

**INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO**

ISO/IEC JTC1/SC29/WG11/ N2461

MPEG 98

October 1998/Atlantic City, USA

Title: MPEG-7 Requirements Document V.7

Source: Requirements

Status: Approved

MPEG-7 Requirements

1. INTRODUCTION.....	1
2. MPEG-7 FRAMEWORK.....	1
3. MPEG-7 TERMINOLOGY.....	2
4. MPEG-7 REQUIREMENTS	6
4.1. MPEG-7 COMMON AUDIO AND VISUAL REQUIREMENTS.....	7
4.1.1. MPEG-7 DDL Requirements.....	7
4.1.2. Descriptors and Description Schemes – General Requirements.....	8
4.1.3. Descriptors and Description Schemes – Functional Requirements.....	9
4.1.4. Descriptors and Description Schemes – Coding Requirements.....	10
4.2. MPEG-7 VISUAL REQUIREMENTS	10
4.3. MPEG-7 AUDIO REQUIREMENTS	12
4.4. MPEG-7 OTHER MEDIA REQUIREMENTS.....	13
5. SYSTEMS REQUIREMENTS.....	13
6. REFERENCES.....	14
ANNEX A - OPEN ISSUES	15
ANNEX B - ONGOING DISCUSSION WITH RESPECT TO THE DDL.....	15
ANNEX C – OPEN ISSUES WITH RESPECT TO THE SYSTEMS REQUIREMENTS.....	15

1. Introduction

This document presents a set of requirements for the MPEG-7 standard.. The requirements presented in this document are likely to undergo further change, both by adding new requirements as well as by improving the current requirements. All contributions (by means of MPEG submissions) are welcome.

2. MPEG-7 Framework

Nowadays, more and more audio-visual information is available, from many sources around the world. Also, there are people who want to use this audio-visual information for various purposes. However, before the information can be used, it must be located. At the same time, the increasing availability of potentially interesting material makes this search more difficult. This challenging situation led to the need of a solution to the problem of quickly and efficiently searching for various types of multimedia material interesting to the user. A second scenario is filtering where the user prefers to receive only those multimedia material which satisfies his/her preferences. Other interesting domains than search or filtering are, for instance, image understanding (surveillance, intelligent vision, smart cameras, etc.), or media conversion (text to speech, picture to speech, speech to picture, etc.). MPEG-7 aims to create a standard for describing the multimedia data which will support these operational requirements.

Since the description of multimedia data is related to the characteristics of a multimedia system, i.e. computer-controlled, integrated production, manipulation, presentation, storage and communication of independent information, different aspects for generating and using content descriptions of multimedia data can be viewed as a sequence of events:

- 1) Extracting of features describing the content.
- 2) Describing the logical organisation of the described multimedia data and placing the extracted values in the framework specifying this structure.
- 3) Manipulating such description frameworks to accommodate them to different needs
- 4) Manipulating instantiated description frameworks to make the description more accessible for human or machine usage

All four steps pose rather different requirements on the expressive power of the formal mechanism to standardise the processing needed during these steps. Most notable is the difference between feature extraction and document structure description: Whereas the first needs powerful encapsulating procedural constructs, the latter needs structures with high declarative power.

The key area of MPEG-7 is the description of AV content. This requires a description formalism which is flexible and expressive enough to represent adequately the content by some formal structure. Moreover, this formalism should allow humans as well as machines — in form of agents — to exchange, retrieve, and re-use relevant material. An agent is understood here as an autonomous computational system acting in the world of computer networks or computer based on a set of goals that it tries to achieve [1].

For efficiently re-using MPEG-7 descriptions and description schemes, users will need to adapt them to their specific needs. This leads to the modification and manipulation of existing structures. Manipulating document structure as well as an instantiated description will benefit from operations that make the traversal and manipulation of trees, linked lists, and webs natural, either to prune or reorganise the structural framework or to transform the values stored in some

nodes to a more user-friendly representation. In order to avoid multiplying ad hoc solutions, a generic way of defining structure transformation and manipulation should be provided.

Since describing AV material is an extremely expensive and time-consuming task (not only for manually performed extraction) it is of importance to avoid as much as possible re-describing data that had been processed before.

It is anticipated, though, that the underlying data structures and their composition are independent from the applied extraction mechanisms. In other words, MPEG-7 structures provide an application independent description framework which is instantiated by extraction mechanisms.

Whatever features are used to describe an AV document, they will be either extracted automatically by an algorithm running on a computer, or be annotated by a human expert. At least for the automatic performance of such a task it is required that a formal specification of the extracted entity is provided. This specification might be atomic, or might represent the weighted sum, or some other derivation, of a few other features. Examples for such features from music are timbre or density, in the visual domains it might be the composition of an image. Finally, since multimedia content is based on temporal and spatial constraints, i.e. presentational constraints, it is obvious that spatial and temporal requirements influence the semantic and syntactic structure of a description.

Taking this broad range of requirements into account, MPEG-7 will not define a monolithic system for content description but a set of methods and tools for the different steps of multimedia description. MPEG-7, formally called "Multimedia Content Description Interface", will standardise:

- A set of description schemes and descriptors
- A language to specify description schemes, i.e. a Description Definition Language (DDL).
- A scheme for coding the description

It may be pointed out that while MPEG-7 aims to standardise a "Multimedia Content Description Interface", the emphasis of MPEG is on audio-visual content. That is, MPEG-7 does not aim to create description schemes or descriptors for text medium. However, MPEG-7 will consider existing solutions for describing text documents (e.g. SGML, and its derivations like XML, RDF, etc.) and support them as appropriate with suitable, necessary interfaces between audio-visual-content descriptions and the textual-content descriptions.

MPEG-4 OCI is an MPEG-4 specific solution for the provision of limited amounts of information about MPEG-4 content. As such, it can be considered to be a subset of MPEG-7.

MPEG is aware of the fact that other standards for the description of multimedia content are under development while MPEG-7 is created. Thus, MPEG-7 will consider other standardisation activities, such as SMPTE/EBU task force, DVB-SI, CEN/ISSS MMI, etc.

For more details regarding the MPEG-7 background, goals, areas of interest, and work plan please refer to "MPEG-7: Context and Objectives" [2] and "Applications for MPEG-7" [3].

3. MPEG-7 Terminology

This section describes the terminology used by MPEG-7.

1. Data

AV information that will be described using MPEG-7, regardless of the storage, coding, display, transmission, medium, or technology. This definition is intended to be sufficiently broad to encompass graphics, still images, video, film, music, speech, sounds, text and any other relevant AV medium. Examples for MPEG-7 data are : a MPEG-4 stream, a video tape, a CD containing music, sound or speech, a picture printed on paper, or an interactive multimedia installation on the web.

2. Feature

A feature is a distinctive part or characteristic of the data which stands to somebody for something in some respect or capacity. Some examples are: colour of an image, pitch of a speech segment, rhythm of an audio segment, camera motion in a video, style of a video, the title of a movie, the actors in a movie etc.

3. Descriptor

A Descriptor (D) defines the syntax and semantics of a representation entity for a feature. The representation entity is composed out of an identifier of the feature and a datatype. An example might be: *Colour: string*. However, the datatype can be composite, meaning that it may be formed by a combination of datatypes. An example might be: *RGB-Colour: [int,int,int]*.

It is possible to have several descriptors representing a single feature, i.e. to address different relevant requirements. Examples for descriptors are: a time-code for representing duration, colour moments and histograms for representing colour, and a character string for representing a title.

4. Descriptor Value

An instantiation of a descriptor is a value assigned to the feature as pertaining to the data. Descriptor values are combined via the mechanism of a description scheme to form a description.

5. Description Scheme

A description scheme (DS) consists of one or more descriptors and description schemes. The DS specifies the structure and semantics of the relationships between them.

A simple description scheme for describing technical aspects of a shot might look like this, where those elements written in bold represent other DSs.

Shot_Technical_Aspects

Lens	providing information about type (wide-angle), movement (e.g. zooms), state (e.g. deep focus), masking, etc.
Camera	providing information about distance (e.g. close-up), angle (e.g. overhead), movement (e.g. pan_left), position (viewpoint of the frame), etc.
Speed	
Colour	
Granularity	
Contrast	

6. Description

A description is the entity describing the data. A description contains or refers to a fully or partially instantiated DS.

7. Coded Description

A coded description is a description that has been encoded to fulfil relevant requirements such as compression efficiency, error robustness, random access, etc.

8. Description Definition Language

This is the language in which description schemes are specified. The DDL will allow the creation of new description schemes and descriptors and the extension of existing description schemes.

To provide a better understanding of the terminology above please find below Figures 1 - 3 and Table 1. The dotted boxes in the figures encompass the normative elements of the MPEG-7 standard. Note that the presence of a box or ellipse in one of this drawings does not imply that the corresponding element shall be present in all MPEG-7 applications.

Figure 1 shows the extensibility of the above concepts. The arrows from DDL to DS signify that the DSs are generated using DDL. Furthermore, the drawing reveals the fact that you can build a new DS using an existing DS.

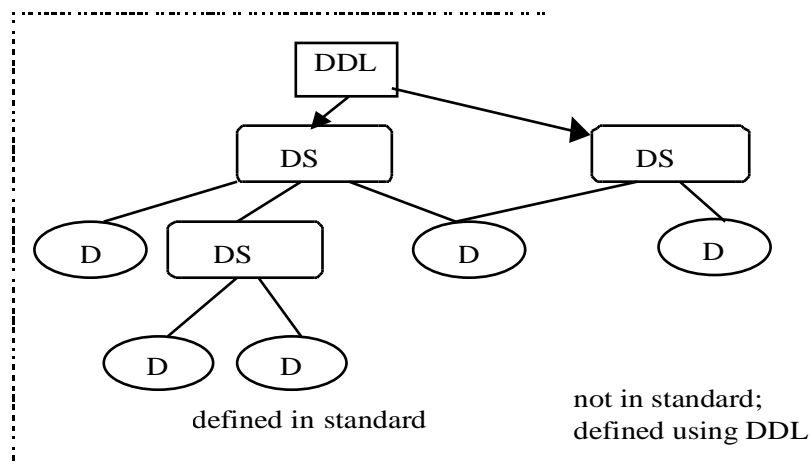


Figure 1: An abstract representation of possible relations between Ds and DSs.

Figure 2 highlights that the DDL provides the mechanism to build a description scheme which in turn forms the basis for the generation of a description. The instantiation of the DS is described as part of Figure 3.

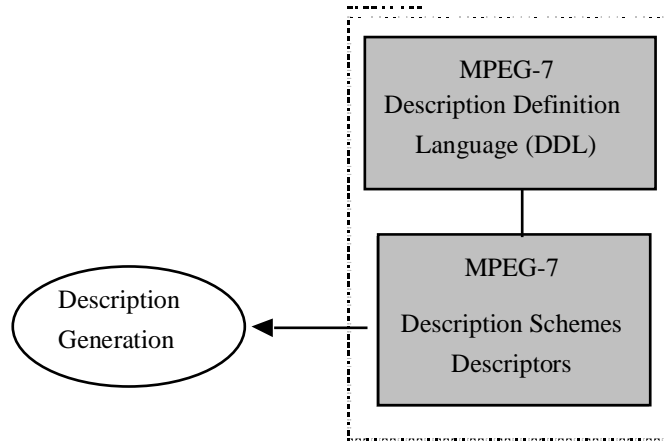


Figure 2: The role of Ds and DSs for the generation of descriptions

Figure 3 explains how MPEG-7 would work in practice. Note: There can be other streams from content to user; these are not depicted here. Furthermore, the use for the encoder and decoder is optional.

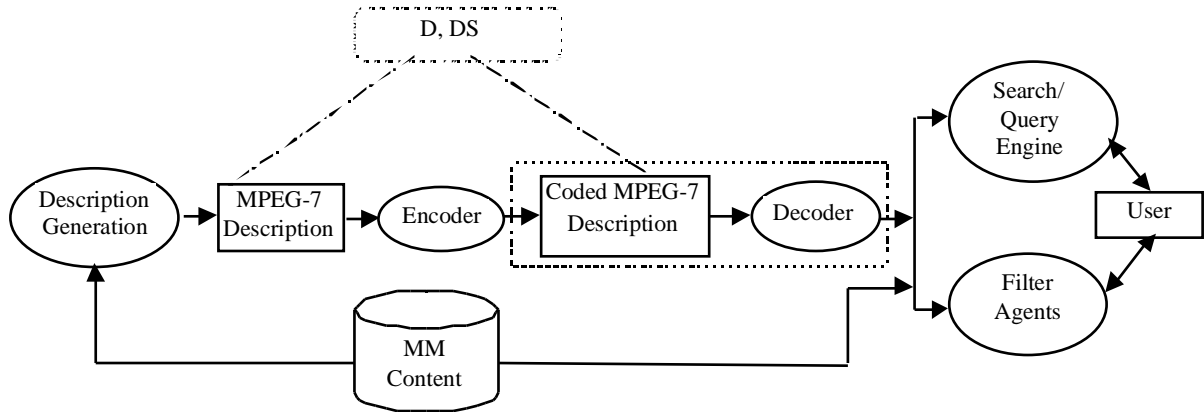


Figure 3: An abstract representation of possible applications using MPEG-7.

Table 1 exemplifies the distinction between a feature and its descriptors. For further details concerning the different feature types please see section 4.1.2 , point 1: Types of features.

Feature types	Descriptor Feature	Datatype
Annotation		text, etc.
N-dimensional spatio-temporal structure	duration of music segments	time code, etc.
	trajectory of objects	chain code, etc.
Statistical Information	colour	colour histogram, etc.
	audio frequency content	average of frequency components, etc.

Objective features	colour of an object	colour histogram, text, etc.
	shape of an object	a set of polygon vertices, a set of moments, etc.
	texture of an object	a set of wavelet coefficients, a set of contrast, coarseness and directionality quantities, etc.
Subjective features	emotion (happiness, angry, sadness, etc.)	a set of eigenface parameters, text, etc.
	style	text, etc.
Production features	Author	text, etc.
	Producer	text, etc.
	Director, etc.	text, etc.
Composition Information	Scene composition	tree graph, etc.
Concepts	event	text, etc.
	activity	text, etc.

Table 1: Typical feature types, features and descriptors

It is understood that, in some situations, the search engine or filter agents (user side) may have to know the exact feature extraction algorithm employed by the description generation process. However, in order to be able to accommodate developments in feature extraction technology as well as in the interest of enabling competition in MPEG-7 application development, the specific extraction algorithm employed by the description generation process is kept outside the scope of MPEG-7 standards. However, MPEG-7 may provide the facility for the DDL to allow code to be embedded or be referenced in description schemes. Note that code is not to be embedded or referenced in the DDL, as the "code" pertains to the description scheme.

4. MPEG-7 Requirements

This section specifies the MPEG-7 Requirements. The requirements are divided in common audio and visual requirements, visual requirements, and audio requirements. The requirements apply, in principle, to both real-time and non real-time as well as push and pull applications.

MPEG will not standardise or evaluate applications. MPEG may, however, use applications for understanding the requirements and evaluation of technology. It must be made clear that the requirements in this document are derived from analysing a wide range of potential applications that could use MPEG-7 descriptions. MPEG-7 is not aimed at any one application in particular; rather, the elements that MPEG-7 standardises shall support as broad a range of applications as possible.

4.1. MPEG-7 Common Audio and Visual Requirements

This section addresses the requirements on DDL, description schemes and descriptors that are common for both the audio and visual media. The DDL requirements are listed first followed by the requirements on descriptors and description schemes (general requirements and functional requirements) and finally the requirements related to coding. Note that while the MPEG-7 standard as a whole should satisfy all requirements, not all requirements have to be satisfied by each individual descriptor or description scheme.

4.1.1. MPEG-7 DDL Requirements

1. **Compositional capabilities:** The DDL shall supply the ability to compose a DS from multiple DSs
2. **Platform independence:** The DDL shall be platform and application independent. This is required to make the representation of content as reusable as possible even on grounds of changing technology.
3. **Grammar:** The DDL shall follow a grammar which is unambiguous, and allows easy parsing (interpretation) by computers.
4. **Primitive data types:** provide a set of primitive data types, e.g. text, integer, real, date, time/time index, version, etc.
5. **Composite datatypes:** The DDL must be able to succinctly describe composite datatypes that may arise from the processing of digital signals (e.g., histograms, graphs, rgb-values).
6. **Multiple media types:** The DDL must provide a mechanism to relate Ds to data of multiple media types of inherent structure, particularly audio, video, audio-visual presentations, the interface to textual description, and any combinations of these.
7. **Partial instantiation:** The DDL shall provide the capability to allow a DS to be partially instantiated by descriptors.
8. **Mandatory instantiation:** The DDL shall provide the capability to allow the mandatory instantiation of descriptors in a DS.
9. **Unique identification:** The DDL shall provide mechanisms to uniquely identify DSs and Ds so that they can be referred to unambiguously.
10. **Distinct name spaces:** The DDL shall provide support for distinct name-spaces. Note: Different domains use the same descriptor for different features or different purposes.
11. **Transformational capabilities:** The DDL shall allow the reuse, extension and inheritance of existing Ds and DSs.
12. **Relationships within a DS and between DSs:** The DDL provides the capability to express the following relationships between DSs and among elements of a DS and express the semantics of these relations
 - a) Spatial relations
 - b) Temporal relations
 - c) Structural relations
 - d) Conceptual relations

13. Relationship between description and data: The DDL shall supply a rich model for links and/or references between one or several descriptions and the described data.

14. Intellectual Property Management: The DDL shall provide a mechanism for the expression of Intellectual Property Management and Protection (IPMP) for description schemes and descriptors.

15. Real time support: The DDL shall desirably provide features to support real time applications (database output like electronic program guides)

4.1.2.Descriptors and Description Schemes – General Requirements

1. Types of features - MPEG-7 shall support multimedia descriptions using various types of features, such as:

Annotation

N-dimensional Spatio-temporal structure

Note: An Example is duration of a music segment

Statistical information

Note: Examples are colour histograms and average audio frequency content.

Objective features

Note: Features, such as the number of beds in a hotel, colour of an object, shape of an object, audio pitch, etc.

Subjective features

Note: Features subject to different interpretations, such as how nice, happy or fat someone is, topic, style, etc.

Production features

Note: This is information about the details of the creation of the data. Examples include the date of data acquisition, producer, director, performers, roles, production company, production history, etc. - essentially any production information that is not necessarily in the IPI field(s).

Composition information

Note: How the scene is composed, editing information, and the like.

Concepts

Examples are event, activity

2. Abstraction levels for Multimedia material – MPEG-7 shall support a means to describe multimedia material hierarchically according to abstraction levels of information to efficiently represent user's information need at different levels.

3. Cross-modality - MPEG-7 shall support audio, visual, or other descriptors which allow queries based on visual descriptions to retrieve audio data and vice versa. *Note: Using an excerpt of Pavarotti's voice for the query, video clips where Pavarotti is singing or video clips where Pavarotti is present are retrieved.*

4. **Multiple Descriptions** – MPEG-7 shall support the ability to handle multiple descriptions of the same material at several stages of its production process, as well as descriptions that apply to multiple copies of the same material.
5. **Description Scheme Relationships** – MPEG-7 description schemes need to express the relationships between descriptors to allow the use of the descriptors in more than one description scheme. The capability to encode equivalence relationships between descriptors in different description schemes shall also be supported.
6. **Feature priorities** - MPEG-7 shall support the prioritisation of features in order that queries may be processed more efficiently. The priorities may denote some sort of level of confidence, reliability, etc
7. **Feature hierarchy** – MPEG-7 shall support the hierarchical representation of different features in order that queries may be processed more efficiently in successive levels where N level features complement (N-1) level features .
8. **Descriptor scalability** - MPEG-7 shall support scalable descriptors in order that queries may be processed more efficiently in successive layers where N-layer description data is an enhancement/refinement of (N-1) layer description data. An example is MPEG-4 shape scalability
9. **Description Schemes with multiple levels of Abstraction.** MPEG-7 shall support DSs which provide abstractions at multiple levels for instance a coarse-to-fine description. An example is a hierarchical scheme where the base layer gives a coarse description and successive layers give more refined descriptions.
10. **Description of temporal range** - MPEG-7 shall support the association of descriptors to different temporal ranges, both hierarchically (descriptors are associated to the whole data or a temporal sub-set of it) as well as sequentially (descriptors are successively associated to successive time periods).
11. **Direct data manipulation** – MPEG-7 shall support descriptors which can act as handles referring directly to the data, to allow manipulation of the multimedia material.
12. **Language of text-based descriptions** – MPEG-7 text descriptors shall specify the language used in description. MPEG-7 text descriptors shall support all natural languages.
13. **Translations in text descriptions** – MPEG-7 text descriptions shall provide the means to contain several translations, and it shall be possible to convey the relation between the description in the different languages.

4.1.3.Descriptors and Description Schemes – Functional Requirements

1. **Content-based retrieval** - MPEG-7 shall support the effective (‘you get what you are looking for and not other stuff’) and efficient (‘you get what you are looking for, quickly’) retrieval of multimedia data based on their contents whatever the semantic involved.
2. **Similarity-base retrieval** - MPEG-7 shall support descriptions allowing to rank-order database content by the degree of similarity with the query.
3. **Associated information** - MPEG-7 shall support the use of information associated to the data, such as text, to complement and improve data retrieval.

Note: As an example, diagnostic medical images are retrieved not only in terms of image contents but also in terms of other information associated *with the images, such as text* describing the diagnosis, treatment plan, etc.

- 4. Streamed and stored descriptions** - MPEG-7 shall support both streamed (synchronised with content) and non-streamed data descriptions.
- 5. Distributed multimedia databases** - MPEG-7 shall support the simultaneous and transparent retrieval of multimedia data in distributed databases.
- 6. Referencing analog data** – MPEG-7 descriptions shall support the ability to reference and describe audio-visual objects and time references of analog format.
- 7. Interactive queries** – MPEG-7 descriptions shall support mechanisms to allow interactive queries.
- 8. Linking** - MPEG-7 shall support a mechanism allowing source data to be located in space and in time using the MPEG-7 data descriptors. MPEG-7 shall also support a mechanism to link to related information.
- 9. Prioritisation of Related Information** - MPEG-7 shall support a mechanism allowing the prioritisation of related information, mentioned under 8. above.
- 10. Browsing** – MPEG-7 shall support descriptions allowing to pre-view information content in order to aide users to overcome their unfamiliarity with the structure and/or types of information, or to clarify their undecided needs.
- 11. Associate Relations** – MPEG-7 shall support relations between components of a description.
- 12. Interactivity support** – MPEG-7 shall support the means allowing to specify the interactivity related to a description. An example for such an interaction is televoting related to broadcast events.

4.1.4. Descriptors and Description Schemes – Coding Requirements

- 1. Description efficient representation** - MPEG-7 shall support the efficient representation of data descriptions.
- 2. Description extraction** – MPEG- shall standardise Descriptors and Descriptions Schemes that are easily extractable from uncompressed and compressed data, according to several widely used formats.
- 3. Intellectual property information** - MPEG-7 shall enable inclusion of copyright, licensing and authentication information related to content and its descriptions. As copyright/licensing may change over time a suitable timestamp or other information may also be required.

4.2. MPEG-7 Visual Requirements

Visual Requirements are related to the retrieval of the visual data classes specified below. The MPEG-7 visual requirements are:

- 1. Type of features** - MPEG-7 shall at least support visual descriptions allowing the following features (mainly related to the type of information used in the queries):

Colour

Visual objects

Texture

Sketch

Shape

Still and moving images (e.g. thumbnails)

Volume

Spatial relations

Note: Related to the spatial and topological relationships among the objects in an image or sequence of images, this means spatial composition information.

Motion

Note: For retrievals using temporal composition information.

Deformation (e.g. the warping of an object)

Source of visual object and its characteristics, e.g., the source object, source event, source attributes, events, event attributes, and typical associated scenes.

Models (e.g. MPEG-4 SNHC)

2. Data visualisation using the description - MPEG-7 shall support a range of multimedia data descriptions with increasing capabilities in terms of visualisation. This means that MPEG-7 data descriptions shall allow a more or less sketchy visualisation of the indexed data.

3. Visual data formats - MPEG-7 shall support the description of the following visual data formats:

digital video and film, such as MPEG-1, MPEG-2 or MPEG-4

analogue video and film

still pictures in electronic such as JPEG, paper or other format

graphics, such as CAD

3D models, notably VRML

composition data associated to video

others to be defined

4. Visual data classes - MPEG-7 shall support descriptions specifically applicable to the following classes of visual data:

natural video

still pictures

graphics

animation (2-D)

three-dimensional models

composition information

Note: For example, the MPEG-4 format includes various data classes, such as natural video, still pictures, graphics, or composition information.

4.3. MPEG-7 Audio Requirements

Audio Requirements are related to the retrieval of the audio data classes specified below. The MPEG-7 audio requirements are:

1. Type of features - MPEG-7 shall support audio descriptions allowing the following features (mainly related to the type of information used in the queries):

Frequency contour (general trend, melodic contour)

Audio objects

Timbre

Harmony

Frequency profile

Amplitude envelope

Temporal structure(including rhythm)

Textual content

Note: This is typically speech or lyrics.

Sonic approximations

Note: A person may vocalise a sonic sketch by humming a melody or by ‘growling’ a sound effect.

Prototypical sound

Note: This is a more typical query-by-example, in which a querent provides an example sound, such as squealing brakes, which would find car-chase scenes.

Spatial structure

Note: This applies to multi-channel sources, with stereo, 5.1-channel, and binaural sounds each having particular mappings.

Source of sound and its characteristics, e.g., the source object, source event, source attributes, events, event attributes, and typical associated scenes.

Models (e.g. MPEG-4 SAOL)

2. Data sonification using the description - MPEG-7 shall support a range of multimedia data descriptions with increasing capabilities in terms of sonification.

3. Auditory data formats - MPEG-7 shall support at least the description of the following types of auditory data:

Digital audio, such as MPEG-1 Audio, Compact Disc

Analog audio, such as vinyl records, or magnetic tape media

MIDI, including General MIDI and Karaoke formats

Model-based audio (e.g. MPEG-4’s Structured Audio Orchestra Language - SAOL)

Production data

- 4. Auditory data classes** - MPEG-7 shall support descriptions specifically applicable to the following sub-classes of auditory data:

Soundtrack (natural audio scene)

Music

Atomic sound effects (e.g. clap)

Speech

Symbolic audio representations (MIDI, SNHC Audio)

Mixing information (including effects)

4.4. MPEG-7 Other Media Requirements

MPEG-7's emphasis is on audio-visual content, hence providing novel solutions addressing text-only documents will not be among the goals of MPEG-7. However multimedia content may include or refer to text in addition to audio-visual information. To accommodate such documents MPEG-7, will consider existing solutions developed by other standardisation organisations for text only documents and support them as appropriate. The requirements on such solutions are:

1. MPEG-7 descriptions of text for text-only documents and composite documents containing text should be the same. That is, using MPEG-7 terminology, the description schemes and descriptors for text only documents and for text in visual documents (for example subtitles) must be the same.
2. The adopted text descriptions and the interface will allow queries based on audio-visual descriptions to retrieve text data and vice versa.

4. Systems Requirements

This section addresses the MPEG-7 systems requirements.

- 1. Robustness to information errors and loss for descriptors** – to allow error resilient audio and visual data descriptors.
Note: The precise error conditions to withstand are to be identified.
- 2. A mechanism for defining Quality of Service (QoS) for MPEG-7 description streams must be provided**
- 3. IPMP mechanisms for the protection of MPEG-7 descriptions must be provided**
- 4. Temporal Synchronisation of content with descriptions** - to allow the temporal association of descriptions with content (AV objects) that can vary over time.
Note: For synchronisation in time absolute and relative time bases would be required .
- 5. Physical location of content with associated descriptions** - to associate descriptions with content (AV objects) that can vary in physical location
Note: Location may be specified in terms of hyperlink, broadcast channel, an object in a scene

- 6. Multiplexing of multiple MPEG-7 descriptions associated with a content item** – to allow flexible localisation of descriptor data with one or more content objects

Note: A variety of descriptors and description schemes could be associated with each content item. Depending on the application, not all of these will be used in all cases. Even though some primitive multiplex may be part of a description scheme, not the complete multiplex must necessarily be specified within MPEG-7, as the following example shows.

In MPEG-7, the multiplex functionality will be similar to database functionality, because the selective access to descriptor data must be much more flexible than in the cases of previous MPEG standards. In pull applications, MPEG-7 data themselves can be kept in a database to manage the access ("multiplex") of descriptor data in a very flexible way. In push applications (e.g. real-time broadcast), the multiplex syntax must allow efficient parsing of MPEG-7 streams.

- 7. Multiplexing of multiple MPEG-7 description streams for transmission over the same connection** - to allow multiple MPEG-7 descriptions to be transmitted over the same connection

Note: It may be necessary to transmit a number of streams containing MPEG-7 data over a single channel whilst maintaining synchronisation of each description stream with the content stream to which it refers.

- 8. Multiplexing of multiple MPEG-7 description streams without content** - to allow multiple MPEG-7 descriptions to be transmitted over the same connection without referenced content.

Note: It may be necessary to transmit a number of streams containing MPEG-7 data over a single channel whilst maintaining synchronisation between the description streams when no content stream is present.

- 9. Transmission mechanisms for MPEG-7 streams** – to allow transmission of MPEG-7 descriptions over a variety of physical media using appropriate protocols.

Note: MPEG-7 descriptions will need to be transmitted over a variety of physical media using a variety of protocols.

- 10. Buffer management within an MPEG-7 capable device** - to provide local storage for MPEG-7 descriptions for as long as necessary.

Note: The requirement for local storage depends on the specific destination of MPEG-7 data, e.g. search engine depends on nature of application: real-time interpretation of MPEG-7 data, delay between content data and description, temporal validity of description data etc.

- 11. File format for MPEG-7** - required for MPEG-7 descriptions.

6. References

[1] Maes, P. (1994) "Modeling Adaptive Autonomous Agents," Journal of Artificial Life, vol. 1, no. 1/2, pp. 135 - 162.

[2] MPEG Requirements Group, "MPEG-7: Context and Objectives", Doc. ISO/MPEG N2460, MPEG Atlantic City Meeting, October 1998

[3] MPEG Requirements Group, "Applications for MPEG-7", Doc. ISO/MPEG N2462, MPEG Atlantic City Meeting, October 1998

Annex A - Open issues

From the many discussions about MPEG-7, a few interesting remarks and questions that may be worthwhile to consider in the future are:

- Requirements for interface between MPEG-7 descriptions and adopted text description schemes
- Set of examples to illustrate the terminology
- Objectivity - "Do we agree on what we mean".
- Registration - what is our position on this ?

Annex B - Ongoing discussion with respect to the DDL

Further studies are being conducted on the DDL within the MPEG-7 requirements process. These studies may result in a separation of DDL execution functionality from DDL descriptive functionality.

Additional study is also required to determine if the execution capability of the DDL should include DS composition operations. Partial DS extraction and subsequent extension of the extracted DS is an area of study within DDL descriptive functionality.

The provision of tools for DS and D development and their relationship with the DDL also needs further investigation.

The need for the DDL to explicitly provide additional capability for the support of real time applications is also being studied.

At present the DDL will not support presentation capabilities. This may become a concern for other parts of the MPEG standardisation process such as MPEG-4. There is capability within MPEG-4 systems activities that may provide ready-made solutions for presentation, multiplexing, streaming and the real time control of the relationships between MPEG-7 descriptions and the related data.

Annex C – Open issues with respect to the Systems Requirements

Open issues regarding the relation between MPEG-4 and MPEG-7

- What is the feasibility/desirability of the attachment of MPEG-7 descriptions to objects in an MPEG-4 BIFS Scene Graph?
- It is believed that an MPEG-7 description stream is an additional MPEG-4 elementary stream type.
- Are there requirements that explicitly drive the linkage between MPEG-4 and MPEG-7?
- Can MPEG-4 systems be regarded as a candidate implementation framework for MPEG-7?

Open issues regarding the scope of the system requirements

There is a need to clarify the scope of MPEG-7 systems. Should this encompass other areas outside what are at present considered to be the normative parts of MPEG-7 (DDL, DS, D etc.) for example:

- Presentation of MPEG-7 descriptions
- Presentation of the media (elementary streams) the descriptions relate to.
- Support for the execution of code embedded in the descriptions:
 - Relating to the presentation of media.
 - Relating to the presentation of descriptions.
 - Relating to the manipulation of description schemes.
 - Relating to the manipulation of descriptions.
 - Relating to manipulation of media.
 - Relating to event and interaction issues.
- Transportation protocol/s.
- Support for an environment for manipulating description schemes.
- Support for feature extraction.
- Support for real time MPEG-7 applications.

Open issues regarding distributed architecture

There have been some references during discussions on system requirements to the use of COM/DCOM or CORBA as candidate implementation technologies for the XM. It is apparent that much of the thinking about MPEG-7 implicitly includes ideas about communications with and the searching of substantial numbers of multimedia databases. This together with the consideration of CORBA/DCOM as implementation technologies for the XM introduces consideration of a distributed architecture for MPEG-7. These questions then arise:

- Are there requirements that indicate the need for an explicitly distributed architecture for MPEG-7?
- What additional requirements would such architecture place on the MPEG-7 systems layer?