporal Data. *Information Systems*, 1994, 19(6):447–465.

[42] Seshadri, P., and Livny, M., and Ramakrishnan, R., Sequence Query Processing. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 430–441.

[43] Stone, H.S., Parallel Querying of Large Databases: A Case Study. *IEEE Computer*, October 1987, 20(10):11–21.

[44] Sung, S.Y., Performance Analysis of Superimposing-Coded Signature Files. *Lecture Notes in Computer Science*, No. 730, David B. Lomet (Ed.). Proceedings of FODO'93, Chicago, Illinois, USA, October 1993, Spriger-Verlag, pp. 115–129.

[45] Tierney, B., Johnston, W.E., Herzog, H., Hoo, G., Jin, G., Lee, J, Chen, L.T., and Rotem, D., Distributed Parallel Data Storage Systems: A Scalable Approach to High Speed Image Servers. *Proceedings of the ACM Multimedia 94*, Second ACM International Conference on Multimedia, October 1994, San Francisco, California, pp. 399-405.

[46] Vingralek, R., Breitbart, Y., and Weikum, G., Distributed File Organization with Scalable Cost/Performance. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 253–264.

[47] Wang, J.T., Chirn, G.W., Marr, T.G., Shapiro, B., Shasha, D., and Zhang, K., Combinatorial Pattern Discovery for Scientific Data: Some Preliminary Results. *Proceedings of the ACM SIGMOD'94*, Minneapolis, Minnesota, USA, May 1994, pp. 1–12.

[48] Yu, C., Sun, W., Bitton, D., Yang, Q., Bruno, R., and Tullis, J., Efficient Placement of Audio Data on Optical Disks for Real-Time Applications. *Communications of the ACM*, July 1989, 32(7):862–871.

http://www.mat.sbg.ac.at/~uhl/wavelet.biblio

```
%
%       BIBLIOGRAPHY OF WAVELET AND WAVE
%
%                          Steven K. Bau
%                  Dept. of Physical Oce
%                      Texas A&M Univers
%                      baum@astra.tamu.e
%                          Version 1.4
%                             6/3/94
%
%    Kindly don't remove this header if
%
%
%AAAA

@techreport{andersson-hall-etal:1993,
 Author =  "Andersson, L., N. Hall, B. J
 Title =  "Wavelets on closed subsets of
 Year =   "1993",
 Ftp =  "maxwell.math.scaroline.edu:pub/
 Size =  "667,500 bytes",
 Pages =  "60",
 Abstract =  "Orthogonal and biorthogona
   a given closed subset of the real lin
   certain types of boundary conditions
   of 'wavelet probing' is introduced wh
   numerical tasks associated with wavel
   This paper is at the wavelet archive

@article{argoul-arneodo-etal:1988a,
```

Fractal Image Encoding Announcements and Questions

This dynamic page contains various announcements and questions related to fractal image encoding. Each section contains a form with which announcements and/or questions can be entered into the document. Please enter only fractal image encoding material here.

- Announcements of on-line papers.
- Announcements of software.
- Questions about fractal image encoding.
  - General questions.
  - Algorithm (code) questions.
  - Looking for....
- Miscellaneous material.
- Return to the main Fractal Image Encoding pages.

---

top Announcements of on-line papers.

- **jim (chosen@stu.pennvalley.cc.mo.us) writes:**

  ```
  who in cyberspace has considered the
  ```

- **Anna Bergström (a94annbe@ikm.his.se) writes:**

  ```
  Real3d(Amiga) to DXF or IGES converte
  ```

- **wayne ponton (pequod@superlink.net) writes:**

- **lee ki tae (s_lkt@cecc-1.gsnu.ac.kr) writes:**

- **Eleonore (ebalbach@southwester.supertrek.k12.mn.us) writes:**

  ```
  I am a teacher in Minneapolis and I w
  and chaos theory to a small group of
  this summer. Does anybody have any su
  or curriculum that might be useful. W
  only. I would appreciate any input.
  ```

- **Christine Koenig (biwi6s9@vision.ee.ethz.ch) writes:**

  ```
  I'm working on segmentation of SAR ima
  I'm looking for good papers about it.
  Thanks
  ```

```
Christine
```

- **Peter W (info@reklamcenter.se) writes:**

```
Converter IGES to DFX format for PC
Win or Dos
```

- **Xiaomin Chen (g_xmchen@mailserv.stu.edu.cn) writes:**

```
I'm looking for material of image com
Who can help me?
```

- **Kate Bailey-Hellen (kate@hn.planet.gen.nz) writes:**

```
How do I insert a JPEG image into an
```

- **Kim (wwatts@pig.net) writes:**

```
I have a few questions.....
1) I have to do a project on weather
2) Any ideas for a fractal project?
3) Is it possible to create your own

Thanks!
```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Alex Galletero (catala@hostmail.com) writes:**

```
Hi! i'm a computer engeenering studen
thanks in advance. Please if anyone kn
```

- **Sean Ramplee (smr94@ecs.soton.ac.uk) writes:**

```
I am implementing a fractal image com
R-Trees to search the domain blocks f
However, R-Trees are doing my head in
a good and simple implementation I co
```

- **Kurt Draven (c/o kathrich@umf.maine.edu) writes:**

```
  I am working on a thesis paper conc
```

- **Edward (glfan@ee.cuhk.edu.hk) writes:**

```
I am studying on new Image Coding Alg
and wavelet will play important role
me an obvious and direct illustration
them?
```

- **Jen (petersjl@acad.etown.edu) writes:**

```
I am writing a paper on fractals in tl
Should fractals be taught as a main ma
appreciate any opinions.
```

- **Surojit Chatterjee (ppdas@cse.iitkgp.ernet.in) writes:**

```
I am working on Fractal Image Compres
(adaptively) of the image for finding
blocks.I am using vector quatization
making a codebook of the domains.The
similarity criteria for the image blo
normalized histograms for this ,any b
i need it really  soon..
```

- **Sundar Dorai-Raj (University of Alabama) (sdorai-3@ua1vm.ua.edu) writes:**

```
I am currently working on a project us
algorithm to generate a fractal. Does
between the strange attractor of my f
dimension?  Is the dimesion of the st
as the overall dimension. Hope to hea
```

- **Jackie Dawes (N13037X@tees.ac.uk) writes:**

```
I am working on a thesis involvong fra
genetic algorithms. I need to represe
as a population of individuals. How ca
```

```
Jackie Dawes
```

- **Hsueh-Ting Chu (mr834342@cs.nthu.edu.tw) writes:**

```
I wrote a fractal coder which can enc
and a 256x256 image in 0.1 sec. Is th
my coder? Please send a mail to me. T
versions.
```

- **Esquieu-Antoine (Supelec, France) (antoine.esquieu@supelec.fr) writes:**

```
  I have the purpose to make the comp
(sound of the sea, traffic jam, ...) 
 by fractals. Unfortunatly, I haven't
about that. Could someone please help
```

- **Robinson Pizzio (rpizzio@ee.pucrs.br) writes:**

```
I am new in the Fractal field.
Could anybody suggest a good book whe
good material about fractal, mainly F

Robinson Pizzio
```

- **chen jiang tao (szcubeJT@public.szptt.net.cn) writes:**

```
Mr fisher:
```

```
 I have read your book one year ago. a
by software. The result is not satisfa
ratio can be 100 or more, I doubt. I
to solve the Fractal problem , Just b
```

- **Radhakrishnan K (parol.radhakrishnan@strath.ac.uk) writes:**

```
I want to work on chaotic systems.
Is it that an image/data set rendering
means it is chaotic?
Is there any restrictions for an image
Request an early reponse
Love
radhakrishnan
```

- **Felix KRause (fkrause@ix.urz.uni-heidelberg.de) writes:**

```
I search for good book where i can fi
 Image compression or compression its
please write me!
Thank you.
```

- **ANGEL (RASTACAT@HOTMAIL.COM) writes:**

```
¡ HELLO !
I AM A MEXICAN STUDENT AND I WANT TO
ABOUT FRACTAL IMAGE COMPRESSION WITH
COULD YUO PLEASE MAIL ME SOMETHING PA
```

```
THANKS...
```

- **Al (bridger@primenet.com) writes:**

```
I'm looking for information on common
how they compare and how they relate.
My is to make a chart of Pros and Con

I'm mostly interrested in Gif, Jpg, T

If you can help, or know where to get
please Email me @  bridger@primenet.c

thank's Al
```

- **Evie Gunawan (beggtem@rad.ned.id) writes:**

```
I am an Indonesia student of informat
Do some of you have papers about IFS
system)?
Where can I find a good introduction
```

- **Armen Balian (abalian@ndu.edu.lb) writes:**

```
I have a project to deliver about Spl
using RAW image files.  The theory is
is real hard, so, if anyone be kind a
split and merge technique in C   or P
thank you in advance ...
```

- **ishan (ishan@broncho.ct.monash.edu.au) writes:**

```
i am looking for different measures i
if yes, please email me..
```

Add an announcement of an on-line paper or result. Feel free to hyperlinks in the announcement. Include both your name and email.
Name: Email:

Enter Announcement

top Announcements of software.

- **Giovambattista Pulcini (mc0878@mclink.it) writes:**

```
I am an Italian student @ "La Sapienz
I would thank you for addressing my f
paper. We found much hints and tips i
The result of our work is an encoding
images which is based on IFS compress
```

```
quadtree partitioning.
We consider this program as a "tutori
shows graphically each step of the en
block, the domain block and the isome
these two.
This sw, called IFSAF, will be soon a

accessible via V.V. WEB home page: 
http://www.webcom.com/~verrando.
```

- **fdsdfsdfs (sdfsdfsd) writes:**

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Ivan G. Zagaynov (sergey@sean.ab.msk.su) writes:**

```
Hi, I am a student of Moscow Institut
For almost two years I studied myself
internet papers - thanks to Y.Fisher!
that can compress Adobe Premiere Film
compression ratio 50-500:1 with encod
```

```
and decoding with over 20 fames per s
Movie frame size 320*256*24bit. 1 min
takes only 1 floppy disk (about 1.4M 
Send me E-mail if you are interested.

I must apologize for my poor English
```

- **q (w) writes:**

```
q
q
t
q
```

- **Hsueh-ting Chu (mr834342@cs.nthu.edu.tw) writes:**

```
I am a Taiwan student.
I received several mails asking me ab
Now I put my coder/decoder into a Win
If you are interested in it, You can 
http://ipsun1.cs.nthu.edu.tw/~mr83434
```

- **Angela Krenz (newtonsapple@ktca.org) writes:**

```
What is the Transference of Curves Th
```

Add an announcement of new software. Feel free to hyperlinks in the announcement. Include both your

name and email.
Name: Email:

Enter Announcement

Questions about fractal image encoding.

General Questions about fractal image encoding.

- **John Q. Public ( ) writes:**

  ```
  Where can I find a good introduction
  compression ?
  ```

  Answers to this question:
  - **Yuval (yfisher@ucsd.edu) writes:**

    ```
    You can find a good introduction or
    ```

```
compression pages at http://inls.uc
on which you can find the SIGGRAPH
These notes are written at a very s
an emphasis on explaining the ideas
subject. These notes are available
(977K) or  without the figures(80K)
```

- **Jean-loup Gailly (gzip@prep.ai.mit.edu) writes:**

```
The second edition of "The Data Com
by Mark Nelson & Jean-loup Gailly
( http://web2.airmail.net/markn/tdc
has a chapter on fractal image comp
source code for a simple compressio
```

- **Martin Chadderton (M.Chadderton@edm.hull.ac.uk) writes:**

```
I may have put this in the wrong se
seem to be having any luck whatsoev
aforementioned web site (web2.airma
the right address? Also, would it b
the source code mentioned above?
Thanks in advance
```

- **mariana (mdelfres@tandil.edu.ar) writes:**

```
What happend with the promise of M.
ratio of 10000:1?
```

- **Xiaomin (g_xmchen@mailserv.stu.edu.cn) writes:**

  ```
  I'm looking for material of image c
  Who can help me?
  ```

- **Xiaomin (g_xmchen@mailserv.stu.edu.cn) writes:**

  ```
  I'm looking for material of image c
  Who can help me?
  ```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

  ```
  Please send e-mail answers,
  Thank you in advance.
  ```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

  ```
  Please send e-mail answers,
  Thank you in advance.
  ```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

  ```
  Please send e-mail answers,
  Thank you in advance.
  ```

- **Michael C Taylor (mctaylor@mta.ca) writes:**

```
For the The Data Compress Book the

-Michael
```

- **Raghavendra Udupa U. (rudupa@mastech.com) writes:**

```
One can find good material on FIC i
"FRACTAL IMAGE COMPRESSION - THEORY
edited by YUVAL FISHER and publishe
```

- **Robinson Pizzio (PUCRS - Brazil) (rpizzio@ee.pucrs.br) writes:**

```
I am complete new in Fractal theory
image compression.
Could any one help? What shoul I st
Fractal looks like so difficult? Ca
I am wrong????
```

- **Cédric (cheval_e@epita.fr) writes:**

```
I'm currently doing a work on fract
and would like to know if there is
fractal compression and artificial
Is it possible to use fractalizatio
find similarities among large datab
compress them ???
```

- **Mel (Mel@charleston.net) writes:**

  ```
  What is a good definition for FRACT
  ```

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **hmahoney@sentex.net (hmahoney@sentex.net) writes:**

```
Will fractal compression provide 200::
GrayScale Image?
Regards
hmahoney@sentex.net
```

Answers to this question:

- **Yuval (yfisher@ucsd.edu) writes:**

  ```
  It depends on the image. Polar bear
  in a snow storm ? No problem. But
  the typical (what ever that is) ima
  no.
  ```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **HOLLY NASTA (ROSA PARKS) writes:**

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Zhihua Liu (liuzh@ie0.ie.ac.cn) writes:**

```
Hi,
   I am working on synthetic aperture
and want to use fractal-based coder.
   Has anyone had experience or advic
   Appreciate it very much!
regards.
Zhihua Liu
```

Answers to this question:

- **Gary Dickson (garyd@ee.bath.ac.uk) writes:**

```
My experience with SAR images is sc
to remember that they are highly te
fractals don't like texture so mayb
be more appropriate.
```

- **Gary Dickson (garyd@ee.bath.ac.uk) writes:**


```
My experience with SAR images is sc
to remember that they are highly te
fractals don't like texture so mayb
be more appropriate.
```


Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Dan Ruderman (dlr@quake.usc.edu) writes:**

```
A few questions...

1. What properites of images from the
   them amenable to fractal image comp

2. What types of images should the pro
```

```
3. What are the "typical" compression
   and color images, and their relati

4. What about "multifractal" image co


Many thanks,
Dan Ruderman
```

Answers to this question:

- **Yuval (yfisher@ucsd.edu) writes:**

```
These answers are my opinions; your
1) Images compress best when then h
    regions and straight edges.
2) The procedure fails for images t
   self-similarity on different sca
   a pixel-sized checkerboard patte
   Highly textured or detailed natu
   difficult for the same reason.
3) There is no typical compression
   is highly image dependent. Resul
   to other techniques, more or les
   don't compare to the original im
   just want a good image that is h
    fractal image compression can d
    than JPEG, for example (which i
    high compression ratios anyway)
4) I don't know what multifractal i
```

```
or what be.
```

- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

  ```
  This is regarding Q.1. The compress
  on the metric used to compare the '
  ```


- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

  ```
  This is regarding Q.1. The compress
  on the metric used to compare the '
  ```


- **Carmen Casalderrey (conocer@redestb.es) writes:**

  ```
  This is regarding Q.4: I don´t know
  http://www-syntim.inria.fr/fractale
  you can find some references.
  ```

- **G.Sriram (sriram@bronto.iitm.ernet.in) writes:**

  ```
  This is regarding Q.3.Unlike JPEG,t
  image compression algorithm.Dependi
  parameters, ratios of 20:1 can be a
  ```

```
"moderate" quality of reproduction.
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Emmanuel (reusens@lts.de.epfl.ch) writes:**

```
The general theory of IFS says that a
in a metric space (M,d) will give a co
Haussdorf space with the Hausdorf met
plane where H is the space of all sub

 Now if we want to apply this theory
we use a collection of mappings contra
images with the distance d such as the
question : what represents the Hausdo
represents the space of images?
```

Answers to this question:

- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

```
If I am right we first have to buil
```

```
usually affine mappings. The mappi
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Sean (sean@tedser.ucg.ie) writes:**

```
Hi, I'm writing a project on fractal
seen use of affine transformations us
everywhere, but no general transforma
the affine transformation is a parall
the general transformation would be a
(or a shape with four vertices)

I've did some work on general transfo
nice results, but theres this nagging
seen general transformations anywhere
transformations in the complex plane.

Any reason why or is it that I don't
```

Answers to this question:

- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

```
First point: It is not true that ge
Second point: If you use general tr
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Fabien Carre (e2fabien@etek.chalmers.se) writes:**

```
What are the main advantages & drawba
compression compared to other recent
and Wavelets ?

What are the application areas and th
of Fractals ?
```

Answers to this question:

  - **Duraid Madina (duraid@fl.net.au) writes:**

```
Fractals are better than JPEG when
compression ratios are called for,
is otherwise better simply because
computational ease. Wavelets are be
```

```
generally the best form of image co
there with compression ratios excee
of fractals (for a given distortion
artifacts look *much* "nicer" than
artifacts, and wavelets share both
resolution-independence and JPEG's
computation.
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Carlo Stazzone (fractal@hermes.dibe.unige.it) writes:**

```
I'm an italian student at "University
I'm working to develop a dedicated ar
fractal image compression algorithm i
Compression", M. F. Barnsley and L. P
I'd like to know if there are previou
a dedicated hardware implementation f
compression algorithm, and ,eventuall
material (papers, http or/and ftp sit
about them.
Thank you in advance.
```

Answers to this question:

- **Bulldog (rcaweb@mbox.vol.it) writes:**

  ```
  Contattami
  ```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Matt Duckham (mjd12@le.ac.uk) writes:**

```
I am a geographer at the University o
interested in the use of FIC as a sur
I have written my own FIC code within
System and am trying to find a way of
dimension of the IFS constructed for
seen a few methods for calculating D
the contraction factors to the power
to hold only for very simple IFS.  Do
where to look for, a way of calculati
used in FIC?
Thanks in anticipation.
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Marty Davidson (pochariot!mdavidson@photon.att.com) writes:**

```
I am interested in any published info
data compression method. Anyone know

Much Thanks.
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Adrian Bolster (A.Bolster@dcs.hull.ac.uk) writes:**

  ```
  Hi, I have this project to do which i
  i) The compression of a heightfield u
  and
  ii)The generation of heightfields usi
  Does anyone out there have any hints,
  algorithms, code ANYTHING??

  Thanks in advance.
  ```

  Answers to this question:

  Add an answer. Include both your name and email.
  Name: Email:

  Enter Answer

- **Dugal Harris (Harris@underdog.ee.wits.ac.za) writes:**

  ```
  Please could someone tell me where to
  introduction in coding fractals, I am
  the code to generate fractal music (n
  ```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Aurelio R. Ramos (aurelior@amadeus.upr.clu.edu) writes:**

```
Has fractal compression ever been con
signals at all? If so, does anybody ha
Any personal opinions about this matt
```

Answers to this question:

- **Dave (nearyd@eeng.dcu.ie) writes:**

```
The WWT (windowed wavelet transforr
that there doesn't really seem much
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Raghavendra Udupa U (rudupa@india.mastech.com) writes:**

```
Can anybody update me on the recent a
Finite Automata based Fractal Image C
Has WFA FIC technique been applied to
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Laurie/IBM AUSTRALIA (SHAWL@VNET.IBM.COM) writes:**

```
Curious: Saw TV show example of Budgi
image to detailed resolution. Serious
frame from,say, a Fred Astair & Ginge
```

```
fractal technique and present a star (
```

Answers to this question:

- **Martin Schaaf (hygore@netcom.com) writes:**

  ```
  No, it won't supply image clarity t
  ```

- **Jeff Harcourt (Not my computer) writes:**

  ```
  Not quite, Martin! Fractal compress
  being compressed has self-symmetry
  the bird's eye was created from the
  image. Hence the fractal compressor
  formula from which the object was c
  ```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Surojit Chatterjee (ppdas@cse.iitkgp.ernet.in) writes:**

  ```
  Hi I am a senior year UG student work
  ```

```
My question is how effective can Neura
quantization of image blocks?
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Dave (nearyd@eeng.dcu.ie) writes:**

```
This may sound strange, but does anyon
can put a front-end? My Prof. wants a
that is fully executable (ie. FracComp
headaches trying to compile enc.c and
allocation errors. Does such code exis

Thanks for the help.
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Zbynek Jurca (jurcaz@risc.upol.cz) writes:**

  How can I compare two images? One is
  second is image which was processed by

  Thank you in advace.

  Answers to this question:

  - **G.Sriram (sriram@bronto.iitm.ernet.in) writes:**

    Find the RMS difference between the
    be the best , but is simple and eas
    more information,read the book - Th
    by Nelson and Gailly.

  Add an answer. Include both your name and email.

  Name: Email:

  Enter Answer

Enter your general questions about fractal image compression below. Include both your name and email.

Name: Email:

Post Question

top Questions about fractal image encoding algorithms.

- **mark whittemore (mark@netheaven.com) writes:**

```
I'd like to see a bit more on these pa
the use of the source codes. What wou
would be if we could get some opinion
who has done this on the PC platform
```

```
that's where many of us are). What sha
can I use ? What sorts of results migl
of seconds vs minutes vs hours vs fil

Please someone explain to be how I ca
8 bit greyscale JPEG images to the RA
required by the fractal program. How a
description of the Raw file format at

I know I'm asking a lot of questions a
bare with me since this page is relati
of my questions appear to be addresse
```

Answers to this question:

- **yuval (yfisher@ucsd.edu) writes:**

```
The enc.c and dec.c code is very va
compiled without problem using seve
on several different platforms - a
PC land is a different world.

In MAC land, there is Photoshop tha
to raw format. I don't how to do th
on PCs. The ras2raw.c program can c
images to raw format; it is availab

Raw files are row orderd 8 bits pe
Each byte represents the grey leve
There is no header information spe
```

```
size or depth. This data has to be
from the image size (the sample ima
at this site are either 512x512=262
or 256x256=65536 bytes long).
```

- **Brendan Ryan (nazpope@freenet.hut.fi) writes:**

```
I have a simple awnser for PC's.. g
variant) or freeBSD or similar for
these are free and come with the gc
at least does..) unfortunately, no-
linux so try watcom or borland - bo
especially (multiple platform suppo
they're not shareware.. it's good,
conversion, try SEA, a popular view
multiple formats and conversion, is
i've seen..
        Rest been answered,
                           Nazlfrag
```

- **Chris Melchior (cmelchio@runet.edu) writes:**

```
From yuval's answer, the Raw format
the Windows bitmap without the head
will convert image formats to Windo
usually easy to see where the heade
code to chop it off.
```

- **Dennis Lim (c107@ee-alta.bham.ac.uk) writes:**

```
The altenative to getting linux for
the DJGPP compiler which is a GNU p
compiler. It is free. I don't know
Unix code straight off but I am to
conversions required are easy.
```

- **Anjan Kr. Dasgupta (adgcal@giascl01.vsnl.net.in) writes:**

```
One Windows based environment that
ones mentioned above is MATLAB unde
```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Andreas Kassler (kassler@informatik.uni-ulm.de) writes:**

```
One very nice solution doing fracta
on a windows PC (Win95) is the prog
which you can find on my fractal-pa
http://www-vs.informatik.uni-ulm.de
It can handle almost all kind of in
used in the PC world and generates
```

```
code. It also produces some analys
Feel free to check it out.
```

- **Martin Schaaf (hygore@netcom.com) writes:**

```
Where can I download the enc.c and
```

- **Ben Schouten (bens@cwi.nl) writes:**

- **Ben Schouten (bens@cwi.nl) writes:**

```
I myself use Fractal Imager from It
For testing we work with a program
The technical University Delft (The
Lot's parameters to change the algo
but on top it gives you beautifull
domainblocks that are selected in t
http://dutiosd.twi.tudelft.nl/~vale
```

- **Julie Phipps (juliep@corel.ca) writes:**

```
If you want a couple of great image

Take a look at these programs if yo
```

Add an answer. Include both your name and email.
Name: ____ Email: ____

Enter Answer

- **Tor Andre Myrvoll (tormy@stud.unit.no) writes:**

  ```
  One problem with the traditional way 
  is that the frequency contents of the
  will differ. This makes the cofing of
  striped suits) difficult.
  Has there been published any papers t
  wonder if theres any research regardi
  different from the usual affine ones?

  Yours Sincerely
  Tor Andre Myrvoll
  ```

  Answers to this question:

  - **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

    ```
    This about Non-linear transforms. 
    maps are non-linear in the z-direct
    3 then people have experimented it
    somewhat attempts this. Also you ca
    Modelling and rendering of nonline
    1994, pp 739-748. He gives a new wa
    ```

```
via distortions on regular grids (2
they should be applied to this part
more stuff but can't remember of ha
```

- **Saeed Asgari (asgari@cae.wisc.edu) writes:**

```
This answer is in regard to the que
Transformations for FIC...I have in
transform (WBFT) on the space of a
and is a generalization of Monro's
I showed the potential results with
for domain block candidates both on
I have found the conditions for (st
the operator under the rms metric a
parameters associated to the transf
form the wavelet-based fractal code
via any interior point method of op
such as LMI toolbox of MATLAB which
the toolbox I used to write all the
why the coder/decoder are slow at t
For a reference to the aformentione
I can email a copy of my thesis whi
the papers, two of, which are still
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Martin Chadderton (M.Chadderton@edm.hull.ac.uk) writes:**

```
Hi. I am trying to compile the enc.c a
but without much luck. I keep getting
I am using both cc and gcc on a UNIX p
compiler errors with both these progra
idea about what I could be doing wrong
```

Answers to this question:

- **Bjoern Bunte (bbunte@usa.net) writes:**

```
At the moment, I`ve got the same pr
with MS Visual C   5.0, Borland C
and gcc on a Spark an Linux. All wi
Now I`ve started to re-write the co
If there is anybody out there, who
it would be very nice if you could
as a help.
Thank you...
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **mary o'neill (oneill@mala.bc.ca) writes:**

```
I am looking for fpu init (shareware)
Thanx
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Brendan Ryan (nazpope@freenet.hut.fi) writes:**

```
Is there any good C archives of fract
around?? I've just gotten enc/dec and
more, esp. video related...
might just have to write my own..:¬)
        Regards,
                Nazlfrag Pope
```

Answers to this question:

- **chul woo park (cwpark@cesb.tongnae.ac.kr) writes:**
- **Maxim Zotov (maxim@inter-soft.com.ru) writes:**

  ```
  Look at ftp://ftp.mei.co.jp/free/Gr
  You can find there cool paper and C
  of single images and sequences of i
  ```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Micahel Baun (baun@valleynet.com) writes:**

  ```
  Since fractal compression methods uti
  ```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Altug Koker (akoker@eng.buffalo.edu) writes:**

```
I am dealing mostly the hardware impl
fractal-based image compression. Do y
about the subject?
```

Answers to this question:

- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

```
Barnsley I believe has a few patent
- on the root page. To start of you
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Centonze Angelo**

(acentonz@adriatica.peoples.it) writes:

```
I am an Italian student in  the Unive
I would like to know more detail abou
Thanks very much.

Angelo Centonze
```

Answers to this question:

- **Dave (nearyd@eeng.dcu.ie) writes:**

  ```
  There is a good paper on this in Yu
  ```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Fan Lixin (fanlixin@iscs.nus.sg) writes:**

```
I am M.Sc. student at the National Un
I would like to know more detail abou
which was mentioned in the Appendix C
compression: Theory and application".
Especially, I want to know if there i
```

```
implementation of hept-tree.
Thanks in advance.

-Lixin
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **G.Satheesh (ens_trn2@hss.hns.com) writes:**

```
With reference to the article "Fracta
 Y.Fisher which appeared in the SIGGR
I don't understand how it takes 7 bit
and 5 bits for determining s(i). I wo
 I can get an answer for it at the ea

                Thanking You,
                Satheesh        .
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **G.Sriram (sriram@bronto.iitm.ernet.in) writes:**

```
I recently downloaded enc/dec softwar
linux machine.But for different value
getting the same encoded image(20,12 
I tried)!If the option is not given,I
good quality,but a comp.ratio of arou
a bug?Why such a low ratio for good q
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Edouard Mercier (merciere@mines.u-nancy.fr) writes:**

```
I'd like to make up my mind about the
device, but I haven't found yet any p
device itself. That is the reason why
help so that a can really make an exp
Moreover, I don't want to buy any boo
I would be very grateful if someone c
information on the device or if someo
find precise informations on the net.
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

Enter your questions about fractal image encoding algorithms below. Include both your name and email.
Name: Email:

Post Question

top Looking for fractal image encoding resources.

- **Marketing Multimedia (mmminfo@mmm.inet.it) writes:**

```
We are interested in a fractal image 
We know that MCI bought the right for
we want to know, if it possible if th
developing something concerning that.

Thank You

Roberto
```

Answers to this question:

  - **jesus sosa (jesus@dec5500.sgia.imp.mx) writes:**

```
iterated function systems, the Barr
has released a plug in for netscape
viewer and you can donwload from:
  http://www.iterated.com/
```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Ivan G. Zagaynov (sergey@sean.ab.msk.su) writes:**

```
Look at announ. of software for det
me your coord. to contact - I think
what you really need
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Scott Webber (webber@norden1.com) writes:**

```
    Hi.  I am interested in getting f
software for an IBM compatable PC.  C
I can find some software that compres
```

```
any standard file format like BMP,PCX
GIF.  Also I heard that fractal techn
scale Pictures without loss of image
anything about that.

Thanks, I appreciate your help,
Scott Webber
```

Answers to this question:

- **Latapy Matthieu (latapy@amertume.ufr-info-p7.ibp.fr) writes:**

  ```
  Hello

  The best place for getting fractal
  programs is certainly Yuval's Fract
  Compression page.
  However, the public-domain program
  developped enough to allow you a co
  a PC environnement. For my part, I
  powerful drawing software that incl
  compression of usual formats files
  However, they'll certainely appear
  mounths... Just wait for us.

  About Zooming on Fractal compressed
  must understand that the methods cr
  just like (or better than) the clas
  methods do. That is, to a certain
  ```

```
zoom, the loss of quality will be h
the fractal scheme. However, you ca
scheme that would create true detai
image that does not conatin them...

Expecting my English's not too bad,
Matthieu.
```

- **Sandeep Sikka (sikka@cse.iitb.ernet.in) writes:**

```
This is about the decoding: Basical
However I remember looking at the c
```

- **Kyle Kline (kkline@quantumlink.net) writes:**

```
This is in regards to your questior
Iterated Systems (http://www.iterat
program (Fractal Imager) that will
(supported are: TGA, BMP, GIF, PCX,
to their FIF format.  It can be dov
only, 30-day trial period.  It is a
program.  Give it a shot!  It's wor
requires Windows95 or the newest re
Try it!!
```

- **TMD (DMA@AntwerpCity.be) writes:**

```
Use a netbrowser to find it!
```

- **Snickers (DMA@AntwerpCity.BE) writes:**

  May the farce be with you!

- **Dennis Aubrey (70541.3621@compuserve.com) writes:**

  Altamira Group in California has a
  product for Photoshop called Genuir
  web site is www.altamira-group.com.

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Ken Call (gbush@upr1.upr.clu.edu) writes:**

  Where could I find Machintosh

Answers to this question:

- **jon tubmen (jon@indelible.com) writes:**

  To Ken Call

```
Hi Ken,

The only decoder I know is the one
available from http://www.iterated.

cheers,                jon

PS. Please mail me direct if you co
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Armin Roehrl (m94asr@ecs.ox.ac.uk) writes:**

```
Hi,

Does anyone know if there's any code
for K. Culik II's (see Fisher's book,
using WFA available anywhere?

Thanks a lot,
Armin
```

Answers to this question:

- **Mentrup Lars (scmentru@fmi.uni-passau.de) writes:**

```
Sorry no answer to your question. (
     for K. Culik II's (see Fisher'
     using WFA available anywhere?)
```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Torsten Fink (tnfink@cip.informatik.uni-erlangen.de) writes:**

```
Hi,
is anybody working on IFS-implementati
SIMD-architectures? We are thinking ab
implementation for the MasPar MP-2.
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Andre Mok (andrem@ddc.aus.net) writes:**

```
Hi
I am trying to find commercialy produ
software. In particular medical image
compression.  All repsonses appreciat

A.
```

Answers to this question:

- **Erik Nilsen (enilsen@engr.latech.edu) writes:**

```
I think you will find www.iterated.
helpful site. Also, I would suggest
Waterloo Fractal Compression Page (
engine to find it).
                      good luck and
```

- **Dennis Aubrey (70541.3621@compuserve.com) writes:**

  ```
  Altamira Group in California has a
  called Genuine Fractals for PC.  Th
  www.altamira-group.com.
  ```

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **William (wjk@phuture.is.net) writes:**

```
At one time I saw software in a catal
compression and the ability to zoom i
fractal images, but regular raster ima
of resolution. Any idea where I can f
```

Answers to this question:

- **Bob (bob@gillies.win.net) writes:**

  ```
  Hello. Try http://www.catalog.com
  ```

Add an answer. Include both your name and

email.
Name: Email:

Enter Answer

- **Becky (101543,3445@compuserve.com) writes:**

```
I am interested in using Fractal comp
Access database. Which compression so
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **EMT AB (lasseb@emt.se) writes:**

```
I have heard the rumor that there is a
available for Borland's Delphi. Borla
so I wonder where I can find this? Pl
```

```
to me ASAP!
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Marty Davidson (md.md@worldnet.att.net) writes:**

```
Anyone know of any resources I may ge
area of using PIFS as a speech/music 
How about as a voice recognition meth
any and all help.
Thanks in advance,
Marty
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Jan M. Rasmussen (c958548@student.dtu.dk) writes:**

```
I am searching for material on triangu
encoding images. Does anyone have book
tation (or whatever!) on the subject?
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Fernando Machado (machado@alma.uc.pt) writes:**

```
I want to make a comparison between di
image compression.
Does anybody know were I can find a go
```

```
Thanks in advance,
Fernando
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **G.Satheesh (ens_trn2@hss.hns.com) writes:**

```
With reference to the article on "Fra
I would like to know how it takes jus
I would be most grateful if anyone ca
        Thanking you,
        Satheesh..
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **G.Sriram (sriram@bronto.iitm.ernet.in) writes:**

```
Suppose I have a 256-level grayscale i
64x64 pixel blocks(ranges?).To achieve
ratio of 50:1, each of these blocks is
of information!How will this suffice t
Even the example codes of Barnsley,Fis
quality of reproduction at more than 1
Please Help me with everything you kno
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Pedro Anton (panton@ctv.es) writes:**

```
Could you tell me about papers on imag
```

```
order to be compared during the searcl
compressing)
Thanks.
```

Answers to this question:

Add an answer. Include both your name and email.

Name: Email:

Enter Answer

- **Estuardo Licona-Nunez (jel15@cornell.edu) writes:**

```
I am looking for the following paper:
R.D. Boss and E.W. Jacobs. Studies of
image compression, and its applicatio
Technical Report 1468, Naval Ocean Sy
Diego CA, 1991.
I have already looked in the ftp site
Fisher's web page and it isn't there.
a way to get a copy or how to contact
Jacobs please let me know.
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Suelika J. Chial (gt2098a@prism.gatech.edu) writes:**

```
Hi.

I am an undergraduate student at Geor
on a project on fractal interpolation
find papers or related websites on fr

Thanks!
```

Answers to this question:

Add an answer. Include both your name and email.
Name: Email:

Enter Answer

- **Marcin Chady (M.Chady-ACS96@cs.bham.ac.uk) writes:**

  ```
  Is it possible to get the spec for th
  (God this page is a mess!)
  ```

  Answers to this question:

  Add an answer. Include both your name and email.
  Name: Email:

  Enter Answer

Enter your queries about fractal image encoding resources below. Include both your name and email.
Name: Email:

Post Question

top Miscellaneous material.

Miscellaneous Postings:

- **Michael C Taylor (mctylr@mailserv.mta.ca) writes:**

  ```
  I am looking for the URL for the UIUC
  The one in the Fractal FAQ is out of
  ```

- **Becky (101543,3445@compuserve.com) writes:**

  ```
  I am interested in using Fractal comp
  ```

- **Becky (101543,3445@compuserve.com) writes:**

  ```
  I am interested in using Fractal comp
  ```

- **Becky (101543,3445@compuserve.com) writes:**

```
I am interested in using Fractal comp
Access database. Which compression so
```

- **Carmen Casalderrey (conocer@redestb.es) writes:**

```
I'm a spanish student working on my F
I'd like to know if there are any ref

                               Thanks
```

- **Susumu Morita (s-m@yk.rim.or.jp) writes:**

```
Please send e-mail answers,
Thank you in advance.
```

- **Don Berliner (exec@usmailbox.com) writes:**

```
I'm trying to compress many images of
to be placed on a CD-ROM and distribu
Fractal compression and have been tol
solution in a production environment.
Question: Is this a good use for frac
2nd Question: Where might I find such
```

- **Raghavendra Udupa U. (rudupa@india.mastech.com) writes:**

```
I and Vinayaka Pandit have developed a
for Weighted Finite Automata ( WFA ) i
Karel Culik and Jarrco Kari. This work
part of our undergraduate project on W
Compression. Those interested can cont
rudupa@india.mastech.com .
       Our algorithm requires signifi
and is faster than Culik_Kari algorith
```

- **Raghavendra Udupa U. (rudupa@india.mastech.com) writes:**

```
Vinayaka D. Pandit can be contacted at
```

- **Mindaugas Satas (satas@santaka.sc-uni.ktu.lt) writes:**

```
Hi, I am a student of Lithuanian Unive
I am interesting in image compression 
about this one. I am waiting messages 
undergraduate works of image compressi
```

- **Ivan G. Zagaynov (sergey@sean.ab.msk.su) writes:**

```
Hi, I am a student of Moscow Institute
During last two years i've studied fra
via internet papers. Finally I wrote m
takes Adobe Premiere Filmstrips and co
```

```
50-500:1 compression ratio with only 
Decoding is possible with 20 frames/s
with size 320*256*24bit. If someone i
please send me E-mail, and I will giv
```

- **Yao Zhao (yzhao@center.njtu.edu.cn) writes:**

```
   I am a image compression researche
image coding. I received my Ph.D degr
I have published a lot of articles ab
coding, here are some of them:
1. Yao Zhao, Baozong Yuan, A Hybrid I
Combining Block-based Fractal Coding 
Image Communication,8(1996),pp.73-78.
2. Yao Zhao, Baozong Yuan, A New Frac
Transformation Theory for Image Codin
& Analysis, B, Vol. 35, No.3. 1995.
3. Yao Zhao, Baozong Yuan, A New Affi
Theory and Application to Image Codin
4. Yao Zhao, Baozong Yuan, Sequence I
 Using Fractal Technique, ICSPAT-95.
5. Yao Zhao, Baozong Yuan, Still Colo
Local Iterated Function Systems(LIFS)
    If you are interested in my paper
to contact me, I will mail them to yo
    By the way, after my graduation, 
my study,but I have no project to sup
have any project and need a co-worker
you will find a qualified young worke
Regards.
```

- **Dr. Itzhak Messing (imessing@dekel.technion.ac.il) writes:**

```
Dear Dr. Yao,
        I wiil apprreciate if you can
        publications.
        My address: Electrical Engine
                Technion, Haifa 32000
```

- **Ted Blackwood ((sorry don't have one)) writes:**

```
If this type of technology is used fo
how can one hide below the grid witho
area?
```

- **Ted Blackwood ((sorry don't have one)) writes:**

```
If this type of technology is used fo
how can one hide below the grid witho
area?
```

- **Pierre Greborio (gsoft@compuserve.com) writes:**

```
I'm a belgian and I'm studing at Univ
I'm working on a thesis involing mult
question:

1) there's somebody that works alread
```

```
2) where may I find information about

Thank you in advance.

Pierre
```

- **G.Sriram (sriram@bronto.iitm.ernet.in) writes:**

```
I'm a first year M.Tech student in CS
of Technology,Madras,India.I plan to
image compression.Dr.Yao,could you pl
I also welcome other related material
demystify the maths behind fractal im

Thanking you all in advance,
G.Sriram
```

- **Diana Dubel (dldubel@earthlink.net) writes:**

```
This is a request for information.
I have a Macintosh Performa 631CD.  T
  I would like to find out what fract
 available to download off the
internet.

I was a math major at San Jose State
and I am very interested in math and
```

- **ANGEL (RASTACAT@HOTMAIL.COM) writes:**

```
¡ HELLO !
```

```
I AM STUDENT OF IPN-MEXICO AND I WANT
ABOUT FRACTAL IMAGE COMPRESSION WITH (
COULD YOU PLEASE MAIL ME YOUR PAPERS /
```

- **Marcos R. Garcia (addsmgt@finred.com.mx) writes:**

  ```
  Quieres dar (give) u obtener (ask)?
  ```

- **Marcos R. Garcia (addsmgt@finred.com.mx) writes:**

  ```
  Quieres dar (to give) u pedir (to ask
  ```

Enter whatever below. Include both your name and email.

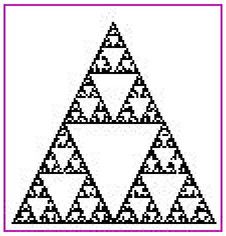
Name: Email:

Enter Stuff

Return to the main Fractal Image Encoding pages.

Yuval Fisher
Institute for Nonlinear Science

Univeristy of California, San Diego
December 15, 1995



Fractal Image Encoding

This site contains links to a variety of information and resources on Fractal Image Encoding and related topics.

This site is updated often, so check out what's new or come again...

Table of Contents

The material in the following sections is not complete - if something is not here, it may still be out there somewhere. The order of information is haphazard, things at the bottom may be "better" than things at the top.

Bibliographies and Other References

Books (**help make my mommy proud; buy my fractal**

image compression book or **the www one**)
Conferences and Announcements
Internet Resources
Papers
Software
About Students, Reprints, INLS, and Yuval Fisher.

Bibliographies and Other References

- biblio.ps.gz is a bibliography focused on fractal image compression particularly. There is also a guide with a brief description of each paper.
- A BibTeX bibliography is available from Brendt Wohlberg, University of Cape Town.
- A nice bibliography (postscript) compiled by John Kominek, from the Department of Computer Science, University of Waterloo, is available by ftp. It contains many papers, focused around fractals in imaging.
- The bibliography from "Fractal Image Compression: Theory and Application to Digital Images" is available in TeX (good), not BiBTeX

(bad); it contains more references than the book's real bibliography (good), but there are also non-fractal-image-compression references (bad).

- Here are several patents: these links are not too stable, unfortunately. More patent information can be found at the U.S. Patent and Trademark Office
    - (4,941,193) **Methods and apparatus for image compression by iterated function system** (Barnsley; Michael F. (Atlanta, GA); Sloan; Alan D. (Atlanta, GA)), Jul. 10, 1990.
    - (5,065,447) **Method and apparatus for processing digital data.** (Barnsley; Michael F. (Atlanta, GA); Sloan; Alan D. (Atlanta, GA)), Nov. 12, 1991.
    - (5,384,867) **Fractal transform compression board** (Barnsley; Michael F. (Atlanta, GA); Sloan; Alan D. (Atlanta, GA); Elton; John H. (Atlanta, GA); Moreman; Charles S. (Lawrenceville, GA); Primiano; Guy A. (Cumming, GA)), Jan. 24, 1995.
    - (5,416,856)**Method of encoding a digital image using iterated image transformations to form an eventually contractive map** (Jacobs; Everett W. (San Diego, CA); Boss; Roger D. (San Diego, CA); Fisher; Yuval (La Jolla, CA), May 16, 1995.
    - (5,430,812) **Fractal transform compression board** (Barnsley; Michael F. (Atlanta, GA);

Sloan; Alan D. (Atlanta, GA); Elton; John H. (Atlanta, GA); Moreman; Charles S. (Lawrenceville, GA); Primiano; Guy A. (Cumming, GA)), Jul. 4, 1995.

top Books

- **Fractal Image Compression: Theory and Application to Digital Images**, Yuval Fisher (Ed.), Springer Verlag, New York, 1995 is a collection of articles on Fractal Image Encoding. About this book:
  - An elementary introduction containing almost no mathematics.
  - Rigorous description of all the relevant mathematics of the subject.
  - Recent theoretical results on fast encoding and decoding methods, various schemes for encoding images using fractal concepts, and theoretical models for the encoding/decoding process.
  - C Code:
    - The code includes an optimized encoding program (enc.c). This program can encode a 256x256 image in a few seconds.

- A non-optimized decoding program (dec.c). The decoding, while slow, is resolution independent.
- A list of possible compilation/runtime problems on various platforms or whatever.
- A manual for the code.
- Some usage notes.
- Please read the copy notice .
- Some raw byte image files are also available.
- Or just download the whole code package.

Ordering information.

- **Fractal Image Encoding and Analysis: A NATO ASI Series Book**, Yuval Fisher (Ed.), Springer Verlag, New York, 1996 contains the proceedings of the Fractal Image Encoding and Analysis Advanced Study Institute held in Trondheim, Norway July 8-17, 1995. This book contains articles by leading researchers in the fields of fractal image encoding and analysis. The book is currently being procuced.

top Conferences and Announcements

Many of these are over, but their web pages still contain useful (or possibly interesting) information.

- Fractals in Engineering will be held in Arcachon, FRANCE, June 25-27, 1997. The goal of the conference is to bring together researchers working in all area of fractal analysis. The scope encompasses recent theoretical advances as well as industrial applications.
- Fractals in the Natural and Applied Sciences will be held in Denver, Colorado, USA, 8-11 April 1997. The conference is intended to provide a forum for the dissemination of the latest research findings in the broad field of fractals. Interdisciplinary submissions are strongly encouraged.
- International Conference on "Future of Fractals" 25th-27th July 1995, Aichi Prefecture Laborer Center, Seto, Aichi, Japan.
- **Fractal Image Encoding and Analysis**, an NATO Advanced Study Institute, was held in Trondheim, Norway July 8-17, 1995. It was a lot of fun, some of which you can see that the conference's. web site.

top Internet Resources

- A dynamic fractal image encoding announcement and question page Leave your questions and announcements there or answer other peoples' questions.
- The Waterloo Fractal Compression Project is part of a general research programme dedicated to the study of fractal analysis and Iterated Function Systems/Fractal Transforms from both theoretical as well as practical perspectives.
- The CWI Fractal Image Compression Page from the Centrum voor Wiskunde en Informatica in Amsterdam, Holland.
- The Groupe Fractales Database, maintained by Paulo Goncalves at INRIA, has links to publications, software, conferences, etc. mostly related to fractal analysis, but with plans to discuss compression also.
- A Fractal Image Compression page is maintained by Brendt Wohlberg.
- Usenet Newsgroups (These are read from your local server, so these links may fail if you have no local news server or are not set up properly):
  - sci.fractals
  - comp.compression
  - bit.listserv.frac-l
  - sci.math

- Iterated Systems, Inc. has a nice WWW site.
- See a A fractal-based photoshop plugin based on Iterated Systems' FIF format.
- There is a nice site related to fractals (not particulary fractal image encoding) at spanky.triumf.ca.
- The Bath group has fractal site related to the Bath Fractal Transform.

top Papers

- For a good introduction to fractal image compression, view the SIGGRAPH '92 Course notes on fractal image compression (977K). The SIGGRAPH '92 Course notes **without the figures** are also available (80K). This paper is also available in French, translated by M. Latapy.
- A postscript version of a paper on A comparison of Fractals with JPEG and Wavelets . This paper (or a version of it) appeared in the 94 San Diego SPIE conference proceedings. This paper is a subset of the data contained in the "comparison of results" section in *Fractal Image Compression*, discussed above.
- Fractal Video Compression material, including a paper, encoded sequences and a decoder.

- View a postscript preprint by F. Dudbridge describing a new technique for fast fractal image coding. If you have any questions please send email to frankd@inls1.ucsd.edu.
- Here is an important paper about the relationship between fractal image encoding, wavelet encoding, mixing the two, and related stuff. It is available from Geoff Davis' home page.
- Get a paper about fractal Image compression, it's connection to the Haar Discrete Wavelet Transform and a fractal coder in the wavelet domain:

  H. Krupnik, D. Malah and E. Karnin, "Fractal Representation of Images via the Discrete Wavelet Transform" IEEE 18th Conv. of EE in Israel, Tel-Aviv, March 1995. There are several other papers at that site also.

- Papers describing parallel implementation of quadtree code are available from Dr. Jeff Jackson's (the author) home page.
- The Bath group has a list of on-line papers on the Bath Fractal Transform.
- Combining Tree and Feature Classification in Fractal Encoding of Images by Behnam Bani-Eqbal, discusses a tree structured classification that speeds encoding times by up to

75%. Code is available at the cite as well.

- An ftp site at ftp.informatik.uni-freiburg.de is maintained by Dietmar Saupe and Raouf Hamzaoui. Here is what is there: (last updated July 12, 1996)

  Guide.ps.gz

  A guided tour of the fractal image compression literature, D. Saupe, R. Hamzaoui, Technical Report 58, Institut f"ur Informatik, July 94. The figures for this paper are in separate files: Saup94a.fig1.ps.gz Saup94a.fig2.ps.gz A first version is contained in the SIGGRAPH'94 course notes.

  AlCl91.ps.gz

  Fractal image compression, M. Ali, T. G. Clarkson, Proc. 1st Seminar on Information Technology and its Applications (ITA'91), Markfield Conf. Centre, Leicester, U.K., 29 Sept., 1991.

  AlCl92.ps.gz

  Survey of Block Based Fractal Image Compression and Its Applications, M. Ali, T. G. Clarkson, Proc. 2nd Seminar on Information Technology and its Applications (ITA'91), Markfield Conf. Centre, Leicester, U.K., Dec. 1993.

  AlCl94.ps.gz

  Using linear fractal interpolation functions to compress video images, M. Ali, T. G. Clarkson,

Fractals 2,3 (1994) 417-421.
AlGeCl93.ps.gz
Analysis, generation and compression of pavement distress images using fractals, M. Ali, M. A. Gennert, T. G. Clarkson, in The Applications of Fractals and Chaos, eds. A.J. Crilly, R.A. Earnshaw and H. Jones, Springer-Verlag, Berlin, 1993, p.147-169. Figures are in AlGeCl93.fig2.ps.gz and AlGeCl93.fig12.ps.gz.
AlPaCl92.ps.gz
The use of fractal theory in a video compression system, M. Ali, C. Papadopoulos, T. G. Clarkson, in Proc. IEEE Data Compression Conference (DCC'92), 24-27 March, 1992.
BSVN94.ps.gz
A new image coding technique unifying fractal and transform coding K. U. Barthel, J. Sch{\"u}ttemeyer, Th. Voy{\'e}, P. Noll, IEEE Int. Conf. on Image Processing (ICIP'94), Austin, Texas.
BaMaKa93.ps.gz
Hierarchical interpretation of fractal image coding and its applications to fast decoding, Z. Baharav, D. Malah, E. Karnin, Intl. Conf. on Digital Signal Processing, Cyprus, 1993.
Bani94.ps.gz
Speeding up fractal image compression, B. Bani-Eqbal, Proceedings from IS\&T/SPIE 1995

Symposium on Electronic Imaging: Science \& Technology Vol.\ 2418: Still-Image Compression, 1995.

Bart95.ps.gz

Entropy constrained fractal image coding, K. U. Barthel, NATO ASI on Fractal Image Coding, Trondheim, Norway, July 1995

BaVo94.ps.gz

Adaptive fractal image coding in the frequency domain, K. U. Barthel, T. Voye, Proceedings of International Workshop on Image Processing, Budapest, June 1994.

BaVo95.ps.gz

Three-Dimensional fractal video coding, K. U. Barthel, T. Voye, IEEE Int. Conf. on Image Processing (ICIP'95), Washington, D.C., USA

BaVoNo93.ps.gz

Improved fractal image coding, K. U. Barthel, T. Voye, P. Noll, Proceedings from Picture Coding Symposium, March 1993.

BoMe92.ps.gz

Kohonen neural network for image coding based on iteration transformation theory, A. Bogdan, H. E., Meadows, Proceedings from SPIE Neural and Stochastic Methods in Image and Signal Processing, Vol.\ 1766, pp.\ 425--436, 1992.

Bogd94a.ps.gz

Multiscale (inter/intra-frame) fractal video

coding, A. Bogdan, Proc. ICIP-94 IEEE International Conference on Image Processing, Austin, Texas, Nov.\ 1994.

CaMo94.ps.gz

Generalized self-similarity, wavelets and image analysis C. Cabrelli, U. Molter, Preprint #78, Dept. of Math., University of Buenos Aires, 1994.

ChDaBe93.ps.gz

Compression fractale par partitionnement de Delaunay, J.-M. Chassery, F. Davoine, E. Bertin, 14th Conference GRETSI, Juan-les-Pins, Sept. 1993.

ChZh95.ps.gz

Multiresolution approximation of fractal transform B. Cheng, X. Zhu, to appear in Signal Processing.

Cisc96.ps.gz

On entropy coding Fisher's fractal quadtree code, G. Ciscar, June 1996.

DaBeCh93.ps.gz

From rigidity to adaptive tessellation for fractal image compression: Comparative studies, F. Davoine, E. Bertin, J.-M. Chassery, IEEE 8th Workshop on Image and Multidimensional Signal Processing, Cannes, Sept. 1993.

DaCh94.ps.gz

Adaptive Delaunay triangulation for attractor image coding, F. Davoine, J.-M. Chassery, 12th

International Conference on Pattern Recognition, Jerusalem, Oct. 1994.

DaSvCh95.ps.gz

A mixed triangular and quadrilateral partition for fractal image coding, F. Davoine, J. Svensson, J.-M Chassery, IEEE Int. Conf. on Image Processing (ICIP'95).

FGHS94.ps.gz

Image compression based on fractal theory, C. Frigaard, J. Gade, T. Hemmingsen, T. Sand, Institute for Electronic Systems, Aalborg University, Denmark, 1994.

FoVr94a.ps.gz

Solving the inverse problem for function/image approximations using iterated function systems, I. Theoretical basis, B. Forte, E. R. Vrscay, Fractals 2,3 (1994) 325--334.

FoVr94b.ps.gz

Solving the inverse problem for function/image approximations using iterated function systems, II. Algorithm and computations, B. Forte, E. R. Vrscay, Fractals 2,3 (1994) 335--346.

FoVr95a.ps.gz

Solving the inverse problem for function and image approximation using iterated function systems, B. Forte, E. R. Vrscay, to appear in Dynamics of Continuous, Discrete and Impulsive Systems 1,2 (1995). Figures are in

FoVr95a.figs.ps.gz.
FoVr95b.ps.gz
Theory of generalized fractal transforms, B. Forte, E. R. Vrscay, to appear in the Proceedings of the NATO ASI on Fractal Image Encoding and Analysis, July 1995, Trondheim.
GoIbGr95.ps.gz
Fractal image coding and magnification using invariant features, D, G\"{o}tting, A., Ibenthal, R.-R., Grigat, NATO ASI Conf. Fractal Image Encoding and Analysis, Trondheim, July 1995, to appear in a special issue of Fractals.
Hamz95.ps.gz
Codebook clustering by self-organizing maps for fractal image compression, R. Hamzaoui, NATO ASI Conf. Fractal Image Encoding and Analysis, Trondheim, July 1995, to appear in a special issue of Fractals.
Hamz96a.ps.gz
A new decoding algorithm for fractal image compression, R. Hamzaoui, to appear in Electronics Letters.
HaMuSa96a.ps.gz
VQ-enhanced fractal image compression, R. Hamzaoui, M. M\"uller, D. Saupe, ICIP-96 IEEE International Conference on Image Processing, Lausanne, Sept. 1996.
HaMuSa96b.ps.gz
Enhancing fractal image compression with

vector quantization, R. Hamzaoui, M. M\"uller, D. Saupe, 1996 IEEE Digital Signal Processing Workshop, Loen, Sept. 1996.

Huer93.ps.gz

Contractivity of fractal transforms for image coding, B. Huertgen, Electronics Letters, 29 (1993) 1749-1750.

Huer95a.ps.gz

Performance bounds for fractal coding, B. Huertgen, Proceedings of ICASSP-1995 IEEE International Conference on Acoustics, Speech and Signal Processing, Vol.\ 4, Detroit, 1995.

Huer95b.ps.gz

Statistical evaluation of fractal coding schemes, B. Huertgen, ICIP'95, Vol. 3 pp. 280-283, Washington, DC, USA, 1995.

HuBu93.ps.gz

Fractal approach to low-rate video coding, B. Huertgen, P. Buettgen, Proceedings from SPIE Visual Communications and Image Processing, Vol.\ 2094, pp.\ 120--131, 1993.

HuHa94.ps.gz

On the convergence of fractal transforms, B. Huertgen, T. Hain, Proceedings of ICASSP-1994 IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 5, pp. 561--564, Adelaide, 1994.

HuMu94.ps.gz

Modelling of fractal coding schemes, B.

Huertgen, F. Mueller, Proceedings of the VIIth European Signal Processing Conference EUSIPCO'94 Vol. 1 Edinburgh, Scotland pp. 600-603, 1994.
HM94.ps.gz
Selbstahnlichkeit als neuartiges Prinzip zur Quellencodierung, B. Huertgen, F. Mueller, Proceedings ITG-Fachtagung 130, Codierung fur Quelle, Kanal und Ubertragung, pp. 277-284, Muenchen, 1994.
HuMuSt93.ps.gz
Adaptive fractal coding of still pictures, B. Huertgen, F. Mueller, C. Stiller, Proceedings of the International Picture Coding Symposium PCS'93 Lausanne, Switzerland, 1993.
HuMoSi94.ps.gz
Fractal transform coding of color images, B. Huertgen, P. Mols, S. F. Simon, Proceedings of the International Conference on Visual Communications and Image Processing, SPIE '94, Vol. 2308, pp. 1683-1691, Chicago, Illinois, USA, 1994.
HuSi94.ps.gz
On the problem of convergence in fractal coding schemes, B. Huertgen, S. F. Simon, ICIP'94, Austin, Texas, 1994.
HuSt93.ps.gz
Fast hierarchical codebook search for fractal coding of still images, B. Huertgen and C. Stiller,

Proceedings EOS/SPIE Visual Communications and PACS for Medical Applications '93 Vol. 1977 Berlin, Germany pp. 397-408, 1993.

KiKiLe95.ps.gz

Novel fractal image compression method with non-iterative decoder, C.-S., Kim, R.-C., Kim, S.-U, Lee, Proc. ICIP-95 IEEE International Conference on Image Processing, Washington, D.C., 1995.

KiLe95.ps.gz

Fractal coding of video sequence by circular prediction mapping, C.-S., Kim, S.-U, Lee, NATO ASI Conf. Fractal Image Encoding and Analysis, Trondheim, July 1995, to appear in a special issue of Fractals.

Komi95a.ps.gz

Algorithm for fast fractal image compression, J. Kominek, Proceedings of SPIE, Volume 2419, 1995.

Komi95b.ps.gz

Convergence of fractal encoded images, J. Kominek, Proceedings DCC'95 Data Compression Conference, J. A. Storer, M. Cohn (eds.), IEEE Computer Society Press, March 1995.

KrMaKa96.ps.gz

Fractal representation of images via the discrete wavelet transform, H. Krupnik, D. Malah, E. Karnin, IEEE 18th Conf. of Electrical

Engineering, Tel-Aviv, March 1995.
Nova93a.ps.tar.gz
Attractor coding of images, M. Novak, Licentiate Dissertation, Linkoping University, May 1993.
Nova93b.ps.gz
Attractor coding of images, M. Novak, Picture Coding Symposium, Lausanne 93.
OBLMK94.ps.gz
A new improved collage theorem with applications to multiresolution fractal image coding, G. E. Oien, Z. Baharav, S. Lepsoy, D. Malah, E. Karnin, Proc. ICASSP, 1994.
Reus93.ps.gz
Sequence coding based on the fractal theory of iterated transformations systems, E. Reusens, Proc. of VCIP, Boston 1993.
Reus94a.ps.gz
Overlapped adaptive partitioning for image coding based on the theory of iterated function systems E. Reusens, Proc. of ICASSP 1994, Adelaide.
Reus94b.ps.gz
Partitioning complexity issue for iterated functions systems based image coding, E. Reusens, Proc. of VII EUSIPCO, Edinburgh, September 1994.
RiCa94.ps.gz
An image coding scheme using block

prediction of the pyramid subband decomposition, R. Rinaldo, G. Calvagno, Proc. ICIP 1994, Austin, Texas.

RiCa95.ps.gz

Image Coding by Block Prediction of Multiresolution Subimages, R. Rinaldo, G. Calvagno, IEEE Trans. on Image Proc., July 1995.

Saup94a.ps.gz

Breaking the time complexity of fractal image compression, D. Saupe, Technical Report 53, Institut f"ur Informatik, 1994 The figures for this paper are in separate files: Saup94a.fig1.ps.gz Saup94a.fig2.ps.gz Saup94a.fig3.ps.gz Note: The paper [Saup94c] entitled "From classification to multidimensional keys" cited as [Saup94c] in "A guided tour ..." is just an excerpt of this paper [Saup94a]. So you do not need a copy of it.

Saup95a.ps.gz

Accelerating fractal image compression by multi-dimensional nearest neighbor search, D. Saupe, Proceedings DCC'95 Data Compression Conference, J. A. Storer, M. Cohn (eds.), IEEE Computer Society Press, March 1995.

Saup95b.ps.gz

Fractal image compression via nearest neighbor search, D. Saupe, in: Conf. Proc. NATO ASI Fractal Image Encoding and Analysis, Trondheim, July 1995, Y. Fisher (ed.), to appear

in Springer-Verlag, New York, 1995.
Saup96a.ps.gz
Lean domain pools for fractal image compression, D. Saupe, in: Proceedings from IS\&T/SPIE 1996 Symposium on Electronic Imaging: Science \& Technology -- Still Image Compression II,} Vol.\ 2669, Jan.\ 1996.
Saup96c.ps.gz
The futility of square isometries in fractal image compression, D. Saupe, IEEE Int. Conf. on Image Processing (ICIP'96), Lausanne, Sept. 1996.
SaHa95a.ps.gz
Complexity reduction methods for fractal image compression, D. Saupe, R. Hamzaoui, in: IMA Conf. Proc. on Image Processing; Mathematical Methods and Applications, Sept. 94, J. M. Blackledge (ed.), to appear in 1996.
SaHar96a.ps.gz
Lossless acceleration of fractal image compression by fast convolution, D. Saupe, H. Hartenstein, ICIP-96 IEEE International Conference on Image Processing, Lausanne, Sept. 1996.
SaRu96a.ps.gz
Evolutionary fractal image compression, D. Saupe, M. Ruhl, ICIP-96 IEEE International Conference on Image Processing, Lausanne, Sept. 1996.

Sign95.ps.gz

Geometrical interpretation of IFS based image coding, J. Signes, NATO ASI Conf. Fractal Image Encoding and Analysis, Trondheim, July 1995, to appear in a special issue of Fractals.

Simo95.ps.gz

Explicit link between local fractal transform and multiresolution transform, B. Simon, IEEE Int. Conf. on Image Processing (ICIP'95).

VeRu94.ps.gz

Fractal compression of single images and image sequences using genetic algorithms, L. Vences, I. Rudomin.

Vine93.ps.gz

Signal Modeling with Iterated Function Systems, G. Vines, PhD Thesis, Georgia Institute of Technology, 1993.

Wall95.ps.gz

Merging fractal image compression and wavelet transform methods, A. van der Walle, NATO ASI Conf. Fractal Image Encoding and Analysis, Trondheim, July 1995, to appear in a special issue of Fractals.

WoJa95.ps.gz

Fast image domain fractal compression by DCT domain block matching, B. E. Wohlberg, G. de Jager, Electronic Letters 31 (1995) 869--870.

ZhPo95.ps.gz

Fractal color image compression using vector

distortion measure, Y. Zhang, L. M., Po, Proc. ICIP-95 IEEE International Conference on Image Processing, Washington, D.C., 1995.
biblio.ps.gz
Updated bibliography, June 1996.

top Software

- Fractal Video Compression material, including a paper, encoded sequences and a decoder.
- Peruse older executable code and sample encoded images. This includes PC executables, but unfortunately the encoded images are not compatible with the slightly newer version available below. This material is also available by ftp from kurosawa.unice.fr:/pub/compression/fractals.
- Newer executable code and sample encoded images. Executables for SUN, SGI, and NeXT are included. This same material is also available by ftp from kurosawa.unice.fr:/pub/compression/fractals-new.
- Many interesting things related to chaos and fractals are available from Matt Kennel's ftp archive now at inls.ucsd.edu.
- FracCompress, written by Jude Sylvestre

(sylvestre@hws3.hws.edu), is a Win-32 application based on Visual C++ class named CFracComp. This class is based on Yuval Fisher's encoder and decoder C code. The class was written as part of Honors Thesis in computer science at Hobart and Williams Colleges. It has been tested on Windows NT 3.5, Windows 95 Pre-release and Windows 3.1 with win32s extensions. This class can be consider as an Alpha version, much of the code is being improved and modified in order to provide support for multiple graphic format, and increasing speed. The class library is able to save images in RAW format(Row-col major ordering). Image can also be saved t-the clipboard; input file must also be in RAW image format. The class library is provided AS-IS and as Freeware. A very useful program to uses with this application is Paint Pro v3.0. It is able to read and write RAW image file. It can be founded at ftp://oak.oakland.edu/pub/simtelnet/win3/graphics/

- Giovambattista Pulcini and Valerio Verrando from "La Sapienza", Rome, Italy, have written a program to encode 24bpp color images. The program is partially based on Yuval Fisher's enc.c/dec.c, but uses a simpler classification scheme. Their site has more information about their Window's 95 code.
- Several students at the Institute for Electronic

Systems at Aalborg University, Denmark have written a quadtree based code called Limbo. You can read about it or just go there. This code is newly (Jan 27, 1995) improved.

- Scott Hollatz, when he was at the Department of Mathematics and Statistics, University of Minnesota-Duluth, put together a package, including an undergraduate level technical report and working code. The report is called "Digital image compression with two-dimensional affine fractal interpolation functions," which sums up the approach (to be brief: encoding is done on scanlines, with an iterative ``chaos game'' decoding). The code outputs ASCII so there is work to do to get compression out of it. You can get it (753090 bytes) or ftp to hp.uwsuper.edu. A hard copy can be received by mail by contacting the mathematics department.
- Andreas Kassler (kassler@informatik.uni-ulm.de) wrote a windows based program based on Yuval Fisher's enc/dec code for his Diplomarbeit. The program features:
  - full true-color compression in rgb- or yuv-modell
  - new fileformat so there's no need for commandline parameters
  - iterative decoding or pyramide-decoder
  - postprocessing with Fisher's heuristic schema or with Ramstad/Lepsoy scheme.

    - there are a few imageprocessing tools implemented like filtering, edge-detection ...
    - program can read and write fileformats like Windows-BMP, TIFF, PCX, GIF, EPS, JPEG, WMF
- A DOS binary for the Bath Fractal Transform is now available.
- Iterated Systems has released a fractal image (for their own format and with a 6K license agreement) decoder for Windows and Macs. Another page contains some image sources.
- Code demonstrating the Combining of Tree and Feature Classification in Fractal Encoding of Images by Behnam Bani-Eqbal, demonstrates a tree structured classification that speeds encoding times by up to 75%. A companion paper is available at the cite as well.

---

top About Reprints, Students, INLS, and me.

I get a a steady trickle of email based on the misconception that I can accept students or influence the acceptance of postdocs at the University of California, San Diego, or possibly at the institute for nonlinear science (INLS). This is not so. I have a

visiting position, and this means that I cannot influence the applications of students, accept students myself, or invite visitors. It also means that I don't have the resources to photocopy and mail reprints (not to mention the inclination). (If you have bags of money to hand out, consider this a shameless plea).

I am at INLS due to the gracious hospitality of Henry Abarbanel, the director of the institute. INLS is home to physicists, mostly, who work on spacio-temporal chaos.

To pass the time, I recently wrote a book:

Yuval Fisher (yfisher@ucsd.edu)
Institute for Nonlinear Science
University for California, San Diego
July 20, 1995.

This site is updated often, so check out what's new or come again...

If you would like to leave your papers here, ftp them to inls.ucsd.edu:pub/uploads/ and email me with the file name and a blurb that you want to

with the link to the paper.

If you have questions about fractal image compression/encoding or if you would like to announce code and/or results, feel free to do so at the Fractal Image Encoding Announcements and Questions page. If you have comments about these pages or personal non-fractal questions, please feel free to mail them to me.

Usage Statistics

Brian C. Smith is assistant professor of Computer Science at Cornell University. His research interests are multimedia applications and development tools, graphical user interfaces, computer graphics, and image processing. His previous research projects include the design and implementation of the CMT multimedia toolkit, a distributed programming extension to Tcl/Tk called Tcl-DP, and the Berkeley MPEG player.

A Survey of Compressed Domain Processing Techniques

Brian C. Smith

Table of Contents

1. Introduction
2. JPEG
3. MPEG Video
4. MPEG Audio
5. Conclusions
6. References

Introduction

Processing video data is problematic due to the high data rates involved. Television quality video requires approximately 100 GBytes for each hour, or about 27 MBytes for each second. Such data sizes and rates severely stress storage systems and networks and make even the most trivial real-time processing impossible without special purpose hardware. Consequently, most video data is stored in a compressed format.

While compression reduces the storage and network costs, it increases the cost of processing since the data must be decompressed before it can be processed. The overhead of decompression is enormous: today's sophisticated compression algorithms, such as JPEG or MPEG, require between 150 and 300 instructions per pixel for decompression. This corresponds to a rate of 2.7 billion instructions for each second of NTSC quality video processed. Furthermore, the data must often be compressed after processing, which adds one to fifty times more overhead.

One way to circumvent these problems is to process the video data in compressed form. This technique

reduces the amount of data that must be processed and avoids complex compression and decompression. Decreasing data volume has the side effect of increasing data locality and thus more efficiently uses processor cache, which further improves performance.

Over the past 4 years, many researchers have applied this idea to MPEG compressed audio, MPEG video, and motion JPEG video. This paper surveys these techniques.

<-- Table of Contents

JPEG

JPEG is a standard for still image compression [10] that can achieve compression ratios as high as 50 to 1. When JPEG is applied to each frame in a video sequence, the result is called "motion JPEG" video compression. The baseline JPEG algorithm is outlined in figure 1. Briefly, JPEG converts an RGB image into the YUV color space, subsampling the chrominance (U and V) components. Each component is divided into 8x8 pixel blocks, which are transformed using a discrete cosine transform (DCT). The post-DCT 8x8 block is encoded as a one

dimensional, 64 element vector using zig-zag scanning, a heuristic that places the low frequency coefficients, which are located in the upper left corner of the block, early in the vector. The vector is scaled by dividing each coefficient by a constant. The constant varies from element to element, and is proportional to the human eye's sensitivity to an error in each coefficient. The scaled vector has the following property: an error of up to one half can be introduced in each element without a noticeable difference in the reconstructed image. Consequently, the elements of the scaled vector are rounded to the nearest integer, a process called quantization. Experiments reveal that typically 55 to 60 of the 64 quantized elements are zero, so they are encoded using run-length encoding (RLE), and the result is entropy coded using either Huffman or arithmetic coding.

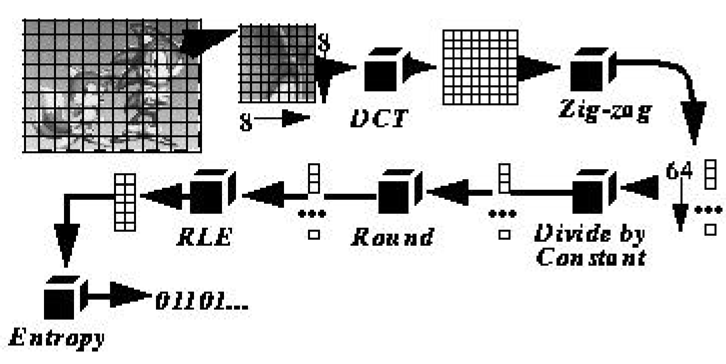


Figure 1: JPEG Image Compression

In summary, JPEG is a seven step process:

1. The color space of the image is converted (RGB to YUV)
2. The discrete cosine transform (DCT) is applied to each 8x8 block of pixels
3. The result is transform into a 64 element vector using zig-zag coding
4. The vector is scaled.
5. The scaled vector is quantized (rounded to the nearest integer).
6. The quantized vector is run-length encoded (RLE)
7. The RLE vector is entropy coded.

Decompression is the reverse of compression. Two important facts about JPEG compression will be needed for the discussion in the next section. First, steps 1-4 are linear operators. Second, the RLE vector is a sparse-matrix representation of the scaled, quantized DCT vector.

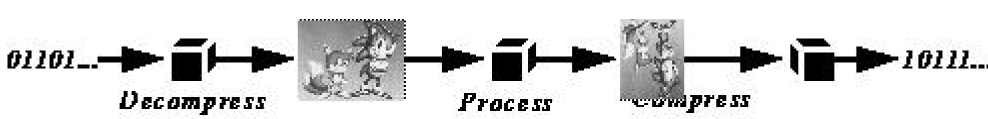


Figure 2: Spatial Domain Processing

We now turn to the problem of processing JPEG data. A naive approach would decompress the data,

process it, and compress the result, as shown in figure 2. This strategy is called *spatial domain* processing. The basic idea of compressed domain processing is to convert the spatial domain process into its frequency domain equivalent. This conversion is accomplished by representing the spatial domain process and the transforms used in compression a linear operators, and form the composite operator. Compressed data is then processed by entropy decoding the bitstream to recover the sparse vector data, applying the compressed domain operator, and quantizing and compressing the result. Figure 3 shows this strategy.

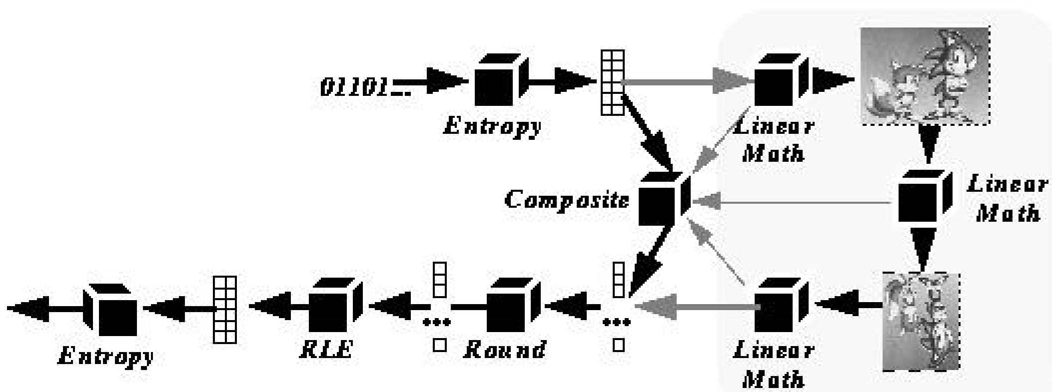


Figure 3: Compressed Domain Processing

Perhaps the simplest type of video processing that has been mapped into the compressed domain (for JPEG) is pixel arithmetic [13]. This type of processing

consists of four operations: scalar addition (add a constant to each pixel in an image), scalar multiplication (multiply each pixel in an image by a constant), pixel-wise addition (add the pixels in two images), and pixel-wise multiplication (multiply the pixels in two images). These operations can be implemented directly on the run-length coded, quantized vectors (the RLE vectors), which can be obtained by Huffman decoding the JPEG bitstream. It can be shown that scalar addition is equivalent to altering the DC component of each block (the first element in the RLE vector), scalar multiplication is equivalent to multiplying each element of the RLE vector by a constant, and pixel-wise addition is equivalent to adding two RLE vectors. Pixel-wise multiplication is more complex: it involves the multiplication all possible pairs in the two RLE vectors. Since the vectors are relatively sparse, this operation is not too expensive in the average case.

Pixel arithmetic can be used to implement a number of operations. For example, a cross-fade between two video sequences is described by the function

$$g(x, y) = \alpha f_2(x, y) + (1 - \alpha) f_1(x, y) \qquad \text{(EQ 1)}$$

where *g* is an image in the output sequence, *f1* and *f2* are images in the source sequences, and α is a parameter that varies from 0 (beginning of the

cross-fade) to 1 (end of the cross-fade). The implementation of this operation on compressed domain data is straightforward, since it involves only scalar multiplication and pixel-wise addition, and the compressed domain operator is about 100 times faster than its spatial domain equivalent [13].

Another application of pixel arithmetic is image composition. A familiar example of this operation is overlaying a forecaster on a weather map. The overlay is accomplished using three images: a background image (the weather map), a foreground image (the forecaster), and a masking image, which is a bi-level image that has a pixel value of 1 if the corresponding pixel in the output should be taken from the foreground image, 0 if it should be taken from the background. The output image is computed using the following function:

$$g(x, y) = m(x, y) f(x, y) + (1 - m(x, y)) b(x, y) \qquad \text{(EQ 2)}$$

where *g* is the output image, *m* is the mask, *f* is the foreground, and *b* is the background. Since the only operations that are used in this processing are pixel-wise multiplication, scalar addition, and pixel-wise addition, the operation can be mapped into the compressed domain. For typically applications, the compressed domain operator is about 10 to 50 times faster than the corresponding spatial domain

implementation [13].

Two problems arise with image composition in the compressed domain. First, there is no known simple way to generate the mask image (called the *matte*) in the compressed domain. In television studios, these images are generated in the spatial domain using chroma-keying: the forecaster is placed in front of a bright green or blue screen, and that color is used to generate the mask image. No one has determined how to generate a matte directly on compressed data.

A second problem is that the edges of the foreground object and the edges in the matte are often highly correlated, since they both lie on the foreground object's boundary. The presence of edges increases the size of the RLE vectors, which makes the masking operation can be relatively expensive for the foreground image, since the cost of implementing a pixel-wise multiply on two RLE vectors is proportional to the product of the size of the two vectors. Often, the compressed domain operation is slower for these blocks than the spatial domain equivalent. But since there are relatively few edge blocks in typical mattes (maybe 10%), operating in the compressed domain is many times faster than in the spatial domain.

The four operations described above, pixel and scalar addition and multiplication, are *local* operations: the

value of a pixel in the output is uniquely determined by the value of the corresponding pixel(s) in the source image(s). A more complicated type of processing uses *global* operations, where the value of a pixel in the output image is a function of many pixels in the source image. Linear global operations, where each output pixel is a linear combination of input pixels, can be implemented in the JPEG domain [14].

The basic idea of compressed domain global processing is to factor JPEG into steps that can be written mathematically as linear operators (scaling, DCT, zigzag scanning) and steps that can not (quantization, run-length and Huffman coding). Once JPEG compression, JPEG decompression, and the spatial domain operation (i.e., the processing) are written as linear operators, they can be combined. Two complications muddy this otherwise simple idea. First, the resulting operators can be more expensive to implement than the spatial domain equivalent, both in memory and operation count. This defect can be corrected using a scheme called *condensation* to approximate the compressed domain operator by a sparse operator. The image that results from using the sparse operator is nearly indistinguishable from the correct result, but its computation is much faster: about twice the time required for Huffman compression and decompression [14].

The second complication is that mapping global operations into the compressed domain is complicated because the spatial domain operators may cross block boundaries. For example, a common way to implement the operation of smoothing an image is to replace each pixel by an average of its immediate neighbors. Since an 8x8 pixel block in the output image is a function of a 10x10 block of pixels in the source image, the coefficients of the corresponding output RLE vector will be a complex linear combination of coefficients from nine RLE vectors in the input. Computing the appropriate linear combination is tractable if the spatial domain operators, and the operators corresponding to the compression and decompression, are written as high dimensional matrices. This technique gives rise to fairly complicated index gymnastics, but the calculation is straightforward.

Global operations in the compressed domain are several times faster than their spatial domain equivalents, but this efficiency comes at the cost of increased complexity of implementation. Moreover, this method of mapping global operations into the compressed domain results in impractically large data structures (tens of megabytes) unless the operation exhibits special symmetry.

These limitations have prompted other researchers to examine special cases that can be implemented efficiently. Two examples are the inner block operations discussed by Shen and Sethi [12], and the scaling operations derived by Natarajan and Bhaskaran [9].

Natarajan and Bhaskaran showed that the operation of shrinking an image by a factor of two could be implemented in the compressed domain using only shifts and adds. His work maps the scaling operator directly into the compressed domain and then factors the result into a sequence of operations with many common sub-expressions that consist of multiplies and divides by powers of two. The speed of the resulting operations is approximately 5:1, and the image quality is very high.

Sethi and Shen examined the special case of inner block transforms (IBTs), which are a form of factoring global operations. In IBTs, the coefficients in an output block is a function a single block in the source, but the combination of pixels in a block is a global operation. When IBTs are combined with pixel addition, IBTs can be used to implement a wide variety of operations, including image rotation.

Manipulating pixel values in the compressed domain provides the basis for implementing special effects,

but does little to address image and video content. For example, it would be advantageous to search an image/video database using only compressed domain operations. Several researchers have pursued this idea. Two notable contributions are Seales work on object recognition [11] and Arman's work on cut detection [2].

Seales examined searching a database of "mug shots" in the compressed domain using an *eigenspace* method. The basic idea of this work is to represent a large-dimensional object (an image) in a lower-dimensional subspace. The basis vectors for the subspace are images, called "eigen-faces", capture the most important features of the images in the database. Once the eigen-faces are computed, a linear combination of the eigen-faces is computed for each image in the database that closely approximates the image. The coefficients of this linear combination are used as an index in the database. To locate an image in the database, the eigen-face coefficients of the target image (the one you are searching for) are computed and compared against the database index. The matches are sorted and returned to the user.

Seales's method computes the eigen-faces directly on the RLE vectors, rather than pixel values. This allows the target image's index to be computed

without decompressing the image, a significant savings. Seales showed that image searching using eigen-faces in the compressed domain was just as effective as in the image domain for moderate quality images, proving the validity of the technique and showing that the information lost during compression was insignificant.

Arman, Hsu, and Chiu examined the problem of cut detection in the compressed domain. The problem of cut detection is this: given a video, segment it into a sequence of scenes or "shots." The shot boundaries are called cuts, and are the results of editing the raw video footage. Arman showed how to perform cut detection directly on JPEG compressed video data. His approach compares a subset of coefficients selected from RLE blocks which are extracted from frames of a video sequence. The position of the extracted RLE blocks, as well as the subset of coefficients chosen, remains constant throughout the sequence. The set of coefficients thus obtained is used to form a normalized vector that represents the frame. Two frames are compared by computing the dot product of their normalized representative vectors. When the dot product is sufficiently small, a scene cut is declared. The procedure is found to be fast, simple, and effective.

<-- Table of Contents

MPEG Video

Although JPEG is very effective at compressing still images, further compression can be achieved using the MPEG video standard [8]. MPEG augments JPEG with motion compensation, a form of differential encoding, and achieves compression ratios as high as 200 to 1. In MPEG, each frame in a video sequence is encoded as either an I-frame, a P-frame, or a B-frames. A typical frame sequence is shown in figure 4, along with the inter-frame dependencies. I-frames are independent of other frames, and their encoding is almost identical to JPEG.

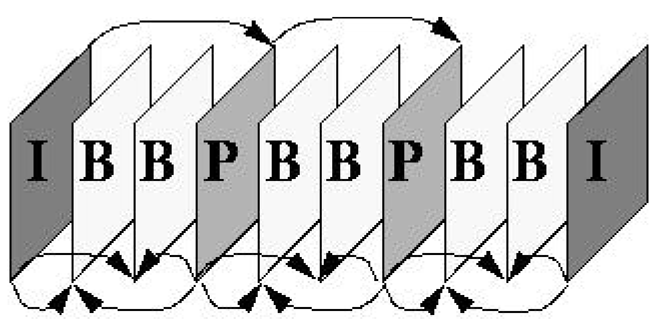


Figure 4: MPEG Dependencies

B-frames and P-frames are encoded as a difference from nearby I- or P-frames (called *reference* frames).

P-frame encoding is shown in figure 5. MPEG partitions a P-frame into 16x16 pixel *macroblocks*. Each macroblock in the P-frame, called the *target* macroblock, is encoded by searching for a macroblock in the reference frame (the most recent I- or P-frame), typically using the mean squared error or mean absolute error as a metric. The difference in coordinates between the target macroblock and its best match is called the *motion vector*, and their pixel-wise difference is called the *error term*. The error term is compressed using JPEG and stored along with the motion vector in the MPEG bitstream. B-frame compression is nearly identical to P-frame compression, except that two reference frames (one from the earlier in the sequence and one from later) can be used. This extra degree of freedom can result in very high compression ratios, as high as 500 to 1.

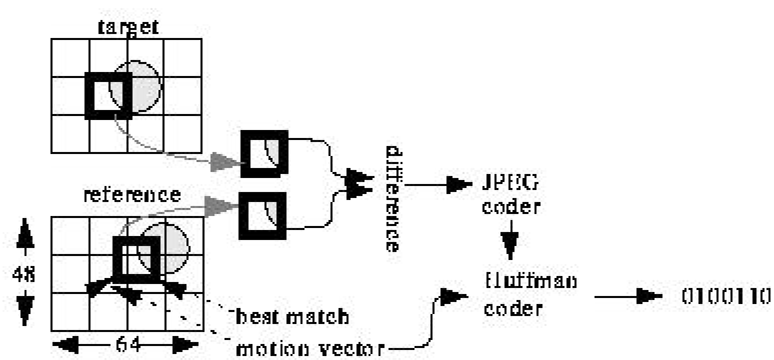


Figure 5: P-frame encoding

To decode an MPEG frame, all relevant reference frames must first be decoded. Then each error term is decompression, and the result is added pixel-wise to the 16x16 pixel block in the reference frame specified by the motion vector.

The processing in the previous section used JPEG compressed image and video data. A natural question is "can these techniques be extended to MPEG video data?" Since MPEG uses a more complex compression strategy than JPEG, the extension is non-trivial. Recall that MPEG defines three types of frames, I-, P-, and B-frames. I-frames are very similar to JPEG frames, while P- and B-frames use motion compensation. This motion compensation poses the most difficult problem for operating on MPEG data, since the motion vector may refer to a block of pixels that is not aligned on a macroblock boundary.

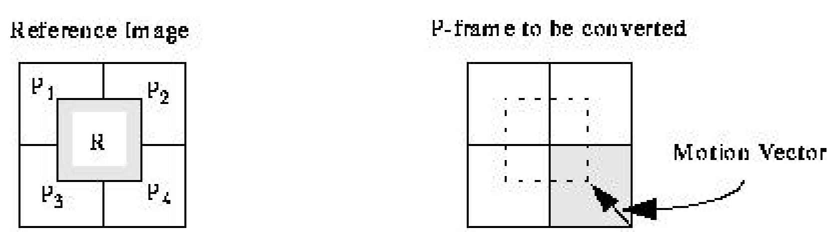


Figure 6: Reference Macroblock (R), motion vector, and original macroblocks

---

Chang's solution to this problem is to convert P- and B- frames into I-frames, then process the resulting frames using strategies similar to those of JPEG frames [4, 5]. Converting a P-or B-frame into an I-frame requires computing DCT coefficients of a motion compensated macroblock from the motion vector and the coefficients in the reference frame. Chang showed how to compute these coefficients by pre-multiplying and post-multiplying the reference macroblocks with appropriate matrices. Suppose in figure 6, *R* is the macroblock in the reference frame that we wish to derive from the four original macroblocks *P1... P4*. In the spatial domain, this conversion is expressed by the following equation:

$$R = \sum_{1}^{4} S_{i1} P_i S_{i2} \qquad \text{(EQ 3)}$$

where *Sij* are matrices like

$$\begin{bmatrix} 0 & 0 \\ I_n & 0 \end{bmatrix} \quad or \quad \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix}$$

Each *In* is an identity matrix of size *n*.

Applying the DCT to each side of this equation, we have

$$DCT(R) = \sum_{1}^{4} DCT(S_{i1})\,DCT(P_i)\,DCT(S_{i2})$$

where DCT(R) denotes the DCT on matrix R. Once the DCT coefficients of R have been calculated, the result can be scaled and added to the error term in the compressed domain using pixel addition.

Once the problem of P- or B- to I-frame conversion has been solved, it becomes possible to transcode MPEG data to JPEG data in the compressed domain. The advantage of this technique would be that existing JPEG hardware could be used to display MPEG data. Smith and Arachya have studied this problem recently [1], with promising results: MPEG to JPEG conversion can be performed at rates four to five times faster in the compressed domain than in the spatial domain. Their method caches converted motion-compensated macroblocks for additional performance.

Yeo and Liu have studied the problem of cut detection on MPEG data. Their approach analyzes the difference image between successive frames in the spatial domain using *DC-images*, which are obtained by shrinking an image by a factor of 8 in each dimension. The DC-image is obtained from I-frames by taking the first coefficient in an RLE block. They use a clever approximation of Chang's

technique to compute the DC-images, for P- and B-frames. The cost to compute each pixel in the DC-image using their technique is at most 4 multiplications.

<-- Table of Contents

MPEG Audio

The MPEG audio encoder achieves compression ratios between 6 and 12 to 1 [6]. MPEG audio converts a group samples into 32 equal width frequency bands using the subband analysis. For example, CD quality audio digitized at 44.1 kHz would be divided into 32 bands, each 689 Hz wide. The result of the analysis is 32 amplitude coefficients, one for each band, that indicates the strength of that band. Each coefficient is then quantized using masking.

Masking is a psychoacoustic effect where larger amplitude signals obscures lower amplitude signals. Studies have shown that masking decreases with distance in the frequency domain. This fact means that, in MPEG audio, an amplitude coefficient with a large value will tend to render a neighboring coefficient with a small value inaudible. MPEG uses

this fact to represent the smaller, neighboring coefficient with fewer bits than its large neighbor. The exact number of bits assigned to each coefficient is determined by a psychoacoustic model. The model also computes a scaling factor for each band, which is essentially an exponent which is used to expand the quantized sample to the necessary range. The quantized sample, bit allocations, and scaling factors comprise the resulting bitstream. The MPEG audio compression process is illustrated in figure 7.

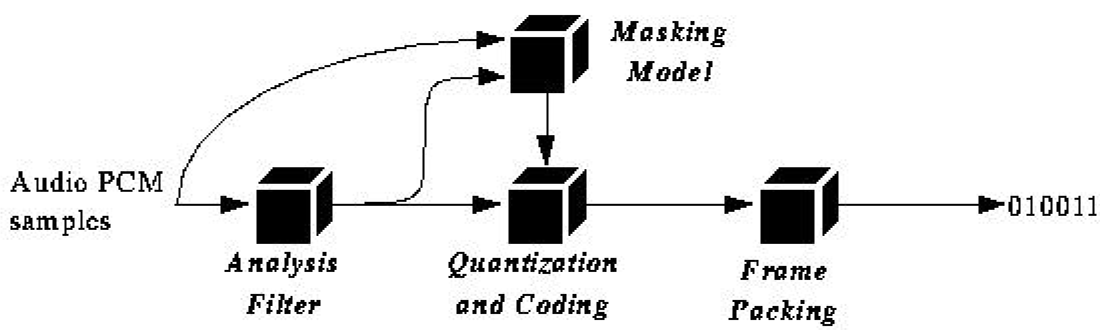


Figure 7: MPEG Audio Encoding

Few researchers have explored the possibility of directly manipulating MPEG compressed audio data. The area has attracted little attention because many researchers believe that less significant gains are possible because audio data rates are lower; CD-quality audio corresponds to a data rate of 192 KBytes/sec. But the sophisticated compression and

decompression techniques used in MPEG audio compression make real-time processing of compressed audio data impossible on current generation machines.

To date, the only work known by the author is that of Broadhead and Owen, who studied the problem of gain control, mixing, and equalization on MPEG audio streams [3]. A detailed discussion is beyond the scope of this survey, but they report that mixing in the compressed domain is five times faster than the equivalent uncompressed domain operation, and that the resulting bitstreams are very high quality.

<-- Table of Contents

Conclusions

This paper has surveyed techniques for processing audio, image, and video data in the compressed domain. The advantage of compressed domain operations is that they can often be performed many times faster than their spatial domain counterparts, and that they process less data, which increases locality and lowers bandwidth requirements.

Several open problems remain. There are no known

techniques for performing non-linear operations in the compressed domain. For example, chroma-keying in the compressed domain is impossible, as is any sort of color map modification such as gamma correction. Another problem is that the techniques that process MPEG produce JPEG data (or MPEG with all I-frames) as a result. The techniques would be significantly more attractive if motion-compensated MPEG data, with B- and P-frames, could be calculated directly in the compressed domain.

The motion compensation data present in MPEG video could be used in other ways. For example, many vision algorithms use motion information in their processing. Since this information is effectively pre-computed in MPEG data, such algorithms might work particularly well in the compressed domain. Finally, direct manipulation of compressed audio data deserves more attention.

<-- Table of Contents

References

[1] Soam Acharya, personal communication

[2] Farshid Arman, Arding Hsu, Ming-Yee Chiu,

Image Processing on Compressed Data for Large Video Databases, Proceedings of the First ACM International Conference on Multimedia, August 1993, Anaheim, Calif.

[3] M. Alexander Broadhead, Charles B. Owen, *Direct Manipulation of MPEG Compressed Digital Audio*, to appear in Proc. of the Third ACM International Conference on Multimedia, San Francisco, CA, November 5-9, 1995

[4] Shih-Fu Chang, Wen-Lung Chen and David G. Messerschmitt, *Video Compositing in the DCT Domain*, IEEE Workshop on Visual Signal Processing and Communications, Raleigh, NC, pp. 138-143, Sep. 1992.

[5] Shih-Fu Chang and David G. Messerschmitt, *A New Approach to Decoding and Compositing Motion Compensated DCT-Based Images*, IEEE Intern. Conf. on Accoustics, Speech, and Signal Processing, Minneapolis, Minnesota, pp. V421-V424, April, 1993.

[6] A.J.M. Houtsma, *Psychophysics and Modern Digital Audio Technology*, Philips Journal Research, Vol 47, 1992, pp 3-14.

[7] K. Fukunaga, *Introduction to Statistical Recognition*, Academic Press, 1990.

[8] D. Le Gall, *MPEG:A Video Compression Standard for Multimedia Applications*, Communications of the ACM, April 1991, vol.34, (no.4):46-58

[9] B. Natarajan, V. Bhaskaran, *A Fast Approximate Algorithm for Scaling Down Digital Images in the DCT Domain*, to appear in IEEE International Conference on Image Processing, Oct. 23-26, 1995, Washington D.C.

[10] William B Pennebaker, *JPEG still image data compression standard*, Van Nostrand Reinhold, New York, 1992.

[11] W. Brent Seales, Matthew D. Cutts, *Vision and Multimedia: Object Recognition in the Compressed Domain*, University of Kentucky Technical Report, Oct 1995, seales@dcs.uky.edu

[12] Bo Shen and Ishwar K. Sethi, *Inner-block operations on compressed images*, to appear in Proc. of the Third ACM International Conference on Multimedia, San Francisco, CA, November 5-9, 1995

[13] Brian C. Smith, Lawrence A. Rowe, *Algorithms for Manipulating Compressed Images*, IEEE Computer Graphics and Applications, Sept. 1993, vol.13, (no.5):34-42.

[14] Brian C. Smith, *Fast Software Processing of Motion JPEG Video*, Proc. of the Second ACM International Conference on Multimedia, San Francisco, CA, October 15-20, 1994.

<-- Table of Contents

MAGAZINE

Image Description on the Internet

A Summary of the CNI/OCLC Image Metadata Workshop
September 24 - 25, 1996
Dublin, Ohio

Stuart Weibel and Eric Miller
Office of Research
OCLC Online Computer Library Center, Inc.
Dublin, Ohio

D-Lib Magazine, January 1997

ISSN 1082-9873

September 24-25, 1996. Seventy practitioners in the area of networked image description attended a two day workshop sponsored by the Coalition for Networked Information (CNI) and the OCLC Online Computer Library Center in Dublin, Ohio. This third in

the series of metadata workshops addressed the application of the Dublin Core element set to image resource description (see the Dublin Core Homepage for more detailed information about this workshop and others in the series).

The two day workshop reached consensus, supporting the notion that the Dublin Core, within the context of the Warwick Framework, affords a foundation for the development of a simple resource description model to support network-based discovery of images. As Charles Rhyne, Chair of Art History at Reed College observed:

> "I was not especially surprised that we concluded that the elements needed to discover text and images on the internet are similar. The text and the images themselves are radically different and require different types of expertise to study and interpret them, but most of the primary categories under which we classify and search for them are similar."

Given the original objective of the Dublin Core element set -- to define a simple, easily understood semantic core for network resource discovery -- satisfying core description requirements for both textual and visual information with a single element set is attractive indeed. The enthusiasm for settling

on a single set was modulated with a strong recommendation to make the labels for existing elements more amenable to the dual purpose of text and image description.

Is an image a document-like-object?

The abstraction of a document-like-object emerged in the first workshop as a way of sidestepping differences in individual notions of what constitutes a discrete object worthy of separate description. One of the first issues addressed in the Image Metadata Workshop was whether an image is a document-like-object or is it so different that an alternative framework for description is required?

Consensus emerged around the idea that images are not so different from the document-like-objects of the first workshop. The expectation that a set of image-specific elements (an Image-Core) would emerge from the workshop gave way to the idea that the application of a slightly modified Dublin Core element set might serve as well. As Jennifer Trant, of the Arts and Humanities Data Service in the UK, wrote after the workshop:

> "That images are 'document-like' was to me one of the more significant contributions of the meeting. We went into the discussion assuming

> that there would be an 'image core', expressed as a separate box within . . . the initial model for our discussions.
>
> "We emerged from our two days of discussion with only one, slightly extended, set of core elements to support the discovery process, a set which seems to me to reflect the various conceptual categories researchers bring to their search for information. These categories did not change based on the media of the information resource (visual or otherwise) that might satisfy the query.
>
> "After spending so long thinking that images were 'special' [to use museum-like assumptions] it was fascinating for me to have a group of image specialists say that in most content terms fixed/static/bounded images really are a lot like text-based document-like objects."

The single, slightly extended, set of core elements for image discovery emerged from two days of discussion as a set which seems to reflect the various conceptual categories researchers bring to their search for information. These categories were judged by the image specialists in attendance not to differ significantly based on the media (visual or otherwise) of the information resources that might satisfy the

query.

The defining characteristic of a document-like-object is not its textual versus graphical content, but rather whether or not the resource is bounded, or fixed, in the sense that the resource looks the same to all users. Thus, images, movies, musical performances, speeches and other information objects which are characterized by being fixed (i.e., having identical content for each user) can also be thought of as document-like-objects.

Non-document-like objects, on the other hand, include such resources as virtual experiences, databases (including ones that generate document-like outputs), business graphics, CAD/CAM or geographic information generated from database values, and interactive applications which might have different content for each user. In the context of image discovery, these sources do not "contain" images as much as they "generate" images. The images they generate may be described as fixed document-like objects, but the metadata required to describe them (the systems doing the generating) are distinct.

Consider the example of the Visible Human Project (described in a workshop plenary talk by Earl Henderson of the National Institutes of Health). More

than a collection of fixed images, the Visible Human Project at the National Library of Medicine is a collection of applications unified by a data set that is nothing if not visual in character. The scope of the project itself is dynamic and evolving rapidly, and the character of the visual outputs of any of the many applications growing up around this data set defy simple description and certainly are not bounded in the sense understood in this workshop. Such applications are systems, rather than collections of images.

A Model for Metadata

Much of the consensus-building surrounding the Dublin Core has involved accommodating pragmatic stakeholder concerns borne of long standing experience with legacy description models. It is helpful to have a conceptual model to guide this pragmatism, and just such a model developed in the course of the workshop. This model is an outgrowth of previous work of Bearman (see A Reference Model for Business Acceptable Communications) and provides conceptual support for both the Dublin Core and Warwick Framework by illustrating the transactional relationship of metadata and the research process.

The research process can be thought of as a series

of interactive processes, which can provisionally be described as including:

- Discovery: the identification of relevant resources
- Retrieval: the transfer of resources to a local site
- Collation: the aggregation and organization of selected resources
- Analysis: the intellectual and/or computational analysis of resources
- Re-presentation: the formulation of derivative intellectual artifacts based on the resources and previous processes in the sequence

These processes involve events and resources distributed among institutions, machines, networks, and the minds of individuals. Metadata, then, become any one set of elements drawn from the many kinds of information necessary for decision-making within this matrix of minds, machines, and networks.

For example, access to *discovery* metadata may lead to the return of terms and conditions elements, necessary for *retrieval*. *Retrieval* metadata might include the network address of a resolver from which the resource may be accessed or the publisher of an item with whom a usage agreement must be transacted. *Collation* metadata might include data about an image collection schema or the provenance of an item. *Analysis* might require a color map for the

item. *Re-presentation* could involve information validating credit to rights holders, and might well require a link to update use history of the source object.

A variety of metadata will be needed to satisfy the requirements of each stage, and hence the functional requirements of metadata packages might well be defined by these requirements. To be used effectively, elements of metadata must be readily available as required by each stage in the research process in which the user is engaged (though different implementations might deliver some metadata at stages prior to its being needed). It is recognized that the pragmatics of collection and management of metadata will likely compromise this ideal, but the model can nonetheless inform our thinking and design.

One need not imagine all possible linkages to recognize the complexity of such a model, nor is it necessary to accommodate at the outset all possible elements, packages, and necessary infrastructure. But in the search for appropriate compromises, it is helpful to see the larger picture that this model attempts to capture. What *is* necessary, though, is an agreement as to the notion of assembling sets of descriptive elements, which enables extensibility and forward compatibility.

Like the Warwick Framework, this model explicitly recognizes that metadata will be created and managed by a variety of agents, for different reasons, at different times in the life of the object. This implies an infrastructure and architecture that does not now exist, but that will evolve, driven by the marketplace of information access. The modest achievement of this workshop is to reaffirm the semantic characteristics of but a single variety of metadata package -- the core elements of a resource discovery element set--and to assert its suitability for both textual and visual resources.

How are Images Different?

It is gratifying that the workshop reached agreement that text and images could be classified using similar categories, but just as clearly, images offer a number of technological and descriptive challenges peculiar to themselves.

Textual materials can be indexed, often simplifying or partially automating the task of description, whereas most of the descriptive elements of images are extrinsic to the work (or are not easily extracted from the work).

Encoding schemes are critical for using images. This

can be true for textual materials as well, but there are fewer varieties of textual representation, and at least for some of them, there is some graceful failure (HTML or SGML, for example, are hard, but possible to read without a suitable rendering program).

Rendering of images is unforgiving and the variant forms are combinatorially overwhelming. Commonly encountered web graphics display by default and presumably tolerate wide differences in display characteristics. As more sophisticated imaging applications populate the Internet, metadata will play a more important role in discovery and selection. Information necessary for rendering may include:

- type (bit-mapped, vector, video)
- format (TIFF, GIF, JFIF, PICT, PCD, Photoshop, EPS, CGM, TGA . . .)
- compression schemes and ratios (JPEG, LZW, QuickTime. . .)
- dimensions
- dynamic range
- color lookup tables and related metrics (CMYK, RGB. . .)

Characteristics of original image capture, while less critical for the casual user, may be of overwhelming importance to the archival or research significance of the image or collection. This sort of information is

also, for the most part, irrecoverable if not recorded at the time of capture.

Categories of information about the scanning process include: light source (full spectrum or infrared, for example) resolution, dynamic range, type of scanner, date of scan, journal/audit trails, and digital signatures for authentication.

Variant forms of the image content are also important (the Versioning Problem writ large): source image, different views of the same object, different scans of the same object, different resolutions of the same image, details of the same image, source ID, responsible institution. All these categories may be critical elements of metadata for a particular image or collection.

The complexity of adequately capturing and encoding such information conflicts with one of the original design goals of the Dublin Core: simplicity. If the Dublin Core is to be applied in some useful way to the domain of images, it will be necessary to isolate the essential core of information appropriate to a simple description record and to identify a graceful extension mechanism that supports encoding of the richer array of descriptors hinted at in the preceding paragraphs.

Modifications to the Dublin Core

The workshop consensus and subsequent animated discussions on the META2 list (the primary forum for discussion of Dublin Core issues) resulted in a number of changes to the Dublin Core element set (see Table 1). Several element names were modified slightly to make them less text-centric, and two elements were added to the original thirteen. The reference description of the elements resides at http://purl.org/metadata/dublin_core_elements.

Subject and Description Separated

SUBJECT and DESCRIPTION are now separate elements in the core, partly because of the judgment among the image specialists that these are quite distinct concepts for images. Other participants in the metadata discussions on META2 agreed that such a distinction is also useful for other media.

Thus, SUBJECT is intended to include keywords, controlled vocabulary terms, and formal classification designators, while DESCRIPTION is to be used for descriptive prose or content description (in the case of images) and affords a natural place for abstracts in the case of textual documents.

A Rights-Management Field

A simple rights-management field is perceived by many as a necessary component of a core description record. While arguably not an intrinsic dimension of discovery, it is of such importance to the use of images that failure to include such an element may hinder wide deployment. This is a good example of the imprecise lines of demarcation between different varieties of metadata that will inevitably blur the idealized functional boundaries one might hope for among metadata packages. Resource description is a messy business -- ask any cataloger.

The digital world requires a sophisticated language for expression and negotiation of intellectual property rights; the evolution of the supporting infrastructure is well underway (some of these have been reported in this journal). This element should not be construed as a substitute for such a language or metadata structure, but rather as a means for communicating simple terms and conditions where they exist or providing a link to more complex information as it evolves.

One proposed application of the field is as follows:

- null - there may or may not be restrictions on use, and users have to figure it out independently, outside the context of this particular collection of

metadata.
- the string "No Restrictions on Reuse" - there are no restrictions on re-use.
- URI or other pointer - there are restrictions on use, and users can follow the link to find out more information .

This approach addresses several implementation issues. The metadata could be used to retrieve materials with no restrictions on use at a top-level search, without getting into any subsidiary packages of metadata. Second-level packages of rights-management metadata could be retrieved automatically or presented to the user as links within the search results. All records in a single collection could share a single value in the rights management field. Additionally, managers that don't fill in the rights-management field, or that have rights issues but have no on-line access to that information, enjoy the presumption that a null response means there might be restrictions.

Open Issues

Surrogates and Objects

Among the most important impediments to coherent deployment of a metadata element set is the confusion between description of the object versus

description of the digital surrogate of that object. This can be a problem with text, but in general, it is the intellectual content rather than its presentation that is of primary importance with text, and increasingly the primary version of a text is its electronic form.

With images, the variety of forms an image may assume in its life cycle is liable to be greater than for a piece of text, and the relationships among these versions are intrinsically more complex. The degree to which such information is captured, and the means of encoding it are difficult problems the solutions to which must evolve in tandem with the pragmatics of implementation.

Collection Versus Item-level Description

Collection descriptions and the schemas that account for the aggregation of images in such collections are essential for effective collection discovery. Early discussions embraced the possibility of a separate element for addressing this, though ultimately consensus emerged around the idea of capturing this sort of information in existing fields. The simplest possibility is to include a Resource Type or RELATION flag (COLLECTION | ITEM). Further explorations are necessary to determine whether this is sufficient or whether there might be other sensible values for this sub-element. This is part of the larger

relation problem, which requires elaboration for visual and textual materials alike.

SOURCE is Dangerously Recursive

SOURCE information is potentially recursive (and probably complex) for any object, but especially with images. How can such object-surrogate-derivative relationships be expressed to both aid in discovery and to explicate intellectual property lineage?

Mapping of Dublin Core to Other Element Sets

One of the first tangible outcomes of the first metadata meeting was the mapping of Dublin Core elements to MARC fields by Rebecca Guenther of the Library of Congress. This discussion paper contributed substantively to the community awareness of the emergence of Dublin Core as a model for network resource description, and fed back into the change process for MARC. Similar mapping between Dublin Core elements and existing image description standards will clarify the role of the various elements and provide guidance for the application of the Dublin Core to image collections. As has been suggested on the workshop discussion list, existing standards or practices can serve as templates for the development of guidelines, thereby jump-starting interoperability and reducing the effort

necessary to develop description standards.

Viewing Requirements.

The bandwidth and time penalty for retrieving images is often high, making it desirable to have some indication of usability prior to retrieval. It was agreed that an existing element (the FORMAT element, previously called FORM) could probably be used to express this information, but a standard of best practice needs to evolve through real-world implementation.

The most problematic aspect of this issue is where to stop. An archival site using the Dublin Core to describe items at a deep level might want to include a large set of image-related descriptors but this would hardly be expected to be the norm for broad deployment. A flexible means for including such information in the FORMAT element should be proposed in user guidelines, with an eye towards the evolution of a Warwick Framework-style package to support such needs in the future.

Future Developments

The Dublin Core is a high-level reference model. In and of itself, it does not provide guidance for cataloging or searching, nor is it a blueprint for

system development. Rather, it provides guidance for the semantic content of a simple resource description model that may profitably be applied to visual as well as textual resources. The consensus developed around this model is the major product of the workshops. This consensus is the result of many people of standing who have used their own experiences and the collective intelligence of the communities they represent to arrive at a common foundation for networked resource description.

These workshops are but tentative steps in bringing this collective intelligence to bear on the difficult problem of resource discovery on the Internet. The achieved consensus is important but incomplete. It requires integration of detail. It requires elaboration and extension. It requires building a small community into a larger one. Most importantly, it requires sharing this vision with the system designers, authors, and information managers that must, through application and use, turn the model into applications that will help real users solve real problems.

The path forward will be charted through the collective means of the workshop mailing lists and subsequent workshops that will refine and elaborate this work-in-progress (see http://www.dstc.edu.au/DC4/ for information on the upcoming fourth Dublin Core Workshop). Slightly less

than two years after the original workshop, prototype applications of Dublin Core are emerging, and recognition of the Dublin Core as a foundation for discovery-oriented resource description is growing. Building upon these prototypes and refining this consensus will provide the foundation for a network-wide body of practice that can help rationalize resource description across domains and make Internet resources more accessible.

Table 1: Abbreviated Description of Dublin Core Elements

(the reference description of the elements resides at http://purl.org/metadata/dublin_core_elements.

Element Descriptions

1. TITLE
   The name given to the resource by the CREATOR or PUBLISHER.
2. AUTHOR OR CREATOR
   The person(s) or organization(s) primarily responsible for the intellectual content of the resource.
3. SUBJECT AND KEYWORDS

The topic of the resource, or keywords, phrases, or classification descriptors that describe the subject or content of the resource.

4. DESCRIPTION
   A textual description of the content of the resource, including abstracts in the case of document-like objects or content descriptions in the case of visual resources.
5. PUBLISHER
   The entity responsible for making the resource available in its present form, such as a publisher, a university department, or a corporate entity.
6. OTHER CONTRIBUTORS
   Person(s) or organization(s) in addition to those specified in the CREATOR element who have made significant intellectual contributions to the resource but whose contribution is secondary to the individuals or entities specified in the CREATOR element.
7. DATE
   The date the resource was made available in its present form.
8. RESOURCE TYPE
   The category of the resource, such as home page, novel, poem, working paper, technical report, essay, dictionary. It is expected that RESOURCE TYPE will be chosen from an enumerated list of types.
9. FORMAT

The data representation of the resource, such as text/html, ASCII, Postscript file, executable application, or JPEG image. FORMAT will be assigned from enumerated lists such as registered Internet Media Types (MIME types).

10. RESOURCE IDENTIFIER
    String or number used to uniquely identify the resource. Examples for networked resources include URLs and URNs (when implemented).
11. SOURCE
    The work, either print or electronic, from which this resource is derived, if applicable.
12. LANGUAGE
    Language(s) of the intellectual content of the resource.
13. RELATION
    Relationship to other resources. Formal specification of RELATION is currently under development.
14. COVERAGE
    The spatial locations and temporal durations characteristic of the resource. Formal specification of COVERAGE is currently under development.
15. RIGHTS MANAGEMENT
    The content of this element is intended to be a link (a URL or other suitable URI as appropriate) to a copyright notice, a rights-management statement, or perhaps a server that would provide

such information in a dynamic way.

Return to text

Further Reading

The following links will provide the interested reader with further information on the topics dealt with in this report.

The Dublin Core Homepage
The definitive source of information on historical and current developments concerning the Dublin Core and related metadata initiatives.
http://purl.org/metadata/dublin_core

The META2 Mailing List, *meta2@mrrl.lut.ac.uk*
The primary mailing list forum for the discussion of Dublin Core issues
To subscribe, send a message to majordomo@mrrl.lut.ac.uk with a body of:
subscribe meta2 *your-name-here*

The Warwick Metadata Workshop: A Framework for the Deployment of Resource Description
Lorcan Dempsey and Stuart L. Weibel
DLib Magazine, July 1996

http://www.dlib.org/dlib/july96/07weibel.html

The Warwick Framework: A Container Architecture for Aggregating Sets of Metadata
Carl Lagoze, Clifford A. Lynch and Ron Daniel Jr.
DLib Magazine, July 1996
http://www.dlib.org/dlib/july96/lagoze/07lagoze.html

A Syntax for Dublin Core Metadata: Recommendations from the Second Metadata Workshop
Lou Burnard, Eric Miller, Liam Quin, and C.M. Sperberg-McQueen.
http://users.ox.ac.uk/~lou/wip/metadata.syntax.html

Metadata for the Masses
Paul Miller
Ariadne, September, 1996
http://www.ukoln.ac.uk/ariadne/issue5/metadata-mass

Dublin Core Qualifiers
Jon Knight, and Martin Hamilton
http://www.roads.lut.ac.uk/Metadata/DC-SubElements

A Proposed Convention for Embedding Metadata in HTML
A position paper from the May, 1996 W3C Workshop on Distributed Indexing and Searching
http://www.oclc.org:5046/~weibel/html- meta.html

Home Magazine Comments

PREV NEXT

hdl:cnri.dlib/january97-weibel

Reconnecting Science and Humanities in Digital Libraries

A Symposium Sponsored by
The University of Kentucky and The British Library

19-21 October 1995
Marriott's Resort at Griffin Gate, Lexington, Kentucky

Objective:

Digital libraries in the humanities pose computationally and methodologically challenging problems that hamper humanists, but offer computer scientists exciting areas for research and development. The objective of this symposium is to improve access to humanistic information in digital form and mechanisms for access by creating closer links between the humanities and the sciences.

Goals:

1. To produce a summary statement of the challenges offered by digital research and by the digital library in the humanities which collaboration between the sciences and the humanities can address.

2. To identify technical and research issues crucial to the humanities that are being only peripherally addressed by current National Science Foundation (NSF) funding for digital libraries.

3. To outline areas of development offering opportunities for partnership with commercial ventures that will assist in the dissemination of or access to digital libraries by scholars and the general

public.

Program schedule:

Thursday, 19 October 1995.

7:30 p.m.: Reception and buffet at the Clock Tower reception area at Griffin Gate. Coat and tie. Welcoming comments by *Charles T. Wethington, Jr.*, President of the University of Kentucky; *Elisabeth Zinser*, Chancellor of the Lexington Campus; *Fitzgerald B. Bramwell*, Vice President for Research and Graduate Studies; and *Andrew Phillips*, Director of Humanities and Social Sciences and Keeper of Printed Books, The British Library.

Friday, 20 October 1995. Open sessions.

8:30 a.m.: Van ride to the University of Kentucky's Lexington campus.
9-9:30 a.m.: Continental breakfast in foyer of King Library North.

9:30 a.m.:

INTRODUCTORY COMMENTS

(Kevin Kiernan, Professor of English, Electronic Beowulf *Project, University of Kentucky)*

10-11:30 a.m.:

1. THE ROLE OF COMPUTER SCIENCE IN UNITING SCIENCE AND HUMANITIES

(Marilyn Deegan, Professor of Humanities Computing, De Montfort University, Chair)

- The Digital Library in theory and practice: a historian's B view, Andrew Prescott, Curator of Manuscripts, The British Library

- Projects and their place in Digital Libraries, Richard Heseltine, University Librarian, University of Hull

11:30 a.m. - 1:30 p.m. Lunch at the Boone Faculty Club, sponsored by Lexmark International.

1:30-5 p.m.:

2. EASE OF ACCESS FOR NON-SCIENTISTS

(Patrick Conner, Centennial Professor of Arts and Sciences, West Virginia University, Chair)

- Humanities needs and expectations for intelligent graphical user interfaces, Seamus Ross, Assistant Secretary for Information Technology, The British Academy

- Content-based searching of large-image databases, Mary Larsgaard, Alexandria Digital Library, University of California, Santa Barbara

- Content-based Multimedia Data Management and Efficient Remote Access, Brent Seales, James Griffioen, and Raj Yavatkar, Assistant Professors of Computer Science, University of Kentucky

3:15-3:45 Tea and Coffee Break

- Electronic librarians, intelligent network agents, and information catalogues, Edward Fox, Professor of Computer Science, Virginia Tech

- Dancing to the telephone: network demands and opportunities, Charles Henry, Director of Libraries, Vassar College

5:30-6:30: Visit to ATM lab.

6:30 p.m.: Vans leave ATM lab for cocktails and buffet dinner at the Kiernans, 627 South Ashland Avenue. Casual attire.

Saturday, 21 October, 1995. Closed sessions.

8:30-9 a.m.: Continental Breakfast outside the Dixiana Room at the Marriott.

9-12:30:

3. MAXIMUM FIDELITY OF IMAGES IN HUMANITIES DATABASES

(Jennifer Trant, Manager, Imaging Initiative, Getty Art History Information Program, Chair)

- Images: quantity is not always quality, B Michael Lesk, Director, Computer Science Research Department, Bellcore Laboratories

- From conversion to presentation: benchmarking image quality requirements, Anne Kenney, Associate Director, Department of Preservation and Conservation, Cornell University

- Digital preservation: a time bomb for Digital Libraries, Margaret Hedstrom, Associate Professor, School of Information and Library Studies, University of Michigan

10:45-11:15 Tea and Coffee Break

- A survey of compressed domain processing techniques, Brian Smith, Assistant Professor, Department of Computer Science, Cornell University

- Bound Images: encoding and analysis, Gerhard Jaritz, Senior Research Fellow, Institut für Realienkunde, Austrian Academy of Sciences, Krems

4. MECHANISMS FOR FUNDING

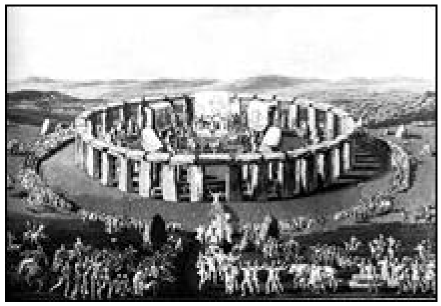
(Deian Hopkin, Dean of Human Sciences, London Guildhall University, Chair)

- Funding the Digital Library: the British Experience, Derek Law, Director of Information Services, Kings College, London

- Digital Libraries and NEH and NSF, Stephen Griffin, Director, Division of Information, Robotics and Intelligent Systems, National Science Foundation

3:15-3:45 Tea and Coffee Break

3:45-6:00 p.m.:



Concluding Roundtable

7:30. Barbecue at Marriott's Pavilion. Casual. Music from Kentucky by the Second String Band.

Sponsored by the College of Arts and Sciences, the College of Engineering, the FACTS Center, the Center for Computational Sciences, the University Libraries, the Department of English, the Department of Computer Science, the Honors Program, and Research and Graduate Studies at the University of Kentucky; the Center for Information Management and Advanced Technology for Scholarship (CIMATS) at London Guildhall University; and the British

Library.

Specifying the PREMO Synchronization Objects

D.J. Duke[1]**, D.A. Duce**[2]**, I. Herman**[3]**, G. Faconti**[4]

[1]Dept of Computer Science, University of York, Heslington, York, YO1 5DD, UK

[2]Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, UK

[3]Centrum voor Wiskunde en Informatica (CWI), Kruislaan 413, 1098 SJ Amsterdam, The Netherlands

[4]CNR - Instituto CNUCE, Via S.Maria 36, 56126 Pisa, Italy

Abstract

This paper describes the formal specification of object types for managing inter-media synchronisation and control within PREMO, an emerging ISO/IEC standard for multimedia systems. Object-Z, an object-oriented extension to the Z specification language, is used for this purpose. Some aspects of PREMO are non-trivial to express in Object-Z, and the paper outlines the reasons for choosing this specific language, and sets out recommendations for further research in the use of formal languages in this area. The work reported here has been carried out by members of the ISO/SC24 committee involved in producing the PREMO standard, and has been informed by a number of workshops sponsored by the ERCIM Computer Graphics Network.

1 Introduction

Maintaining the presentation of a continuous media data stream at a sufficient rate and quality for human perception represents a significant challenge for multimedia systems, and may impose significant resource requirements on the multimedia computing environment. Aside from this inherent constraint (sometimes referred to as the problem of intra-media synchronization) a further difficulty arises from the fact that multimedia applications often wish to use several instances of continuous media data at the same time: an animation sequence with some accompanying sound, a video sequence with textual annotations, etc. The difficulty here is that not only should the individual media data be presented with an acceptable quality, but well-defined portions of the various media content should appear, at least from a perceptual point of view, simultaneously: some parts of a sound track belong to a specific animation sequence, subtitles should appear with specified frames in a video sequence, etc. This problem is usually referred to as inter-media synchronization. The specific problems raised by intra-media synchronization will not be addressed in this paper; in what follows, the term synchronization is always used to refer to inter-media synchronization.

Synchronization has received significant attention in the multimedia literature, see, for example, the recent book by Gibbs and Tsichritzis[11] or the article of Koegel Buford[2] for further information and references on the topic. An efficient implementation of inter-media synchronization represents a major load on a multimedia system, and it is one of the major challenges in the field. What emerges from the experience of recent years is that, as is very often the case, one cannot pin down one specific place among all the computing layers (from hardware to the application) where the synchronization problem should be solved. Instead, the requirements of synchronization should be considered across all layers, i.e., in network technology, operating systems, software architectures, programming languages, etc. and user interfaces. This paper describes and formalises a model for inter-media synchronization which is contained in the PREMO specification [16], an ISO/IEC standard under development for multimedia programming. Being part of an upcoming ISO/IEC standard, the model represents a synthesis of the various synchronization techniques used in practice.

PREMO standardization is still at a development stage, hence a short overview of the main goals of this Standard are given below. The remainder of the paper concentrates on the formal specification of these objects. The formal notation used is Object-Z [3, 7, 8], an object-oriented extension to the Z notation [20]. Use of Object-Z to specify PREMO was one of the recommendations of a Special Rapporteur's Report on Formal Description Techniques in PREMO [19] prepared for ISO/IEC JTC1/SC24, and the specification of aspects of PREMO using the Z and Object-Z notations have appeared as [6] and [5]. One contribution of this paper is demonstrating use of a convention for describing exceptions and error handling, and details of the approach are to be found

in Appendix A. An overview of Z and Object-Z is beyond the scope of this paper, and the reader is directed to [12, 20] (Z) and [3, 7, 8] (Object-Z) for details.

1.1 A short overview of PREMO

This section gives a very short overview of PREMO; for a more detailed presentation the interested reader should consult [13] or [14].

Today's application developers needing to realize high-level multimedia applications which go beyond the level of multimedia authoring do not have an easy task. There are only a few programming tools that allow an application developer the freedom to create multimedia effects based on a more general model than multimedia document paradigms, and these tools are usually platform specific. In any case, there is currently no available ISO/IEC standard encompassing these requirements. A standard in this area should focus primarily on the presentation aspects of multimedia, and much less on the coding, transfer, or hypermedia document aspects, which are covered by a number of other ISO/IEC or de-facto standards (for example, MHEG[15]). It should also concentrate on the programming tool side, and less on, e.g., the (multimedia) document format side. These are exactly the main concerns of PREMO.

It is quite natural that the initiative for a standardization activity aiming at such a specification came from the group which has traditionally concentrated on presentation aspects over the past 15 years, namely ISO/IEC JTC1/SC24 (Computer Graphics). Indeed, this is the ISO subcommittee whose charter has been the development of computer graphics and image processing standards in the past. The Graphical Kernel System was the first standard for computer graphics published in this area; it was followed by a series of complementary standards, addressing different areas of computer graphics and image processing. Perhaps the best known of the application program interface (API) standards are PHIGS, PHIGS PLUS, and IPS (see, e.g., Arnold and Duce[1] for an overview of all these Standards). The subcommittee has now turned its attention to presentation media in general as a way of augmenting traditional graphics applications with continuous media such as audio, video, or still image facilities, in an integrated manner. The need for a new generation of standards for computer graphics emerged in the past 4-5 years to answer the challenges raised by new graphics techniques and programming environments and it is extremely fortunate that the review process to develop this new generation of presentation environments coincided with the emergence of multimedia. In consequence, a synergistic effect can be capitalized on.

The JTC1 SC24 subcommittee recognised the need to develop such a new line of standards. It also recognised that any new presentation environment should include more general multimedia effects to encompass the needs of various application areas. To this end, a project was started in SC24 for a new standard called PREMO (Presentation Environment for Multimedia Objects) and is now a major ongoing activity in ISO/IEC JTC1 SC24 WG6. The subcommittee's goal is to reach the stage of a Draft International Standard in 1997.

The major features of PREMO can be briefly summarised as follows.

- PREMO is a Presentation Environment. PREMO, as well as the SC24 standards cited above, aims at providing a standard "programming" environment in a very general sense. The aim is to offer a standardized, hence conceptually portable, development environment that helps to promote portable multimedia applications. PREMO concentrates on the application program interface to "presentation techniques"; this is what primarily differentiates it from other multimedia standardization projects.
- PREMO is aimed at a Multimedia presentation, whereas earlier SC24 standards concentrated either on synthetic graphics or image processing systems. Multimedia is considered here in a very general sense; high-level virtual reality environments, which mix real-time 3D rendering techniques with sound, video, or even tactile feedback, and their effects, are, for example, within the scope of PREMO.
- PREMO is Object Oriented. This means that, through standard object-oriented techniques, a PREMO implementation becomes extensible and configurable. Object-oriented technology also provides a framework to describe distribution in a consistent manner.

A precise object model constitutes a major part of PREMO. The object model is fairly traditional, and is based on the concepts of subtyping and inheritance. It is also very pragmatic in the sense that it includes, for efficiency

reasons, the notion of non-object (data) types, as is the case with a number of object-oriented languages, such as C++ or Java, and in contrast to 'pure' object-oriented models, such as SmallTalk. The PREMO object model originates from the object model developed by the OMG consortium for distributed objects, but some aspects of the OMG model have been adapted to the needs of PREMO. A strong emphasis is placed in the model on the ability of objects to be active. That is, PREMO uses objects that have, conceptually, their own thread of control. However, if every object contains a thread of control, objects become 'heavyweight' and unsuitable for use as record-like structures. Having to develop a separate (non-object) data type for structures would be unfortunate, since treating structures as objects bring benefits in the form of inheritance, subtyping etc. The PREMO object model reconciles the tension between the desire for active objects on the one hand, and efficient structure objects on the other, by splitting the object type hierarchy into separate branches for 'simple' objects that can be used efficiently as structures, and 'enhanced' objects that have their own thread of control. The top level of the PREMO object type hierarchy is shown in Figure 1.

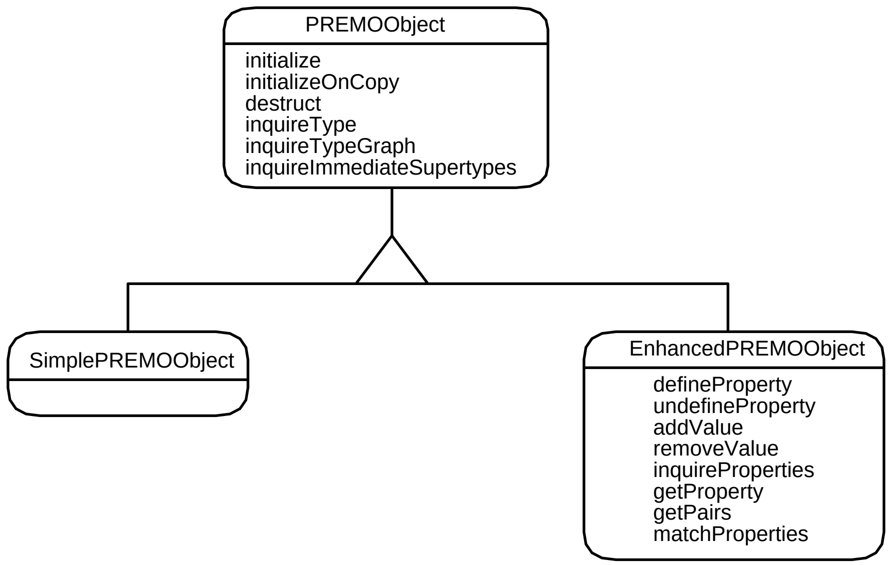


Figure 1: The top of the PREMO Object Type Hierarchy.

Enhanced PREMO objects can communicate with one another through messages, i.e., through the operations defined on the object types. Objects can become suspended either by waiting for an operation invocation to return, or by waiting on the arrival of an operation request. Consequently, operations on objects serve as a vehicle to synchronize various activities (note that this concept of object synchronization is not the same as media synchronization although, of course, the concepts are related). Whether the concurrent activity of active objects is realized through separate hardware processors, through distribution over a network, or through some multithreaded operating system service, is oblivious to PREMO and is considered to be an implementation dependency.

The emphasis on the activity of objects stems primarily from the need for synchronization in multimedia environments and forms the basis of the synchronization model described in this paper. Using concurrency to achieve synchronization in multimedia systems is not specific to PREMO. Other models and systems have taken a similar approach (see, for example, [4] and PREMO, whose task is to provide a synthesis for standardization, has obviously been influenced by these models.

2 Synchronizable Objects

As described above, the PREMO synchronization model is based on the fact that objects in PREMO can be active. Different continuous media (e.g., a video sequence and corresponding sound track) are modelled as concurrent activities that may have to reach specific milestones at distinct and possibly user definable synchronization points. This is the event-based synchronization approach, which forms the basic layer of synchronization in PREMO. Although a large number of synchronization tasks are, in practice, related to synchronization

in time, the choice of an essentially ”timeless” synchronization scheme as a basis offers greater flexibility. While time-related synchronization schemes can be built on top of an event-based synchronization model, it is sometimes necessary to support purely event-based synchronization to achieve special effects required by some application. Examples of how the various synchronization objects may be used can be found in [14].

In line with the object-oriented approach of PREMO, the synchronization model defines abstract object types that capture the essential features of synchronization. For the event-based synchronization scheme two major object types are defined:

- synchronizable objects, which form the supertypes of, e.g., various media object types;
- synchronization points, which may be used to manage complex synchronization patterns among synchronizable objects.

These objects are described in somewhat more detail below.

2.1 Supporting Synchronization in PREMO

Synchronizable objects in PREMO are autonomous objects, which have an internal progression along an internal one dimensional coordinate space. This space can be:

- extended real ($\mathbb{R}_\infty$), or
- extended integer ($\mathbb{Z}_\infty$), or
- extended time (Time_∞);

where “extension” means the inclusion of positive and negative “infinity” to the real and integer numbers, respectively. (The symbol “C” is used in this section to denote either an extended real, an extended integer, or extended time.) The obvious extension of the notions “greater than”, “smaller than”, etc., on these types allows the behaviour of synchronizable objects to be defined more succinctly. $Time$ is used here as an abstract type, with no commitment made either to a discrete, dense, continuous or discontinuous foundation. Subtypes of synchronizable objects may add a semantic meaning to this coordinate space. For example, media objects may represent time, or video frame numbers along this space. Attributes that define the extent of the progression space can be set through operations defined on these objects.

Technical Note: The formal representation of these extended types within Object-Z is a non-trivial, though solvable, problem. For example, Z does not define a type for the real numbers, and constructing a model of the reals within the type theory of Z is a complex undertaking[1]. Also, the symbol ‘∞’ is overloaded in the definitions given above. As it happens, we do not need to utilise specific properties of the real numbers, and it could be argued that a specification of PREMO would benefit from a more abstract description than $\mathbb{R}_\infty$ etc, for example by the introduction of a single abstract type to cover the three cases mentioned above. The trade-off here would be the loss of direct contact between the definition of the standard and its specification. The use of infinity, for example, is useful within the PREMO standard in describing media streams originating from ‘live’ sources such as microphones, where the temporal extent of the stream is unbounded.

Reference points are points on the internal coordinate space of synchronizable objects where synchronization elements can be attached. Synchronization elements contain information on an event instance (which is, essentially, a structure containing the object reference of the sender, a unique event type identity, and some event-dependent data), a reference to a PREMO object, a reference to one of the operations of this object, and, finally, a boolean Wait flag. When a reference point is reached, the synchronizable object makes a message call to the object stored in the reference point, using the operation reference to identify which operation it has

[1]The same is true actually in many specification languages.

to call, and using the event instance as an argument to the call. Finally, it may suspend itself if the Wait flag is set to TRUE. Through this mechanism, the synchronizable object can stop other objects, restart them, suspend them, etc. Operations are defined on synchronizable objects to add and delete reference points, and to add and delete synchronization elements associated with reference points.

A synchronizable object is a finite state machine that controls the position and progress through an ordered collection of coordinates, some of which contain *synchronization elements* that can be used to organise the behaviour within a system based on such objects. The intention is that object types representing different kinds of media (video, sound etc) will inherit from this class and specialise the coordinate system and state machine in an appropriate way. For example, a video might use frame numbers as coordinates. The synchronizable object type defines operations for moving the machine between four different modes:

$$SyncMode == \mathbb{N}$$

$$STOPPED, PLAY, PAUSED, WAITING : SyncMode$$

$$STOPPED = 0 \land PLAY = 1 \land PAUSED = 2 \land WAITING = 3$$

Technical Note: The data type $SyncMode$ might better be represented in Z as a disjoint union; the use of explicit numerical tags is used to reflect the approach taken in the Standard. The use of natural numbers to represent modes allows later subtypes to extend the set of modes; the cost is that operations that accept as input a natural number representing a mode have to check that this input describes a mode that is valid for objects of that type.

In more precise terms, a Synchronizable object type is defined in PREMO as a supertype for all objects which may be subject to synchronization. This object is defined to be a finite state machine. The possible states, the state transitions, and the operations resulting in state transitions, are shown in Figure 2. The initial state is STOPPED. Note that no *observable* operation is defined for a transition into state WAITING; the only way a Synchronizable object can go into the WAITING state is through an operation internal to its processing cycle.

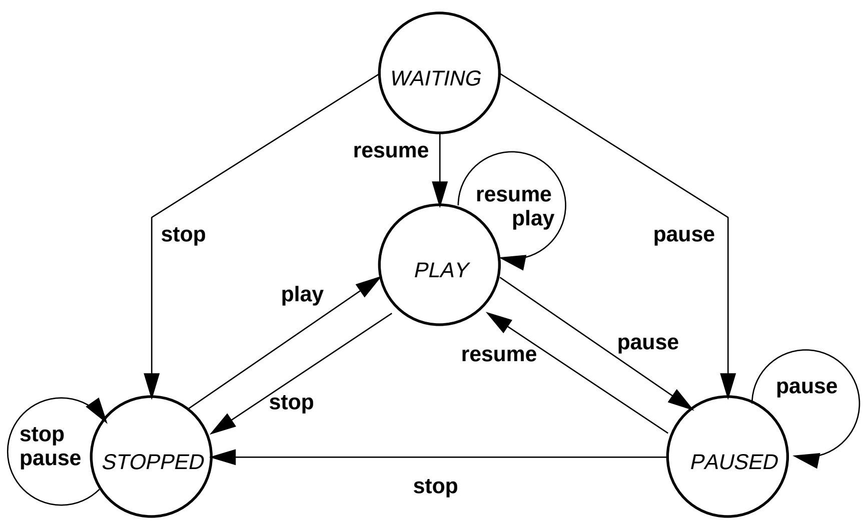


Figure 2: The States and Transitions of the *Synchronizable* Object Type.

The Synchronizable object type places no particular interpretation on these modes, except that certain operations can only be performed in certain modes. Apart from mode, the type also makes use of a notion of direction, which defines the meaning of progress through the coordinate space:

$$Direction ::= Forward \mid Backward$$

The positions within a synchronizable object at which some action needs to take place is marked by a synchronization element. Synchronization is defined in terms of the PREMO event system. The basic unit of this is the event structure, defined below.

$$
\begin{array}{|l}
\hline
\textit{Event} \\
\textit{SimplePREMOObject} \\
\begin{array}{|l}
\hline
\textit{eventName} : \textit{String} \\
\textit{eventData} : \mathrm{seq}(\textit{Key} \times \textit{Value}) \\
\textit{eventSource} : \textit{EnhancedPREMOObject} \\
\hline
\end{array} \\
\hline
\end{array}
$$

Technical Note: The object type $Event$ inherits only from $SimplePREMOObject$; it is used as a structure type to encapsulate information about events. It uses the following data types:

- $String$ is a given type representing character strings;
- Key is a given type representing names that can be used as the first component of key-value pairs, where the $Value$ component is a disjoint union over the non-object data types of PREMO;
- $EnhancedPREMOObject$ is the root of the active object hierarchy, and thus the source of an event can be any active PREMO object.

In the ISO documents, the type of $eventSource$ is $Ref\ EnhancedPREMOObject$, since the standard is explicit about the use of object references. In Object-Z, an instance of an object type is implicitly a reference to an object of that type, i.e. references are part of the semantic model underlying Object-Z, and are not mentioned explicitly in the text of a specification.

A synchronization element contains a reference to an event handler, plus the event that should be signalled to that handler when the element is activated; the $EventHandler$ object type is described in Section 5. The third component of the element is a boolean flag, $wait$. When a synchronization object encounters an element with this flag set to true, it enters pause mode immediately upon signalling the event to the handler. This is necessary to enable synchronization between multiple media objects. Here it may be important to ensure that the position within the media does not advance once the event has been signalled, and because of message delays, network latency, or contention between objects, this control cannot be guaranteed through the use of normal messages.

$$
\begin{array}{|l}
\hline
\textit{SyncElement} \\
\textit{SimplePREMOObject} \\
\begin{array}{|l}
\hline
\textit{eventHandler} : \textit{Callback} \\
\textit{syncEvent} : \textit{Event} \\
\textit{waitFlag} : \mathbb{B} \\
\hline
\end{array} \\
\hline
\end{array}
$$

Technical Note: $Callback$ is the name of an object type that defines an operation called $callback$ that accepts an input $callbackValue?$ of type $Event$. By inheriting from $Callback$, the interface of an object type will contain the $callback$ operation, and instances of the object type can then be used as targets for event notification. The object type, $CallbackByName$, a subtype of $Callback$, is similar, except that the $eventName$ field of the event passed as input to $callback$ is taken to be the name of an operation, and in processing the event the receiver invokes that operation on itself.

An action element, like a synchronization element, contains a reference to a $Callback$ object, but in this case only the name of an event is provided. The event data will be defined by the context in which the action element is used.

$$
\begin{array}{|l}
\hline
\textit{ActionElement} \\
\textit{SimplePREMOObject} \\
\end{array}
$$

$eventHandler : Callback$
$eventName : String$

2.2 The 'Synchronizable' Object Type

A specification of the synchronizable object type appears below. The generic parameter C represents the coordinate system restricted to be an extended integer, extended real, or extended time.

$Synchronizable\,[C :: \mathbb{Z}_\infty \mid \mathbb{R}_\infty \mid TIME_\infty]$

$EnhancedPREMOObject$ **redef** $(initialize, initializeOnCopy)$
$CallbackByName$

The following declarations are read-only attributes, i.e. each comes with an implicit operation for getting its value, but the value can only be changed by the action of specific operations in the interface of the type.

$currentDirection : Direction$
$loopCounter : \mathbb{N}$ [number of loops completed]
$currentState : SyncMode$ [playing, paused etc]
$currentPosition : C$
$maximumPosition : C$
$minimumPosition : C$ [fixed bounds of the span]

Technical Note: The match between Object-Z access concepts and PREMO access concepts, for example read-only, is not trivial. PREMO uses concepts of read-only attributes, read/write attributes, and internal (protected) operations. The latter is derived in part from the access control facilities provided by C++ and Java [10]. Although Object-Z does have a concept of (externally visible) attribute, there is no direct counterpart to protected operations. The PREMO standard defines a number of conventions for 'decorating' operation names to indicate the accessibility of the operation. These are not used in the specification, although we do utilise the Z convention of naming an operation schema ΦS if S is not a final operation but a 'frame' that captures a common pattern of processing that will be used as the basis for defining subsequent operations.

The declarations in this fragment of the specification are readable and writable attributes, i.e. each comes with implicit operations for setting and getting its value. The first pair of attributes identify the user-selectable subset of the coordinate space that is to be processed or presented.

$startPosition : C$
$endPosition : C$ [user-definable boundary]

$minimumPosition \leq startPosition$
$startPosition \leq stopPosition$
$stopPosition \leq maximumPosition$

Technical Note: Within the standard, setting the values of these attribute may cause exceptions. This is difficult to document formally in Object-Z, as the set and get methods are not an explicit part of the text. This is a good example of one of the trade-offs involved in specification, in this case between explicitness and readability. It could be argued that specification should prompt and assist a designer to be explicit about the behaviour of a system, and this in turn

requires being explicit about the components of the system that contribute to its behaviour. The cost is that a model can quickly become cluttered with detail.

Two further read/write attributes are provided:

$repeatFlag : \mathbb{B}$ [should the presentation cycle?]
$nloop : \mathbb{N}$ [total number of loops required]

The remainder of the state contains variables introduced to support the specification of the intended behaviour of this object type. They are not mentioned explicitly in the functional part of the Standard. The ability of a client to either set the presentation into an infinite cycle, or to specify a number of iterations, means that a given location within the coordinate space may be visited multiple times. It is useful in the specification to distinguish between 'visits' to a given coordinate. To achieve this a type is defined that combines a position in the space with a visit number,

$$Location == C \times \mathbb{N}_\infty$$

and we define a total order over this type.

$$\_ prec \_ : Location \leftrightarrow Location$$

$$\forall c_1, c_2 : C;\ n_1, n_2 : \mathbb{N} \bullet (c_1, n_1)\ prec\ (c_2, n_2) \Leftrightarrow n_1 < n_2 \lor (n_1 = n_2 \land c_1 < c_2)$$

No invariant is given at this point to link the coordinates visited during traversal with the parameters that determine traversal behaviour. Operations defined later in this object type can update these parameters, and it simplifies the specification if the relationship between these variables is captured as part of a 'framing' schema that is then used to define the effect of such operations. Here $refpoints$ defines the synchronization elements that have been associated with specific reference points; $loopStart$ is the coordinate that progression will start from initially. The locations that will be passed during traversal define the $span$, while the relation $\prec$ defines the order in which locations will be traversed.

$refpoints : C \rightarrowtail SyncElement$ [the sync. points]
$loopStart : C$ [the starting coord for loops]
$span : \mathbb{P}\ Location$ [locations to be traversed]
$\_ \prec \_ : Location \leftrightarrow Location$ [order of traversal]

$$\mathrm{dom}\ refpoints \subseteq minimumPosition \ldots maximumPosition$$
$$currentDirection = forward \Rightarrow loopStart = startPosition \land (\prec) \subseteq prec$$
$$currentDirection = backward \Rightarrow loopStart = endPosition \land (\prec) \subseteq prec^{-1}$$

Three further variables are used to define how progress is made during play, while moving from processing one item of data in the coordination space to the next.

$stepping : \mathbb{B}$ [true while moving from current to new]
$requiredPosition : Location$ [determined by progressPosition]
$point : Location$ [location in span while moving to requiredPosition]

The final state variable represents actions that may be associated with specific pairs of states, with the semantics that the event handler component of the action will have its 'callback' method invoked with an event containing the name stored in action whenever the synchronizable object makes a transition from the first state to the second.

$$actions : SyncMode \times SyncMode \rightarrowtail ActionElement$$

$$\forall s, t : SyncMode \bullet (s, t) \in \mathrm{dom}\ actions \Rightarrow s \neq t$$

In the initialisation section below, the initial position is set to the 'front' of the media, and the direction to forward. There are no refpoints, and the state of the object is 'STOPPED'.

INIT

$$
\begin{aligned}
&startPosition = minimumPosition \\
&endPosition = maximumPosition \\
&currentPosition = loopStart \\
&loopCounter = 0 \\
&repeatFlag = False \\
&nloop = 1 \\
&refpoints = \varnothing \\
&currentState = STOPPED \\
&currentDirection = Forward \\
&point = (loopStart, loopCounter)
\end{aligned}
$$

Technical Note: Object-Z officially does not accept parameters to the init description, since init is a state, rather than an operation. This means that some initialisations in PREMO may need to be modelled by explicit operations. More problematic from a specification viewpoint, PREMO objects may also have an *initializeOnCopy* operation. Copying is part of the operational machinery involved in executing a program on a machine, and is not a concept that can be described (easily) within a specification, without making details of the underlying machine (for example, locations) explicit in the model.

From quite early on in the specification, we will be describing operations that can change the state of a synchronizable object and therefore may cause the invocation of an action associated with the corresponding state change. Rather than repeat the relevant fragment of specification text in every such operation, it is more convenient for us first to define a *framing schema* that captures this common situation, and to then include this as part of the description of those operations that might result in a change to $currentState$.

$\Phi DoAction$

$$
\Delta(currentState)
$$

$$
\begin{aligned}
&(currentState, currentState') \in \operatorname{dom} actions \\
&\Rightarrow \\
&\exists\, callbackValue? : Event \bullet \\
&\quad callbackValue?.eventName = actions(currentState, currentState').eventName \\
&\quad callbackValue?.eventSource = self \\
&\quad actions(currentState, currentState').eventHandler.callback
\end{aligned}
$$

The callback operation will be invoked with the event data given in the event structure.

The next part of the specification addresses the meaning of progression through the coordinate space. If the object's state is PLAY, the object carries out its internal processing in a loop of processing stages. Each stage consists of the following steps:

1. The value of the current position is advanced using a (protected) operation $progressPosition$ which returns the required next position. Here 'protected' means that the operation is not accessible to clients of the object via its interface, but can be modified within object types that inherit from this class.

2. This required position is compared with the current position and the end position, and the following actions are performed:

   (a) If there are reference points lying between the current position and the newly calculated position, then any associated synchronization actions are performed (in the order in which they are defined on C). This means:

- perform data presentation for any data identified by the points on the progression space between the current position or the previous reference point and the next reference point or the end point;
- invoke the operation, whose description is stored in the reference point, on the object whose reference is stored at the reference point, using the stored event as an argument;
- if the `Wait` flag stored in the synchronization element belonging to the reference point is set to TRUE, the object's state is changed to WAITING. This internal transition is the only means by which the WAIT state can be entered. If the state of the object is set back, eventually, to PLAY, the stage continues at this point.

(b) If the required position is smaller than the end position, then this becomes the local position and the processing stage is finished.

The details of these steps will be explained through a series of operation definitions. Note however that two aspects of this processing cycle are left unspecified in the definition of $Synchronizable$:

- what 'data presentation' means, and
- the detailed semantics of the $progressPosition$ operation.

Both these aspects should be specified in the appropriate subtypes of $Synchronizable$. The abstract specification of a synchronizable object is such that no media specific semantics are directly attached to it. Subtypes, realizing specific media control should, through specialization, attach semantics to the object through their choice of the type of the internal coordinate system, through a proper specification of what data presentation means, and through a proper specification of the $progressPosition$ operation. The latter defines what it really means to "advance" along the internal coordinate system. For example, this progression may mean the generation of the next animation frame, decoding the next video frame, advance in time, etc.

The following operation calculates the next location required; it is expected that it will be specialised by subclasses to address behaviour specific to various types of media. It is a 'protected' operation in the sense that clients of a $Synchronizable$ object are not supposed to invoke this. The point in the coordinate space that will be visited next is returned as an output.

$progressPosition$

$$\Delta(requiredPosition, stepping)$$
$$newPosition! : C$$

$$currentState = PLAY \land \neg\, stepping$$
$$\exists\, count : \mathbb{N} \mid count < nloop \bullet$$
$$\quad (newPosition!, count) \in span$$
$$\quad requiredPosition' = (newPosition!, count)$$
$$stepping'$$

Technical Note: Once a new location has been calculated, the object is placed into a 'stepping' mode. This is not a mode behaviour described in the standard, but rather is an artefact of the specification introduced to model the sequence of operations that are assumed to take place internally. While it could be argued that such detail is better captured in a process oriented language, a better alternative (in principle) would be to find a more declarative means of specifying the behaviour of the object, perhaps through invariants over behaviours.

The technical point above raises an important issue about PREMO, which has implications for other standards and systems, for example VRML [17]. In developing a standard, particularly in an area where performance is a non-trivial concern, there may be implicit assumptions about the execution model that will be used to realise the system. In the case of PREMO for example, the specification of $progressPosition$ given above, and the

semantics of the internal *stepping* mode that will be presented shortly, involve a level of operational detail that normally one would associate more with a design or implementation. The problem is that a more abstract description of the intended behaviour may be rather more difficult to understand; state machines are after all a well understood engineering concept, a fact that has been borne out by the experience of others in designing languages to document requirements and specifications [18].

We return to the semantics of a synchronizable object in *stepping* mode. This is captured through four conceptual operations which represent the aspects of the processing cycle, and which are later composed into a description of stepping behaviour.

> The $\Phi step$ operation represent the selection of the next location in the span for which the related data will be presented. Some of the locations between the current point and the new point may be skipped. However, no location for which the underlying coordinate has a reference point can be skipped.
> In the formal text, the functions *fst* and *snd* select the first and second components of a tuple respectively.

$\Phi step$

$$
\begin{array}{l}
\Delta(point, span, loopCounter) \\
\hline
stepping \land point \neq requiredPosition \\
point \prec point' \\
\neg\ requiredPosition \prec point' \\
\textbf{let}\ skipped == \{loc : span \mid loc \prec point'\} \bullet \\
fst(\!|\ skipped\ |\!) \cap \mathrm{dom}\ refpoints = \varnothing \\
loopCounter' = loopCounter + |\,snd(point') - snd(point)\,| \\
span' = span \setminus skipped
\end{array}
$$

Technical Note: PREMO itself describes an ideal situation where the progression space may be continuous. The stepping mechanism introduced here has a discrete flavour. There is a correspondence with the sampling and quantization processes that occur when continuous media are processed by digital systems.

> The $\Phi signal$ schema is a framing schema that describes the relationship between the location of the current point in the coordination space and the synchronisation point, if any, that is located at that coordinate. If a synchronisation point has been set, then an event is signalled to the appropriate event handler, and further, if the wait flag has been set, the object enters WAITING mode. If no synchronisation element is present, the state of the object is unchanged.

$\Phi signal$

$$
\begin{array}{l}
\Delta(currentState) \\
\Phi DoAction \\
\hline
\textbf{let}\ coord == fst(point) \bullet \\
coord \in \mathrm{dom}\ refpoints \\
\quad \Rightarrow \left(\begin{array}{l}
\exists\ callbackValue? : Event \bullet \\
\quad callbackValue = refpoints(coord).syncEvent \\
\quad refpoints(coord).eventHandler.callback \\
refpoints(coord).wait \Rightarrow currentState' = WAITING \\
\neg\ refpoints(coord).wait \Rightarrow currentState' = currentState
\end{array}\right) \\
coord \notin \mathrm{dom}\ refpoints \\
\quad \Rightarrow currentState' = currentState
\end{array}
$$

The callback operation will be invoked with the event data given in the event structure.

The "$eventHandler$" object in the synchronization element (i.e., the object which has to be notified that the synchronizable object has reached a reference point) can be any PREMO object that inherits from the $Callback$ object type. PREMO offers several different types of objects which fulfil this criterion, for example

- so-called "controller" objects, which are essentially finite state machines;
- event handler objects, which can dispatch events among several registered targets;
- other synchronizable objects.

The next two parts of the behaviour description deal with termination of stepping mode. This will happen when:

- the value of $point$ reaches $requiredPosition$ as established by $progressPosition$, or
- while in $WAITING$ mode, the object undergoes a $STOP$ or $PAUSE$ operation, or
- the object completes playing.

The first two cases are described in $\Phi\, doneStepping$, the latter in $\Phi\, donePlay$.

$\Phi\, doneStepping$

$\Delta(stepping, currentPosition)$

$point = requiredPosition \land \neg\, stepping' \land currentPosition' = fst(point)$

$\Phi\, donePlay$

$\Delta(stepping)$

$stop$ [Defined on next page]

$span = \varnothing \land currentState' = STOPPED \land \neg\, stepping'$

Each cycle of processing is then defined by the progression to a new position followed by a sequence of step and signal operations. As a consequence of the moding induced by the $stepping$ flag, this processing is forced to continue until the required position has been reached.

$\Phi behaviour \mathrel{\widehat{=}} (\Phi step \land \Phi signal) \lor \Phi\, doneStepping \lor \Phi\, donePlay$

A number of operations are provided to move a synchronizable object from one current state to another, and to modify various attributes that define the behaviour in PLAY mode. As a result of these operations the set of points and the successor relation between points can change. Rather than include the relationships between these parameters as a state invariant, a framing schema called $\Phi\, UpdateSpan$ defines the relationships as a postcondition. By using this schema as an initial framework for operations that can affect playback parameters, the postcondition of these operations will ensure that the state of the object is left in a state consistent with the most recent settings.

The following framing schema defines the value of the specification variable $span$ that is used to describe how progress occurs through the coordinate space of an object when the current state of the object is 'PLAY'. The schema captures a general invariant that will be required to hold after certain other operations, defined below.
In general the play operation can involve making several iterations through the coordinate space, and thus any coordinate may be visited a number of times. This is captured by the use of the $Location$ type rather than the generic coordinate space parameter C to represent the span.
An auxiliary function is defined to determine the target number of iterations based on the value of the 'repeatFlag' and 'nloop' attributes. Note that, as the value of the 'loopCounter' attribute represents the number of completed loops, its target value in the case that the repeat flag is false is one less than the number of loops required.

$$target : \mathbb{B} \times \mathbb{N} \rightarrow \mathbb{N}_\infty$$

$$\forall\, loops : \mathbb{N} \bullet$$
$$\quad target(true, loops) = \infty$$
$$\quad target(false, loops) = loops - 1$$

The span is updated to the set of all locations where the coordinate lies between the start and end positions and the loop number is between zero and the target computed by the function above, **minus** those locations that have already been visited. This restriction is necessary as a span may need to be recomputed when a synchronizable object has been paused.

$\Phi\, UpdateSpan$

$$\Delta(span)$$

$$span' = (startPosition \mathrel{..} endPosition) \times (0..target(repeatFlag, nloop)) \setminus \{loc : Location \mid loc \prec (currentPosition, loopCounter)\}$$

The operations that modify the main parameters of a synchronizable object are defined next. Some of these are only applicable in certain 'modes', and therefore come with preconditions.

Technical Note: The notation used in this part of the specification for describing exceptions and error handling is described in Appendix A.

The first pair of operations allow the object to be placed into the state 'PLAY' or 'STOPPED'. The former can only be achieved when the media object is stopped; if this condition is not met, an error is raised. A media object can however be stopped when it is in any state. Stopping an object causes its position and loop counter to be reset to their initial values, and therefore requires 'internal' variables to be updated.

$play$

$$\Delta(currentState)$$
$$\Phi DoAction$$
$$\Phi\, UpdateSpan$$

$currentState \in \{PLAY, STOPPED\} \longrightarrow$ exc WrongState

$$currentState' = PLAY$$

WrongState $\longrightarrow$ $currentState' = currentState$

$stop$

$$\Delta(currentState, loopCounter, currentPosition, stepping)$$
$$\Phi\, UpdateSpan$$
$$\Phi DoAction$$

$$currentState' = STOPPED$$
$$loopCounter' = 0$$
$$currentPosition' = loopStart$$
$$\neg\, stepping'$$

If the object is PAUSED or WAITING, then it can only react to a very restricted set of operation requests: the attributes of the object may be retrieved (but not set) and the resume, pause or stop operations may be invoked, which may result in a change in state. The difference between PAUSED and WAITING is that, in the latter case, the object returns to the place where it had been suspended by a *Wait* flag, whereas, in the former case,

a complete new processing stage begins. The differentiation between these two states, i.e., the usage of the *Wait* flag, is essential; this mechanism ensures an instantaneous control over the behaviour of the object at a synchronization point. If the object could only be stopped by another object via a pause call, an unwanted race condition could occur.

pause

$\Delta(currentState, stepping)$
$\Phi DoAction$

$currentState \neq STOPPED \wedge currentState' = PAUSED \wedge \neg\, stepping'$
$\vee$
$currentState = STOPPED \wedge currentState' = currentState$

resume

$\Delta(currentState)$
$\Phi DoAction$

$currentState \in \{PLAY, PAUSED, WAITING\} \quad \longrightarrow$ [exc] WrongState

---

$currentState' = PLAY$
WrongState $\longrightarrow$ $currentState' = currentState$

The remaining operations are relatively straightforward, and involve setting or retrieving the values of state components. Again, some of these involve preconditions.

The 'current state' of an object can be inquired.

getState

$currentState! : SyncMode$

$currentState! = currentState$

Both the start and end positions within an object can be set to new coordinates, provided that the start position remains strictly less than the end position.

setStartPosition

$\Delta(startPosition)$
$\Phi UpdateSpan$
$startPosition? : C$

$startPosition? < endPosition \quad \longrightarrow$ [exc] WrongValue
$currentState = STOPPED \quad \longrightarrow$ [exc] WrongState

---

$startPosition' = startPosition?$
WrongValue $\longrightarrow$ $startPosition' = startPosition$

setEndPosition

$\Delta(endPosition)$
$\Phi UpdateSpan$
$endPosition? : C$

$endPosition? > startPosition \quad \longrightarrow$ [exc] WrongValue
$currentState = STOPPED \quad \longrightarrow$ [exc] WrongState

---

$endPosition' = endPosition?$
WrongValue $\longrightarrow$ $endPosition' = endPosition$

The current value of both the start and end positions can be inquired.

getStartPosition

$$startPosition! : C$$

$$startPosition! = startPosition$$

getEndPosition

$$endPosition! : C$$

$$endPosition! = endPosition$$

In addition to the progression caused in 'PLAY' state, the position within the coordinate space can also be changed explicitly, using the *jump* operation. The Standard requires that the object is in 'PLAY' or 'WAITING' state, and that the new position must be within the bounds of the object.

jump

$$\Delta(currentPosition)$$
$$\Phi\, UpdateSpan$$
$$newPosition? : C$$

$currentState \in \{PAUSED, STOPPED\} \longrightarrow$ exc WrongState
$newPosition? \in startPosition \ldots endPosition \longrightarrow$ exc WrongValue

$$currentPosition' = newPosition?$$

WrongState $\vee$ WrongValue $\longrightarrow currentPosition' = currentPosition$

At any time, the current position can be inquired.

getPosition

$$currentPosition! : C$$

$$currentPosition! = currentPosition$$

The *repeat* flag determines what will happen when the position reaches the end of the extent defined by the start and end locators. If true, the position will be reset and playback continues. The value of this flag is set and queried by the pair of operations given below.

setRepeatFlag

$$\Delta(repeatFlag)$$
$$\Phi\, UpdateSpan$$
$$r? : \mathbb{B}$$

$currentState = STOPPED$
$\longrightarrow$ exc WrongState

$$repeatFlag' = r?$$

getRepeatFlag

$$repeatFlag! : \mathbb{B}$$

$$repeatFlag! = repeatFlag$$

A bounded number of repetitions can be requested by setting the 'number of loops' variable through the operation defined below left; the value of this variable can be enquired through the operation on the right.

setNumberOfLoops

$$\Phi\, UpdateSpan$$
$$\Delta(nloop)$$
$$numberOfLoops? : \mathbb{N}$$

$currentState = STOPPED$
$\longrightarrow$ exc WrongState

$$nloop' = numberOfLoops?$$

getNumberOfLoops

$$numberOfLoops! : \mathbb{N}$$
$$loopCounter! : \mathbb{N}$$

$$numberOfLoops! = nloop$$
$$loopCounter! = loopCounter$$

While in play mode, a loop counter keeps track of the number of loops that have been completed through the media object. This count can be reset to the number of loops specified by $nloop$.

resetLoopCounter

$\Delta(loopCounter)$

$currentState = STOPPED \longrightarrow$ exc WrongState

$loopCounter' = nloop$

The extreme bounds of a given media object can be discovered through the $getBounds$ operation given below.

getBounds

$minimumPosition! : C$
$maximumPosition! : C$

$minimumPosition! = minimumPosition$
$maximumPosition! = maximumPosition$

The remaining operations are those that manipulate the synchronization elements, for example setting a new one at a reference point or deleting an existing one. The first group of these set and delete individual points, and allow enquiry about the set of points between two given coordinates.

A synchronization element can be set at a given coordinate within the extent of a media object. This cannot be done while the object is in 'PLAY' state, and the specified coordinate must be valid for that object, i.e. between the minimum and maximum bounds.

setSyncElement

$\Delta(refpoints)$
$refpoint? : C$
$syncData? : SyncElement$

$currentState \in \{PAUSED, STOPPED\} \longrightarrow$ exc WrongState
$refpoint? \in minimumPosition \ldots maximumPosition \longrightarrow$ exc WrongValue

---

$refpoints' = refpoints \oplus \{refpoint? \mapsto syncData?\}$
WrongState $\vee$ WrongValue $\longrightarrow refpoints' = refpoints$

A synchronization element at a given coordinate can be deleted, by passing the coordinate to the $deleteSyncElement$ operation. As before, this cannot be done while the object is in 'PLAY' state, and the specified coordinate must be valid for that object.

deleteSyncElement

$\Delta(refpoints)$
$refpoint? : C$

$currentState \in \{PAUSED, STOPPED\} \longrightarrow$ exc WrongState
$refpoint? \notin \mathrm{dom}\ refpoints \longrightarrow$ exc WrongValue

---

$refpoints' = \{refpoint?\} \lhd refpoints$
WrongState $\vee$ WrongValue $\longrightarrow refpoints' = refpoints$

The synchronization elements occurring between two specified coordinates can be obtained through the operation *getSyncElements* given below. The sequence of values returned does not necessarily preserve the order of the points; instead, each value in the sequence is a pair consisting of the synchronization element, plus the coordinate at which it occurs.

getSyncElements

$$refpoint_1? : C$$
$$refpoint_2? : C$$
$$syncData! : \mathrm{seq}(SyncElement \times C)$$

$$minimumPosition \leq refpoint_1? \leq refpoint_2? \leq maximumPosition$$
$$\longrightarrow \boxed{\text{exc}}\ \text{WrongValue}$$

$$\textbf{let}\ inrange == \{s : SyncElement;\ c : C \mid refpoint_1? \leq c \leq refpoint_2? \wedge s = refpoints(c)\} \bullet$$
$$\mathrm{ran}\ syncData! = inrange \wedge \#syncData! = \#inrange$$
$$\text{WrongValue} \longrightarrow syncData! = \langle\rangle$$

It is also possible to set and delete *periodic* synchronization elements, in terms of a base coordinate plus offsets from that base.

Periodic events are specified by two coordinates, one giving the time of the first event, the other giving the time that should elapse between events (the periodicity). The common synchronization element that should be invoked on each occurrence of an event is also given as an argument to the operation.

setPeriodicSyncElement

$$\Delta(refpoints)$$
$$startRefPoint? : C$$
$$periodicity? : C$$
$$syncData? : SyncElement$$

$$currentState \in \{PAUSED, STOPPED\} \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$
$$startRefPoint? \in minimumPosition \ldots maximumPosition \longrightarrow \boxed{\text{exc}}\ \text{WrongValue}$$

$$\textbf{let}\ points == \{n : \mathbb{N} \bullet (startRefPoint? + n \times periodicity?)\} \bullet$$
$$refpoints' = refpoints \oplus \left(\begin{array}{l}(minimumPosition \ldots maximumPosition) \\ \lhd \\ \{p : points \bullet p \mapsto syncData?\}\end{array}\right)$$
$$\text{WrongState} \vee \text{WrongValue} \longrightarrow refpoints' = refpoints$$

Technical comment: To describe the operation, we first construct the (infinite) set of all the time points that are related by the start point and periodicity. Each point in this set is mapped to the common sychronization element, and this mapping, restricted to the bounds of the object's coordinates, is written into the reference points. Any existing synchronization point that happens to occur at the same time as a periodic event is thus replaced by the new event.

deletePeriodicSyncElement

$$\Delta(refpoints)$$
$$startRefPoint? : C$$
$$periodicity? : C$$

$$currentState \in \{PAUSED, STOPPED\} \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$
$$startRefPoint? \in minimumPosition \ldots maximumPosition \longrightarrow \boxed{\text{exc}}\ \text{WrongValue}$$

$$\textbf{let}\ points == \{n : \mathbb{N} \bullet (startRefPoint? + n \times periodicity?)\} \bullet$$
$$refpoints' = points \lhd refpoints$$
$$\text{WrongState} \lor \text{WrongValue} \longrightarrow refpoints' = refpoints$$

The next two operations set and remove actions that are to be invoked on transitions between states.

Technical Note: The handling of actions invoked by state changes is orthogonal to other aspects of the synchronizable object type, and probably would be better defined in a separate super type of synchronizable, both in the specification presented here, and in the Standard.

setActionOnPair

$$\Delta(actions)$$
$$stateOld? : SyncMode$$
$$stateNew? : SyncMode$$
$$action? : ActionElement$$

$$stateOld? \in SyncMode \land stateNew? \in SyncMode \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$

$$actions' = actions \oplus \{(stateOld?, stateNew?) \mapsto action?\}$$
$$\text{WrongState} \longrightarrow actions' = actions$$

removeActionOnPair

$$\Delta(actions)$$
$$stateOld? : SyncMode$$
$$stateNew? : SyncMode$$

$$stateOld? \in SyncMode \land stateNew? \in SyncMode \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$

$$actions' = \{(stateOld?, stateNew?)\} \lhd actions$$
$$\text{WrongState} \longrightarrow actions' = actions$$

The final operation clears all of the synchronization elements.

clearSyncElements

$$\Delta(refpoints)$$

$$currentState \in \{PAUSED, STOPPED\} \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$

$$refpoints' = \varnothing$$
$$\text{WrongState} \lor \text{WrongValue} \longrightarrow currentPosition' = currentPosition$$

Before concluding this section, it is worth mentioning one further weakness in the specification which is a consequence of the use of the 'internal' $stepping$ mode to define the meaning of progression. From a client's view, the progression of an object consists of atomic steps, and in particular an operation request to an object should only be received and processed outside of 'stepping' activities. In an implementation, progression through the coordinate space would be achieved through the internal invocation of the $progressPosition$ operation, and other operation requests would either be blocked or buffered until that processing was accomplished. In other words, the existence of the $stepping$ mode should be invisible to the environment of an object. However, in the formal specification, operations such as changing the current mode or setting various attributes should have $\neg\ stepping$ as a precondition. Including this would result in a more cluttered specification, and may be misleading to an implementor. In a sense, the approach taken to the specification has forced us to posit a mechanism for explaining the concept of progression. However, it is the behaviour that emerges from this mechanism, rather than its components, that is important. Finding a way of describing such behaviour without introducing implementation details or bias into the specification appears to be a challenge. A more abstract model perhaps based on traces or behaviours might be suitable, but it is unclear how well such an approach would cope with the structural complexity of the $Synchronizable$ object type, or how directly it would map to the text of the Standard.

3 The 'Clock', and Related Object Types

The clock object type provides PREMO with an interface to whatever notion of time is supported by its environment. Specifically, the clock object type supports an operation, $inquireTick$, that returns the number of ticks elapsed. This start of era is defined for all PREMO systems in the Standard. However, the accuracy and units with which a particular PREMO implementation can describe the elasped duration since the start of era will vary, and for this reason the clock object type provides two inquiry functions for determining the performance of the local object. The clock object type assumes the existence of the following two non-object types, one to measure elapsed ticks (realised for example as a 64-bit integer), and the other, an enumerated type, to define the unit represented by each clock tick, for example a year or micro-second.

$[Time]$

$[TimeUnit]$

Clock

$tickUnit : TimeUnit$
$accuracyUnit : TimeUnit$
$accuracy : Time$ [read-only]

$ticks : Time$

$\Phi Tick$

$\Delta(ticks)$

$ticks' > ticks$

$inquireTick$

$ticks! : Time$

$ticks! = ticks$

Suppose that the output of $inquireTick$ is T, and of $inquireAccuracy$ is A. Suppose the start of era is E. Then mathematically, the actual time in the world when $inquireTick$ is called will be between $E + (T - f(A)/2)$ and $E + (T + f(A)/2)$, i.e $E + T \pm f(A)/2$, where $f(A)$ is a function which converts the accuracy value from its own units to the units of T.

Technical Note: The operation $\Phi Tick$ is intended to represent the progression of time by requiring that the number of ticks after the operation occurs is larger than the number of ticks before. How occurrences of the operation should be related to the notion of time in the environment of PREMO is an open problem. Note that this operation is only defined in the specification, and does not appear in the Standard, where the semantics of time are conveyed informally.

A system clock is a clock for which the $inquireTick$ operation returns the number of ticks elapsed since the start of the PREMO era, which in the Standard is defined to be 00:00am, 1st January 1995, UTC.

$$
\begin{array}{l}
\textbf{SysClock} \\
\hline
Clock\ \textbf{redef}\ (inquireTick) \\
\quad \underline{\ inquireTick\ } \\
\quad ticks! : TIME \\
\quad \hline
\quad ticks! = ticks
\end{array}
$$

PREMO also provides an object type called a $Timer$, which intuitively behaves like a combination of a clock and a state machine similar to that governing the modes of a synchronizable object. There are three timer states, identified by distinct natural numbers:

$$
\begin{array}{l}
TRUNNING, TSTOPPED, TPAUSED : \mathbb{N} \\
\hline
\#\{TRUNNING, TSTOPPED, TPAUSED\} = 3
\end{array}
$$

The specification of the timer object type appears below; in many ways it is quite similar to the Synchronizable object type, raising a question of whether these two types should be derived from a common supertype in the Standard.

$$
\begin{array}{l}
\textbf{Timer} \\
\hline
Clock\ \textbf{redef}\ (inquireTick, \Phi Tick) \\
\quad timerCurrentState : \mathbb{N} \\
\quad \underline{\ inquireTick\ } \\
\quad ticks! : Time \\
\quad \hline
\quad ticks! = ticks
\end{array}
$$

Technical Note: The standard indicates that the $inquireTick$ operation of $Timer$ redefines that of $Clock$, since the implementation of the former may be substantially different. However, in terms of the abstract model, the operations are the same, and so the use of '**redef**' is strictly unnecessary.

$$
\begin{array}{l}
\quad \underline{\ run\ } \\
\quad \Delta(timerCurrentState) \\
\quad \hline
\quad timerCurrentState \in \{TRUNNING, TSTOPPED\} \\
\quad timerCurrentState' = TRUNNING
\end{array}
$$

$$
\begin{array}{l}
\quad \underline{\ pause\ } \\
\quad \Delta(timerCurrentState) \\
\quad \hline
\quad timerCurrentState' = TPAUSED
\end{array}
\qquad
\begin{array}{l}
\underline{\ stop\ } \\
\Delta(timerCurrentState) \\
\hline
timerCurrentState' = TSTOPPED
\end{array}
$$

resume

$\Delta(timerCurrentState)$

$timerCurrentState \in \{TRUNNING, TPAUSED\}$
$timerCurrentState' = TRUNNING$

reset

$\Delta(ticks)$

$ticks' = 0$

$\Phi Tick$

$\Delta(ticks)$

$timerCurrentState = TRUNNING$
$ticks' > ticks$

Note that the internal $\Phi Tick$ operation has been redefined to indicate that ticks should only be recorded when an instance of the object type is in $TRUNNING$ mode. The technical comment following the $Clock$ object type still applies, however.

4 The 'TimeSynchronizable', and Related Object Types

A $TimeSynchronizable$ object type is a synchronizable object type enriched with a timer interface and an attribute called $speed$ which relates progress through the coordinate space inherited from $Synchronizable$ with the progression of time as measured by the $Timer$. In other words, speed defines the number of coordinate space units that an instance of the object type will progress through in one tick.

$TimeSynchronizable\,[T]$

$Synchronizable\,[T]$ **redef** $(play, stop, resume, pause)$

$Timer$ **redef** $(play, stop, resume, pause)$

$speed : C$

Speed represents the number coordinate space units in C that are processed in PLAY mode per unit tick.

The operations for moving a $TimeSynchronizable$ object instance from one mode to another redefine the like-named operations from the two inherited classes. In the new operations, the two machines state machines are slaved, so that a change in mode at the outer level is realised by changing the mode of both component machines.

$$run \mathrel{\widehat=} Timer :: run \mathbin{⨟} Synchronizable :: run$$
$$play \mathrel{\widehat=} Timer :: play \mathbin{⨟} Synchronizable :: play$$
$$stop \mathrel{\widehat=} Timer :: stop \mathbin{⨟} Synchronizable :: stop$$
$$resume \mathrel{\widehat=} Timer :: resume \mathbin{⨟} Synchronizable :: resume$$
$$pause \mathrel{\widehat=} Timer :: pause \mathbin{⨟} Synchronizable :: pause$$

The $reset$ operation inherited from $Timer$ is extended to place a marker on the corresponding point of the coordinate space. Some aspects of the marker concept are unclear, and therefore this part of the specification is tentative. Since the value of the $marker$ is persistent, it is introduced as a component of the internal state. From the text, it seems likely that the initial value of $marker$ is $startPosition$.

$marker : C$

Init

$marker = startPosition$

$$\begin{array}{|l} \textit{reset} \\ \hline \Delta(\textit{marker}) \\ \hline \textit{marker}' = \textit{currentPosition} \\ \hline \end{array}$$

Two operations are defined for converting between values of time relative to the marker, and positions within the coordinate space.

$$\begin{array}{|l} \textit{timeToSpace} \\ \hline \textit{positionTime}? : \textit{Time} \\ \textit{positionSpace}! : C \\ \hline \textit{positionSpace}! = \textit{marker} + \textit{positionTime}? \times \textit{speed} \\ \hline \end{array}$$

$$\begin{array}{|l} \textit{spaceToTime} \\ \hline \textit{positionSpace}? : C \\ \textit{positionTime}! : \textit{Time} \\ \hline \textit{positionTime}! = (\textit{positionSpace}? - \textit{marker}) / \textit{speed} \\ \hline \end{array}$$

The $\textit{TimeSynchronizable}$ object type overloads a number of operations inherited from $\textit{Synchronizable}$ to allow time values to be used as alternatives to positions within the coordinate space.

Technical Note: Object-Z, like many specification languages, does not allow such overloading, so the following block of specification text would be rejected by a type checker.

$$\textit{jump} \\ \quad \widehat{=} \ \textit{timeToSpace} \fatsemi \textit{Synchronizable} :: \textit{jump} \ [\textit{positionSpace}! / \textit{refPoint}?]$$

$$\textit{setSyncElement} \\ \quad \widehat{=} \ \textit{timeToSpace} \fatsemi \textit{Synchronizable} :: \textit{setSyncElement} \ [\textit{positionSpace}! / \textit{refPoint}?]$$

$$\textit{deleteSyncElement} \\ \quad \widehat{=} \ \textit{timeToSpace} \fatsemi \textit{Synchronizable} :: \textit{deleteSyncElement} \ [\textit{positionSpace}! / \textit{refPoint}?]$$

$$\textit{setPeriodicSyncElement} \\ \quad \widehat{=} \ \textit{timeToSpace} \fatsemi \textit{Synchronizable} :: \textit{setPeriodicSyncElement} \ [\textit{positionSpace}! / \textit{refPoint}?]$$

$$\textit{deletePeriodicSyncElement} \\ \quad \widehat{=} \ \textit{timeToSpace} \fatsemi \textit{Synchronizable} :: \textit{setPeriodicSyncElement} \ [\textit{positionSpace}! / \textit{refPoint}?]$$

The final part of this object type is a promoted variant of $\textit{getSyncElements}$ which uses values of type $\textit{Time}$ rather than C to define the boundaries for obtaining the synchronization points.

Technical Note: The specification of this operation uses a framing schema to convert inputs of type $\textit{Time}$ into values of type C. This can then be composed with the $\textit{getSyncElements}$ operation from the $\textit{Synchronizable}$ class, though the previous comment concerning overloading still applies.

$$\begin{array}{|l} \Phi\textit{convert} \\ \hline \textit{refPoint}1? : \textit{Time} \\ \textit{refPoint}2? : \textit{Time} \\ \textit{rp}1! : C \\ \textit{rp}2! : C \\ \hline \textit{self}.\textit{timeToSpace}[\textit{rp}1! / \textit{positionSpace}!] \\ \textit{self}.\textit{timeToSpace}[\textit{rp}2! / \textit{positionSpace}!] \\ \hline \end{array}$$

$$\begin{array}{l} getSyncElements \mathrel{\widehat{=}} \\ \qquad \Phi convert \gg Synchronizable :: getSyncElements\,[rp1?/refPoint1?, rp2?/refPoint2?] \end{array}$$

The $TimeLine$ object type extends the $TimeSynchronizable$ object type with a constraint that each coordinate represents one unit of time; in other words, the speed is fixed at '1'.

$$\begin{array}{|l} \hline TimeLine \\ TimeSynchronizable\,[Time] \\ \quad speed = 1 \\ \hline \end{array}$$

The remaining time-related object type is called $TimeSlave$, and allows progression of an instance through its coordinate space to be related to that of a 'master' object of type $TimeSynchronizable$.

$$\begin{array}{|l} \hline TimeSlave\,[C] \\ TimeSynchronizable\,[C] \\ \quad master : TimeSynchronizable \\ \quad speed : C \end{array}$$

The value of $master$ is either a reference to another $TimeSynchronizable$ object or the $NULLObject$. In the latter case, no synchronization with an external master is done. In the former case, the value of $speed$ is measured in terms of the $master$ object's ticks.

A master-slave relationship allows the calculation of time discrepancies between the clocks of the two objects involved. This requires the slave to have access to a function, called $masterToSlave$ in the specification, that maps ticks of the master clock into ticks of the slave clock. In the specification, this function is defined as a state variable. Also defined in the state are a set of alignment thresholds, consisting of a mapping from $Time$ to $Callback$ object references. The meaning of these are, that when the difference between the internal time of the slave, and the time of the master converted via $masterToSlave$ exceeds a particular threshold, the $callback$ operation of the corresponding object is invoked.

$$\begin{array}{l} masterToSlave : Time_\infty \rightarrow Time_\infty \\ thresholds : \mathrm{seq}(Time \times Callback) \end{array}$$

In the Standard, the $reset$ operation is redefined to allow data for time discrepencies to be requested from the master. This is not needed in the specification, since the definition of discrepancy assumes the existence of a suitable conversion function, and operations in the specification are timeless.

The current alignment between the slave and master can be inquired using an operation called $inquireAlignment$. It is not clear how this operation should behave in the event that no master has been set. Callbacks for time discrepencies can be set via $setSyncEventHandlers$, and are cleared by providing an empty sequence as input. From this, it is assumed that it is not possible to update the set of callbacks, other than by removing all and asserting a new set.

$$\begin{array}{|l} \hline inquireAlignment \\ positionSpace! : Time \\ \hline positionSpace! = |\,ticks - masterToSlave(master.ticks)\,| \\ \hline \end{array}$$

$$\begin{array}{|l} \hline setSyncEventHandlers \\ \Delta(thresholds) \\ syncEventHandlers? : \mathrm{seq}(Time \times Callback) \\ \hline thresholds' = syncEventHandlers? \\ \hline \end{array}$$

The remaining part of the specification is an extension to the $\Phi\,Tick$ operation that was introduced to model the behaviour of time-related objects. It is unclear from the Standard exactly when time discrepencies are checked; here it is assumed that the check takes place after each tick. It is also unclear whether, once callback objects have been notified of a discrepancy, continued deviation should result in further notifications.

$$
\begin{array}{l}
\Phi\,Tick \\
\hline
\textbf{let}\ dev == |ticks - masterToSlave(master.ticks)| \bullet \\
\forall\, t : Time;\ c : Callback \mid (t, c) \in \operatorname{ran} thresholds \land dev > t \bullet \\
\quad \neg\ \exists\, u : Time \mid u \in \operatorname{dom} \operatorname{ran} thresholds \land dev > u > t \\
\quad \Rightarrow \quad \exists\, callbackValue? : Event \mid \\
\qquad callbackValue?.eventName = \text{'OutofSync'} \\
\qquad callbackValue?.eventData = \langle \text{'Discrepency'} \mapsto dev \rangle \\
\qquad callbackValue?.eventSource = self \bullet \\
\qquad c.callback
\end{array}
$$

5 The 'EventHandler' Object Type

The $EventHandler$ object type provides basic support for allowing objects to register interest in particular events, and for those objects to be notified, via the handler, when such events occur. Some preliminary definitions are required. When an object registers interest in an event, the registration is given a unique $EventId$ that can be used subsequently to unregister interest. Also, a sequence of constraints on the key-value properties of events can be provided as an additional filter. A constraint matching mode ($AndOr$) determines whether all constraints, or only some, must match for success. Rather than represent all the details of such constraints, we make use of the abstraction provided by specification languages and simply indicate that there is a satisfaction relationship (sat) between constraints and key-value pairs such as those that appear as event data.

$[EventId]$ - Identifiers for event registration.
$[Constraint]$ - The constraint description language.
$AndOr ::= And \mid Or$ - Constraint matching mode

$$\_\,sat\,\_ : \operatorname{seq}(Key \times Value) \leftrightarrow Constraint$$

The $EventHandler$ object type is an enhanced PREMO object that also inherits from $Callback$ to enable instances of $EventHandler$ to be themselves the targets of event notification.

$$
\begin{array}{l}
EventHandler \\
\hline
EnhancedPREMOObject \\
Callback\ \textbf{redef}\ (callback) \\
\hline
registered : \mathbb{F}\, EventId \\
notify : EventId \rightarrowtail\!\!\!\!\!\!+\ Callback \\
constraints : EventId \rightarrowtail\!\!\!\!\!\!+\ \operatorname{seq} Constraint \\
matchMode : EventId \rightarrowtail\!\!\!\!\!\!+\ AndOr \\
eventType : EventId \rightarrowtail\!\!\!\!\!\!+\ String \\
\hline
registered = \operatorname{dom} notify = \operatorname{dom} constraints = \operatorname{dom} matchMode = \operatorname{dom} eventType
\end{array}
$$

The (internal) state of an event handler consists of the set of event id's representing registration, together with the information associated with each registration. This consists (in order) of the object that registered (and should therefore be notified), the constraints on the event data required to trigger notification, whether all or any constraints must be met, and the type of event on which notification is to occur.

The operations of registering and unregistering interest in an event are quite straightforward.

register

$$\Delta(registered, notify, constraints, matchMode, eventType)$$
$$eventType? : String$$
$$constraints? : \text{seq } Constraint$$
$$fullConstraintMatchMode? : AndOr$$
$$objectRef? : Callback$$
$$id! : EventId$$

$$id! \notin registered$$
$$registered' = registered \cup \{id!\}$$
$$eventType' = eventType \oplus \{id! \mapsto eventType?\}$$
$$constraints' = constraints \oplus \{id! \mapsto constraints?\}$$
$$matchMode' = matchMode \oplus \{id! \mapsto fullConstraintMatchMode?\}$$
$$notify' = notify \oplus \{id! \mapsto objectRef?\}$$

unregister

$$\Delta(registered, notify, constraints, matchMode, eventType)$$
$$id? : EventId$$

$id? \in registered \longrightarrow$ exc InvalidEventId

$$registered' = registered \setminus \{id?\}$$
$$eventType' = \{id?\} \lhd eventType$$
$$constraints' = \{id?\} \lhd constraints$$
$$matchMode' = \{id?\} \lhd matchMode$$
$$notify' = \{id?\} \lhd notify$$

InvalidEventId $\longrightarrow registered' = registered \land eventType' = eventType$
$constraints' = constraints \land matchMode' = matchMode$
$notify' = notify$

Dispatching an event to the event handler invokes the callback operation of all objects that have registered interest in the event and for which the associated constraint is satisfied by the event instance.

dispatchEvent

$$newEvent? : Event$$

$$\forall\, e : registered \bullet$$
$$eventName(e) = newEvent?.eventName$$
$$matchMode(e) = AND \Rightarrow \forall\, c : \text{ran } constraints(e) \bullet newEvent?.eventData\ sat\ c$$
$$matchMode(e) = OR \Rightarrow \exists\, c : \text{ran } constraints(e) \bullet newEvent?.eventData\ sat\ c$$
$$\Rightarrow$$
$$notify(e).callback\ [newEvent?/callbackValue?]$$

Technical Note: The universal quantifier in the predicate part of this operation effectively suggests that the notifications are performed in parallel. In practice, notification may well be sequential, and as a result of execution time and propagation delays (particularly for remote objects), race conditions are a possibility. For example, a callback to one object may trigger processing that sends an 'unregister' request back to the event handler for an object that has not (yet) been notified of the current event. It should also be noted that the description of this and other operations in which an operation is invoked on some object reference relies critically on the semantics of object identity developed recently for Object-Z.

The callback routine of an event handler is just the dispatch event operation, with an appropriate renaming of the input variables.

$$callback \mathrel{\widehat=} dispatchEvent\ [callbackValue?/newEvent?]$$

6 The ‘SynchronizationPoint’ Object Type

A synchronization point is an event handler, specialised so that a set of objects can be associated with each event name. Subclasses of this type can specialise the behaviour of event dispatching so that conditions of this set are checked before the operation associated with the event is invoked. One such specialization is defined later in this section. For brevity, we introduce a type name to represent the key-value pairs that are used as the data component of events:

$$EventData == \mathrm{seq}(Key \times Value)$$

The base type for synchronization points is introduced below.

SynchronizationPoint

A Synchronization point is a kind of event handler, but we redefine the dispatchEvent operation as the default behaviour - invoking the operation associated with the event - is no longer legitimate.

$EventHandler$ **redef** $(dispatchEvent)$

The state of an object contains a register of source objects that are synchronizing via an event. This variable relates ‘signals’, defined as the combination of an event name and event data, with references to the objects that have registered as sources for this event. For example, if $en : String, ed : EventData$ and $o : objref$, then $(en, ed) \mapsto o \in sources$ means that the object o has registered an interest on synchronizing on those en events that carry ed data.

$$events : \mathbb{F}\, Event$$
$$sources : (String \times EventData) \leftrightarrow EnhancedPREMOObject$$

$$\mathrm{dom}\, sources = \{e : events \bullet e.eventName \times e.eventData$$

Initially, the registries of sources and events are empty.

initialise

$$events = \varnothing$$

A new synchronization event can be defined by passing an event containing the event name and data of interest to a synchronization point. The name and data components of the event are used to identify the kind of event being registered, while the $eventSource$ field of the event represents an object that can notify the handler that the event has occurred; notifications from unregistered sources are ignored. In the Standard, attempting to add an event twice results in an exception being generated.

addSyncEvent

$$\Delta(sources, events)$$
$$syncEvent? : Event$$

$$syncEvent? \notin events \quad \longrightarrow \quad \boxed{\text{exc}}\ \text{RepeatedEvent}$$

$$\textbf{let} \left[\begin{array}{l} signal == (syncEvent?.eventName, syncEvent?.eventData) \\ source == syncEvent?.eventSource \end{array} \right]$$
$$\bullet$$
$$events' = events \cup \{syncEvent?\}$$
$$sources' = sources \cup \{signal \mapsto source\}$$
$$\text{RepeatedEvent} \longrightarrow sources' = sources \wedge events' = events$$

An event can be removed from the set of registered events; attempting to delete a non-existent event raises an exception.

$deleteSyncPoint$

$\Delta(sources, received, events)$
$syncEvent? : Event$

$syncEvent? \in events \longrightarrow$ exc UnknownEvent

let $\left[\begin{array}{l} signal == (syncEvent?.eventName, syncEvent?.eventData) \\ source == syncEvent?.eventSource \end{array} \right]$
•
$events' = events \setminus \{syncEvent?\}$
$sources' = sources \setminus \{signal \mapsto source\}$
UnknownEvent $\longrightarrow$ $sources' = sources \wedge received' = received$

The dispatch operation, at this point in the object type hierarchy, extends the corresponding operation inherited from $EventHandler$ by checking the validity of the input event, and raising an exception if this event has not previously been registered.

$dispachEvent$

$newEvent? : Event$

$newEvent? \in events \longrightarrow$ exc UnknownEvent

UnknownEvent $\longrightarrow$ $received' = received$

7 The 'ANDSynchronizationPoint' Object Type

A particular form of synchronization, representing a common need in multimedia applications, is defined by a subtype of $SynchronizationPoint$ called $ANDSynchronizationPoint$. Instances of this object type wait until all objects that have registered as event sources have signalled the event, before the object then invokes the callback operation on objects that have registered interest in being notified on the event. The description of this behaviour in the Standard is somewhat unclear, for example on the relationship between the registration facilities defined in $EventHandler$ and the $addSyncEvent$ operation inherited from $SynchronizationPoint$. To model the behaviour of this object type, the state is extended with a variable that records event notifications.

$ANDSynchronizationPoint$

$SynchronizationPoint$ **redef** $(dispatchEvent)$

The Standard indicates all of the operations from $SynchronizationPoint$ are redefined in this object type. However, from a specification viewpoint, only the behaviour of the $dispatchEvent$ operation need be redefined explicitly; other operations can either be extended, or in the case of $addSyncElement$, inherited without any change.

$received : (String \times EventData) \leftrightarrow EnhancedPREMOObject$

$received \subseteq sources$

The invariant indicates that any notification of an event must correspond to an event that has been registered.

$initialise$

$received = \varnothing$

The $deleteSyncEvent$ operation can remove an event from the set of registered events, and in this case any partial notifications (i.e. cases where some, but not all, of the sources have notified the object of this event) must be

cleared. Note that this operation extends the description of $deleteSyncEvent$ from the $SynchronizationPoint$ object type.

$$
\begin{array}{l}
\textbf{deleteSyncEvent} \\
\Delta(received) \\
syncEvent? : Event \\
\hline
\textbf{let } signal == (syncEvent?.eventName, syncEvent?.eventData) \bullet \\
received' = \{signal\} \lhd received
\end{array}
$$

The ‘dispatch event’ operation in this type checks whether the object that signalled the latest event completes the set of objects that have registered interest in that event. If so, the $received$ register is reset so that the objects can again synchronize on the event, and the $callback$ action is invoked on objects that have registered interest in this event, subject to the same filtering mechanism as described in the version of this operation defined in $EventHandler$. Note again that this aspect of the behaviour of this object type is not clear from the Standard. If all required objects have not yet signalled the event, the latest event notification is added to the $received$ register.

$$
\begin{array}{l}
\textbf{dispatchEvent} \\
\Delta(received) \\
newEvent? : Event \\
\hline
newEvent? \in events \quad \longrightarrow \quad \boxed{\text{exc}}\ \text{UnknownEvent} \\
\hline
\textbf{let} \left[\begin{array}{l} signal == (newEvent?.eventName, newEvent?.eventData) \\ source == newEvent?.eventSource \end{array} \right] \\
\bullet \\
received(\!|\ \{signal\}\ |\!) \cup \{source\} = sources(\!|\ \{signal\}\ |\!) \Rightarrow \\
\qquad received' = \{signal\} \lhd received \\
\qquad EventHandler :: dispatchEvent \\
\qquad \forall\, e : registered \bullet \\
\qquad\quad eventName(e) = newEvent?.eventName \\
\qquad\quad matchMode(e) = AND \Rightarrow \\
\qquad\qquad \forall\, c : \mathrm{ran}\ constraints(e) \bullet newEvent?.eventData\ sat\ c \\
\qquad\quad matchMode(e) = OR \Rightarrow \\
\qquad\qquad \exists\, c : \mathrm{ran}\ constraints(e) \bullet newEvent?.eventData\ sat\ c \\
\qquad\quad \Rightarrow \\
\qquad\quad notify(e).callback\ [newEvent?/callbackValue?] \\
received(\!|\ \{signal\}\ |\!) \cup source \neq sources(\!|\ \{signal\}\ |\!) \Rightarrow \\
\qquad received' = received \cup \{signal \mapsto source\} \\
\text{UnknownEvent} \ \longrightarrow \ received' = received
\end{array}
$$

8 Conclusion

A substantial part of this document was written at the same time as the material that was entered into the Working Draft of the PREMO document. The result was that a number of issues, both minor and non-trivial, were detected before they became embedded into a Standards document and therefore the subject of formal commenting. In this way, the use of formal specification has been of significant help in supporting the development of the PREMO Standard. It was fortunate in that at the time the Working Draft was produced, both the editor of the relevant part of the PREMO document (PREMO Part 2) and the first author of this paper were at the same institution, and it was possible to develop both documents in parallel, with the two authors sitting in a room and exchange comments directly. Following further meetings of the PREMO Rapporteur Group, the Time and Synchronization aspects of PREMO underwent several changes, and the original specification became inconsistent with the base document. In preparing this paper, the two have been realigned, though in the process we

have discovered a significant number of issues and questions concerning the technical content of PREMO that will need to be resolved within the Rapporteur Group. In summary, the value of a document such as this is twofold. First, it contains a precise statement of the expected behaviour of a portion of the PREMO standard. Second, the process of writing the document has, we believe, contributed to improving the overall quality of that Standard.

With respect to technical content of the specification, we have found that Object-Z provides a good starting point for building a description of PREMO in a way that mirrors the structure of the ISO document. This is important, since not all members of the Rapporteur Group are experts in formal specification or Object-Z, and the close mapping simplifies the task of explaining consequences and issues identified from the formal model in terms of the material as presented in the normative document. The process of mapping the PREMO object types into Object-Z classes is not however straightforward. In an earlier paper [5] we reported on issues related to differences in the object models of PREMO and Object-Z. The concerns identified here are wide ranging, but a general theme has been that the state-operation style of Object-Z can necessitate encoding aspects of PREMO object type behaviour in a rather operational style. This is perhaps best illustrated by the model of internal progression within the $Synchronizable$ object type.

The concerns raised in this paper are not criticisms of Object-Z. There is a tension in specification language design between providing an expressive language while maintaining a simple underlying semantic model. Languages that attempt to handle all aspects of systems, for example concurrency, synchronisation, real time, and error handling are likely to become difficult to understand and to use. We believe that progress in taming the intellectual complexity of systems like PREMO is likely to come, not from new and more complex specification logics, but from developing approaches that support the integration of partial specifications that capture particular aspects of a system in an appropriate representation. A basis for relating these representations may then be found by examining the underlying mathematical structures. The idea of formal specification, is after all, to utilise the simplicity, elegance and richness of mathematics to understand the behaviour of systems such as PREMO. Specification languages are simply one means to this end.

Acknowledgements

The work presented in this paper was carried out under the auspices of the ERCIM Computer Graphics Network, funded under the CEC Human Capital and Mobility Programme (contract CHRX-CT93-0085). The authors are grateful to Dr Mieke Massink for valuable comments on the work reported here.

References

[1] D.B. Arnold and D.A. Duce. *ISO Standards for Computer Graphics: The First Generation*. Butterworths, 1990.

[2] J.F. Koegel Buford. Architecture and issues for distributed multimedia systems. In *Multimedia Systems*. ACM Press/Addison Wesley, 1994.

[3] D.A. Carrington, D.J. Duke, R.W. Duke, P. King, G.A. Rose, and G. Smith. Object-Z: An object-oriented extension to Z. In S. Vuong, editor, *Formal Description Techniques (FORTE'89)*. North Holland, 1990.

[4] R.B. Dannenberg, T. Neuendorffer, J. Newcomer, D. Rubine, and D. Anderson. Tactus: Tolkit-level support for synchronized interactive multimedia. *Multimedia Systems Jounral*, 1(2):77–86, 1993.

[5] D.A. Duce, D.J. Duke, P.J.W. ten Hagen, I. Herman, and G.J. Reynolds. Formal methods in the development of premo. *Computer Standards and Interfaces*, 17(5-6):491–509, 1995. Special Issue on Formal Methods and Standards.

[6] D.A. Duce, D.J. Duke, P.J.W. ten Hagen, and G.J. Reynolds. Premo - an initial approach to a formal definition. *Computer Graphics Forum*, 13(3), 1994. Conference Issue: Proc. Eurographics'94, Oslo, Norway.

[7] R. Duke, P. King, G. Rose, and G. Smith. The Object-Z specification language: Version 1. Technical Report 91–1, Software Verification Research Centre, University of Queensland, 1991.

[8] R. Duke, G. Rose, and G. Smith. Object-Z: A specification language advocated for the description of standards. *Computer Standards and Interfaces*, 17(5-6):511–533, 1995. Special Issue on Formal Methods and Standards.

[9] D.J. Duke (editor). Time and synchronisation in PREMO: A formal specification of the NNI proposal. Technical Report OME-116, ISO/IEC JTC1 SC24/WG6, 1995. ftp://ftp.cwi.nl/pub/Premo.

[10] A. Freeman and D. Ince. *Active Java: Object-Oriented Programming for the World Wide Web*. Addison-Wesley, 1996.

[11] S.J. Gibbs and D.C. Tsichritzis. *Multimedia Programming*. ACM Press/Addison-Wesley, 1995.

[12] I.J. Hayes. *Specification Case Studies*. Series in Computer Science. Prentice Hall International, second edition, 1992.

[13] I. Herman, G.S. Carson, J. Davy, P.J.W. ten Hagen, D.A. Duce, W.T. Hewitt, K. Kansy, B.J. Lurvey, R. Puk, G.J. Reynolds, and H. Stenzel. Premo: An ISO standard for a presentation environment for multimedia objects. In D. Ferrari, editor, *Proceedings of the Second ACM International Conference on Multimedia (MM'94)*. ACM Press, 1994.

[14] I. Herman, G.J. Reynolds, and J. Van Loo. PREMO: An emerging standard for multimedia presentation. *IEEE Multimedia*, 1996. To appear.

[15] Geneva ISO Central Secretariat. Information processing systems, information technology, coding of multimedia and hypermedia information (mheg). ISO/IEC 13522-1, 1994.

[16] Geneva ISO Central Secretariat. Information processing systems, computer graphics, presentation environment for multimedia objects (PREMO). ISO/IEC 14478, 1996. http://www.cwi.nl/Premo/docs.html.

[17] Geneva ISO Central Secretariat. Information processing systems, computer graphics, virtual reality modelling language (VRML). ISO/IEC 14722, 1996.

[18] N.G. Leveson, M.P.E. Heimdahl, H. Hildreth, and J.D. Reese. Requirements specification for process control systems. *IEEE Transactions on Software Engineering*, 20(9):684–707, 1994.

[19] G.J. Reynolds, D.A. Duce, and D.J. Duke. Report of the ISO/IEC JTC1/SC24 special rapporteur group on formal description techniques. Technical Report ISO/IEC JTC1/SC24 N1152, ISO, 1994.

[20] J.M. Spivey. *The Z Notation: A Reference Manual*. Prentice Hall International, second edition, 1992.

A Exception Handling in a Specification

The formal specifications of PREMO Object types published to date have not included explicit descriptions of exceptions. The PREMO Standard does define an explicit model of exceptions, and indicates the circumstances in which particular exceptions should be raised. A preliminary approach to modelling exceptions within Object-Z specifications appeared in an internal report by members of the PREMO Rapporteur Group [9]. The approach taken in this paper is a development of that first attempt.

We introduce the following type to define the exception values that are discussed in the formal specification. The list given below is not complete, but it should be clear that it could be enumerated if required.

$$
\begin{array}{lll}
Exception ::= & \mathrm{Okay} & \text{– Used in the specification to indicate no exception} \\
\quad | & \mathrm{WrongState} & \\
\quad | & \mathrm{WrongValue} & \\
\quad | & \mathrm{EndPosition} & \\
\quad | & \ldots & \text{– to be completed}
\end{array}
$$

A common problem with dealing with exceptions in any representation is that the details of exceptional cases, and the corresponding actions, can easily obscure the normative behaviour of an operation. However, one of the values of a formal specification is that it helps to identify possible causes for failure, so it is highly desirable to document both normal behaviour and the behaviour that results when an operation is invoked inappropriately. To overcome the problem of detail, we introduce notation to hide some of the detail. For example, the full specification of $setSyncElement$ operation from the class $Synchronizable$ appears below. What is of concern here is the amount of material needed to describe the two exceptions.

$setSyncElement$

$$\Delta(refpoints)$$
$$refpoint? : C$$
$$syncData? : SyncElement$$
$$exceptions! : Exception$$

$$currentState \in \{PAUSED, STOPPED\}$$
$$refpoint? \in minimumPosition \ldots maximumPosition$$
$$refpoints' = refpoints \oplus \{refpoint? \mapsto syncData?\}$$
$$exceptions! = Okay$$
$$\vee$$
$$currentState \notin \{PAUSED, STOPPED\} \wedge exceptions! = WrongState$$
$$\quad \wedge\ refpoints' = refpoints$$
$$\vee$$
$$refpoint? \notin minimumPosition \ldots maximumPosition$$
$$exceptions! = WrongValue \wedge refpoints' = refpoints$$

Another specification of the same operation is given below, this time using some conventions to structure the description of exceptions and related behaviour.

$setSyncElement$

$$\Delta(refpoints)$$
$$refpoint? : C$$
$$syncData? : SyncElement$$

$$currentState \in \{PAUSED, STOPPED\} \qquad \longrightarrow \boxed{\text{exc}}\ \text{WrongState}$$
$$refpoint? \in minimumPosition \ldots maximumPosition \qquad \longrightarrow \boxed{\text{exc}}\ \text{WrongValue}$$

$$refpoints' = refpoints \oplus \{refpoint? \mapsto syncData?\}$$
$$\text{WrongState} \vee \text{WrongValue} \longrightarrow refpoints' = refpoints$$

The points to note are that:

1. Preconditions whose failure should raise an exception are gathered into a special section of the operation, and each such condition is tagged by the name of the exception (e.g. 'WrongState') that should be raised if that precondition is not satisfied when the operation is invoked.

2. We assume that all exceptions are returned through a variable $exceptions!$ of type $Exception$, which is implicitly declared in any operation that defines exceptions.

3. The 'normal processing' of an operation is written as usual, with the return of an 'okay' value through $exceptions!$ being implicit.

4. A gap is used to separate the 'normal' behaviour of an operation from predicates that describe required behaviour in exceptional cases. In this section of the operation, exception names are related to the predicate that should hold in the operation if that exception is raised.

Nordisk samarbeidsorgan for vitenskapelig informasjon
Tieteellisen informaation pohjoismainen yhteistyöelin
Norræna samvinnunefndin um vísindalegar upplýsingar
Nordisk samarbejdsorgan for videnskabelig information
Nordiska samarbetsorganet för vetenskaplig information
The Nordic Council for Scientific Information

- News
- Presentation
- Secretariat
- Board of directors
- Projects
- Publications
- Conferences and meetings
- Nordic and European links

- NORDINFO NYTT
- NORDGUIDE
- NordELIB
- Nordic Web Index

NORDINFO
c/o Helsinki University of Technology library
Otnäsvägen 9, FIN-02150 Esbo, Finland
Tel. +358-9-4552633
Fax +358-9-4552576
Email nordinfo@hut.fi

Page last updated May 12 1997

```html
<HTML>
<HEAD>
<TITLE>Links for ACM DL'97 Tutorial</TITLE>
</HEAD>
<BODY>
<H1>Links for ACM DL'97 Tutorial</H1>
<HR>
<H3>Museums and Images</H3>
<UL>

<LI><A HREF="http://www.ahip.getty.edu/mesl/">
 The Museum Educational Site Licensing Project</A> and
 <A HREF="MESL.htm">local cache</A>

<LI><A
HREF="http://jefferson.village.virginia.edu/uvamesl/content_selection/index.html">
 MESL Content Selection Cumulative Content Summary</A> and
 <A HREF="MESLcont.htm">local cache</A>

<LI><A HREF="http://www.io.org/~jtrant/papers/mesl.dpc.illinois.html">
 New Models for Distributing Digital Content,
 A paper for the Data Processing Clinic, University of Illinois,
 March 25, 1996, J. Trant, Getty Art History Information Program</A> and
 <A HREF="MESLpres.htm">local cache</A>

<LI><A HREF="http://www.io.org/~jtrant/papers/jt.eva.florence.9602.html">
 Enabling Educational Use of Museum Digital Materials:
   The Museum Educational Site Licensing (MESL) Project,
                              by J. Trant,
                 Getty Art History Information Program
 </A> and
 <A HREF="MESL9602.htm">local cache</A>

<LI><A HREF="http://www.io.org/~jtrant/Overview.html">
 Jennifer Trant, esp. Recent Papers and Publications</A> and
 <A HREF="Trant.htm">local cache</A>

<LI><A HREF="http://www.ahip.getty.edu/intro_imaging/0-Cover.html">
 Introduction to Imaging,
 Issues in Constructing an Image Database,
 Howard Besser and Jennifer Trant</A> and
 <A HREF="IntrImg0.htm">local cache</A>

<LI><A HREF="http://www.ahip.getty.edu/intro_imaging/Tbl.html">
 Introduction to Imaging,
 Issues in Constructing an Image Database: Contents,
 Howard Besser and Jennifer Trant</A> and
 <A HREF="IntrImg1.htm">local cache</A>

<LI><A HREF="http://www.ahip.getty.edu/mesl/resources.html">
 MESL Related Resources</A> and
 <A HREF="MESLres.htm">local cache</A>

<LI><A HREF="http://www.cni.org/pub/CIMI/www/framework.html">
```

Standards Framework for the Computer Interchange of
Museum Information</A> and
<A HREF="CIMI.htm">local cache</A>

<LI><A HREF="http://www-personal.umich.edu/~jweise/quality/NetworkedImagesTOC.HTML">
Preparing Quality Images for Computer Networks</A> and
<A HREF="Prep.htm">local cache</A>

<LI><A HREF="http://sunsite.berkeley.edu/Imaging/Databases/">
Image and Multimedia Database Resources (Besser)</A> and
<A HREF="BessRes.htm">local cache</A>
</UL>

<H3>Singapore</H3>
<UL>

<LI><A HREF="http://www.iss.nus.sg/">
Institute of Systems Science, NUS, Singapore</A> and
<A HREF="ISS.htm">local cache</A>

<LI><A HREF="http://stoat.iss.nus.sg:8080/">
ISS Visual Information Processing Group</A> and
<A HREF="ISSVIP.htm">local cache</A>

<LI><A HREF="http://stoat.iss.nus.sg:8080/project.html">
ISS VIPG projects:
Robust Face Recognition System,
Robust Fingerprint Recognition System,
Web-based Multimedia Spatial Information System with query language,
FACEit - a content-based face image archival and retrieval system,
STAR - System for Trademark Archival and Retrieval,
Video Processing facilitates manipulation, indexing and retrieval
of MPEG encoded video,
Personalized Information Network delivers personalized multimedia news
</A> and
<A HREF="ISSproj.htm">local cache</A>

</UL>

<H3>NSF/ARPA/NASA DLI</H3>
<UL>
<LI><A HREF="http://dli.grainger.uiuc.edu/national.htm">
DIGITAL LIBRARIES INITIATIVE,
Funded through a joint initiative of
NSF/DARPA/NASA</A> and
<A HREF="DLI.htm">local cache</A>

<LI><A HREF="http://alexandria.sdc.ucsb.edu/">
The Alexandria Project</A> and
<A HREF="ADLhome.htm">local cache</A>

<LI><A HREF="http://alexandria.sdc.ucsb.edu:3366/doc/tutorial/index.html">
Alexandria Project Tutorial Table of Contents</A> and
<A HREF="ADLtutTC.htm">local cache</A>

<LI><A HREF="http://alexandria.sdc.ucsb.edu/public-documents/annual-report/">
Alexandria Digital Library Annual Report Table of Contents</A> and
<A HREF="ADLrep.htm">local cache</A>

<LI><A HREF="http://alexandria.sdc.ucsb.edu/public-documents/papers/adr2/">
The Second Alexandria Design Review,
Report on the Workshop,
19-21 February, 1997,
Santa Barbara, CA</A> and
<A HREF="ADLrep2.htm">local cache</A>

<LI><A HREF="http://www.dlib.org/dlib/august95/alexandria/08manjunath.html">
Image Browsing
in the Alexandria Digital Library (ADL) Project </A> and
<A HREF="ADLimage.htm">local cache</A>

<LI><A HREF="http://alexandria.sdc.ucsb.edu/other-sites/">
Universe: Spatial Data Available Over the Web</A> and
<A HREF="ADLdata.htm">local cache</A>

<LI><A HREF="http://www.dlib.org/dlib/july96/07wactlar.html">
Informedia Digital Video Library
Technology Outreach</A> and
<A HREF="CMUout.htm">local cache</A>

<LI><A HREF="http://www.dlib.org/dlib/september96/informedia/09christel.html">
Informedia Goes to School:
Early Findings from the Digital Video Library Project </A> and
<A HREF="CMUsch.htm">local cache</A>

<LI><A HREF="http://www.informedia.cs.cmu.edu/">
CMU's Informedia Digital
Video Library</A> and
<A HREF="CMUhome.htm">local cache</A>

<LI><A HREF="http://www.ius.cs.cmu.edu/demos/edge.html">
Edge Detector Demo</A> and
<A HREF="CMUedge.htm">local cache</A>

<LI><A HREF="http://www.ius.cs.cmu.edu/demos/facedemo.html">
CMU's Face Detector Demo</A> and
<A HREF="CMUface.htm">local cache</A>

<LI><A HREF="http://www.cs.cmu.edu/afs/cs.cmu.edu/project/cil/www/v-demos.html">
Computer Vision Demos - worldwide</A> and
<A HREF="CMUimage.htm">local cache</A>

<LI><A HREF="http://www.ius.cs.cmu.edu/idb/">
CMU Image Database - data sets</A> and
<A HREF="CMUimgDB.htm">local cache</A>

<LI><A HREF="http://www.informedia.cs.cmu.edu/html/description.html">
Informedia: News-on-Demand</A> and

<A HREF="CMUnews1.htm">local cache</A>

<LI><A HREF="http://www.eecs.uic.edu/~yaslando/dbis.html">
U. Illinois Chicago Database and Information Systems Laboratory</A> and
<A HREF="YuDISL.htm">local cache</A>

<LI><A HREF="http://www.eecs.uic.edu/research/disl.htm">
U. Illinois Chicago Database and Information Systems Laboratory</A> and
<A HREF="YuDISL2.htm">local cache</A>

</UL>

<H3>Search</H3>
<UL>

<LI><A HREF="http://wwwqbic.almaden.ibm.com/~qbic/qbic.html">
QBIC(TM) -- IBM's Query By Image Content</A> and
<A HREF="qbic.htm">local cache</A>

<LI><A HREF="http://www.virage.com/online/">
Virage Inc. VIR Technology Demo - find similar images</A> and
<A HREF="virage.htm">local cache</A>

<LI><A HREF="http://www.virage.com/wpaper/">
Visual Information Retrieval: A Virage Perspective</A> and
<A HREF="http://www.virage.com/wpaper/vir_engi.pdf">PDF version</A> and
<A HREF="VIRpaper.htm">local cache</A> and
<A HREF="../pdf/vir_engi.pdf">local PDF</A>

<LI><A HREF="http://www.dlib.org/dlib/february97/columbia/02chang.html">
Finding Images/Video in Large Archives:
Columbia's Content-Based Visual Query Project</A> and
<A HREF="Columbia.htm">local cache</A>

<LI><A HREF="http://www.itnm.columbia.edu/webseek/">
Webseek: A Content-Based Image and Video Search and Catalog Tool
for the Web </A> and
<A HREF="WebSeek.htm">local cache</A>

<LI><A HREF="http://www.uky.edu/~kiernan/DL/brent.html">
Content-based Multimedia Data Management and Efficient Remote Access,
James Griffioen, Brent Seales, Raj Yavatkar</A> and
<A HREF="CBRbrent.htm">local cache</A>

<LI><A HREF="http://www-ercim.inria.fr/www-ercim.inria.fr/publication/technical_reports/035-abstract.html">
Storage and Retrieval Techniques for Multimedia Data
by P. Zezula and C. Galindo-Legaria (IEI-CNR)</A> and
<A HREF="ftp://ftp.inria.fr/associations/ERCIM/research_reports/pdf/0495R035.pdf">PDF format</A> and
<A HREF="ERCIMsr.htm">local cache</A> and
<A HREF="../pdf/0495R035.pdf">local PDF</A>

</UL>

<H3>Compression</H3>
<UL>

<LI><A HREF="http://www.mat.sbg.ac.at/~uhl/wavelet.biblio">
BIBLIOGRAPHY OF WAVELET AND WAVELET-RELATED DOCUMENTS;
see also:
Subject: Wavelet Digest, Vol. 4, Nr. 1.
From: wavelet@math.scarolina.edu (Wavelet Digest)
</A> and
<A HREF="wavelet.txt">local cache</A>

<LI><A HREF="http://www.met.inpe.br/pub-html/wavelet/wa-project.html">
WAVELET PROJECT figures</A> and
<A HREF="BRwave.htm">local cache</A>

<LI><A HREF="http://inls.ucsd.edu/y/Fractals/frac-an.html">
Fractal Image Encoding Announcements and Questions</A> and
<A HREF="fractal.htm">local cache</A>

<LI><A HREF="http://inls3.ucsd.edu/y/Fractals/">
Fractal Image Encoding - resources</A> and
<A HREF="fractalr.htm">local cache</A>

<LI><A HREF="http://www.uky.edu/~kiernan/DL/bsmith.html">
A Survey of Compressed Domain Processing Techniques,
Brian C. Smith</A> and
<A HREF="CDbsmith.htm">local cache</A>

</UL>

<H3>Misc.</H3>
<UL>
<LI><A HREF="http://www.dlib.org/dlib/january97/oclc/01weibel.html">
Image Description on the Internet:
A Summary of the CNI/OCLC Image Metadata Workshop,
September 24 - 25, 1996,
Dublin, Ohio</A> and
<A HREF="DublinIm.htm">local cache</A>

<LI><A HREF="http://www.uky.edu/~kiernan/DL/symp.html">
Reconnecting Science and Humanities in Digital Libraries,
A Symposium Sponsored by
The University of Kentucky and The British Library,
19-21 October 1995,
Marriott's Resort at Griffin Gate, Lexington, KY</A> and
<A HREF="UKYsymp.htm">local cache</A>

<LI><A HREF="http://www-ercim.inria.fr/www-ercim.inria.fr/publication/technical_reports/049-abstract.html">
ERCIM Computer Graphics Network Report and Recommendations from the
VRML Workshop,
29/30 January 1997,The Coseners House, Abingdon, UK</A> and
<A HREF="ERCIMvr.htm">local cache</A>

<LI><A HREF="http://www-ercim.inria.fr/www-ercim.inria.fr/publication/technical_reports/048-abstract.html">
Specifying the PREMO Synchronization Objects,
D.J. Duke, D.A.Duce, I. Herman, G.Faconti</A> and
<A HREF="ftp://ftp.inria.fr/associations/ERCIM/research_reports/pdf/0297R048.pdf">PDF version</A> and
<A HREF="ERCIMpr.htm">local cache</A> and
<A HREF="../pdf/PREMO.pdf">local PDF</A>

<LI><A HREF="http://www.hut.fi/NORDINFO/english/index.html">
The Nordic Council for Scientific Information</A> and
<A HREF="NORDINFO.htm">local cache</A>

<LI><A HREF="http://www.macromedia.com/software/director/index.html">
Macromedia Director</A> and
<A HREF="Director.htm">local cache</A>

</UL>

<H3>Education</H3>
<UL>
<LI><A HREF="http://www.cs.cornell.edu/Info/Faculty/bsmith/mmsyl.htm">
Multimedia Educational Materials</A> and
<A HREF="mmsyl.htm">local cache</A>

<LI><A HREF="http://www.si.umich.edu/cristaled/">
Coalition
on
Reinventing
Information
Science,
Technology
and
Library
Education</A> and
<A HREF="CRISTAL.htm">local cache</A>

<LI><A HREF="http://www.si.umich.edu/cristaled/curriculum/core.html">
CRISTAL-ED Core Curriculum</A> and
<A HREF="UMcore.htm">local cache</A>

<LI><A HREF="http://www.si.umich.edu/cristaled/specializations.html">
CRISTAL-ED Specializations</A> and
<A HREF="UMspec.htm">local cache</A>

<LI><A HREF="http://www.si.umich.edu/hp/institutes/dtk/tools.htm">
U Mich. courses on Digital Tools</A> and
<A HREF="UMtools.htm">local cache</A>

<LI><A HREF="http://www.sims.berkeley.edu/impact/Winter95/HTML/class.html">
Impact of New Information Resources:
Multimedia and Networks,
Winter 1995 Homepage (Besser)</A> and

<A HREF="Impact.htm">local cache</A>

<LI><A HREF="http://www.sims.berkeley.edu/impact/Winter95/HTML/studentessays.html#cdrom">
Evaluation:
Besser Course - Student Essays</A> and
<A HREF="Essays.htm#cdrom">local cache</A>

</UL>

</BODY>
</HTML>

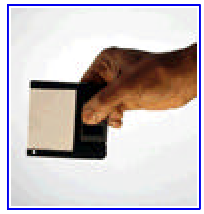
CRISTAL-ED

Coalition on Reinventing Information Science, Technology and Library Education

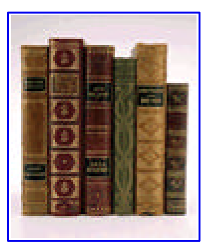
A collaborative project of the School of Information at the University of Michigan with the

Background and General Information: Charting a New Direction; Objectives; Timeline; Participants

Planning for a New Curriculum: Curriculum Planning; Prototyping



Linking Beyond the Boundaries of SI: Electronic Mail Discussion; Distance-Independent Learning; Invitational Conference; Pilot Projects



Kellogg CRISTAL-ED

support of the W.K. Kellogg Foundation HRISM initiative

Courses for the Core Curriculum and Specializations; Educational Approaches; Task Force Reports

[Interim Report and Original Proposal:](#) Text of the first Interim Report to the W.K. Kellogg Foundation, and the full proposal of the CRISTAL-ED project as submitted in 1993

[Reaching Us:](#) Names and addresses of Kellogg CRISTAL-ED

staff at SI

Core Curriculum

Planning

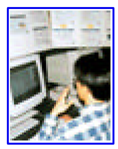
- The Mission Statement of the "new school"
- Status Reports
  - It's official! The new curriculum of the School of Information
  - Planning for a new curriculum
  - The Enhanced School
    - Faculty
    - School name
    - Project structure
- Report of the first CRISTAL-ED planning retreat

Prototyping

- SI 523/529/530: "Core Knowledge"
- SI 523: "Organization of Information Resources"
- SI 526: "Technologies for Information Management"
- SI 529: "Sources of General Information"
- SI 609: "Foundations"

Technical Education

- Background: Technology and the humanities: courses for students and faculty
  - Humanities workshop
  - Technology workshops
  - Introductory workshops
  - Computer hardware and software availability
  - DIAD Classroom/Laboratory
- Background: Management in the SI curriculum

Home Discussion

CRISTAL-ED

Specializations

Planning

The following documents detail the planning that went into the new curriculum of the School of Information, and therefore, may not reflect the current details of the curriculum. On February 26, 1997, the new curriculum received final approval from the Horace H. Rackham School of Graduate Studies.

- Introduction
- Archives and Records Management
- Digital Documents/Digital Publishing
  - Background: Digital Documents Curriculum Report
- Human-Computer Interaction
  - Background: Human-Computer Interaction Report
- New Systems Architecture
  - Background: New Systems Architecture Report
- Library and Information Services

    - Background: Library and Information Services Report
- Information Systems Management
- Health Sciences Informatics

Prototyping

- SI 603: "Image Databases"
- SI 604: "Copyright for Librarians and Information Science Professionals"
    - A list of course resources is available
- SI 605: "The Making of Digital Libraries"
- SI 605: "Ethics and Values Dilemmas in Use of Information Technology: Policy Implications for Institutions of Higher Learning
- SI 606: "Future Visions of the Information Age"
- SI 606: "Internet: Resource Discovery, Organization, and Design"
- SI 606 and 607: "Structured Information: Principles and Enabling Technology" and "Structured Information: Principles and Applications"
- SI 607: "Entrepreneurship in the Information Industry"
- SI 608: "Archives Administration of Electronic Records"
- SI 608: "Medical Informatics -- Introduction to Health Sciences Informatics: Theory and Practice"

- SI 609: "Foundations"
- SI 609: "Impact of New Information Resources: Multimedia and Networks"
- SI 611: "Future Opportunities in Educational Systems for Information Professionals"
- SI 613: "Human Computer Interaction Software Projects Lab"
- SI 614: "Object Oriented Design and Programming"
- SI 614: "Visual Communications: System, Interfaces and Applications"
- SI 725: "Community Networking"
- SI 726: "Information Technology, Impacts and Implications"
  - Internet Public Library, an outcome of this course
- SI 888: "Information Visualization"

Pedagogy

- Project-based Learning
- Distance-independent learning
  - Outreach and Distance-independent Learning
  - The Concept of Distance-independent Learning
  - Distance-independent Learning at SI
  - SI 529: "Sources of General Information"
  - SI 609: "Impact of New Information Resources: Multimedia and Networks"

    - M-Vision
    - Vision 2010
- Laboratory facilities
    - Technology workshops
    - Media and Advanced Projects Laboratory (and software availability)
    - DIAD Classroom/Laboratory
- Pilot projects
  Pilot projects give students real opportunities to reinforce, master, and apply skills and knowledge learned in the classroom to solving problems in real-world settings, and addressing the specific needs of practitioners required to lead their communities in the digital environment. Although pilot-project development began in the project's first year, several news projects were initiated this year.

    - Kellogg Community Networking Initiative
        - Background: Community Networking Initiative Report
        - Greater Flint Community Networking Initiative
        - Dedication of the Flint Community Networking Initiative Training Center
        - Media coverage of the Flint project

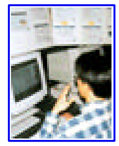
    - Lapeer Community Schools Internet Training Seed Program

            - Life and Works of Newbery Award-Winner Marguerite DeAngeli

        - Internet Public Library
            - Background: Internet Public Library Report
            - Web site: Internet Public Library
            - IPL in action: Teaching Web skills to librarians
            - Funding: IPL receives grant
            - Collaboration: IPL/HTI join forces

        - Cultural Heritage Initiative for Community Outreach
            - Web site: Cultural Heritage Initiative for Community Outreach

- Student professional training experiences
    - Assistantships
    - University of Michigan Digital Library Project
    - Digital Information Associates
    - The Directed Field Experience Program
    - University Library Associates
    - Head Librarian Program of the Residence Hall Libraries at the University of Michigan
    - Residence Hall Libraries at the University of Michigan
    - SI Art Image Project

Home Discussion

Digital Tools

Select from the following courses; fee for each is $295/$250 (EB). "EB" denotes Early Bird rate. See Fee Information and Policies for details.

Mastering Basic HTML Skills
This hands-on workshop for the Web neophyte covers all the basics of authoring documents in the Hypertext Markup Language (HTML) used by the World Wide Web. Focus on the basic elements necessary to create a simple, functional Web page, learn general structure and tagging, build a troubleshooting repertoire, and review basic Web authoring tools.

(TK03) June 4-6
(TK06) Aug 18-20

Kristen Garlock, M.I.L.S., and Sherry Piontek, M.I.L.S./Coordinators, JSTOR User Services, University of Michigan

---

Advanced Web Development
Master advanced HTML tools, including tables, frames, forms, and image maps; specialized graphical software tools (Photoshop and Illustrator); CGI scripting to create guestbooks; the newest Web innovations like VRML, JavaScript, ActiveX, Shockwave, RealAudio, and others. Completion of "Mastering Basic HTML Skills" or basic HTML experience recommended.

(TK01) May 12-14
(TK07) Aug 20-22

Craig Strickland, B.S./Computer Systems Consultant, University of Michigan Media Union

---

Introduction to Java and Internet Programming

Geared for the non-programmer, this course covers fundamental programming concepts (primitive variables, classes, objects, and control structures) as well as concepts behind event-driven programming, Java Applets, image and audio content, GUI widgets, animation, and other types of interactivity. Plenty of hands-on practice. No prior programming experience required.

(TK02) May 14-16

Craig Strickland, B.S./Computer Systems Consultant, University of Michigan Media Union

---

Introduction to SGML

The problem: databases, spreadsheets, images, and other types of information are

created in a growing variety of incompatible formats! The answer: SGML, or Standard Generalized Markup Language, which embeds standardizing tags within documents so they can be read, combined, retrieved, and managed regardless of the hardware or software being used. Learn how to mark up documents and how to build a DTD (document type definition). Also discuss software that takes advantage of the SGML difference.

(TK08) Aug 25-27

Craig Strickland, B.S./Computer Systems Consultant, University of Michigan Media Union

---

Introduction to MacroMedia Director
Learn the fundamentals of multimedia authoring with Macromedia Director. Create a presentation enhanced by visual aids, including slides, interactivity, and simple

animation sequences. Basic computing skills helpful.

(TK04) June 18-20

Nathan Eriksen, B.A./Administrative Manager, University of Michigan School of Information

---

Collaborative Tools at Work

Take an in-depth look at the organizational aspects of Lotus Notes and other shared technology tools. Hands-on practice with Lotus Notes and discussions of organizational issues provide insight into the process, benefits, and barriers of using and implementing Lotus Notes in your organization.

(TK05) July 21-23

Lisa Covi, Ph.D./Research Fellow, University of Michigan Collaboratory for Research on

Electronic Work

Home Search Feedback Help