The Problem

- There is an increased availability of digital audio-visual (AV) content.
- If you want to use this content, you must first find it.
- How do you do this easily and efficiently?

- Note: this problem is not restricted to database applications.
  - E.g., how do you select a broadcast channel (radio or TV) that is potentially interesting?
The Solution: MPEG-7

- In September 2001, Moving Picture Experts Group (MPEG) specified an international standard called Multimedia Content Description Interface, or MPEG-7.
- MPEG-7 is a standardized description, based on XML Schema, of various types of AV content.
- Material (e.g., still pictures, graphics, audio, video) with associated MPEG-7 data can be efficiently indexed and searched for.

Note: the standard does not include the extraction of descriptions, or the search engine for making use of the descriptions (which are both left to the industry).

MPEG-7 Terminology (1)

- Data.
  - AV content to be described using MPEG-7, regardless of storage, coding, display, transmission, medium, or technology.
  - E.g., MPEG-4 stream, a speech, a picture printed on paper.
- Feature.
  - Distinctive characteristic of the data with some significance.
  - E.g., title of a movie, pitch of a speech segment, color of an image.
- Descriptor (D).
  - Representation of a feature that defines the syntax and semantics of the feature representation.
  - It is possible to have several Ds representing a single feature.
  - E.g., for the color feature, possible Ds are the color histogram, the average of the frequency components.
MPEG-7 Terminology (2)

- Descriptor value.
  - Instantiation of a D for a given data set.
- Description Scheme (DS).
  - Specifies the structure and semantics of the relationships between its components, which can be Ds and DSs.
  - The distinction between D and DS is Ds 1) contain only basic data types and do not refer to another D or DS; 2) are meant to be automatically-extracted; and 3) are typically not human-readable since they provides a compact binary description.
  - E.g., a radio segment DS may note the recording date, the broadcast date, the producer, the talent, and include pointers to a transcript.

MPEG-7 Terminology (3)

- Description.
  - A DS and a set of descriptor values that describe the data.
  - These are encoded to fulfill relevant requirements such as compression efficiency and error resilience.
  - The coded descriptions may be physically located with the associated AV content (in the same data stream or same storage system), or may reside somewhere else.
- Description Definition Language (DDL).
  - Language that allows the creation of new DSs, and the extension and modification of existing DSs (since not all AV content will fit into a prescribed structure).
  - It is an extension of the XML Schema Language.
MPEG-7 Terminology (4)

MPEG-7 and Other Standards

- MPEG-7 provides complementary functionality to the other MPEG standards: MPEG-1, -2, and -4 make content available; MPEG-7 allows finding the content one needs.
- Yet MPEG-7 can be used independently of them (e.g., applied to an analog movie).
- There are some ties between MPEG-7 and MPEG-4:
  - The representation defined in MPEG-4 (i.e., the representation of AV content in terms of objects) is well suited to the MPEG-7 standard.
  - Some MPEG-7 tools may be similar to those of MPEG-4.
- MPEG-7 is interoperable with other leading standards such as SMPTE Metadata Dictionary, Dublin Core, EBU P/Meta, and TV-Anytime.
What Makes MPEG-7 Industrially Applicable?

- MPEG-7 aims to enable the creation of tools for coping with an “outbreak” of generic content.
- MPEG-7 is harmonizing with other standards that have demonstrated success and acceptance in both traditional media and new media businesses like W3C (XML, XML Schema), IETF (URI, URN, URL), etc.
- MPEG-7 also makes searching the web for AV content as easy as searching for text-only files. The description tools can be based on the following features of the content:
  - Catalog (e.g., title, creator, rights);
  - Semantic (e.g., the who, what, when and where information about objects and events); and
  - Structural (e.g., the measurement of the color associated with an image, or the timbre of a recorded instrument).
- It also helps in semi-automatic multimedia editing

Target Application Domains

- Search engines, digital libraries, broadcast networks.
- Entertainment and news distributors, streaming businesses.
- Governmental, educational, law, medical & remedial services, and non-profit organizations looking for digital media solutions.
  - E.g., the U.S. Library of Congress receives over 10,000 multimedia items each week, and is committed to the long-term preservation of these multimedia items in digital format.
Some Industrial Tools: IBM MPEG-7 Annotation Tool

- The IBM MPEG-7 Annotation Tool helps in annotating video sequences with MPEG-7 metadata. These annotations can also be customized.
- The annotated descriptions are associated with each video shot and are put out and stored as MPEG-7 descriptions in an XML file.
- This tool can also open MPEG-7 files in order to display the annotations for the corresponding video sequence.

Ricoh MovieTool

- MovieTool is again a tool for describing video content.
- It is intended for use by researchers and designers of MPEG-7 applications.
- MovieTool generates MPEG-7 descriptions based on the structure of a video. By using MovieTool, the user can create the structure while watching the video. Alternatively, MovieTool's editing functions can be used to edit the XML file that contains the MPEG-7 description.
- A major advantage of using MovieTool is that the user can quickly and easily see the correspondence between the MPEG-7 descriptions and the video structure of each scene.
Canon MPEG-7 Speech Recognition Engine

- This tool allows one to create an MPEG-7 Audio “SpokenContent” description file from an audio file in “wav” format.
- This software is based on speech recognition technology being developed by Canon.
  - The “SpokenContent” descriptor represents the output of a speech recognizer as a mixed phoneme and word lattice. This enables retrieval of files or file segments even if word recognition rates are relatively low due to out-of-vocabulary words or high background noise levels.

Market Potential for MPEG-7 Applications

- According to a Goldman Sachs projection, the market for content management tools will grow from US$378M in 2000 to US$4.5B by 2005, and for content management, interoperable tools sell better than non-interoperable tools.
- MPEG-7 is the gold standard for content management interoperability--not just for entertainment companies, but for every company and every industry.
References

- MPEG-7 FAQ
- MPEG-7
- mpeg-industry.com <http://www.mpeg-industry.com/>.
- EE Times - MPEG-7 tackles multimedia content
- alphaworks : IBM MPEG-7 Annotation Tool
- MovieTool Home
  <http://www.ricoh.co.jp/src/multimedia/MovieTool/>.
- MPEG-7 Home <http://www.cre.canon.co.uk/mpeg7asr/>.