

MPEG-7 MULTIMEDIA TECHNIQUE

Ms. S.G. Dighe¹

tiya.10588@gmail.com, Mobile No.: 9881003517

¹ Amrutvahini College of Engineering, Sangamner,

Abstract- MPEG-7, formally known as the Multimedia Content Description Interface, includes standardized tools (descriptors, description schemes, and language) enabling structural, detailed descriptions of audio-visual information at different granularity levels (region, image, video segment, collection) and in different areas (content description, management, organization, navigation, and user interaction). It aims to support and facilitate a wide range of applications, such as media portals, content broadcasting, and ubiquitous multimedia. We present a high-level overview of the MPEG-7 standard. We first discuss the scope, basic terminology, and potential applications. Next, we discuss the constituent components. Then, we compare the relationship with other standards to highlight its capabilities.

Keywords— MPEG-7, multimedia asset management, audiovisual indexing, content-based image retrieval, speech Recognition, signal segmentation

Introduction

How many times have you seen science fiction movies such as 2001: A Space Odyssey and thought, “Wow, we’re so far away from having any of the fancy gadgets depicted in these movies!” In 2001, Hal, the talking computer intelligently navigates and retrieves information or runs complex operations instigated by spoken input. Or how about using an image-based query, say an image of the motorbike used by Arnold Schwarzenegger in the movie T2, to find images of similarly looking motorbikes. Dreams or reality?

As more and more audiovisual information becomes available from many sources around the world, many people would like to use this information for various purposes. This challenging situation led to the need for a solution that quickly and efficiently searches for and/or filters various types of multimedia material that’s interesting to the user.

For example, finding information by rich-spoken queries, hand-drawn images, and humming improves the user-friendliness of computer systems and finally addresses what most people have been expecting from computers. For professionals, a new generation of applications will enable high-quality information search and retrieval. For example, TV program producers can search with “laser-like precision” for occurrences of famous events or references to certain people, stored in thousands of hours of audiovisual records, in order to collect material for a program. This will reduce program production time and increase the quality of its content.

MPEG-7 is a multimedia content description standard, (to be defined by September 2001), that addresses how humans expect to interact with computer systems, since it develops rich descriptions that reflect those expectations.

1.1 What Are the MPEG Standards?

The Moving Picture Coding Experts Group (MPEG) is a working group of the Geneva-based ISO/IEC standards organization, (International Standards Organization/International Electro-technical Committee) in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio, and a combination of the two. MPEG-7 then is an ISO/IEC standard being developed by MPEG, the committee that also developed the Emmy Award-winning standards known as MPEG-1 and MPEG-2, and the 1999 MPEG-4 standard.

- MPEG-1:** For the storage and retrieval of moving pictures and audio on storage media.
- MPEG-2:** For digital television, it’s the timely response for the satellite broadcasting and cable television industries in their transition from analog to digital formats.
- MPEG-4:** Codes content as objects and enables those objects to be manipulated individually or collectively on an audiovisual scene. MPEG-1, -2, and -4 make content available. MPEG-7 lets you to find the content you need.

Besides these standards, MPEG is currently also working in MPEG-21 a Technical Report about Multimedia Framework.

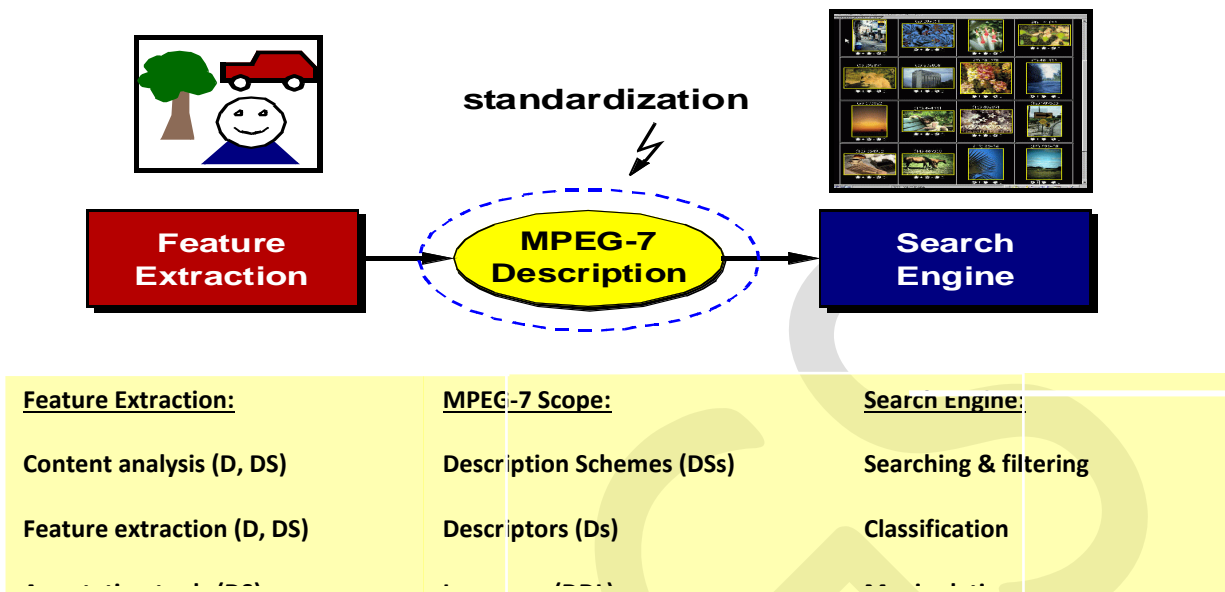


Figure 1: The Scope of MPEG-7

1.2 Defining MPEG-7

MPEG-7 is a standard for describing features of multimedia content.

1.2.1 Qualifying MPEG-7

MPEG-7 provides the world's richest set of audio-visual descriptions.

These descriptions are based on catalogue (e.g., title, creator, rights), semantic (e.g., the who, what, when, where information about objects and events) and structural (e.g., the colour histogram - measurement of the amount of colour associated with an image or the timbre of an recorded instrument) features of the AV content and leverages on AV data representation defined by MPEG-1, 2 and 4.

Comprehensive Scope of Data Interoperability.

MPEG-7 uses XML Schema as the language of choice for content description MPEG-7 will be interoperable with other leading standards such as, SMPTE Metadata Dictionary, Dublin Core, EBU P/ Meta, and TV Anytime.

1.3 The Key Role of MPEG-7

MPEG-7, formally named "Multimedia Content Description Inter-face," is the standard that describes multimedia content so users can search, browse, and retrieve that content more efficiently and effectively than they could using today's mainly text-based search engines. It's a standard for describing the features of multimedia content.

However MPEG-7 will not standardize the (automatic) extraction of AV descriptions/features. Nor will it specify the search engine (or any other program) that can make use of the description.

2.1 Organization of MPEG-7 Description Tools

Over 100 MPEG-7 Description Tools are currently being developed and refined. The relationships between the MPEG-7 Description Tools are outlined in Figure 2.

The basic elements, at the lower level, deal with basic data types, mathematical structures, schema tools, linking and media localization tools, as well as basic DSs, which are elementary components of more complex DSs. The Schema tools section specifies elements for creating valid MPEG-7 schema instance documents and description fragments.

The first three elements primarily address information that's related to the management of the content (content management), whereas the last two are mainly devoted to the description of perceivable information (content description).

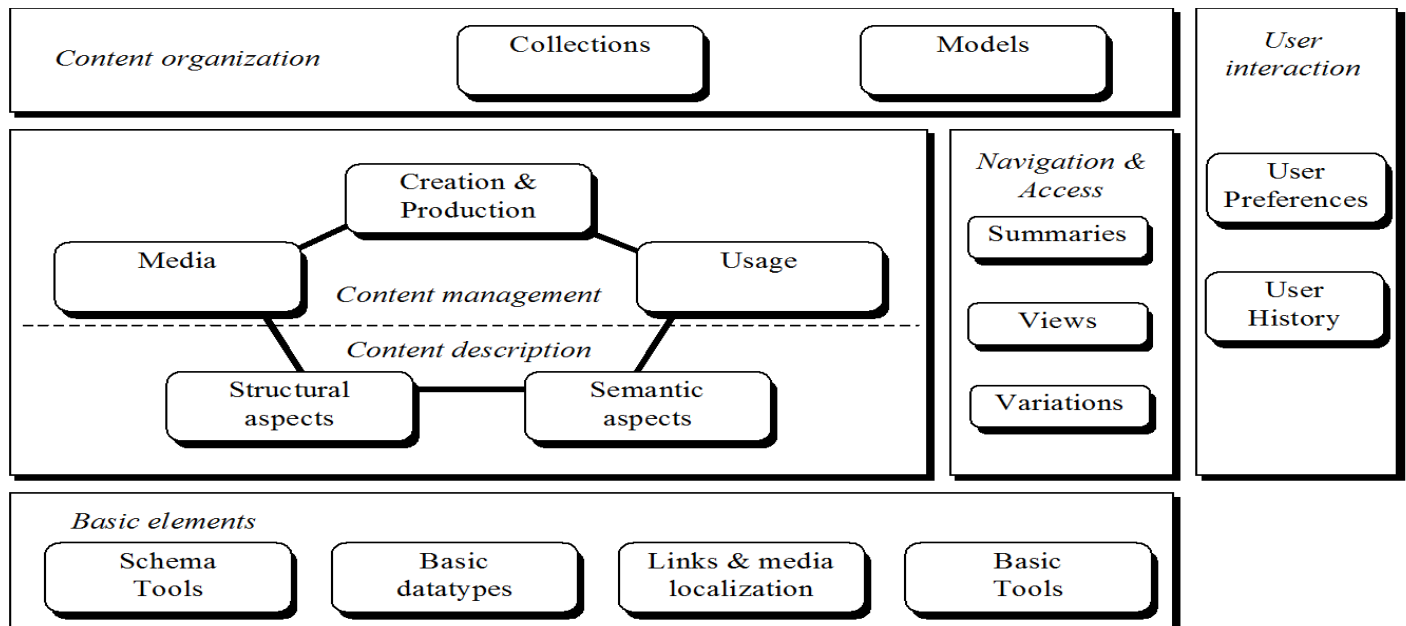


Figure 2: Overview of MPEG-7 Multimedia Description Schemes (DSs)

- **Creation and Production:** Contains meta information that describes the creation and production of the content; typical features include title, creator, classification, and purpose of the creation. Most of the time this information is author-generated since it can't be extracted from the content.
- **Usage:** Contains meta information that's related to the usage of the content; typical features involve rights holders, access rights, publication, and financial information. This information may be subject to change during the lifetime of the AV content.
- **Media:** Contains the description of the storage media; typical features include the storage format, the encoding of the AV content, and elements for the identification of the media. *Note:* Several instances of storage media for the same AV content can be described.
- **Structural aspects:** Contains the description of the AV content from the viewpoint of its structure. The description is structured around segments that represent physical, spatial, temporal, or spatio-temporal components of the AV content. Each segment may be described by signal-based features (color, texture, shape, motion, audio) and some elementary semantic information.
- **Conceptual Aspects:** Contains a description of the AV content from the viewpoint of its conceptual notions.

The five sets of Description Tools are presented here as separate entities, however, they are interrelated and may be partially included in each other. For example, Media, Usage or Creation & Production elements can be attached to individual segments involved in the structural description of the content. Tools are also defined for *navigation and access* and there is another set of tools for *Content organization* which addresses the organization of content by classification, by the definition of collections and by modeling. Finally, the last set of tools is *User Interaction* which describes user's preferences for the consumption of multimedia content and usage history.

2.2 MPEG-7 Working Groups

Currently MPEG-7 concentrates on the specification of description tools (Descriptors and Description Schemes), together with the development of the MPEG-7 reference software, known as XM (eXperimentation Model). The XML Schema Language was chosen as the base for the Description Definition Language (DDL).

The **MPEG-7 Audio group** develops a range of Description Tools, from generic audio descriptors (e.g., waveform and spectrum envelopes, fundamental frequency) to more sophisticated description tools like Spoken Content and Timbre. Generic Audio Description tools will allow the search for similar voices, by searching similar envelopes and fundamental frequencies of a voice sample against a database of voices. The Spoken Content Description Scheme (DS) is designed to represent the output of a great number of state of the art Automatic Speech Recognition systems, containing both words and phonemes representations and most likely transitions. This alleviates the problem of out-of-vocabulary words, allowing retrieval even when the original word was wrongly decoded. The Timbre descriptors (Ds) describe the perceptual features of instrument sound, that make two sounds having the same pitch and loudness appear different to the human ear. These descriptors allow searching for melodies independently of the instruments.

The **MPEG-7 Visual group** is developing four groups of description tools: Color, Texture, Shape and Motion. Color and Texture Description Tools will allow the search and filtering of visual content (images, graphics, video) by dominant color or textures in some (arbitrarily shaped) regions or the whole image. Shape Description Tools will facilitate “query by sketch” or by contour similarity in image databases, or, for example, searching trademarks in registration databases. Motion Description Tools will allow searching of videos with similar motion patterns that can be applicable to news (e.g. similar movements in a soccer or football game) or to surveillance applications (e.g., detect intrusion as a movement towards the safe zone).

The **MPEG-7 Multimedia Description Schemes group** is developing the description tools dealing with generic and audiovisual and archival features. Its central tools deal with content management and content description as outlined in section 2.1.

The **MPEG-7 Implementation Studies group** is designing and implementing the MPEG-7 Reference Software known as XM.

The **MPEG-7 Systems group** is developing the DDL and the binary format (known as BiM), besides working in the definition of the terminal architecture and access units.

3. MPEG-7 Application Domains

The elements that MPEG-7 standardizes will support a broad a range of applications (for example, multimedia digital libraries, broadcast media selection, multimedia editing, home entertainment devices, etc.). MPEG-7 will also make the web as searchable for multimedia content as it is searchable for text today. This would apply especially to large content archives, which are being made accessible to the public, as well as to multimedia catalogues enabling people to identify content for purchase. The information used for content retrieval may also be used by *agents*, for the selection and filtering of broadcasted "push" material or for personalized advertising. Additionally, MPEG-7 descriptions will allow fast and cost-effective usage of the underlying data, by enabling semi-automatic multimedia presentation and editing. All domains making use of multimedia will benefit from MPEG-7 including,

- Digital libraries, Education (image catalogue, musical dictionary, Bio-medical imaging catalogues...)
- Multimedia editing (personalised electronic news service, media authoring)
- Cultural services (history museums, art galleries, etc.),
- Multimedia directory services (e.g. yellow pages, Tourist information, Geographical information systems)
- Broadcast media selection (radio channel, TV channel,...)
- Journalism (e.g. searching speeches of a certain politician using his name, his voice or his face),
- E-Commerce (personalised advertising, on-line catalogues, directories of e-shops,...)
- Surveillance (traffic control, surface transportation, non-destructive testing in hostile environments, etc.),
- Investigation services (human characteristics recognition, forensics),
- Home Entertainment (systems for the management of personal multimedia collections, including manipulation of content, e.g. home video editing, searching a game, karaoke,...)
- Social (e.g. dating services),

3.1 Typical applications enabled by MPEG-7 technology include:

- **Audio:** I want to search for songs by humming or whistling a tune or, using an excerpt of Pavarotti’s voice, get a list of Pavarotti’s records and video clips in which Pavarotti sings or simply makes an appearance. Or, play a few notes on a keyboard and retrieve a list of musical pieces similar to the required tune, or images matching the notes in a certain way, e.g. in terms of emotions.

- **Graphics:** Sketch a few lines on a screen and get a set of images containing similar graphics, logos, and ideograms.
- **Image:** Define objects, including color patches or textures, and get examples from which you select items to compose your image. Or check if your company logo was advertised on a TV channel as contracted.
- **Visual:** Allow mobile phone access to video clips of goals scored in a soccer game, or automatically search and retrieve any unusual movements from surveillance videos.
- **Multimedia:** On a given set of multimedia objects, describe movements and relations between objects and so search for animations fulfilling the described temporal and spatial relations. Or, describe actions and get a list of scenarios containing such actions.

3.2 Examples of MPEG-7 Applications

The following applications are examples of the type of solutions that MPEG-7 can solve. These application examples represent development work in progress. There are many more applications being developed around the world,

3.2.1 Content Retrieval using Image as the Query

Figure 3 shows possible ways to search for visual content using the inherent structural features of an image. In this example there are four image features detailed. The color histogram feature (1) of an image allows me to search for images that have the same color. Note, the position of the colors is not important but rather the amount of similar color in the image is important. The next feature, spatial color distribution (2) allows me to search for images where the location of the same color is important. You can see that the added object in the right-bottom flag does not affect this type of search. You can additionally search for images that have a similar edge or contour profile as in the spatial edge distribution (3) search technique. Note, color does not make a difference to this type of search. Finally, you can see an example of searching by object shape (4). Here, the color and edge profiles are not important.

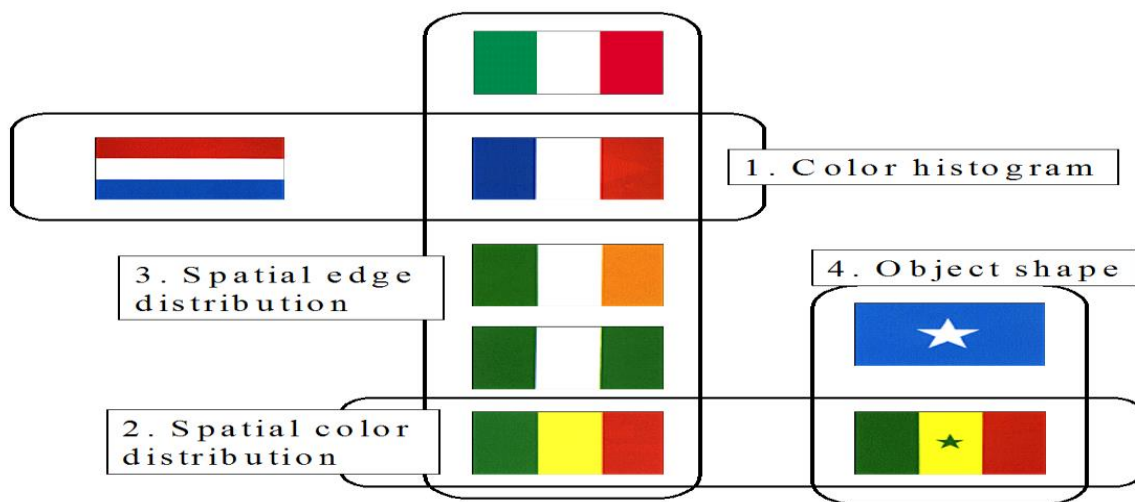


Figure 3: Search using image features

3.2.2 Movie Tool

This is an MPEG-7 description tool for video with easy-to-use visual interface (see figure 5). It is possible to compose a logical structure of the target content, and to also edit and output an MPEG-7 instance file using this tool. Automatic segmentation of video content is done by detecting scene changes. Manual annotation is also possible to allow users provide additional information about the content. The content is arranged in a hierarchy based on topic and sub-topics where visual clips are summarized using thumbnails. In the ideal world, automatic description and organization of content is most desirable but the variety of possible meanings associated with semantic content make it a difficult task. This Movie Tool, though, helps speed up the manual annotation process because of its friendly visual interface environment. Currently, this tool operates on MPEG-1 input content. Since the logical structure of the content is mapped directly to its MPEG-7 instance in the editor, users can easily see the relationships between content and its related MPEG-7

description. This feature provides is very useful when trying to understand the usage of MPEG-7 description tools and their relationship to content.

4. ADVANTAGES:-

- MPEG Standards have been Successful in the Marketplace.
- The contributors to MPEG-7 include experts in every portion of the content value chain: production, post-production, delivery, and consumption.
- Interoperability: rapid uptake of MPEG-7, as it is built on enabling technologies and standards.
- Data Exchange between Subsidiaries.
-

5. APPLICATION:-

- MPEG-7 will enable a New Generation of Multimedia Applications
- MPEG-7 uniquely provides comprehensive standardised multimedia description tools for content.
- Search Engines, Digital Libraries, Broadcast Networks, Entertainment and News Distributors, Streaming Businesses
- Dynamic start-up companies, searching for cutting edge technologies.
- MPEG-7 Intellectual Property and Management Protection
- MPEG-7 Makes Content More Valuable
- MPEG-7 provides a seamless path towards increasingly intelligent content management systems

ACKNOWLEDGMENT

The goal of this paper is to design the advanced multimedia and communication techniques using methods of “**MPEG-7 Multimedia technique**” which can be used easily in day to day life. The function has been realized successfully.

I want to give my whole sincere to my supervisor and grateful appreciation to my colleagues. They tried their best to help me. Without their help and guidance I cannot bring the theories into practice.

On the other hand, I want to thank all my family members and friends for their always support and spiritual motivation.

Thank you very much!

CONCLUSION

Since, we had studied the **MPEG-7 Multimedia technique**. We live in the age of convergence, from the level of production through to distribution and consumption. The technical hardware and communication infrastructure is evolving and will soon reach the point where computing and communications will become embedded in everyday objects and environments. Media will also then become ubiquitous. Ubiquitous media will create a huge demand for new content, and meeting this demand must involve fundamental changes to all stages of media production, management and delivery.

Media archives will become vast and interconnected pools of content, too large to be managed manually. Customization of content within programs, e.g. substitution of structural elements (characters, music, voices) according to viewer desires, content scaling for PDA, cell phones, will be not only possible, but easy and pleasant. MPEG-7 will enable the creation of tools, (through its structured combination of low level features and high-level meta-data), for coping with this "outbreak" of generic content.

REFERENCES:

1. C. G. M. SNOEK, M. WORRING, MULTIMODAL VIDEO INDEXING: A REVIEW OF THE STATE-OF-THE-ART, INTELLIGENT SENSORY INFORMATION SYSTEMS, UNIV. OF AMSTERDAM, ISIS TECHNICAL REPORT SERIES, VOL. 2001-20, DECEMBER 2001
2. Y. WANG, Z. LIU, J. HUANG, MULTIMEDIA CONTENT ANALYSIS USING BOTH AUDIO AND VISUAL CLUES, IEEE SIGNAL PROCESSING MAGAZINE, VOL. 17, 12-36, 2000
3. R. BRUNELLE, O. MICH, C. M. MODENA, A SURVEY ON THE AUTOMATIC INDEXING OF VIDEO DATA, JOURNAL OF VISUAL COMMUNICATION AND IMAGE REPRESENTATION, VOL. 10, 78-112, 1999

4. B. FURHT, S. W. SMOLIAR, H. J. ZHANG, VIDEO AND IMAGE PROCESSING IN MULTIMEDIA SYSTEMS, KLUWER ACADEMIC PUBLISHERS, NORWELL, USA, 2TH EDITION, 1996
5. A. A. ALATAN, A. N. AKANSU, W. WOLF, MULTIMODAL DIALOGUE SCENE DETECTION USING HIDDEN MARKOV MODELS FOR CONTENT-BASED MULTIMEDIA INDEXING, MULTIMEDIA TOOLS AND APPLICATIONS, VOL. 14, 137-151, 2001
6. N. BABAGUCHI, Y. KAWAI, T. KITAHASHI, EVENT BASED INDEXING OF BROADCASTED SPORTS VIDEO BY INTERMODAL COLLABORATION, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 4, 68-75, 2002
7. M. R. NAPHADE AND T. S. HUANG, A PROBABILISTIC FRAMEWORK FOR SEMANTIC VIDEO INDEXING, FILTERING AND RETRIEVAL, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 3, 141-151, 2001
8. SEE FOR INSTANCE RETRIEVALWARE (WWW.CONVERA.COM);VIDEOLLOGGER (WWW.VIRAGE.COM);IMAGEMINE (WWW.IMAGEPRODUCTS.COM) AND MEDIA ARCHIVE (WWW.TECMATH.COM)
9. INTRODUCTION TO MPEG-7: MULTIMEDIA CONTENT DESCRIPTION INTERFACE, EDITED BY B. S. MANJUNATH, P. SALEMBIER, T. SIKORA, JOHN WILEY & SONS, 2002
10. INTERNATIONAL ORGANIZATION FOR STANDARD, MPEG-7 PROJECTS AND DEMOS, DRAFT DOCUMENT, MARCH 2001
11. [HTTP://MPEG-INDUSTRY.COM](http://MPEG-INDUSTRY.COM)
12. L. GAGNON, S. FOUCHER, V. GOUAILLIER, ERIC7: AN EXPERIMENTAL TOOL FOR CONTENT-BASED IMAGE ENCODING AND RETRIEVAL UNDER THE MPEG-7 STANDARD, PROCEEDING OF THE WINTER INTERNATIONAL SYMPOSIUM ON INFORMATION AND COMMUNICATION TECHNOLOGIES (WISICT2004), CANCUN, MEXICO, JAN. 2004 (TO APPEAR)
13. G. BOULIANNE, J. BROUSSEAU, P. OUELLET, P. DUMOUCHEL, FRENCH LARGE VOCABULARY RECOGNITION WITH CROSS-WORD PHONOLOGY TRANSDUCERS, PROCEEDINGS IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP 2000), ISTANBUL, TURKEY, JUNE 5-9, 2000
14. J. VERMAAK, P. PÉREZ, M. GANGNET, RAPID SUMMARIZATION AND BROWSING OF VIDEO SEQUENCES, BMVC'2000, 2000