A Rich Model for Moving Image Databases

James M Turner
École de bibliothéconomie et des sciences de l'information
Université de Montréal
james.turner@umontreal.ca

ABSTRACT

This paper is related to the conference theme of Technologies. It discusses how new theories of information organisation developed from the results of recent research into moving image organisation point to the need for better information systems to manage the rich information that is available but which is underused. Specifically, an information system is proposed which offers access using text, images or sound. XML is proposed as the coding language, in order to foster universal, multilingual systems.

RÉSUMÉ

Cette étude se rapporte au thème du colloque des Technologies. Elle discute comment les nouvelles théories de l'organisation d'information dévéloppées des résultats de la recherche récente de l'organisation des images animées montrent le besoin des meilleurs systèmes d'information pour gérer l'information riche qui est disponible mais qui est sous-utilisée. Expressément, un système d'information est proposé qui offre l'accès en utilisant du texte, des images et du son. XML est proposé comme le langage de code, pour encourager des systèmes universels et multilingues.

INTRODUCTION

Moving image databases represent a kind of last frontier in the automation of information systems. Because digitised moving images mean huge file sizes, even when the files are compressed, affordable information systems for managing moving images remain elusive, contrary to those available for managing text, still images and museum objects. Existing moving image database systems are expensive, and de-emphasize some of the information-rich aspects important to users, often making do with simple keyword access using vocabularies that often are not even controlled. Affordable moving image databases contain only the metadata, while the images themselves are stored offline. A viewing list can be generated by the system, but consulting the images is complicated and time-consuming. This is because each clip on the list can be stored on a different cassette, reel, or other support. In order to see the moving image, the support must be retrieved from a shelf or a vault and mounted on some viewing apparatus. Then the particular clip needs to be located on the support. Clips on a hard disk, a CD-ROM or a DVD-ROM can be accessed randomly, but clips on tape supports are ordered sequentially and it is necessary to wind to the desired sequence before the sequence can be viewed. Thus, in addition to the viewing time that moving images innately require, there is a great deal of overhead involved in getting the image into a position where it can be viewed.

BACKGROUND

This paper considers two principal aspects of the rich model proposed for moving-image databases: text-based metadata and user interfaces. The ideas for text-based metadata are based on the notion of "recycling" text created during the production processes by using it as the raw material for generating indexing to the moving image either automatically or semi-automatically. The ideas for the user interface are borrowed from existing editing systems and are based on the groundwork laid by Brian O'Connor in his article entitled "Access to moving image documents: background concepts and proposals for surrogates for film and video works" (1985).

Research over the last several years has revealed certain characteristics of text produced in the context of the production of moving images which strongly suggest that it can be used as a source for generating high-quality indexes to the finished product. Thus shot-by-shot indexing to news material, films, or historical archive material could be provided if we can learn how to generate this indexing automatically or semi-automatically. As with other kinds of documentation, user behaviours are varied. Users looking for material typically either name persons, objects or events that can be seen in a shot (Turner 1994), remember something that was said somewhere in a production, or perhaps have only a vague notion of what they seek. Furthermore, close correspondence between the names users put to objects in images and the names indexers put to the same images has been established (Turner 1995). Ongoing research projects in this area are studying the use of three main types of pre-existing text that are thought to be the richest sources for generating indexing for moving images: closed-captioning subtitles (Srihari 1995, 1997), audio description text (Turner 1998), and production scripts.

The automatic generation of cross-language indexing for moving images also seems promising (Turner 1996, Turner et Roulier 1999). Ongoing research in this area will confirm or deny prevously-obtained research results. Thus it is hoped that future systems can not only be multi-dimensional in their scope but also multilingual, permitting their use in a broad variety of contexts.

The user interface is of critical importance in system design. Because of the complexity of moving-image metadata and of that of multimedia objects in general, and because of the wide variety of user behaviours that are still not well understood, a multi-dimensional approach to the user interface for future systems is desirable. For some searches, the user will wish to use text strings. For others, a visual approach will be better. Since browsing is so useful in image databases, user interfaces should provide strong support for this activity. Queries in the form of sound also need to be considered because in some situations they are the only useful approach. The system proposed has multiple interfaces, and users can adopt any of them, depending on the particular search at hand. In addition, an overview or system map needs to be available for browsing and to show the parts of the system.

FILM, VIDEO OR MOVING IMAGES?

The distinction between film and video is already blurred by the co-existence of such technologies as video copies of films and productions shot on video then transferred to film. Films distributed in DVD format further blur the distinction between media. The term "video" is also widely used to describe digital moving images, such as a sequence in a multimedia product or a clip transferred from a consumer digital video camera to a computer. The organisation of institutions concerned with moving image documentation also reflects the supports , such as FIAF and FIAT (International Federation of Film Archives, International Federation of Television Archives) and in the U.S., the National Film Preservation Foundation and the initiative to create an equivalent for television, and the cinemathè ques. Other organisations such as AMIA (Association of Moving Image Archivists) and the INA (Institut national de l'audiovisuel) in France avoid distinguishing on the basis of the support, focussing their activities on moving images in general, whatever the support. The arrival of all-digital production of movies and television will eventually eliminate these distinctions once and for all, and in building information systems we will be concerned uniquely with digital moving images rather than with the supports on which these are stored. However, for many years to come, archives will need to distinguish between the various existing supports to accommodate the needs of the collections they hold.

FILE SIZES NEEDED FOR DIGITAL MOVING IMAGES

In the domain of moving images, the move from analogue to digital documents is not trivial. System designers have long recognized the advantage of having the moving image online with the metadata, but thus far they have had to be satisfied with representative frames from the moving image, or samples representing a few frames of the moving image. The few existing terabyte and petabyte systems which permit storing the image online are available only to those with a great deal of money, and the image needs to be compressed. These can be considered for distribution versions of moving images, but archiving is another matter because of the information loss when compression algorithms are used such as MPEG 2, an ISO standard widely used in the industry (e.g. for multimedia productions and films on DVD). Display resolutions are given as follows:

The DVD NTSC resolution is given as 720 x 480 and the DVD PAL/SECAM resolution as 720 x 576 (twentieth century commercial television) with an aspect ratio of 4:3 (Archive Builders 2001, 4). In building systems now, and especially those of an archival nature, we need to move to high definition television. The resolution is given as 1920 x 1200 (Sun Microsystems) and the aspect ratio is 16:9. However, because the MPEG2 standard requires the number of lines to be in multiples of 16, the actual resolution of transmitted HDTV streams will usually be 1920 x 1088 (Archive Builders 2001, 4). Of course, for archival purposes the full resolution will be required.

File sizes for digital multimedia formats for video and audio can be given as follows:

Color video compressed MPEG 2	1 hour	2	GB
Audio (answering machine, dictaphone)	1 hour	10	MB
Audio (CD quality)	1 hour	500	MB

With present technology as well as that of the foreseeable future, compression of moving images is unavoidable. However, the archived master copy should never be compressed. If the file sizes given above are impressive, the storage and processing requirements for uncompressed video are staggering.

Archival images require at least film quality. In the industry, figures for achieving this are generally given as 350 MB/frame, while some advocate a figure higher than 500 MB/frame (Heitmann 1998, 105). Calculating on the conservative figure, we get:

- 350 x 24 frames = 8.4 GB/second for film
- 350 x 30 frames = 10.5 GB/second for video

We need to use the video figures for our calculations, since they represent the digital norm widely used.

- 10.5 Go/sec x 60 seconds = 630 GB/minute
- =630 GB/minute x 60 minutes = 37.8 TB/hour

To allow a little overhead for the system-generated metadata, we calculate on 40 TB/hour for film-quality uncompressed moving images. This base figure is still much closer to the figure of 350 MB/frame than to that of 500+MB/frame. In addition, we don't yet have any of the high-level metadata in our calculation. Although its memory requirement is relatively very modest compared to that of the moving image file, we know there will be lots of it, as several layers have been identified as being necessary parts of the digital file. The need for it has been identified as an urgent area for development (Heitmann 1998, 105).

Keeping to conservative figures, let's review these orders of measurement in computing, using the commercial measurements (1000 Bytes = 1 KB) rather than the computing measurement (1024 Bytes = 1 KB):

1000 Bytes = 1 KiloByte (KB)
1000 KB = 1 MegaByte (MB)
1000 MB = 1 GigaByte (GB)

1000 GB = 1 TeraByte (TB)
1000 TB = 1 PetaByte (PB)
1000 PB = 1 ExaByte (EB)
1000 EB = 1 ZettaByte (ZB)
1000 ZB = 1 YottaByte (YB)

A YottaByte is a trillion TeraBytes. "To put this in some perspective, the entire contents of the Library of Congress would consume a mere 10 terabytes." And "Given good conditions(!), downloading a 1 yottabyte file over a 28.8 Kbps connection would take about 140 billion years." (Word Spy 2001). To further get an idea of this order of magnitude, "in astronomy, a single pixel can include an entire earth type planet (ten-thousand-kilometer pixels = 10 Megameter pixel), a sun type star (one-million-kilometer pixels = 1 Gigameter pixel), or a galaxy (one-hundred-thousand-light-year pixels = 1 Zettameter pixel). The largest practical pixel is a 400 Yottameter pixel, the diameter of the observable universe." (Archive Builders 2001, 5)

The figure given for the "entire contents of the Library of Congress" obviously refers to text files. If we consider even medium-sized moving image collections, the figures increase dramatically. If it were able to digitise its collection of Hearst newsreels alone, the UCLA Film and Television Archive would need 200 PetaBytes of space to store the 5000 hours of viewing time this collection contains. Similarly, the nearly 116 000 hours of video testimony held by Steven Spielberg's Shoah Foundation would need 4,64 ExaBytes of storage space, again just for the image without the high-level metadata. Housing the uncompressed digital masters of the large national archives of moving images or those of the large movie studios would certainly require ExaByte systems if not ZettaByte systems. The first PetaByte systems have only recently been installed in the moving image milieu.

The vast majority of systems used in collections today are a far cry from these expensive and largely theoretical systems. In addition, the few systems that do have the image available online have only relatively modest retrieval capacities because these systems do not take advantage of rich metadata available for providing automated indexing to the image. Clearly, archivists will have to wait for affordable storage technology before they can expect to digitise large quantities of moving images. While we are waiting, however, we can work on modelling databases for housing the moving image itself along with rich accompanying metadata. Such metadata being text-based, we can include it easily in the databases we build, while providing a logical place for the moving image in the structure once we have enough equipment to store it. We can also work on building metadata standards for description and indexing of moving images.

GENERAL MODEL AND OVERVIEW OF THE SYSTEM

The general model of the system encompasses three principal components: picture databases in institutions around the world, user approaches to searching these databases, and the metadata attached to the images which mediates between the user and the images sought. Since the system is Web-based and multilingual, translation layers are part of the system as well. These need to be elaborated in some detail because they are concerned not only with translation between languages but also between types of system information, image formats, metadata schemes, and so on.

Picture databases. Although the system as we imagine it can manage still and moving pictures, in this paper we consider only moving image databases. These in turn are restricted to those which catalogue and index moving images shot by shot, such as might be found in television news archives, in the historical collections of national archives, or in stockshot libraries. Shot- and scene-level descriptions (cataloguing) are of course related to information from the source production, but the metadata components that concern us here are the running descriptions (visual synopses, logs) which usually accompany this kind of material and information about the playing time, technical quality of the image, and so on.
The world in which this material can be used in the context of our system is one in which the material is organised with at least minimal adherence to some kind of metadata standard, such as a core description using some recognised public format or a proprietary format known to metadata shells and crosswalks. This is an area for development, but some groundwork has already been laid. As managers recognize the advantages of being able to communicate information about collections, this in turn will foster the move to adopting metadata standards.

User approaches. In representing the system from the user point of view, one of the first choices is a language to work in. Once this is determined, the system appears to the user as a unilingual system, although this is far from the reality behind the scenes. In addition, the user can choose text, image or sound as the avenue of approach for any query.

In formulating a query, users can use natural language to ask the system for pictures. In a first pass of a text query, the system refines the query by returning to the user the semantic network related to the input string. The user indicates the sense of the input text to the system, pinpointing the concept sought in the system. This important step clarifies what the user is seeking exactly (it is roughly analogous to an online reference interview), and if the semantic network is organised well enough, it guarantees the quality of the match. The translation layer then takes the concept pinpointed and sends it out on the Net in multiple languages, talking with the multilingual metadata layer which mediates between the user and the cataloguing and indexing data of the remote databases. The results of the query across these multiple databases are then gathered, ranked, and presented to users in the language in which they have chosen to work with the system.

Searches using pictures as an interface would take place in cases where users seek images of objects whose names they do not know. The visual dictionary serves as an intermediary by offering some context as an avenue of approach to information sought.

Users can browse in some general area, following leads that seem promising until they reach their target. The labels attached to the images provide the link with the metadata attached to images they are seeking. Thus users search through images to find words (whether they use these or not) which provide matches with other images.

In the case of sound searching, either words or music can be used. Words spoken lead to text matches and these in turn connect to image indexing metadata. Fragments of melodies used as a query look for matches in music located in the system.

Metadata. At the indexing end of the system, high-level metadata is created by humans or generated automatically or semi-automatically using the techniques we have described elsewhere. Each system can use its own language for the creation of this metadata. For example, the databases in Italy contain Italian-language metadata and those in Japan contain Japanese-language metadata. The indexing is filtered through the same Web-based multilingual concept dictionary resource that filters the user query, and it is here that matches are found and processed. The shots requested are gathered and presented to the user for viewing. Several views of the data are available, from text supplying a running description of the content of the shot to keyframes giving an indication of the visual aspects to low-quality moving images to high-quality moving images. Each view offers unique information to the user, and there are trade-offs in time and memory requirements between the various views. However, the system offers all the views, and users make the decisions about which views they wish to use in any given situation.

SOME OF THE COMPONENTS

The proposed system calls on a number of online communities for the use of interdisciplinary resources. The Web-based nature of the system means that ultimately it can be seen as a software shell with ever-changing content, or as an elaborate algorithm for searching multiple databases using a number of languages and employing varied interfaces. Individual tools that can be used as nodes of the system are developed independently, often for other purposes, and their development and upkeep is outside of our control. However, the particular tools called on by the system can be chosen for their stability and for authoritative sponsorship. A number of online communities are re-inventing the wheel with the development of what they call ontologies and what we in information studies call classification, such as the CYC ontology and WordNet (Soergel 1999). In addition, Euro WordNet (2001) provides semantic networks for a number of European languages. There is also a Global WordNet Association which "builds on the results of Princeton WordNet and Euro WordNet" (Global 2001) and coordinates worldwide efforts to build semantic networks for other languages. Because of efforts such as these, we hope to see tools that can perform well for our purposes.

In addition, a number of translation software programs are available on the Web. These are being evaluated for their usefulness in the multilingual indexing of moving image documents in the context of ongoing research projects at the École de bibliothéconomie et des sciences de l'information at the Université de Montréal. Although these tools are mostly rather simplistic and do not yet perform very well in their attempts to translate continuous text, the problem which we face in the design of indexing systems for moving images is not the same. With removal of the aspect of coordinating grammar and idioms across languages, translating isolated names of objects and events becomes a much simpler problem for translation software, and equivalents of the names of things such as chairs, tigers, and mountains can be readily found and automatically generated in many languages.

Another approach to the information system proposed is the use of pictures to find other pictures. Specifically, a visual dictionary of the type published by Québec-Amérique has several characteristics of use in searching moving image databases. The multilingual book version (Corbeil et Archambault 1994) labels objects and parts of objects in four languages, English, French, Spanish, and German. The CD-ROM version (Le Visuel 1996) published in North America uses three languages for its labels: English, French, and Spanish. Another version of the CD-ROM, published in Germany, has English, French, and German (Bildwörterbuch 1997). The CD-ROM versions of the dictionary include animated sequences illustrating certain concepts as well as spoken versions of the labels of the images. Such an interface to the information system could allow users not only to use pictures to search for other pictures (e.g. the user searching for images of specific automobile parts without knowing their names can find these in the visual dictionary) but also as a sound interface since the names of objects are spoken.

Although voice-recognition technology has some way to go before it becomes very useful, we can imagine the day when users will be able to speak the name of an object of which they are seeking pictures, and the system will be able to match the sound with the sounds of names of objects stored in the database. These names in turn would be converted to text and could be be matched with all the available indexing text in the system to find points in productions where certain words or phrases are uttered, as in audio description, closed captioning or dialogue spoken by anyone appearing on or off camera which forms part of the sound track. As with the issue of translating individual words in relation to translating continuous text, the use of isolated words facilitates greatly the work of voice-recognition algorithms, whose difficultes are largely related to decoding continuous sound and identifying boundaries between words.

In addition to words, music is also found as a sound component of moving image documents. Good algorithms are already in use for matching snippets of tunes to whole works found in databases (Downie 2000, Rolland et al. 1999, Uitdenbogerd and Zobel 1999). Incorporating such algorithms in the system or using them as another filter through which queries can be passed, musical information such as theme music from films, songs included in films, musical themes from television programs or jingles from tv commercials all become searchable in our system.

VIEW OF THE INTERFACE

One of the fundamental characteristics that distinguish still images from moving images is that the latter take place in time. Thus using a timeline as the basic organizing principle for information systems for moving images is a logical choice. Building on groundwork laid by O'Connor (1985), and borrowing ideas from existing software interfaces such as those used by Premiere (Adobe Systems), Flash (Macromedia) and iMovie (Apple Computer), we can adapt these familiar interface concepts to our information system. The model from which we will build our system features many kinds of information displayed simultaneously along a timeline, in the manner of parts for different instruments in a musical score, providing an overview of the content from which the user can zoom in on individual components for a closer look or back out for a broader view. As described above, the model also proposes a variety of rich search approaches using text and pictures in order to provide different ways to retrieve information in response to varied and changing kinds of users and queries. It proposes to use XML as a coding language, in order to offer a wide variety of possibilities for manipulation of the data, as well as to offer universality and to promote widespread use.

Features that participating databases can be expected to have include textual metadata such as running descriptions of shot-by-shot content, naming persons, objects and events seen in each shot. Theoretically, even such uncontrolled vocabulary using a variety of nonstandard description methods can is good enough for generating useful indexing to the content of the shots. Collections that use more normalised methods including in-house vocabularies or Web-based cataloguing and indexing systems such as the Dublin Core can be expected to be more useful because the metadata will be more tightly controlled.

The following elements, presented in the form of tags for a hypothetical news database, give an idea of some of the kinds of information users can expect to find in our information system. Each provides a direct link to the corresponding moving image. Of course the system will need to provide for the addition of any number of other tags needed to accommodate information found in all kinds of local situations.

Running text:
<News anchor>
<Reporter1, 2, etc.>
<Interviewee 1, 2, etc.>
<Witness 1, 2, etc.>
<Narration>
<Closed captioning>
<Audio description>
<Intertitles>
<Text of signs, banners, placards etc.>

Automatic or semi-automatic indexing:
<Name of Person1, 2, etc.>

<Description of clothing>
<Gestures>
<Emotions>
<Object>
<Event>
<Time period real>
<Time period represented>
<Geographical area real>
<Geographical area represented>
<Colour>
<Weather>

Tags for sounds:
<Human voice female>
<Human voice male>
<Animal female>
<Animal male>
<Vehicle>
<Weather>
<Sound effect>
<Ambiant noise>
<Music>

These sample tags are only examples of what can be offered, and a great deal of thought and effort building on existing knowledge and practice needs to be invested in order to create the library of tags needed for building actual systems. However, by offering a public library of metadata tags for use in system buiilding, some level of standardisation can be offered as individual databases are created to suit the needs of each collection. Although using these tags may not be required for the actual construction of any participating database, the advantage of taking as many as possible from a common list will be obvious to users as these tags will provide as many bridges between systems. In addition, coordination with existing metadata cores and sets should be fostered.

A WORD ON THE LOW-LEVEL APPROACHES

Although discussion of the low-level approaches to retrieval of moving images on which our colleagues in computer science are working is outside the scope of this paper, we are aware that a great deal of work is going on in this area. In addition, we expect that some of these approaches will be integrated as modules of the large system we hope to build. An important advantage of Web-based systems is that they can be modular, flexible, and subject to constant improvement. Since our own endeavours are in the area of information studies, these are naturally the focus of our work. However, we recognize the usefulness of both the high-level and low-level techniques and the advantages of integrating them in information systems. As the low-level approaches come closer to achieving their goals, there will be more potential for collaboration among researchers building information systems, and the integration of modules will become more and more possible. It is clear that both the high-level and the low-level approaches are useful in the broad picture of storage and retrieval of multimedia data.

THE PROSPECTS FOR BUILDING SYSTEMS USING THIS MODEL

The time is ripe for undertaking a vast initiative to build the system we have only begun to describe in this paper. Many components are involved, a great many skills are required, and the background research is in varying states of progression. But enough is now known to move ahead, and the soil for doing so is especially fertile in view of the general state of anarchy in the management of moving image databases. Proprietary, ad hoc systems abound, and coordination and collaboration among systems are the exception rather than the rule. Although the private sector of the movie and television industries can not be expected to collaborate in the development, there are very many public collections which could contribute data and expertise, perhaps eventually attracting the secretive private sector with usable systems that perform well and which are publicly available.

Processes to be used in the system need to be automated as much as possible since its success as a working product is dependent on as little human intervention as possible in the creation of searchable metadata information. Detailed specifications need to be worked out, and teams of researchers and research assistants to work on the various aspects need to be organised. Fortunately there is a great deal of interest in networked resources, and they are a major focus of research in information studies. Thus we can hope to build a team of knowledgeable workers. It is hoped that activities in this area can be coordinated and carried out in the context of new research resources being developed at our university and elsewhere. With some luck and a great deal of work, we may be able to bring our research efforts over the last several years to fruition.

REFERENCES

Archive Builders. 2001. Digital image sizes. Handout, 12 p. Also available at http://www.ArchiveBuilders.com (document consulted 2001 03 10).

Das große Bildwörterbuch multimedial: Deutsch, Englisch, Französisch. 1997. München: Tewi Verlag GmbH. ISBN 3828112579.

Corbeil, Jean-Claude et Ariane Archambault. 1994. Le visuel multilingue : dictionnaire thématique : français, anglais, espagnol, allemand. Montréal: Les Éditions Québec-Amérique.

The Cyc® Ontology Guide. 2001. Available at http://www.cyc.com/cyc-2-1/intro-public.html (document consulted 2001 04 15).

Downie, J. Stephen. 2000. Access to music information: the state of the art. *Bulletin of the American Society for Information Science* 26, no. 5 (June/July). Available at http://www.asis.org/Bulletin/June-00/downie.html (document consulted 2001 04 16).

Euro WordNet. 2001. Available at http://www.hum.uva.nl/~ewn/ (document consulted 2001 04 15).

The Global WordNet Association. 2001. Available at http://www.hum.uva.nl/~ewn/gwa.htm (document consulted 2001 04 15).

Heitmann, Jürgen. 1998. User requirements and technologies for automated storage and retrieval: interim report of the SMPTE Study Group on User Requirements for Future Automated Storage and Retrieval Systems (V16.09). *SMPTE Journal* (February), 100–105.

O'Connor, Brian C. 1985. Access to moving image documents: background concepts and proposals for surrogates for film and video works. *Journal of Documentation* 41, no. 4 (December): 209-20.

Rolland, Pierre-Yves, Gailius Raskinis, and Jean-Gabriel Ganascia. 1999. Musical content-based retrieval: an overview of the Melodiscov approach and system. In *Proceedings of the ACM Multimedia '99* (Orlando, FL, USA, October 30–November 5, 1999). New York: ACM, 81-84.

Soergel, Dagobert. 1999. The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science* 50, no. 12 (October), 1119-1120.

Srihari, R.K. 1995. Automatic indexing and content-based retrieval of captioned images. *IEEE Computer* 28:9, 49–56.

Srihari, Rohini K. 1997. Using speech input for image interpretation, annotation, and retrieval. In *Digital image access & retrieval,* ed. P. Bryan Heidorn and Beth Sandore. Urbana-Champaign, IL: Graduate School of Library and Information Science, University of Illinois, 140–156

Sun Microsystems. 2001. (http://sun.com).

Turner, James. 1994. Determining the subject content of still and moving image documents for storage and retrieval: an experimental investigation. PhD thesis, University of Toronto.

Turner, James M. 1995. Comparing user-assigned terms with indexer-assigned terms for storage and retrieval of moving images: research results. *Proceedings of the 58th ASIS Annual Meeting, Chicago, Illinois, October 9–12, 1995,* vol. 32, 9–12.

Turner, James M. 1996. Cross-language transfer of indexing concepts for storage and retrieval of moving images: preliminary results. *Global complexity: information, chaos and control: proceedings of the 59th ASIS Annual Meeting, Baltimore, Maryland, October 21–24 1996,* vol. 33. Medford, NJ: Information Today, 214–217.

Turner, James M. 1998. Some characteristics of audio description and the corresponding moving image. *Information access in the global information economy: proceedings of the 61st ASIS Annual Meeting, Pittsburgh,, Pennsylvania, October 24–29 1998,* vol. 35. Medford, NJ: Information Today, 108–117.

Turner, James M. et Jean-François Roulier. 1999. La description d'images fixes et en mouvement par deux groupes linguistiques, anglophone et francophone, au Québec. *Documentation et bibliothèques* 45, no. 1 (janvier-mars), 17–22.

Uitdenbogerd, Alexandria, and Justin Zobel. 1999. Melodic matching techniques for large music databases. In *Proceedings of the ACM Multimedia '99* (Orlando, FL, USA, October 30–November 5, 1999). New York: ACM, 57–66.

Le Visuel : dictionnaire multimédia : français, anglais, espagnol. [CD-ROM]. 1996. Montréal: Les Éditions Québec-Amérique. ISBN 2890378659.

The Word Spy. 2001. Available at http://www.logophilia.com/WordSpy/yottabyte.html (document consulted April 9, 2001).

WordNet. 2001. Available at http://www.cogsci.princeton.edu/~wn (document consulted 2001 04 15).