

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

**ISO/IEC JTC1/SC29/WG11
MPEG99/M5605
December 1999, Maui, USA**

Title: Validation Experiment on the Ordering Key DS and an Unified Syntax for the Weight DS.
Status: Report and Proposal
Source: University of Brescia
Authors: Nicola Adami, Alessandro Bugatti, Riccardo Leonardi and Lorenzo Rossi

1. Introduction

This document presents the experimental results for validating the Ordering Key DS (DS5) in the context of the core experiment of the Weight DS [6]. At the Melbourne MPEG meeting, in October 1999, the aforementioned core/validation experiment was planned in order to show the validity of a set of proposals (Weight DS [5], Descriptor Usage DS [8], Fidelity DS [9], Pointofview DS [4] and Ordering Key DS [7]). In a few words, all these DSs play the role of highlighting, by means of some kinds of weights, description information (DSs or Ds) relevant to user queries. They can provide, for example, confidence measure, priority, fidelity, relevance feedback, information for ordering etc. in order to facilitate user queries and browsing.

As we said, the document focuses on the VE of the Ordering Key DS. Besides it presents an unified DS, in MPEG-7 DDL syntax, that addresses all the different functionalities proposed by the several DSs involved in the CE. In our case, the provided functionality deals with the concept of ordering, as we consider the need of providing users with ordering mechanisms a very relevant issue for MPEG-7.¹ Such ordering mechanisms are derivable from descriptors (e.g. a set of key – frames ordered on the basis a color descriptor or a set of sounds ordered by means of an audio loudness D). However a possible large variety in the types of descriptors composing a description could lead to a consequent high number of ordering criteria to arrange description items. Therefore we propose that the description provider (it could be different from the content provider) should also select a reduced set of descriptors allowing to order a subset of description elements (e.g. key frames, events etc.) pertinent to the MM document being described [7].

This contribution is organized as follows. In Section 2, the motivations for introducing Ordering Key DS in the Generic AV DS are given. Section 3 briefly presents the structure of the Ordering Key DS by means of UML notation and MPEG-7 DDL as well. Moreover the

¹ The requirements for ordering description items was introduced, among the functional requirements (# 12), at the MPEG Vancouver meeting [2].

section discusses about the possible locations of the DS within the Generic AV DS. In Section 4 is shown the output of the experiment: a smart browsing of ordered elements belonging to the test set of MPEG-7 video material. Finally, in Section 5, is proposed a unified DDL structure for the DSs involved in the CE.

2 Motivation

The possibility to have items ordered and orderable is very important for facilitating every kind of query in a high variety of contexts. A traditional example is the index of a technical book: it consists in a list of alphabetically ordered key words allowing readers to retrieve parts of interest. Another example is given by a software interface for reading the content of a computer directory. Typically the contained files have several attributes (name, size, type, date of last modification etc.). Thanks to these attributes an user can search files of interests, even without an exact knowledge of their names, by ordering these elements according to different criteria: the date of last modification if he/she is looking for a recent file, the size if he/she has an idea of that and so on. Thus the possibility of ordering, if properly adopted, can also facilitate queries of unknown elements by restricting the search field, allowing users to browse in a reduced number of items.

The context of the description of MM documents presents some analogies with the aforementioned one. The described objects (images, sounds, video sequences etc.) present several attributes (i.e. descriptors) which can lead to several ordering criteria (e.g. a set of key frames of a violent movie ordered according to the level of the underlying audio loudness). These criteria could be adopted by a software interface so that to allow users to access ordered data. However in MPEG-7 descriptions, there can be so many kinds of Ds making quite difficult the selection of ordering criteria for the items to be arranged, also depending on the kind of item itself. Hence we believe that it would be important to have a mechanism able to highlight some of the available ordering criteria (i.e. descriptors) which can be relevant to user needs according to the involved application context. For instance, an audio loudness descriptor could be relevant as ordering mechanism in documents dealing with football games but probably less important for documentaries about, e.g., ancient furniture.

Thus we propose to incorporate in the Generic AV DS a very simple kind of DS (the Ordering Key DS) which would have the role to highlight descriptor whose values are useful for ordering some description items (in a way relevant to the application scenario), so that to facilitate selection processes (smart browsing).

3 Ordering Key DS

The Ordering Key DS carries the information about the object to order and the set of descriptor suggested for ordering. The structure of the Ordering Key DS is very simple (see Figure 1). It consists of a set of Reference Ds which refer to a corresponding set of Ds. The attribute "Idref" indicates the class of the ordered objects. The attribute "Purpose" of the Reference D indicates the purpose to whom that particular ordering key is useful. For instance, an audio loudness descriptor can be useful for facilitating the selection of shots corresponding to goals in a football match document. Thus in this case, a Reference should correspond to the "Audio Loudness" D and the attribute "Purpose" could be a string such as "goal". This DS can be attached to several points of the actual Generic AV DS (e.g. Segment DS and it can order both visual and audio segments). It can also be applied to Summary DS for providing a representation of Key-frames with an order not chronological.

```

<!-- MPEG-7 DDL Structure of the Ordering Key DS -->
<!-- In the current MPEG-7 AV DS, Ordering Key DS can be attached to any DS
      node which is part of a hierarchy, e.g. Segment DS and Summary DS -->

<DSType name='OrderingKeyDS'>
  <attribute name='Idref' type='IDREF' required='true' />
  <DTypeRef name='Reference' minOccurs='1' maxOccurs='*' />
</DSType>

<DType name='Reference'>
  <attribute name='Idref' type='IDREF' required='true' />
  <attribute name='Purpose' type='string' required='true' />
  <empty/>
</DType>

```

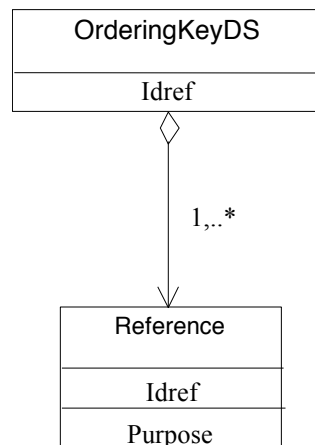


Figure 1: The UML structure of the Ordering Key DS.

3.1 Key Item DS

In the Melbourne proposal M5267 [7], was presented a DS called Key Item DS, which enclosed the Ordering Key DS. Hence we believe that the ordering Key DS could be very useful if enclosed within such a DS. The *Key Items DS* can be seen like a generalization of the concept of key-words in the context of MM documents. It is composed of several DSs (*Key-events DS*, *Key-objects DS*, *Key-images DS*, *Key-sounds DS*), each of them consisting in a set of key-items representative for certain type of description element (e.g., events, objects etc.).

In Figure 2, the structure of the *Key-images DS* is shown. The other key-items DS (*Key events DS*, etc.) reflects the represented structure. As we can see, each *Key-images DS* can contain an arbitrary number of sub-key-images, and therefore forms a hierarchy (tree) of key-items. For example, an AV document of a soccer match can contain, at a higher level of this hierarchy (then at a higher level of abstraction), a key-image representing *goals*, one representing *penalties*, another one representing *corners*, etc. At a lower level of the hierarchy a goal key-image can contains a sets of key-frames, each of them representing the shots describing that goal.

The *Key-images DS* consists of the *Ordering Key DS*, the *Attribute DS*, the *Links to segment DS*, and the *Representation DS* (see Figure 2). The *Ordering Key DS* presents a list of applicable ordering mechanisms for the listed key-items (e.g., having defined a set of key-images in a violent movie; the associated ordering key may be the level of underlying audio loudness descriptor). According to the tree structure of the *Key-images DS*, we can assign different sets of ordering-keys at different key-items pertaining at different level of the

hierarchy. Obviously these ordering keys can be applied to order other description items (for example, all key-frames and mosaics). The *Attribute DS* characterizes the key-item itself. Note that it could be (at least partially) accessed thanks to the *Links to segment DS*. The *Links to segment DS* identifies the parts (e.g., temporal segments, K-frames, ...) in the sub-DS the key-item refers to. *Links to segment DS* allows clearly to refer to descriptors associated with such parts of the sub-DS. The purpose of the *Representation DS* is the visualization and the presentation of the key-items.

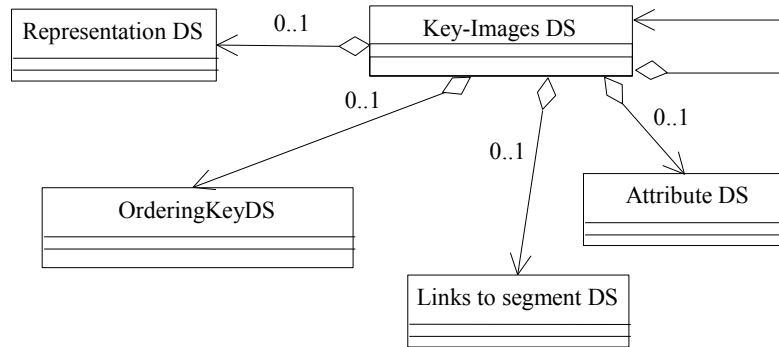


Figure 2: Key-images DS.

4 Validation experiment

The validation experiments for the Ordering Key DS have been conducted with the purpose of demonstrating its effective usefulness for efficient browsing of AV material. In particular it important to show that the highlight of the key descriptor for ordering items (i.e. the Ordering Keys) can help users in smart browsing by allowing an easier retrieval of elements of interest. Thus we will show how the mechanism of selecting suitable ordering criteria works, for navigating throughout the adopted MM documents criteria.

4.1 Content set

The experiments has been conducted over the following material of the official MPEG-7 content set. The table below indicates the file names and their corresponding CD numbers. The adopted materials belong to several categories (sport, news and documentaries) in order to have a large variety of conditions and thus a variety of possible ordering keys.

| ID | FileName | MPEG CD # | Title |
|----|-----------------------------|-----------|----------------------------|
| 1 | portugese\jornaldanoite.mpg | 14 | Jornal da noite |
| 2 | misc1.mpg | 20 | Basketball |
| 3 | news2.mpg | 18 | Soccer: Spain - Sweden |
| 4 | Samsung\foot1022.mpg | 28 | Samsung soccer |
| 5 | kbs\basketball.mpg | 26 | Basketball |
| 6 | kbs\golf.mpg | 26 | Golf |
| 7 | news2.mpg | 18 | Cycling |
| 8 | rtve\culture.mpg | 33 | Documentary: Santillana... |

4.2 Experiments

Example 1

This experiment simply shows a set of segments from the documentary “Santillana de Mar” having a textual label and ordered according to alphabetical order. This is a very obvious but useful ordering criteria for textual material. Another important well know criteria is the chronological order.

Thus if we provide the DDL instantiation we have:

```
<OrderingKeyDS>  
  <Reference idref= “Segment label ” Purpose= “Alphabetical order”/>  
  <Reference ..../>  
</OrderingKey DS>
```

| ID | Start frame # | End frame # | Title |
|----|---------------|-------------|----------------------------|
| 2 | 866 | 4093 | Agricultural jobs |
| 11 | 28005 | 32005 | Ancient houses |
| 5 | 8471 | 9256 | Botanical garden |
| 8 | 12919 | 16597 | Downtown tour |
| 14 | 42498 | 44019 | End titles |
| 12 | 32007 | 40355 | Furnishing |
| 13 | 40357 | 42496 | Horse tour I |
| 6 | 9258 | 9520 | Horse tour II |
| 3 | 4095 | 6013 | Horse tour III |
| 1 | 0 | 864 | Introduction |
| 4 | 6015 | 8469 | Prehistoric Graffiti |
| 10 | 22650 | 28003 | San Cristobal Convent |
| 9 | 16599 | 22648 | San Sebastian Church |
| 7 | 9522 | 12917 | Typical jobs in Santillana |

Example 2

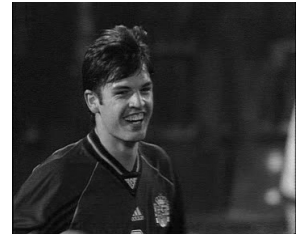
In the following table, we show how an ordering key based on audio loudness descriptor could facilitate the retrieval of goals and other salient events in a football match. The examined document has a duration of 15 minutes and was partitioned by our software in 78 shots. The audio descriptor value adopted consists in the average value of the audio peak value associated to the shots.

```
<OrderingKeyDS>  
  <Reference idref = “Audio loudness” Purpose = “Remarkable game events”/>  
  .....  
  <Reference ..../>  
</OrderingKey DS>
```

| ID | Average audio I. | Event |
|----|------------------|--------------------------|
| 10 | 4099366666666,67 | Morientes's goal – 10' |
| 1 | 4017675000000 | |
| 39 | 3978800000000 | Morientes's goal – 38' |
| 9 | 3307650000000 | Kick-off |
| 12 | 3235400000000 | Exultance after 1st goal |
| 43 | 3012700000000 | Exultance after 2st goal |
| 2 | 2901600000000 | |
| 25 | 2852800000000 | |
| 32 | 2802100000000 | |
| 16 | 2477600000000 | |
| 44 | 2208000000000 | |
| 3 | 2127300000000 | |
| 51 | 2074100000000 | |
| 31 | 2048300000000 | |
| 18 | 1997300000000 | |
| 53 | 1900566666666,67 | |
| 36 | 1884900000000 | |
| 8 | 1874500000000 | |
| 47 | 1871866666666,67 | |
| 6 | 1713100000000 | |
| 76 | 1473900000000 | |
| 59 | 1473200000000 | |
| 69 | 1416300000000 | |
| 34 | 1349200000000 | |
| 20 | 1330300000000 | |
| 28 | 1230600000000 | |
| 61 | 1180782500000 | |
| 27 | 1173877500000 | |
| 38 | 1092100000000 | Morientes's goal – 38' |
| 56 | 1061430000000 | |
| 62 | 9850300000000 | |
| 48 | 9695900000000 | |
| 35 | 9132200000000 | |
| 67 | 8801850000000 | |



13826-14957



22109-22318



13156-13825



15089-15204



22451-22716

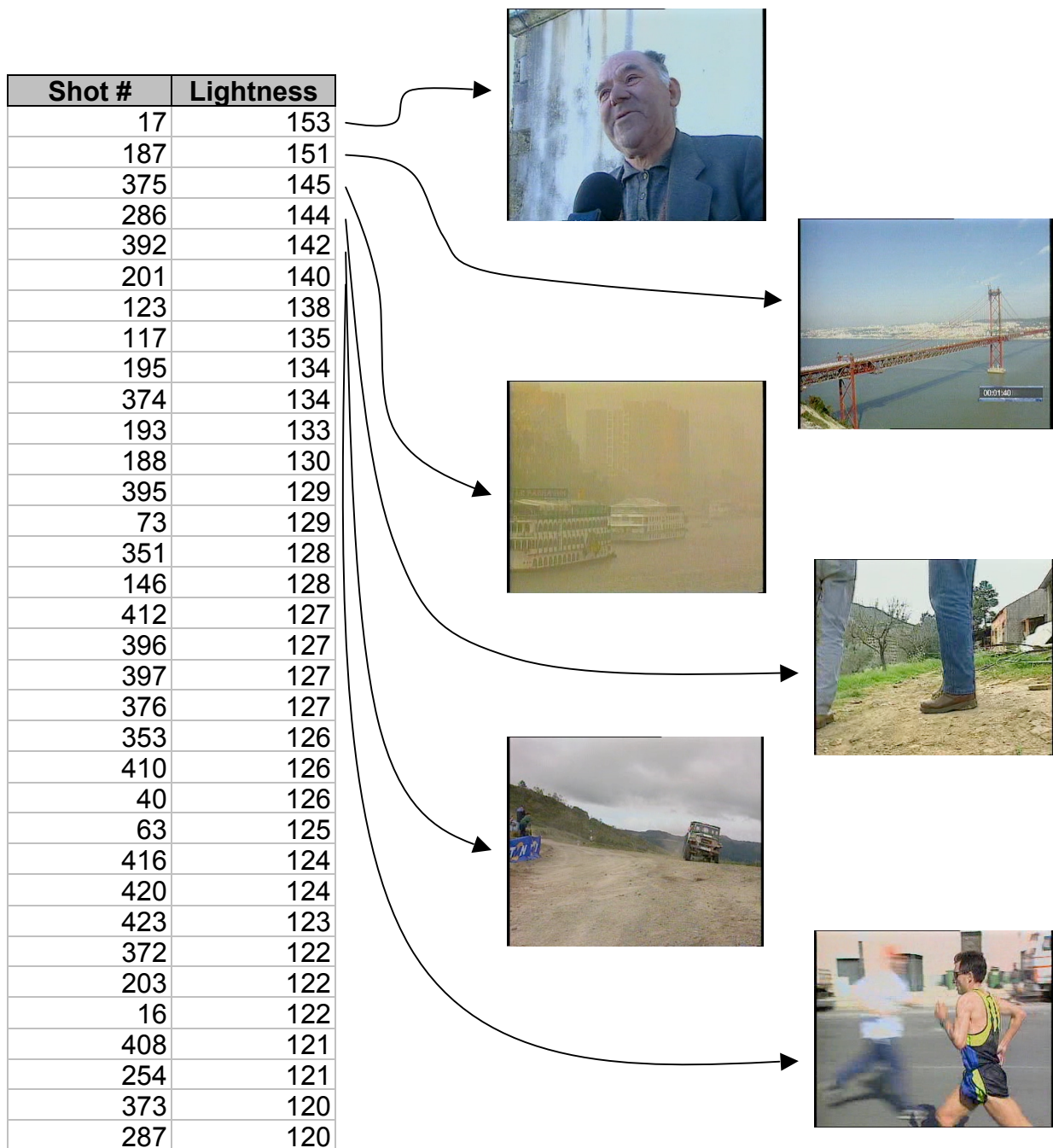


21724-22108

Example 3

In this set of experiments carried on the “Journal da Noite” (CD #), we extracted the Lightness value from the HSV color space. This simple parameter allows, e.g., to easily select key-frames associated to daily outdoor camera records. This information could be easy to retrieve

```
<OrderingKeyDS>
  <Reference idref= “HSVLightness” Purpose= “Retrieve daily outdoor scenes”/>
  <Reference .../>
</OrderingKey DS>
```



5 Generalized Weight DS: Hint DS

The table below quickly summarizes the functionalities of the of the DSs involved in the CE.

| | Functionality | Where does it apply? | Content | Comment |
|------------------------------|--|---|--|--|
| Descriptor usage DS M5003 | Priority between D's associated to one segment | Segment | List of (Priority value / Confidence value / Reference to D) | Priority = weight value of weight DS in case 2. Confidence = estimate of the validity of the descriptor |
| Fidelity DS M5101 | How well a father element represents his children. | HighlightLevel DS and Segment DS. | Fidelity value [0,1]. | No reference to cluster needed |
| Ordering Key DS M5267 | Allow to order elements according to the criteria associated to a list of descriptors. The purpose is smart browsing. | Any DS node which is part of a hierarchy (e.g., segment, event, object, ...). | List of references to Ds with associated purposes. | |
| Point of view DS M5200 | Weight associated with a segment | Segment | List of (point of view 'string' / priority values). | Many point of view can be associated to each segment |
| Segment hint D | Important/not important attribute | Segment | Binary | Exists in current G. AV DS |
| Weight DS M5005 | <ol style="list-style-type: none"> Weight between different sub segments under a segment decomposition Weight between D's associated to one segment. | <ul style="list-style-type: none"> Syntactic DS Segment DS ColorTexture DS, SemanticDS Object DS Event DS | <ul style="list-style-type: none"> Type field (feature,...) Value field Reliability value field | <ol style="list-style-type: none"> Includes reference to the sub-segments or something else to avoid ambiguity Includes reference to D |

We have defined a unified MPEG-7 DDL syntax in order to have a DS (Hint DS) able to address all the functionalities provided by the several DSs. Koichi Emura from Matsushita Corporation provided the XML DTD for the DDL Hint DS (see Annex A) and thus validated the syntax of *Hint DS* thanks to the XML Parser for Java 2.0.15 (IBM Alpha Works). However this DS presents some slight modifications with respect to the one submitted to the AHG on MPEG-7 DS CE reflector. The attributes *reliability* and *id* have been enclosed in the *Hint DS*. Moreover the *Reference D* now contains also the attribute *Purpose*, since it is needed by Ordering Key DS. Other modifications deal with the comments of the subDSs.

```
<!-- ##### -->
<!-- Definition of basic building blocks for HintDS 1.1 -->
<!-- ##### -->

<DSType name='HintDS'>
  <attribute name='HintDSname' ype='HintDStype' required='true' />
  <attribute name='reliability' type='real' />
  <attribute name='id' type='ID' />
  <attribute name='idref' type='IDREF' />
  <DSTypeRef name='HintValueDS' minOccur='1' maxOccur='*' />
</DSType>

<DSType name='HintValueDS'>
  <DSTypeRef name='ReferenceDS' minOccur='0' maxOccur='*' />

```



```

        <DSTypeRef name='ReferenceValueDS' minOccurs='0' maxOccurs='1' />
</DSType>

<DSType name='ReferenceDS'>
  <choice>
    <DTypeRef name='Reference' />
    <DtypeRef name='Keyword' />
  </choice>
</DSType>

<DType name='Reference'>
  <attribute name='idref' type='IDREF' required='true' />
  <attribute name='purpose' type='String' />
  <empty />
</DType>

<DType name='Keyword'>
  <attribute name='Refname' type='string' required='true' />
  <empty />
</DType>

<DSType name='ReferenceValueDS'>
  <choice>
    <DTypeRef name='ReferenceValue' />
    <DtypeRef name='FidelityValue' />
    <DtypeRef name='PriorityValue' />
  </choice>
</DSType>

<DType name='ReferenceValue'>
  <attribute name='Value' type='real' required='true' />
  <empty />
</DType>

<DType name='FidelityValue'>
  <attribute name='Value' type='Fidelity' required='true' />
  <empty />
</DType>

<DType name='PriorityValue'>
  <attribute name='Value' type='Priority' required='true' />
  <empty />
</DType>

<datatype name='HintDSType'>
  <basetype name='string' />
  <enumeration>
    <literal>DescriptorUsageDS</literal>
    <literal>FidelityDS</literal>
    <literal>OrderingKeyDS</literal>
    <literal>PointOfViewDS</literal>
    <literal>WeightDS</literal>
  </enumeration>
</datatype>

<datatype name='Fidelity'>
  <basetype name='real' />
  <minInclusive>0.0</minInclusive>
  <maxInclusive>1.0</maxInclusive>
</datatype>

```

```

<datatype name='Priority'>
  <basetype name='real' />
  <minInclusive>1.0</minInclusive>
  <maxInclusive>5.0</maxInclusive>
</datatype>

```

Below are the DSs participating to the CE on Weight DS. They are specified as subDSs of the Hint DS. The comments help to understand how the specialization is done.

```

<DSType name='FidelityDS'>
  <subDSOf name='HintDS' />
  <!-- Attached to HighlightLevelDS and SegmentDS -->
  <!-- HintDSname = 'FidelityDS' -->
  <!-- Occurrences of ReferenceDS = 0 -->
  <!-- ReferenceValueDS contains FidelityValue D -->
</DSType>

<DSType name='PointOfViewDS'>
  <subDSOf name='HintDS' />
  <!-- Attached to SegmentDS -->
  <!-- HintDSname = 'PointOfViewDS' -->
  <!-- ReferenceDS contains Keyword D -->
  <!-- ReferenceValueDS contains PriorityValue D -->
</DSType>

<DSType name='DescriptorUsageDS'>
  <subDSOf name='HintDS' />
  <!-- Attached to SegmentDS -->
  <!-- HintDSname = 'DescriptorUsageDS' -->
</DSType>

<DSType name='WeightDS'>
  <subDSOf name='HintDS' />
  <!-- Attached to Syntactic DS, Segment DS, ColorTextureDS,
  SemanticDS, Object DS, Event DS -->
  <!-- HintDSname = 'WeightDS' -->
</DSType>

<DSType name='OrderingKeyDS'>
  <subDSOf name='HintDS' />
  <attribute name='idref' type='IDREF' />
  <!-- Attached to Key-ItemDS, SummaryDS -->
  <!-- HintDSname = 'OrderingKeyDS' -->
  <!-- Occurrences of ReferenceValueDS = 0 -->
  <!-- ReferenceDS contains Reference D -->
</DSType>

```

6 Summary

This document presents the motivation for including the Ordering Key DS in the Generic AV DS. We specified the syntax of the DS in MPEG-7 DDL. Besides we have also presented a description scheme (Hint DS) which allows to address all the functionalities provided by the DSs participating to the CE (i.e. Weight DS, Descriptor Usage DS, Fidelity Ds we have carried some experiments in several browsing scenarios. More details about the experiment as well as a demo will be provided during the Maui meeting.

As a result of these experiments, we observed that the availability of ordering key descriptors thanks to the Ordering Key DS allows smart browsing throughout multimedia material by helping users to better access subparts of interest of the multimedia material.

Hence we recommend the inclusion of this simple DS in the MPEG-7 Generic AV DS by means of the incorporation of the Hint DS.

7 References

- [1] AHG on MPEG-7 DS, “MPEG-7 Generic AV Description Schemes (V0.7)”, ISO/IEC JTC1/SC29/WG11 MPEG99/N2966, Melbourne, Australia, October 1999.
- [2] AHG on MPEG-7 Requirements, “MPEG-7 Requirements Document V.9”, ISO/IEC JTC1/SC29/WG11 MPEG99/N2859, Vancouver, Canada, July 1999.
- [3] DDL Group, “MPEG-7 Description Definition Language V 2”, ISO/IEC/JTC1/SC29/WG11 MPEG99/N2997, Melbourne, Australia, October 1999.
- [4] K. Emura and T. Munetsugu, “Proposal of Pointofview DS”, ISO/IEC/JTC1/SC29/WG11 MPEG99/M5200, Melbourne, Australia, October 1999.
- [5] J.-S. Lee et al., “Weight DS in Generic Audio Visual DS proposal”, ISO/IEC/JTC1/SC29/WG11 MPEG99/M5005, Melbourne, Australia, October 1999.
- [6] J.-S. Lee, S. Sull, S. Sekiguchi, K. Emura and R. Leonardi, “MPEG-7 Core/Validation Experiment on the Weight DS”, ISO/IEC/JTC1/SC29/ WG11 MPEG99/N2972, Melbourne, Australia, October 1999.
- [7] R. Leonardi, G. Paltenghi and L. Rossi, “Limitations of the MPEG-7 Generic DS: Reorganizing the Syntactic/Semantic DS’s”, ISO/IEC/JTC1/SC29/WG11 MPEG99/M5267, Melbourne, Australia, October 1999.
- [8] S. Sekiguchi et al., “Proposal of the descriptor usage DS”, ISO/IEC/JTC1/SC29/WG11 MPEG99/M5003, Melbourne, Australia, October 1999.
- [9] S. Sull et al. “Efficient and effective search and browsing using fidelity”, ISO/IEC/JTC1/SC29/WG11 MPEG99/M5101, Melbourne, Australia, October 1999.

Annex A

```
<!-- ##### -->
<!-- Definition of HintDS and its basic building blocks          -->
<!-- Document Type Definition(DTD) Ver.0.1 by K.Emura          -->
<!-- ##### -->

<?xml version="1.0"?>

<!ENTITY % HintDSType
      "(DescriptorUsageDS|FidelityDS|OrderingKeyDS|PointOfViewDS|WeightDS)"
>
<!ENTITY % ReferenceValueType
      "(ReferenceValue|FidelityValue|PriorityValue)">

<!ELEMENT HintDS (HintValueDS+)>
<!ATTLIST HintDS
      HintDSname %HintDSType;      "WeightDS">

<!ELEMENT HintDSValueDS (ReferenceDS*, ReferenceValueDS*)>

<!ELEMENT ReferenceDS (Reference|Keyword)>
<!ELEMENT Reference      EMPTY>
<!ATTLIST Reference
      idref IDREF #REQUIRED>
<!ELEMENT Keyword      EMPTY>
<!ATTLIST Keyword
      Refname      CDATA #REQUIRED>

<!ELEMENT ReferenceValueDS %ReferenceValueType;>
<!ELEMENT ReferenceValue      EMPTY>
<!ATTLIST ReferenceValue
      Value CDATA #REQUIRED>

<!ELEMENT FidelityValue      EMPTY>
<!ATTLIST FidelityValue
      Value CDATA #REQUIRED>

<!ELEMENT PriorityValue      EMPTY>
<!ATTLIST PriorityValue
      Value CDATA #REQUIRED>
```