MPEG-7
Multimedia Content Description Standard
Abstract

- The purpose of this presentation is to provide a better understanding of the objectives & components of the MPEG-7, "Multimedia Content Description Interface" standard, & an overview of the current state of its development.
Introduction

- It is **clearly** much **more fun** to **develop multimedia content** than to **index it**!

- The **amount of multimedia content** available in **digital archives**, on the **World Wide Web**, in **broadcast data streams**, in **personal & professional databases** is **growing out of control**.

- The **trend** is clear: this **enthusiasm** has led to **increasing difficulties** in **accessing, identifying & managing** such **resources** due to their **volume & complexity**, & a **lack of adequate indexing standards**.
MPEG-7 is being developed by the Moving Pictures Expert Group (MPEG).

Unlike the preceding MPEG standards (MPEG-1, MPEG-2, MPEG-4) which have mainly addressed coded representation of audio-visual content, MPEG-7 focuses on representing information about the content, not the content itself.
Introduction (cont.)
Introduction (cont.)

- A single **standard** which can provide a **simple**, **flexible**, **interoperable** solution to the problems of **indexing**, **searching & retrieving** multimedia resources will be extremely valuable & widely deployed.

- **Resources** described using such a standard will **acquire** enhanced value.

- Compliant **hardware & software** tools capable of efficiently **generating & interpreting** such standardized descriptions will be in **great demand**.

- **But** will **MPEG-7** be **able** to deliver such a standard one which satisfies its **formidable goals & widely heterogeneous scope** whilst concurrently providing **simplicity, flexibility, interoperability & usability**?
Context of MPEG-7

- Everyday, more & more audiovisual information is available from many sources around the world & represented in various forms of media, such as still pictures, graphics, 3-D models, audio, speech, video, & various formats.

- While audio & visual information used to be consumed directly by the human being, there is an increasing number of cases where the audiovisual information is created, exchanged, retrieved, & reused by computational systems.
  - image understanding (surveillance, intelligent vision, smart cameras, etc.)
Context of MPEG-7 (cont.)

- A **code** in a **television program triggers** a suitably programmed *personal video recorder* (PVR) to **record** that program.

- An **image sensor triggers** an **alarm** when a certain **visual event** happens.

- This makes it **necessary** to **develop** forms of **audiovisual information representation** that go beyond the simple **waveform** or **sample-based**, **compression-based** (such as **MPEG-1** & **MPEG-2**) or even **objects-based** (such as **MPEG-4**) representations.

- Forms of **representation** that **allow** some degree of **interpretation of the information’s meaning** are necessary.
# MPEG-7 Work Plan


<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Call For Proposals</td>
<td>October 1998</td>
</tr>
<tr>
<td>Evaluation</td>
<td>February 1998</td>
</tr>
<tr>
<td>First Version of Working Draft</td>
<td>December 1998</td>
</tr>
<tr>
<td>Committee Draft</td>
<td>October 2000</td>
</tr>
<tr>
<td>Final Committee Draft</td>
<td>February 2001</td>
</tr>
<tr>
<td>Draft International Standard</td>
<td>July 2001</td>
</tr>
<tr>
<td>International Standard</td>
<td>September 2001</td>
</tr>
</tbody>
</table>

MPEG work strategy:
definition, competition, & collaboration!
Harmonization of Multimedia Meta-data Standards

- **MPEG-7**: Moving Picture Experts Group:
  - Infrastructure standard for Multimedia Meta-data.
  - Supports interpretation of the information’s meaning.
  - Supports broad range of applications.

- **SMEF** – Standard Media Exchange Framework:
  - **BBC** developed data **models** for information involved in the production: development, use, & management of **media assets**.

- **P/Meta** – EBU P/Meta Project:
  - Exchange of program **content** between **high-level business functions** of EBU members: production, delivery/broadcast, & archive.

- **SMPTE** – Meta-data dictionary & **MXF**:
  - Addresses program **interchange independent of format**.

- **Dublin Core** Meta-data Initiative:
  - Interoperable online meta-data standards supporting broad range of purposes & business models.

- **TV-Anytime** – TV-Anytime Meta-data:
  - **Attractors/descriptors** used *e.g.* in **electronic program guides** (EPG), or in web pages to describe content.

- **Indecs** – Indecs Meta-data Framework:
  - An international initiative of **rights owners creating meta-data** standards for **e-commerce**.

MPEG-7 is the answer!
Scope & Applications

- MPEG-7 is intended to describe audiovisual information regardless of storage, coding, display, transmission, medium, or technology.

- All application’s domains making use of multimedia will benefit from MPEG-7.

- Considering that at present day it is hard to find one not using multimedia, please extend the list of the examples below using your imagination:
Scope & Applications (cont)

- **Architecture**, real estate, & interior design (*e.g.*, searching for ideas).
- **Broadcast media selection** (*e.g.*, radio channel, TV channel).
- **Cultural services** (history museums, art galleries, etc.).
- **Digital libraries** (*e.g.*, image catalogue, musical dictionary, biomedical imaging catalogues, film, video & radio archives).
- **E-Commerce** (*e.g.*, personalized advertising, on-line catalogues, directories of e-shops).
- **Education** (*e.g.*, repositories of multimedia courses, multimedia search for support material).
- **Home Entertainment** (*e.g.*, systems for the management of personal multimedia collections) including manipulation of content (*e.g.*, home video editing, searching a game).
Scope & Applications (cont)

- **Investigation services** (*e.g.*, human characteristics recognition).
- **Journalism** (*e.g.*, searching speeches of a certain politician using his name, his voice or his face).
- **Multimedia directory** services (*e.g.*, yellow pages, Tourist information, geographical information systems (GIS)).
- **Multimedia editing** (*e.g.*, personalized electronic news service, media authoring).
- **Shopping** (*e.g.*, searching for clothes).
- **Social** (*e.g.*, meeting services).
- **Surveillance** (*e.g.*, traffic control, surface transportation, non-destructive testing in hostile environments).
MPEG-7 Objectives

- MPEG-7 aims to standardize:
  - A core set of *descriptors* (Ds) that can be used to describe the various features of *multimedia content*.
  - Pre-defined structures of descriptors & their relationships, called *description schemes* (DSs).
  - A language to define description schemes & descriptors, called the *description definition language* (DDL).
  - Coded representations of descriptions to enable efficient storage & fast access.
MPEG-7 Objectives (cont.)

- MPEG-7 descriptions (a set of instantiated description schemes) will need to be linked to the content itself to allow fast & efficient searching for material of a user's interest.

- The descriptions may be physically located with the associated AV material, in the same data stream, on the same storage system, or the descriptions could be stored remotely.

- Hence mechanisms that can link the AV material to their MPEG-7 descriptions (& vice versa), regardless of where the content & its descriptions are located, are required.
MPEG-7 Parts

- **MPEG-7 Systems** – the tools needed to prepare MPEG-7 descriptions for efficient transport & storage & the terminal architecture.

- **MPEG-7 description definition language** - the language for defining the syntax of the MPEG-7 description tools & for defining new description schemes.

- **MPEG-7 Visual** – the description tools dealing with (only) visual descriptions.

- **MPEG-7 Audio** – the description tools dealing with (only) audio descriptions.
MPEG-7 Parts (cont.)

- **MPEG-7 Multimedia Description Schemes** - the description tools dealing with *generic features* & multimedia descriptions.
- **MPEG-7 Reference Software** - a *software implementation* of relevant parts of the MPEG-7 standard with *normative status*.
- **MPEG-7 Conformance Testing** - *guidelines & procedures* for testing conformance of MPEG-7 implementations.
- **MPEG-7 Extraction & Use of Descriptions** – *informative material* (in the form of a technical report) about the *extraction & use* of some of the description tools.
MPEG-7 Descriptors

- A *descriptor* (D) **defines** the *syntax & the semantics* of one representation of a *particular feature* of audiovisual content.

- A *feature* is a *distinctive characteristic* of the data which is of *significance* to a *user*.

- Each descriptor is defined by *normative & non-normative* parts.

- The *normative* parts consist of the descriptor's *syntax, semantics & binary representations* of these.

- The optional, *non-normative* parts are the recommended extraction & similarity matching methods [6].
### MPEG-7 Descriptors (cont.)

<table>
<thead>
<tr>
<th>Visual Descriptors Samples</th>
<th>Color</th>
<th>Color space</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Dominant color</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Color histogram</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Color quantization</td>
</tr>
<tr>
<td>Texture</td>
<td></td>
<td>Spatial image intensity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Homogeneous texture</td>
</tr>
<tr>
<td>Motion</td>
<td></td>
<td>Camera motion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object motion trajectory</td>
</tr>
</tbody>
</table>
Description Schemes

- A *description scheme* (DS) specifies the *structure & semantics* of the *relationships* between its *components*, which may be both *descriptors* & *description schemes*.

- The *generic audiovisual DS* represents the *integration* of all of the *DS proposals & submissions* within a single DS. At the top level it consists of:
  - A collection of *syntactic structure* DSs (*i.e.*, physical features such as *segments*, *regions*, *color*, *texture*, & *motion*) are described here.
  - A collection of *semantic structure* DSs (*i.e.*, semantic features such as *objects*, *actors*) or *events* (*e.g.*, "goal", "advertisement“).
Description Schemes

Audiovisual DS
- MediaInfo DS
- MetaInfo DS
- Summary DS
- Model DS

Syntactic DS
- Segment DS
- Region DS
- Segment/Region relation graph DS

Syntactic-semantic Link DS
- Event DS
- Object DS
- Event/Object relation graph DS

Semantic DS
Description Definition Language

- The *description definition language* (DDL) is the language that allows the creation of new description schemes & descriptors.

- It also allows the extension & modification of existing description schemes.

- MPEG-7 has a XML-Base DDL.
MPEG-7 Multimedia Indexing & Searching

- **MPEG-7 Indexing & Searching:**
  - Semantics-based (people, places, events, objects, scenes).
  - Content-based (color, texture, motion, melody, timbre).
  - Meta-data (title, author, dates).

- **MPEG-7 Access & Delivery:**
  - Media personalization.
  - Adaptation & summarization.
  - Usage environment (user preferences, devices, context).
The following example gives an MPEG-7 description of a car that is depicted in an image:

```xml
<Mpeg7>
  <Description xsi:type="SemanticDescriptionType">
    <Semantics>
      <Label>
        <Name> Car </Name>
      </Label>
      <Definition>
        <FreeTextAnnotation>
          Four wheel motorized vehicle
        </FreeTextAnnotation>
      </Definition>
    <MediaOccurrence>
      <MediaLocator>
        <MediaUri> image.jpg </MediaUri>
      </MediaLocator>
    </MediaOccurrence>
    </Semantics>
  </Description>
</Mpeg7>
```
The following example gives an MPEG-7 description of the event of handshake between people:

```xml
<Mpeg7>
  <Description xsi:type="SemanticDescriptionType">
    <Semantics>
      <Label>
        <Name>Shake hands</Name>
      </Label>
      <SemanticBase xsi:type="AgentObjectType" id="A">
        <Label href="urn:example:acs">
          <Name>Person A</Name>
        </Label>
      </SemanticBase>
      <SemanticBase xsi:type="AgentObjectType" id="B">
        <Label href="urn:example:acs">
          <Name>Person B</Name>
        </Label>
      </SemanticBase>
      <SemanticBase xsi:type="EventType">
        <Label><Name>Handshake</Name></Label>
        <Definition>
          <FreeTextAnnotation>Clasping of right hands by two people</FreeTextAnnotation>
        </Definition>
      </SemanticBase>
    </Semantics>
  </Description>
</Mpeg7>
```
IBM MPEG-7 Annotation Tool

- The VideoAnnEx annotation tool assists authors in the task of annotating video sequences with MPEG-7 metadata.

- Each shot in the video sequence can be annotated with static scene descriptions, key object descriptions, event descriptions, & other lexicon sets.

- The annotated descriptions are associated with each video shot & are stored as MPEG-7 descriptions in an output XML file. VideoAnnEx can also open MPEG-7 files in order to display the annotations for the corresponding video sequence.

- The annotation tool also allows customized lexicons to be created, saved, downloaded, & updated.
The End