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

ISO/IEC JTC1/SC29/WG11
MPEG00/M6768
January 2001, Pisa

Source: **Ordering Key DS Core Experiment participants**
Status: **Report**
Title: **CE report on OrderingKey D**
Author: **Alessandro Bugatti (University of Brescia)**
Contributors: **Riccardo Leonardi (University of Brescia)**
Benoit Mory (Philips)
Yoshiaki Shibata (Sony)

1 Introduction

In this document, we report the results of our core experiment on the Ordering Key D. The concept of the Ordering Key D has been proposed for the first time in Beijing meeting [m6227] as a part of the generic Weight DS. According to the discussions during that meeting the Weight DS has been specialized in order to obtain exclusively ordering functionalities [w3473]. At the same time the demo showed in Beijing seemed to demonstrate the useful of the ordering beetwen entities of various types. At the following meeting a specialized and improved semantic and syntax have been proposed. In spite of positive results about its utility, as shown in the demo [m6588] and also in the document [m6491], it was not clear the use of XPath syntax as mechanism to obtain reference to entities. At the La Baule meeting the DDL group had not yet decided about the use of XPath as reference datatype. In conclusion the aim of this CE was not to demonstrate the OrderingKey effectiveness but to investigate about its syntax [w3630].

2 Outline of CE activity

2.1 Participants and functionalities to be validated

The main activities and participants are summarized in the below table

Name	Company	Functionality
Alessandro Bugatti, Riccardo Leonardi	University of Brescia	Follow the evolution of the XPath datatype, evaluate the validity of the OrderingKey D definition in XM 5.0, provide the software for the XM integration.
Benoit Mory	Philips	Cross-check of results of Brescia experiments

3 Results

3.1 The XPath datatype

As first step we used the new definition of XPath datatype, as it has defined in the current version of Text of the CD [w3705]. The `XPath` mechanism has the advantage, respect to `uriReference` or `IDREF,` that the reference target has not to instantiate an `ID` attribute.

```
<!-- ################################################ -->
<!--   Definition the Reference Datatype              -->
<!-- ################################################ -->

<complexType name="ReferenceType">
      <choice minOccurs="1" maxOccurs="1">
            <element   name="HRef"   type="uriReference"   minOccurs="1"
maxOccurs="1"/>
            <element     name="IDRef"     type="IDREF"     minOccurs="1"
maxOccurs="1"/>
            <element  name="XPath"  type="mpeg7:xPathType"  minOccurs="1"
maxOccurs="1"/>
      </choice>
</complexType>
<!-- ################################################ -->
```

```
<!--   Definition of the xPath Datatype                -->
<!-- ############################################### -->

<simpleType name="xPathType">
      <restriction base="string">
            <pattern value="(/|//|\.|\.\.|([\i-[:]][\c-[:]]*)
            (\[\d\])?|)+(@[\i-[:]][\c-[:]]*)?"/>
      </restriction>
</simpleType>
```

3.2 The OrderingKey D

The OrderingKey D is a tool that allows to represent the descriptors (or part of descriptors) that can be used to order DS descriptions of the same type (e.g, segment, objects, or events). An example of this functionality could be the ordering key "Dominant color" applied to a set of key frames that represent shots. Using this ordering key is possible to obtain an effective ordering in which all the key frames containing the speaker are clustered together (see Appendix A)

3.2.1 OrderingKey D Syntax

```
<!-- ############################################### -->
<!-- Definition of OrderingKey D                     -->
<!-- ############################################### -->

<complexType name="OrderingKeyType">
      <complexContent>
            <extension base="mpeg7:DType">
                  <sequence minOccurs=”1” maxOccurs=”1”>
                        <element name="Selector" type="mpeg7:xPathType"
                                    minOccurs=”1” maxOccurs”1”/>
                        <element name="Field" type="mpeg7:xPathType"
                                    minOccurs="1"
maxOccurs="unbounded"/>
                  </sequence>
                  <attribute name="name" type="string" use="optional"/>
                  <attribute        name="semantics"        type="string"
use=”optional”/>
                  <attribute        name="direction"        use="default"
value="descending">
                        <simpleType>
                              <restriction base="string">
                                    <enumeration value="descending"/>
                                    <enumeration value="ascending/"/>
                              </restriction>
                        </simpleType>
                  </attribute>
            </extension>
      </complexContent>
</complexType>
```

3.2.2 OrderingKey D Semantics

Semantics of OrderingKey D:

Name	*Definition*
`OrderingKey`	Useful descriptors (or part of descriptors) that can be used to order DS descriptions of the same type.
`Selector`	xPath expression that specifies the DS descriptions (e.g., Segment , Object, and Event DS descriptions) that need to be ordered.
`Field`	xPath expression that specifies the ordering key, i.e., the descriptors (or parts of descriptors) that can be used to order the DS descriptions specified by the Selector element. The descriptors (or parts of descriptors) used as ordering keys have to be either strings or single numeric values (e.g., integer or floats) so that an alphabetical or magnitude criterion can be used in the ordering process. It is not possible to use complex descriptors (or part of descriptors) as ordering keys to avoid ordering ambiguities (e.g. the color histogram is not a valid ordering key because it is a vector of integers and it would not be clear for the applications how to order).
`name`	String that specifies a label for the instance of the OrderingKey D. This attribute is optional.
`semantics`	String that specifies the semantics of the ordering. It can be a simple word indicating the descriptor used (e.g. if it has been used the id of video segments as ordering key semantics could be "accordingtoSceneID") or a more complex textual description of the ordering results (e.g. if PanSpeed inside video segments of a soccer match is used as ordering key a possible semantics could be "Shots containing a close-up of a player")
`direction`	String that specifies the direction of the ordering. The possible directions of the ordering can be “ascending” and “descending”. These are defined as follows: • “ascending” indicates that the DS descriptions need to be ordered in ascending order. • “descending” indicates that the DS descriptions need to be ordered in descending order. The attribute value is “descending” by default.

In the example below, two camera motion descriptors (the track left and zoom in descriptor) are used to obtain some useful ordering of video segments. In the first OrderingKey D description, the ordering is based on the left track part of the CameraMotion D so the shots containing the most amount of “action” would appear first in the ordered list. In the second OrderingKey D description, the ordering is based on zoom part of the CameraMotion D description so the shots containing close-ups would appear first in the ordered list. In this case XPath is the only mechanism in order to individuate the OrderingKey, because TRACK_LEFT_F and ZOOM_IN_F have not an ID attribute. Using an absolute path within the Field element it is possible to select the Ordering Key without ambiguity.

```
<?xml version="1.0" encoding="UTF-8"?>
<Mpeg7Unit
    xmlns=http://www.mpeg7.org/2001/MPEG-7_Schema
    xmlns:xsi=http://www.w3.org/1999/XMLSchema-instance
    xsi:schemaLocation="http://www.mpeg7.org/2001/MPEG-7_Schema
        C:\users\Bugatti\UPMSAnnotator\mds4CE.xsd">

    <!--  Details of the description of the video segments have  -->
    <!--  been omitted for simplicity.                           -->
    <VideoSegment id="id1">
        <SegmentDecomposition DecompositionType="temporal">
            <VideoSegment id="id2">
                <MediaTime>
```

```
                    <MediaTimePoint/>
                </MediaTime>
                <CameraMotionNumSegmentDescription="3"
DescriptionMode="1">
                    <CameraMotionSegment>
                        <FractionalPresence>
                            <TRACK_LEFT_F>0.23</TRACK_LEFT_F>
                            <ZOOM_IN_F>0.56</ZOOM_IN_F>
                        </FractionalPresence>
                        <Amount/>
                    </CameraMotionSegment>
                </CameraMotion>
            </VideoSegment>
            <VideoSegment id="id3">
                <MediaTime>
                    <MediaTimePoint/>
                </MediaTime>
                <CameraMotion NumSegmentDescription="3"
DescriptionMode="1">
                    <CameraMotionSegment>
                        <FractionalPresence>
                            <TRACK_LEFT_F>0.1</TRACK_LEFT_F>
                            <ZOOM_IN_F>0.8</ZOOM_IN_F>
                        </FractionalPresence>
                        <Amount/>
                    </CameraMotionSegment>
                </CameraMotion>
            </VideoSegment>
            <VideoSegment id="id4">
                <MediaTime>
                    <MediaTimePoint/>
                </MediaTime>
                <CameraMotion NumSegmentDescription="3"
DescriptionMode="1">
                    <CameraMotionSegment>
                        <FractionalPresence>
                            <TRACK_LEFT_F>0.67</TRACK_LEFT_F>
                        </FractionalPresence>
                        <Amount/>
                    </CameraMotionSegment>
                </CameraMotion>
            </VideoSegment>
        </SegmentDecomposition>
        <MediaTime>
            <MediaTimePoint/>
        </MediaTime>
    </VideoSegment>

    <OrderingKey name="ActionOrdering"
                 semantics="The shots containing more action">
        <Selector>
            /VideoSegment/SegmentDecomposition/VideoSegment
        </Selector>
        <field>
/VideoSegment/SegmentDecomposition/VideoSegment/CameraMotion/CameraMoti
onSegment/FractionalPresence@TRACK_LEFT_F
        </Field>
```

```
    </OrderingKey>

    <OrderingKey name="Close-up Ordering"
                 semantics="The shots containing close-up of a player">
        <Selector>
            /VideoSegment/SegmentDecomposition/VideoSegment
        </Selector>
        <field>
/VideoSegment/SegmentDecomposition/VideoSegment/CameraMotion/CameraMoti
onSegment/FractionalPresence@ZOOM_IN_F
        </Field>
    </OrderingKey>

</Mpeg7Unit>
```

4 Software for the XM

As final step of this CE the software for the XM has been developed and sent to the XM repository. This software contains a Search Tool and a Client Application. In case of OrderingKey D the application is a little bit different from the usual matching applications. In fact our application reads a query in a DDL file containing the OrderingKey D and uses this OrderingKey to try to order the selected entities in the input DDL file containing an MPEG-7 description. As output the application gives a set of entities ordered by the OrderingKey.

5 Conclusions and recommendations

Based on the previous CE on the OrderingKey D, we have inserted the use of XPath datatype according to its definition in CD. This mechanism to indicate the ordering key and the objects that have to be ordered has been experimented in the previous CE [m6588]. Finally we have developed the software of the XM.
We recommend to insert the OrderingKey in the Study of CD at the end of Pisa meeting.

6 References

[m6227] ISO/IEC JTC1/SC29/WG11/M6227, CE report on Weight DS, A. Bugatti et al., Beijing, July 2000.

[w3473] ISO/IEC JTC1/SC29/WG11/N3473, Workplan for CE on OrderingKey DS, Beijing, July 2000.

[m6491] ISO/IEC JTC1/SC29/WG11/M6491, Internet Streaming Media Metadata Interchange with MPEG-7 MDS, Eric Rehm, La Baule, October 2000.

[m6588] ISO/IEC JTC1/SC29/WG11/M6588, CE report on OrderingKey DS, A. Bugatti et al., La Baule, October 2000.

[w3630] ISO/IEC JTC1/SC29/WG11/N3630, Workplan for CE on OrderingKey D, La Baule, October 2000.

[w3705] ISO/IEC JTC1/SC29/WG11/N3705, Text of ISO/IEC 15938-5/CD, Peter van Beek, Ana B. Benitez, Joerg Heuer, Jose Martinez, Philippe Salembier, John Smith, Toby Walker, La Baule, October 2000.

[XPath] XML Path Language (XPath) - Version 1.0 - W3C Recommendation - 16 November 1999

7 Appendix A

7.1 OrderingKey D Examples

The OrderingKey D is a tool that allows to order DS descriptions of the same type (e.g., segments, objects, or events). Another example of OrderingKey D descriptions is included below. The example describes an ordering based on the value of the attribute id of video segments

In the example below, the resulting ordering would be Scene1→ Scene2 → Scene3 according to the alphabetical order applied to the selected ordering key (the direction is descending by default).

```
<?xml version="1.0" encoding="UTF-8"?>
<Mpeg7Unit
    xmlns=http://www.mpeg7.org/2001/MPEG-7_Schema
    xmlns:xsi=http://www.w3.org/1999/XMLSchema-instance
    xsi:schemaLocation="http://www.mpeg7.org/2001/MPEG-7_Schema
        c:\users\bugatti\mpeg7\WeightDS\mds4CE.xsd">

    <!--  Details of the description of the video segments have  -->
    <!--  been omitted for simplicity.                           -->
    <VideoSegment id="program">
        <SegmentDecomposition DecompositionType="temporal">
            <VideoSegment id="scene1">
                <MediaTime>
                    <MediaTimePoint/>
                </MediaTime>
            </VideoSegment>
            <VideoSegment id="scene2">
                <MediaTime>
                    <MediaTimePoint/>
                </MediaTime>
            </VideoSegment>
            <VideoSegment id="scene3">
                <MediaTime>
                    <MediaTimePoint/>
                </MediaTime>
            </VideoSegment>
        </SegmentDecomposition>
        <MediaTime>
            <MediaTimePoint/>
        </MediaTime>
    </VideoSegment>

    <OrderingKey name="SceneOrdering" semantics="accordingToSceneId">
        <Selector>
            /VideoSegment/SegmentDecomposition/VideoSegment
        </Selector>
        <field>@id</Field>
    </OrderingKey>
</Mpeg7Unit>
```

7.2 Validation experiments

Two demos were built to show ordering key functionalities and instantiations of these two examples are contained in files SoccerMatch.xml and News.xml. Both the files was validated using XML Spy 3.0 and a MDS WD and XM schema from Yoshiaki with adds in order to validate the above examples (mds4CE.xsd).
In the first demo is possible to order the shots based on their camera movement Pan speed obtaining in first positions the shots containing close-up of players.
In the second one a cluster of the shots containing the anchor man is obtained using the dominant color of the key frame representing the shots as ordering key.

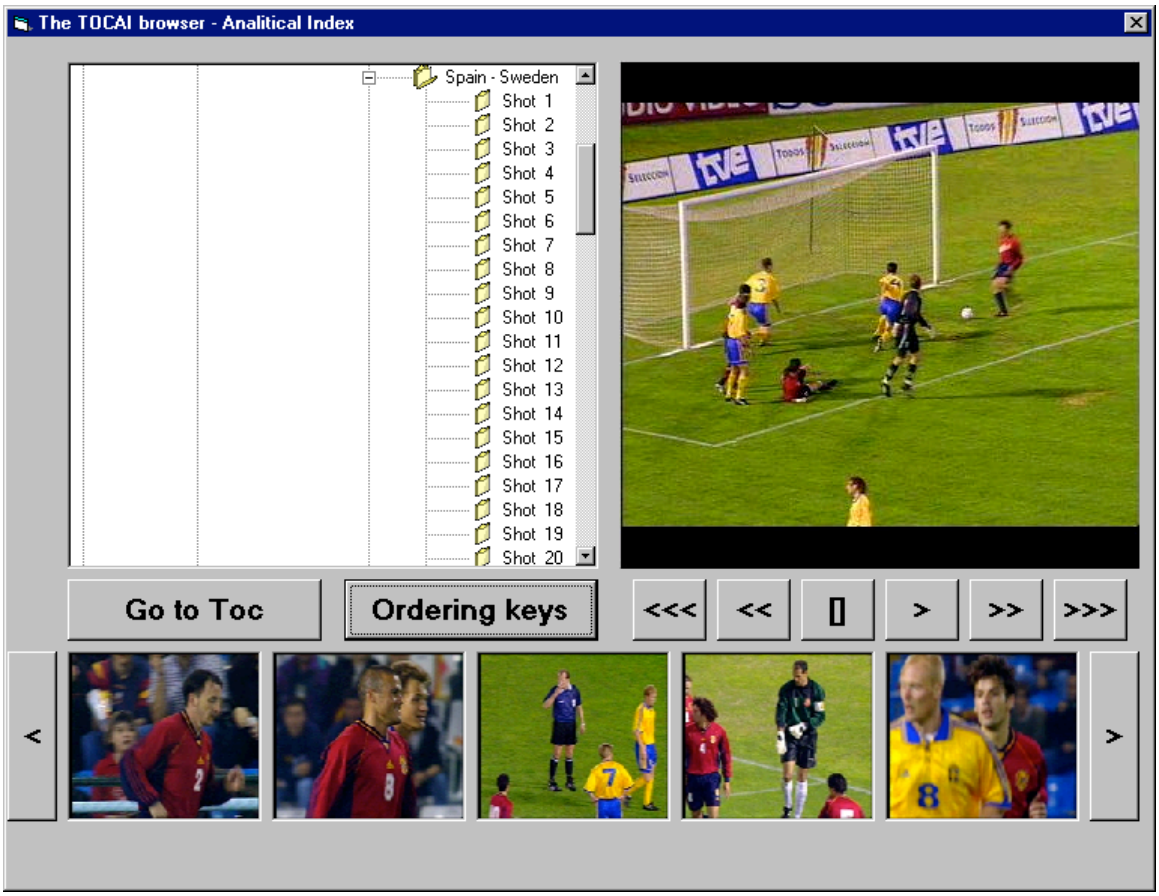

Figure1 - The shots contained in the video of the soccer match Spain-Sweden are ordered according on Pan speed ordering key. As you see the first shots contain close-up of players.

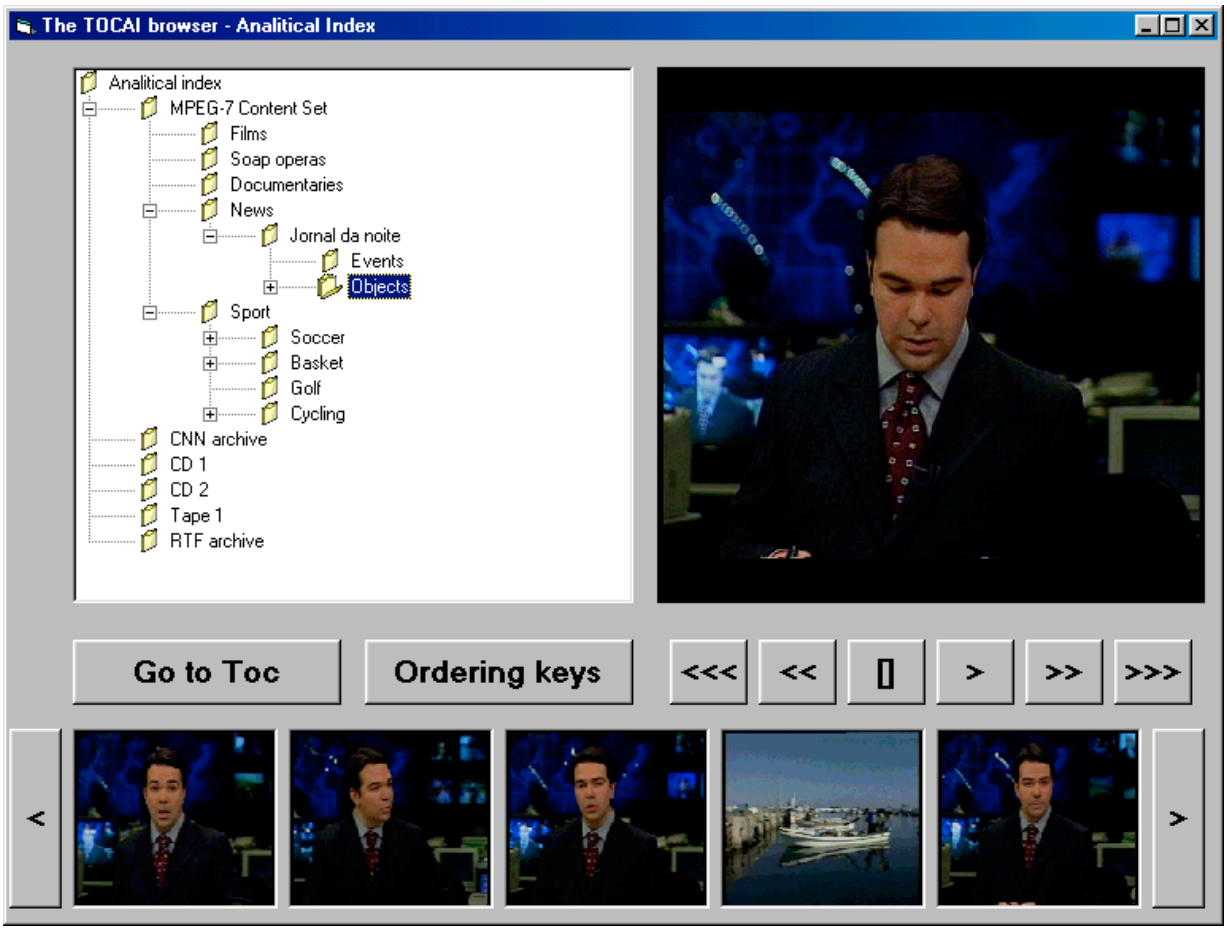

Figure 2 - In this case it is possible to order the segment according to their dominant color associated to the key frame representing the shot. This ordering key can be applied on shots of 'Jornal da Noite' portughese news: in this way it is possible for example to cluster all the shots containing the speaker in spite of their temporal position in the multimedia stream, as you see in the figure above.