

Digital Video Archives: Managing Through Metadata

*Howard D. Wactlar and Michael G. Christel
Computer Science Department
Carnegie Mellon University*

Executive Summary

As analog video collections are digitized and new video is created in digital form, computer users will have unprecedented access to video material—getting what they need, when they need it, wherever they happen to be. Such a vision assumes that video can be adequately stored and distributed with appropriate rights management, as well as indexed to facilitate effective information retrieval. The latter point is the focus of this paper: how can metadata be produced and associated with video archives to unlock their contents for end users?

Video that is "born digital" will have increasing amounts of descriptive information automatically created during the production process, e.g., digital cameras that record the time and place of each captured shot, and tagging video streams with terms and conditions of use. Such metadata could be augmented with higher-order descriptors, e.g., details about actions, topics, or events. These descriptors could be produced automatically through ex-post-facto analysis of the aural and visual contents in the video data stream. Likewise, video that was originally produced with little metadata beyond a title and producer could be automatically analyzed to fill out additional metadata fields to better support subsequent information retrieval from video archives.

As digital video archives grow, both through the increasing volume of new digital video productions and the conversion of the analog audiovisual record, the need for metadata similarly increases. Automatic analysis of video in support of content-based retrieval will become a necessary step in managing the archive; a recent editorial by the director of the European Broadcasting Union Technical Department notes that "Efficient exploitation of broadcasters' archives will increasingly depend on accurate metadata" (Laven 2000). He offers the challenge of finding an aerial shot of the Sydney Harbour Bridge at sunset. Given a small collection of Sydney videos, such a task is perhaps tractable, but as the volume of video grows, so does the importance of better metadata and supporting indexing and content-based retrieval strategies.

Digital library research has produced some insights into automatic indexing and retrieval. For example, it has found that narrative can be extracted through speech recognition; that speech and image processing can complement each other; that metadata need not be precise to be useful; and that summarization strategies lead to faster identification of the relevant information. The purpose of this chapter is to discuss these findings. Particular emphasis is placed on the Informedia Project at Carnegie Mellon University and the new National Institute of Standards and Technology Text Retrieval Conference (NIST TREC) Video Retrieval Track, which is investigating content-based retrieval from digital video.

Introduction

We are faced with a great opportunity as analog video resources are digitized and new video is produced digitally from the outset. The video itself, once encoded as bits, can be copied without loss in quality and distributed cheaply and broadly over the ever-growing communication channels set up for facilitating transfer of computer data. The great opportunity is that these video bits can be described digitally as well, so that producers' identities and rights can be tracked and consumers' information needs can be efficiently, effectively addressed. The "bits about bits" (Negroponte 1995), referred to as "metadata" throughout this paper, allow digital video assets to be simultaneously protected and accessed. Without metadata, a thousand-hour digital video archive is reduced to a terabyte or greater jumble of bits; with metadata, those thousand hours can become a valuable information resource.

Metadata for video are crucial when one considers the huge volume of bits within digital video representations. When digitizing an analog signal for video, the signal needs to be sampled a number of times per second, and those samples quantized into numeric values that can then be represented as bits. Only with infinite sampling and quantization could the digital representation exactly reproduce the analog signal. However, human physiology provides some upper bounds on differences that can actually be distinguished. For example, the human eye can typically differentiate at most 16 million colors, and so representing color with 24 bits provides as much color resolution as is needed for the human viewer. Similar visual physiological factors on critical viewing distance and persistence of vision establish other guidelines on pixel resolution per image and images per second playback rate. For a given screen size and viewer distance, 640 pixels per line and 480 lines per image provide adequate resolution, with 30 images per second resulting in no visible flicker or break in motion. Digital video at these rates requires $640 \times 480 \times 30 \times (24 \text{ bits per pixel}) = 221$ megabits per second, or 100 gigabytes per hour. The number of bits increases if higher resolution (such as high-density TV [HDTV] resolution of 1920 by 1080) is desired (for example, to allow for larger displays viewed

at closer distances without distinguishing the individual pixels). Hence, even a single hour of video can result in 100 gigabytes of data. Associating metadata with the video makes these gigabytes of data more manageable.

Numerous strategies exist to reduce the number of bits required for digital video, from relaxed resolution requirements to lossy compression in which some information is sacrificed in order to reduce significantly the number of bits used to encode the video. Motion Picture Experts Group-1 (MPEG-1) and MPEG-2 are two such lossy compression formats; MPEG-2 allows higher resolution than MPEG-1 does. Because preservationists want to maintain the highest-quality representation of artifacts in their archives, they are predisposed against lossy compression. However, the only way to fit more than a few seconds of HDTV video onto a CD-ROM is through lossy compression. The introduction to scanning by the Preservation Resources Division of OCLC Online Computer Library Center, Inc., reflects this tension between quality and accessibility:

Although traditional preservation methods have ensured the longevity of endangered research materials, it has sometimes been at the cost of reduced access. With digital technology, images are used to reproduce rare items, allowing for virtually universal copying, distribution, and access. The technology also makes it possible to bring collections of disparate holdings together in digital form, making resource sharing more feasible (OCLC 1998).

Hence, for long-term preservation, digital video presents a number of challenges. What should the sampling and quantization rates be? What compression strategies should be used—lossy or lossless? What media should be used to store the resulting digital files—optical (such as digital video disc [DVD]) or magnetic? What is the shelf life for such media, i.e., how often should the digital records be transferred to new media? What are the environmental factors for long-term media storage? What decompression software needs to exist for subsequent extraction of video recordings? These challenges are not discussed further here, as they warrant their own separate treatments. Regardless of how these challenges are addressed, digital video has huge size, but also huge potential, for facilitating access to video archive material.

Digital technology has the potential to improve access to research material, allowing access to precisely the content sought by an end user. This implies full content search and retrieval, so that users can get to precisely the page they are interested in for text, or precisely the sound or video clip for audio or video productions. Creating such metadata by hand is prohibitively expensive and inappropriate for digital video, where much of the metadata is a by-product of the way in which the artifact is generated. Current

research will extend the automated techniques for contemporaneous metadata creation.

To realize this potential, video must be described so that its production attributes are preserved and so users can navigate to the content meeting their needs. Video has a temporal aspect, in which its contents are revealed over time, i.e., it is isochronal. Finding a nugget of information within an hour of video could take a user an hour of viewing time. Delivering this hour of video over the Internet, or perhaps over wireless networks to a personal digital assistant (PDA) user, would require the transfer of megabytes or gigabytes of data. Isochronal media are therefore expensive both in terms of network bandwidth as well as user attention. If, however, metadata enabled surrogates to be produced or extracted that either were nonisochronal or significantly shorter in duration, then both bandwidth and the user's attention could be used more efficiently. After checking the surrogate, the user could decide whether access to the video was really necessary. A surrogate can also pinpoint the region of interest within a large video file or video archive.

As video archives grow, metadata become increasingly important: "In spite of the fact that users have increasing access to these [digitized multimedia information] resources, identifying and managing them efficiently is becoming more difficult, because of the sheer volume" (Martinez 2001). The capability of metadata to enrich video archives has not been overlooked by research communities and industry. For example, a number of workshops addressed this topic as part of digital asset management (DAM) (USC 2000). Artesia Technologies (Artesia 2001) and Bulldog (Bulldog 2001) are two corporations offering DAM products. Digital asset management refers to the improved storage, tracking, and retrieval of digital assets in general. Our focus here is on digital video in particular, beginning with a discussion of relevant metadata standards and leading to the automatic creation of video metadata and implications for the future.

Metadata for Digital Video

As noted in a working group report on preservation metadata (OCLC 2001), metadata for digital information objects, including video, can be assigned to one of three categories (Wendler 1999):

1. *Descriptive*: facilitating resource identification and exploration
2. *Administrative*: supporting resource management within a collection
3. *Structural*: binding together the components of more complex information objects

The same working group report continues that of these categories,

"descriptive metadata for electronic resources has received the most attention—most notably through the Dublin Core metadata initiative" (OCLC 2001, 2). This paper likewise will emphasize descriptive metadata, while acknowledging the importance of the other categories, as descriptive metadata can be automatically derived in the future for added value to the archive. Further details on administrative and structural metadata are available in the 2001 OCLC white paper and its references.

Various communities involved in the production, distribution, and use of video have addressed the need for metadata to supplement and describe video archives. Librarians are very concerned about interoperability and having standardized access to descriptors for archives. Producers and content rights owners are greatly interested in intellectual property rights (IPR) management and in compliance with regulations concerning content ratings and access controls. The World Wide Web Consortium (W3C) produces recommendations on XML, XPath, XML-Schema, and related efforts for metadata formatting and semantics. Special interest groups such as trainers and educators have specific needs within particular domains, e. g., tagging video by curriculum or grade level. This section outlines a few key standardization efforts affecting metadata for video.

Dublin Core

The Dublin Core Metadata Initiative provides a 15-element set for describing a wide range of resources. While the Dublin Core "favors document-like objects (because traditional text resources are fairly well understood)" (Hillman 2001), it has been tested against moving-image resources and found to be generally adequate (Green 1997). The Dublin Core is also extensible, and has been used as the basis for other metadata frameworks, such as an ongoing effort to develop interoperable metadata for learning, education, and training, which could then describe the resources available in libraries such as the Digital Library for Earth System Education (DLESE) (Ginger 2000). Hence, Dublin Core is an ideal candidate for a high-level (i.e., very general) metadata scheme for video archives. An outside library service, with likely support for Dublin Core, would then be able to make use of information drawn from video archives expressed in the Dublin Core element set.

Video Production Standardization Efforts

Professional video producers are interested in tagging data with IPR, production and talent credits, and other information commonly found in film or television credits. In addition, metadata descriptors from the basic Dublin Core set are too general to adequately describe the complexity of a video. For example, one of the Dublin Core elements is the instantiation

date (Hillman 2001), but for a video, date can refer to copyright date, first broadcast date, last broadcast date, allowable broadcast period, date of production, or the setting date for the subject matter.

Producers are especially interested in defining metadata standards because video production is becoming a digital process, with new equipment such as digital cameras supporting the capture of metadata such as date, time, and location at recording time. The Society of Motion Picture and Television Engineers (SMPTE) has been working on a universal preservation format for videos, the SMPTE Metadata Dictionary (SMPTE 2000). For born-digital material, many of the metadata elements can be filled in during the media creation process.

The SMPTE Metadata Dictionary has slots for time and place, further resolved into elements such as time of production and time of setting, place of production and place setting, where place is described both in terms of country codes and place names as well as through latitude and longitude. The SMPTE effort is often cited by other video metadata efforts as a comprehensive complement to the minimalist Dublin Core element set.

In 1999, the European Broadcasting Union (EBU) launched a two-year project named "EBU Project P/Meta" designed to develop a common approach to standardizing and exchanging program-related information and embedded metadata throughout the production and distribution life cycle of audiovisual material. According to 1999 press releases, the project began by identifying and standardizing the information commonly exchanged between broadcasters and content providers, using the BBC's Standard Media Exchange Framework (SMEF) as the reference model. They then were to assess the feasibility of applying new SMPTE metadata standards within Europe to support the agreed exchange framework, and move toward implementation.

The TV Anytime Forum is an association of organizations that seeks to develop specifications to enable audiovisual and other services based on mass-market, high-volume digital storage.

MPEG-7 and MPEG-21

A number of professional industry and consortia standardization efforts are in progress to provide more detailed video descriptors. The new member of the MPEG family, Multimedia Content Description Interface, or MPEG-7, aims at providing standardized core technologies allowing description of audiovisual data content in multimedia environments. It will extend the limited capabilities of proprietary solutions in identifying content that exist today, notably by including more data types. An overview of MPEG-7 by

Martinez (2001) acknowledges the diversity of standardization efforts and notes the purpose of MPEG-7:

MPEG-7 addresses many different applications in many different environments, which means that it needs to provide a flexible and extensible framework for describing audiovisual data. Therefore, MPEG-7 does not define a monolithic system for content description but rather a set of methods and tools for the different viewpoints of the description of audiovisual content. Having this in mind, MPEG-7 is designed to take into account all the viewpoints under consideration by other leading standards such as, among others, SMPTE Metadata Dictionary, Dublin Core, EBU P/Meta, and TV Anytime. These standardization activities are focused to more specific applications or application domains, whilst MPEG-7 tries to be as generic as possible. MPEG-7 uses also XML Schema as the language of choice for the textual representation of content description and for allowing extensibility of description tools. Considering the popularity of XML, usage of it will facilitate interoperability in the future.

Because the descriptive features must be meaningful in the context of the application, they will be different for different user domains and different applications. This implies that the same material may be described using different types of features, tuned to the area of application. To take the example of visual material, a lower abstraction level would be a description of shape, size, texture, color, movement (trajectory), and position (where in the scene can the object be found?). For audio, a description at this level would include key, mood, tempo, tempo changes, and point of origin. The highest level would give semantic information, e.g., "This is a scene with a barking brown dog on the left and a blue ball that falls down on the right, with the sound of passing cars in the background." Intermediate levels of abstraction may also exist.

The level of abstraction is related to the way in which the features can be extracted: many low-level features can be extracted in fully automatic ways, whereas high-level features need human interaction.

Next to having a continuous description of the content, it is also required to include other types of information about the multimedia data. It is important to note that these metadata may also relate to the entire production, segments of it (e.g., as defined by time codes), or single frames. This enables granularity that can describe a single scene's action, limit that scene's redistribution because of its source, or classify that scene as inappropriate for child viewing because of its content.

- *Form*: An example of the form is the coding scheme used (e.g., Joint Photographic Experts Group [JPEG], MPEG-2), or the overall data size. This information helps in determining whether the material can be "read" by the user.
- *Conditions for accessing the material*: This includes links to a registry with IPR information, including such entries as owners, agents, permitted usage domains, distribution restrictions, and price.
- *Classification*: This includes parental rating and content classification into a number of predefined categories.
- *Links to other relevant material*: The information may help the user speed the search.
- *The context*: In the case of recorded nonfiction content, it is important to know the occasion of the recording (e.g., the final of 200-meter men's hurdles in the 1996 Olympic Games).

In many cases, it will be desirable to use textual information for the descriptions. Care will be taken, however, that the usefulness of the descriptions is as independent from the language area as is possible. A clear example where text comes in handy is in giving names of authors, films, and places.

Therefore, MPEG-7 description tools will allow a user to create, at will, descriptions (that is, a set of instantiated description schemes and their corresponding descriptors) of content that may include the following:

- information describing the creation and production processes of the content (director, title, short feature movie)
- information related to the usage of the content (copyright pointers, usage history, broadcast schedule)
- information about the storage features of the content (storage format, encoding)
- structural information on spatial, temporal, or spatio-temporal components of the content (scene cuts, segmentation in regions, region motion tracking)
- information about low-level features in the content (colors, textures, timbres, melody description)
- conceptual information of the reality captured by the content (objects and events, interactions among objects)
- information about how to browse the content in an efficient way (summaries, variations, spatial and frequency subbands)
- information about collections of objects
- information about the interaction of the user with the content (user preferences, usage history)

There is room for domain specialization within the metadata architectures, whether by audience and function (education vs. entertainment), genre

(documentary, travelogue), or content (news vs. lecture), but there is also a risk of overspecificity. Because the technology continues to evolve, MPEG-7 is intended to be flexible.

The scope of MPEG-21 could be described as the integration of the critical technologies enabling transparent and augmented use of multimedia resources across a wide range of networks and devices to support functions such as content creation, content production, content distribution, content consumption and usage, content packaging, intellectual property management and protection, content identification and description, financial management, user privacy, terminals and network resource abstraction, content representation, and event reporting.

Standards for Web-Based Metadata Distribution

The W3C is a vendor-neutral forum of more than 500 member organizations from around the world set up to promote the World Wide Web's evolution and ensure its interoperability through common protocols. It develops specifications that must be formally approved by members via a W3C recommendation track. These specifications may be found on the W3C Web site.

A number of key W3C recommendations, published in 1999 and referenced below, enabled the separation of authoring from presentation in a standardized manner. For video archives, these recommendations allow the separation of video metadata from the library interface and from the underlying source material. This enables the interface to be customized for the particular application or audience (adult entertainment vs. secondary school education) and to the communication medium or device specifications (desktop PC vs. PDA), even though the same underlying data will be accessible to each use. The W3C recommendations useful for accessing, integrating, exploring, and transferring digital video metadata through the Web and Web browsers include the following:

- XML (Extensible Markup Language): the universal format for structured documents and data on the Web, W3C Recommendation February 1998 (<http://www.w3.org/XML/>)
- XML Schema: express shared vocabularies for defining the semantics of XML documents, W3C Recommendation as of May 2001 (<http://www.w3.org/XML/Schema>)
- XSLT (XSL Transformations): a language for transforming XML documents, W3C Recommendation November 1999 (<http://www.w3.org/TR/xslt>)
- XPath (XML Path Language): a language for addressing parts of an XML document, used by XSLT, W3C Recommendation November 1999 (<http://www.w3.org/TR/xpath.html>)

Case Study: Informedia

The Informedia Project at Carnegie Mellon University pioneered the use of speech recognition, image processing, and natural language understanding to automatically produce metadata for video libraries (Wactlar et al. 1999). The integration of these techniques provided for efficient navigation to points of interest within the video. For example, speech recognition and alignment allows the user to jump to points in the video where a specific term is mentioned, as illustrated in figure 1.



Fig. 1. Effects of seeking directly to a match point on "Lunar Rover," courtesy of tight transcript to video alignment provided by automatic speech processing

The benefit of automatic metadata generation is that it can perform a post-facto analysis for video archives that were produced in analog form and later digitized. Such archives will not have the benefit of a rich set of metadata captured from digital cameras and other sources during a digital production process. The speech, vision, and language processing are imperfect, so the drawback of automatic metadata generation, compared with hand-edited tagging of data, is the introduction of error in the descriptors. However, prior work has shown that even metadata with errors

can be very useful for information retrieval, and that integration across modalities can mitigate errors produced during the metadata generation (Witbrock and Hauptmann 1997; Wactlar et al. 1999).

More complex analysis to extract named entities from transcripts and to use those entities to produce time and location metadata can lead to exploratory interfaces and allow users to directly manipulate visual filters and explore the archive dynamically, discovering patterns and identifying regions worth closer investigation. For example, using dynamic sliders on date and relevance following an "air crash" query shows that crashes in early 2000 occurred in the African region, with crash stories discussing Egypt occurring later in that year, as shown in figure 2.

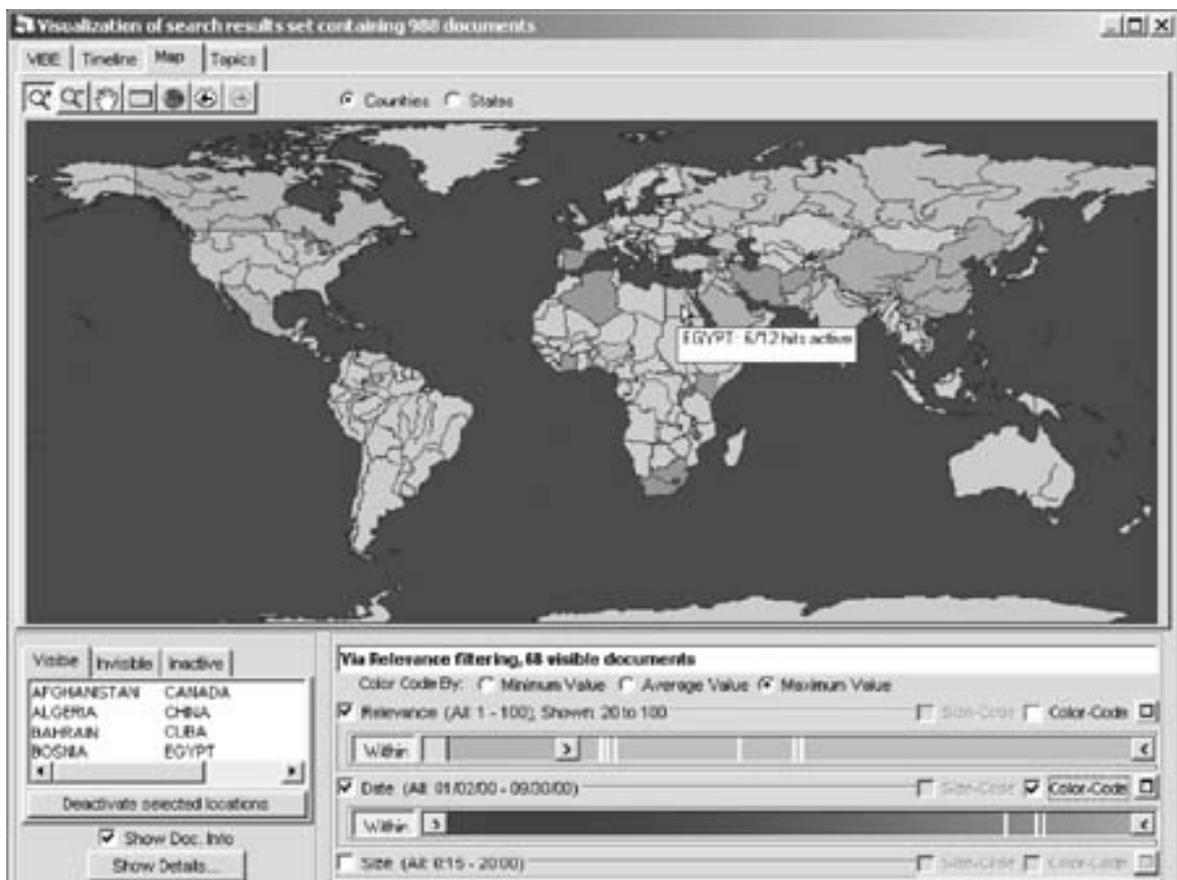


Fig. 2. Map visualization for results of "air crash" query, with dynamic query sliders for control and feedback

The goal of the CMU Infromedia-II Project is to automatically produce summaries derived from metadata across a number of relevant videos, i.e., an "autodocumentary" or "autocollage," and thereby facilitate more efficient information access. This goal is illustrated in figure 3, where visual cues can be provided to allow navigation into "El Niño effects" and quick discovery that forest fires occurred in Indonesia and that such fires corresponded to a time of political upheaval. Such interfaces make use of

metadata at various grain sizes. For example, descriptions of video stories can produce a story cluster of interest, with descriptions of shots within stories leading to identification of the best shots to represent a story cluster, and descriptions of individual images within shots leading to a selection of the best images to represent the cluster within collages such as those shown in figure 3.

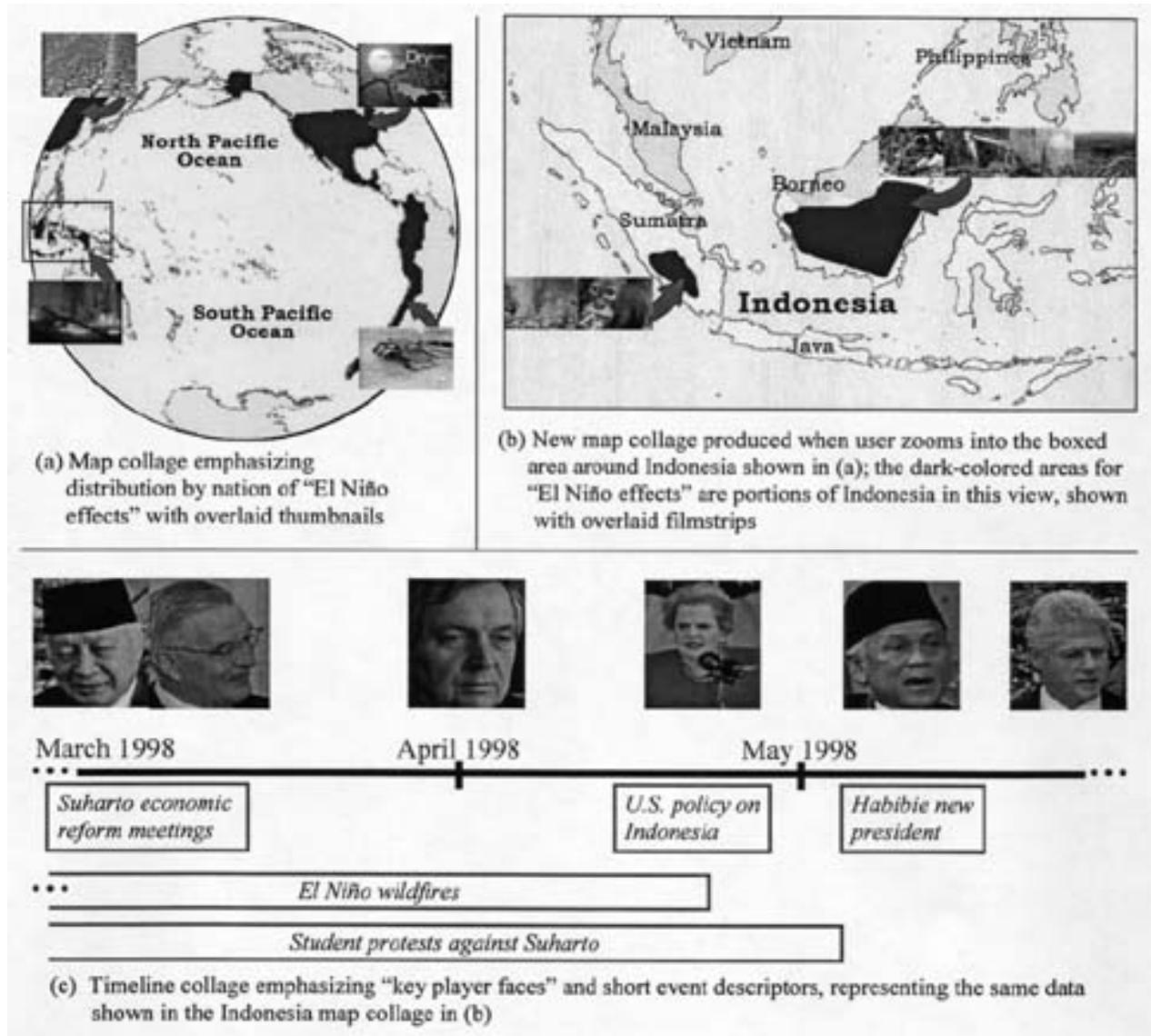


Fig. 3. Prototype of Informedia-II collage summaries built from video metadata

Preserving Digital Data

Librarians and archivists have priorities that go beyond the agenda of content access, distribution, and payment systems for consumers and producers. Archivists and preservationists are vested with selecting a medium that will survive the longest and a system that will transcend the

most generations of "player" hardware and software. Content that will be created digitally has both advantages and disadvantages over conventional analog film and video content. The National Film Preservation Board (NFPB) serves as a public advisory group to the Library of Congress (LC). Led by William J. Murphy, the LC produced a comprehensive report in 1997 that reviews the various facets of television and video preservation and surveys the various elements relevant to retention of all digitally produced content (LC 1997).

Media longevity problems exist both for analog and for digital content. Magnetic tapes will lose signal strength and stretch on stored reels. There are no standardized systems or methodologies for evaluating the physical or data-loss effects of tape aging. Digital video discs can delaminate, and many compact discs (CDs) with inadequate protective layers may be vulnerable to the effects of temperature, humidity variation, and pollution in less than five years. Such degradation can render digital data unreadable. On the positive side, digital media can be created with data redundancy, error-detection, and even error-correcting codes that detect and compensate for dropped bits. These techniques have long been used in digital communication and storage systems. Furthermore, digital content can be inexpensively recorded, or cloned, without generational loss, providing cheap and practical physical redundancy (there is no single master copy). Data that are kept online in disc-based systems can have data loss minimized by redundant array of inexpensive discs (RAID) storage systems. Such systems can also continuously or periodically refresh their data, thus sustaining their integrity.

Perhaps of greater concern is the rapid obsolescence of digital media formats and encoding schemes as advancing technology out-modes recording and playback devices in time frames much shorter than the media life. For example, two digital recording formats, D-1 and D-2, have been available to the industry since the late 1980s. Early generations of Sony's D-1 and D-2 equipment are already obsolete in production environments. The last few years have seen the introduction of numerous new video formats such as D-5 (for studio production), D-6 (for HDTV), DCT, Digital Betacam, DV, DVC, and Digital-S. Some new recording equipment also digitizes directly into digitally compressed formats, MPEG-1 (VHS quality) and MPEG-2 (studio-to-HDTV quality). The emerging standard for MPEG-7 will also allow for embedded metadata generated contemporaneously or following production. What is required is a format-independent cloning solution that will enable the digital content to be transparently interchanged, regardless of storage system, media type, encoding format, or transport mechanism, and without loss of data quality and fidelity.

DAM systems can separate the indexing and cataloging information that

enable access from the underlying format of the medium. A database archive may be architecturally layered to render it medium-independent, thereby enabling access from one system to storage on another. This facilitates rapid and independent refreshing or conversion of the underlying data, data formats, and media. Modern systems should allow multiple types of archive storage media data banks to operate simultaneously through a common access interface. Thus, the lifetime of the metadata that index the content can far exceed that of the original media.

Conclusion

Content-based video retrieval is getting more attention as the volume of digital video grows dramatically. The Association for Computing Machinery (ACM) Multimedia Conference, started in 1994, has included a workshop dealing with multimedia information retrieval since 1999, and TREC started a new track on indexing and retrieval from digital video in 2001. TREC is an annual benchmarking exercise for information retrieval applications that has taken place at the National Institute for Standards and Technology for the last nine years (<http://trec.nist.gov>). TREC has been instrumental in fostering the development of effective information retrieval on large-scale corpus collections, and with the new digital video track signifies the emergence of digital video as an information resource.

These forums and others hosted by the Institute of Electrical and Electronics Engineers, Inc. (IEEE), the Audio Engineering Society, and other technical societies examine ways in which metadata can be generated for video through an automated analysis of the auditory and visual data streams. Evaluations are under way (for example, the TREC digital video track) to determine what metadata have value for identifying known items and exploring within a video archive. Metadata in the future should be more carefully tagged as to the confidence of the descriptor and producer to help the user direct the information search and exploration process. For an item known to be in the corpus, for example, the user might start by specifying that only metadata produced at the time the video was first recorded should be used. Another user exploring a topic may be willing to see all shots that might contain a face; an automated face detector returns a match in the shot but perhaps with low confidence. Through an appropriate interface, the user can quickly filter out those shots that truly contain faces from those that contain other images that only look like faces. Hence, along with an increased use of automatic metadata generators, these generators will also produce "metadata about the metadata," including production credits and confidence metrics. MPEG-7 recognizes the value of metadata and provides intellectual property protection for the descriptors themselves as well as for the video content.

Digital video will remain an expensive medium, in terms of broadcast/

download time and navigation/seeking time. Surrogates that can pinpoint the region of interest within a video will save the consumer time and make the archive more accessible and useful. Of even greater interest will be information-visualization schemes that collect metadata from numerous video clips and summarize those descriptors in a cohesive manner. The consumer can then view the summary, rather than play numerous clips with a high potential for redundant content and additional material not relevant to his or her specific information need. Metadata standards efforts discussed earlier can help with the implementation of such summaries across documents, allowing the semantics of the video metadata to be understood in support of comparing, contrasting, and organizing different video clips into one presentation.

Metadata will continue to document the rights of producers and access controls for consumers. Combined with electronic access, metadata enable remuneration for each viewing or performance down to the level of individual video segments or frames, rather than of distributions or broadcasts. Metadata can grow to include specific usage information; for example, which portions of the video are played, how often, and by what sorts of users in terms of age, sex, nationality, and other attributes. Of course, such usage data should respect a user's privacy and be controlled through optional inclusion and specific individual anonymity.

Metadata provide the window of access into a digital video archive. Without metadata, the archive could have the perfect storage strategy and would still be meaningless, because there would be no retrieval and hence no need to store the bits. With appropriate metadata, the archive becomes accessible. Furthermore, the window need not be fixed, i.e., the metadata should be capable of growing in richness through added descriptors for domain-specific needs of new user communities, unforeseen rights management strategies, or advances in automatic processing. By enhancing the metadata, the archive can remain fresh and current and accessible efficiently and effectively; there is no need to reformat or rehost the video contents to accommodate the metadata. Only the metadata are enhanced, which in turn enhances the value of the video archive.

References

Artesia Technologies. 2001. What Is Digital Asset Management (DAM)? Available at http://www.artesiotech.com/what_dam.html.

Bulldog. 2001. Welcome to Bulldog. Available at: <http://www.bulldog.com/>

[view.cfm](#).

Bormans, J., and K. Hill, eds. 2001. *MPEG-21 Overview*. ISO/IEC JTC1/SC29/WG11/N4318 (July). Available at: <http://www.csel.it/mpeg/standards/mpeg-21/mpeg-21.htm>.

Ginger, K. Web page maintainer. 2000. DLESE Metadata Working Group Homepage. (November 6). Available at: <http://www.dlese.org/Metadata/index.htm>.

Green, D. 1997. Beyond Word and Image: Networking Moving Images: More Than Just the "Movies." *D-Lib Magazine* (July-Aug.). Available at: <http://www.dlib.org/dlib/july97/07green.html>.

Hillman, D. 2001. Using Dublin Core, DCMI Recommendation (April 12). Available at <http://dublincore.org/documents/usageguide/>.

Laven, P. 2000. Confused by Metadata? *EBU Technical Review*, No. 284 (September). Available at http://www.ebu.ch/trev_home.html.

Li, F., et al. 2000. Browsing Digital Video. *CHI Letters: Human Factors in Computing systems, CHI 2000* 2(1) 169-176.

Library of Congress. 1997. *Television and Video Preservation: A Report of the Current State of American Television and Video Preservation*. vol. 1. Report of the Librarian of Congress (October). Edited by W. Murphy. Available at: <http://lcweb.loc.gov/film/tvstudy.html>.

Martinez, J. M., ed. 2001. Overview of the MPEG-7 Standard (Version 5.0). ISO/IEC JTC1/SC29/WG11 N4031 (March). Available at: <http://mpeg.telecomitalia.com/standards/mpeg-7/mpeg-7.htm>.

Negroponte, N. 1995. *Being Digital*. New York: Knopf.

National Institute of Standards and Technology Text Retrieval Conference. Video Retrieval Track. 2001. Available at: <http://www-nlpir.nist.gov/projects/t01v/>.

OCLC. 1998. Preservation Resources Digital Technology. Available at: <http://www.oclc.org/oclc/presres/scanning.htm>.

OCLC/RLG. 2001. Preservation Metadata Working Group Issues White

Paper, *Preservation Metadata for Digital Objects: A Review of the State of the Art* (January 31). Available at: http://www.oclc.org/digitalpreservation/presmeta_wp.pdf.

Society of Motion Picture and Television Engineers. 2000. SMPTE Metadata Dictionary RP210a, Trial Publication Document, Version 1.0 (July). Available at: <http://www.smpte-ra.org/mdd/Rp210a.pdf>.

University of Southern California, Annenberg Center for Communication. Digital Asset Management Conferences I, II, and III, 1998-2000. Available at: <http://dd.ec2.edu/>.

Wactlar, H., et al. 1999. Lessons Learned from the Creation and Deployment of a Terabyte Digital Video Library. *IEEE Computer* 32(2): 66-73. See also: <http://www.informedia.cs.cmu.edu/>.

Wendler, R. 1999. LDI Update: Metadata in the Library. *Library Notes*, no. 1286 (July/August): 4-5.

Witbrock, M. J., and A. G. Hauptmann. 1997. Using Words and Phonetic Strings for Efficient Information Retrieval from Imperfectly Transcribed Spoken Documents. In *Proceedings of the Association for Computing Machinery DL '97*. New York: Association for Computing Machinery.

Web sites noted:

World Wide Web Consortium. 2002. Available at <http://www.w3.org>

Informedia research at Carnegie Mellon University. 2002. Available at <http://www.informedia.cs.cmu.edu>

[Return to Start](#) [Previous](#)

[Return to CLIR Home Page >>](#)