INTERNATIONAL STANDARD

ISO/IEC 23000-17

First edition
2018-10

Information technology — Multimedia application format (MPEG-A) —

Part 17: Multiple sensorial media application format

Technologies de l'information — Format pour application multimédia (MPEG-A) —

Partie 17: Format pour application multimédia sensorielle

Reference number
ISO/IEC 23000-17:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23000 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Along with the sensation associated with the 3D film industry, the development of MulSeMedia (multiple sensorial media), or 4D media, has received much attention from the public. 4D movies generally add sensory effects to 3D and/or IMAX movies, allowing audiences to immerse themselves more deeply in the movie viewing experience. Along with the two human senses of sight and hearing, sensory effects such as wind, vibration, scent and others can stimulate other senses, such as the tactile and olfaction senses.

The multiple sensorial media application format (MSMAF) defines a file format for multiple sensorial media services. It specifies core structures of multiple sensorial media application format being organized by the combination of related information for multiple sensorial media applications.

Services using the MSMAF can be realized in two ways: offline and online. The offline services include DVDs and movie plays in theatres, whereas the online services include streaming services on IP-TV or broadcasting with MPEG-2 transport streams.

Information technology — Multimedia application format (MPEG-A) —

Part 17: Multiple sensorial media application format

1 Scope

This document specifies a file format which is capable of storage, interchange, management, editing and presentation of multiple sensorial media contents based on the ISO base media file format. The file format provides the overall structure for storing multiple sensorial media contents.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purpose of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at https://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
sensory effect metadata
record that defines the description schemes and descriptors to represent sensory effects

3.2
sensory effect
effect to augment perception by stimulating human senses in a particular scene of a multimedia application

Note 1 to entry: Combinations of tactile display may also provide directional, shape information.

EXAMPLE Scent, wind, light, haptic (kinesthetic-force, stiffness, weight, friction, texture, widget (button, slider, joystick), tactile: air-jet, suction pressure, thermal, current and vibration.

4 Abbreviations

BiM	binary MPEG format for XML (ISO/IEC 23001-1)
DIA	digital item adaptation (ISO/IEC 21000-7)
MPEG-21	multimedia framework (ISO/IEC 21000)
MPEG-7	multimedia content description interface (ISO/IEC 15938)
SEDL	sensory effects description language (ISO/IEC 23005-3)

SEM	sensory effect metadata (ISO/IEC 23005-3)
SEV	sensory effects vocabulary (ISO/IEC 23005-3)
XML	extensible mark-up language
XSI	XML streaming instructions (ISO/IEC 21000-7)

5 Overview

5.1 Overview of sensory effect metadata (SEM)

It is necessary to play multiple sensorial media that contains metadata for representing real-sense effects according to the scenes of the media. In order to satisfy this requirement, sensory effect metadata (SEM) was standardized in ISO/IEC 23005-3. The SEM is metadata for representing 4D real-sense effects and consists of two main parts: effect property and effect variable. The effect property can be transferred in front of a media stream or transferred periodically for preparing channel switching. The effect variables can also easily fragmented and transferred by the unit of a time slice. The effect property contains the definition of each sensory effect applied to multiple sensorial media. After analysing the effect property, representing devices like a home server can map each sensory effect to a proper 4D device in the user's environment. The effect variables contain control variables to operate 4D real-sense devices. Figure 1 shows the SEM processing architecture.

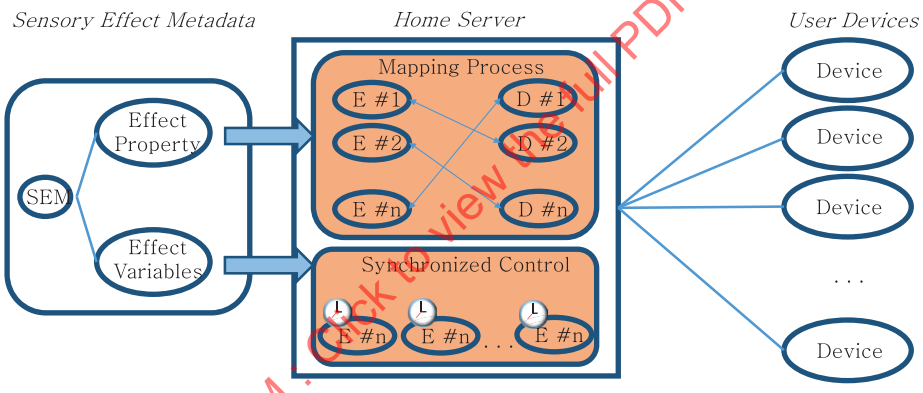


Figure 1 — SEM processing architecture for representing 4D effects

Table 1 shows some examples of effects and corresponding properties and Table 2 shows variables and mapped device commands. This system can support wind, shading, heating, motion, light, flash, diffusion, sting, tickler, air jet and water sprayer effects to define properties and control devices.

Table 1 — Examples of effects and corresponding properties

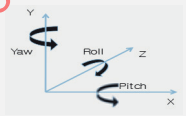
Effects	SEMs
Wind	<SEM><Effect xsi:type="sev:WindType" id="WindEffect_1" activate="true" pts="0" duration="1" intensity-value="10.0" intensity-range="0.0 12.0"/></SEM>
Incline	<SEM><Effect xsi:type="sev:RigidBodyMotionType" id="MotionEffect_1" activate="true" pts="42" duration="5"><sev:Incline pitch="45" roll="0" yaw="0" pitchSpeed="10"/></Effect></SEM>
Wave	<SEM><DescriptionMetadata><ClassificationSchemeAlias alias="WAVE" href="urn:mpeg:mpeg-v:01-SI-WaveDirectionCS-NS"/><ClassificationSchemeAlias alias="WAVESTR" href="urn:mpeg:mpeg-v:01-SI-WaveStartDirectionCS-NS"/></DescriptionMetadata> <Effect xsi:type="sev:RigidBodyMotionType" id="MotionEffect_2" activate="true" pts="25" duration="9"><sev:Wave direction=":WAVE:front_rear" startDirection=":WAVESTR:up" count="3" distance="10"/></Effect></SEM>

Table 1 *(continued)*

Effects	SEMs
Shake	<SEM><DescriptionMetadata><ClassificationSchemeAlias alias="SHAKE" href="urn:mpeg:mpeg-v:01-SI-ShakeDirectionCS-NS"/></DescriptionMetadata> <Effect xsi:type="sev:RigidBodyMotionType" id="MotionEffect_3" activate="true" pts="67" duration="2"><sev:Shake direction=":SHAKE:Heave" count="1" distance="20"/></Effect></SEM>
LED	<SEM><Effect xsi:type="sev:LightType" id="LEDEffect_1" activate="true" pts="81" duration="30" intensity-value="100.0" intensity-range="0.00001 32000.0" color="#FFFFFF"/></SEM>
Scent	<SEM><DescriptionMetadata><ClassificationSchemeAlias alias="SCENT" href="urn:mpeg:mpeg-v:01-SI-ScentCS-NS"/></DescriptionMetadata> <Effect xsi:type="sev:ScentType" id="ScentEffect_1" activate="true" pts="23" duration="4" intensity-value="9.0" intensity-range="0.0 10.0" scent=":SCENT:lilac"/></SEM>

Table 2 shows an example of SEM. Each SEM data will be parsed, analysed and then changed into proper control values.

Table 2 — Effect variables and selected device commands

Effects	Variables	Device command
Wind	SetWindSpeedRatio=100	Turn on fan with maximum level
Incline	SetIncline pitch="45" roll="0" yaw="0", pitchSpeed=10	 Incline motion chair to front angle 45 degrees with speed 10 cm/sec
Wave	SetWave direction="**front_rear**" startDirection**="up"** count="**3**" distance="**10**"	 Wave motion chair from front-up to rear-down, 3 times, 10 cm
Shake	SetShake direction="**Heave**" count="**1**" distance="**20**"	 Shake motion chair from up to down, 1 time, 20 cm.
LED	SetColor r="**255**" g="**255**" b="**255**"	Turn on led light “white”
Scent	SetOnOff=**True,** SetDensityRatio=**100,** SetDiffusionSourceID=**1**	Turn on diffuser value number “1” to make strong (100 %) smell

5.2 Fragmentation of SEM

5.2.1 SEM streaming instruction

The SEM root element, the GroupOfEffects element, Effect element, and ReferenceEffect element in ISO/IEC 23005-3 can contain the siAttributeList that is the XML streaming instruction specified in ISO/IEC 21000-7. The XML streaming instructions allow firstly to identify process units in an XML document and secondly to assign time information to them.

A process unit is a set of connected XML elements. It is specified by one element, named an anchor element, and by a process unit mode indicating how other connected elements are aggregated to this anchor to compose the process unit. Depending on the mode, the anchor element is not necessarily the root of the process unit. The content provider may require that a given process unit be encoded as a random access point.

The main functionalities of the XML streaming instruction are as follows:

— fragmenting an XML document into process units;

— indicating which process units shall be encoded as random access point;

— assigning time information (i.e., processing time stamp) to these process units.

5.2.2 Properties related to the fragmentation

The properties related to the fragmentation are specified as follows.

— `anchorElement` is one of `{undefined, false, true}`. When set to true, this property indicates an anchor element. This property is not inherited and its default value is `undefined`.

— `puMode` is one of `{undefined, self, ancestors, descendants, ancestorsDescendants, preceding, sequential}`. The process unit mode property specifies how elements are aggregated to the anchor element to compose a process unit. This property is inherited and its default value is `undefined`.

  — `self`: the process unit contains only the anchor element.

  — `ancestors`: the process unit contains the anchor element and its ancestors stack, i.e., its ancestors elements.

  — `descendants`: the process unit contains the anchor element and its descendant elements.

  — `ancestorsDescendants`: the process unit contains the anchor element, its ancestor, and descendant element (Figure 2).

  — `preceding`: the process unit contains the anchor element, its descendant and parent elements, and all the preceding-sibling elements of its ancestor elements and their descendants.

  — `precedingSiblings`: the process unit contains the anchor element, its descendant and parent elements, and all the preceding-sibling elements (and their descendants) of its ancestor elements.

  — `sequential`: the process unit contains the anchor element, its ancestors stack, and all the following elements (i.e., descendants, following siblings, and their ancestors) until a next element is flagged as an anchor element.

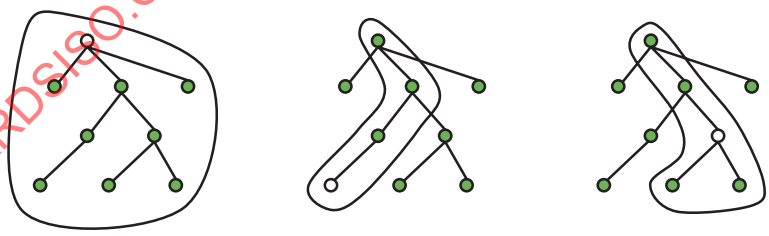


Figure 2 — Example of process unit – puMode = ancestorsDescendants

— `encodesAsRAP` is one of `{undefined, false, true}`. When set to `true`, this property indicate that the process unit shall be encoded as a random access point. This property is inherited and its value is `undefined`.

5.2.3 Properties related to time information

The processing time stamp of a process unit defines the point in time where it is available to the application for consumption. It is specified by the set of time properties of the PU's anchor element. There are two defined time modes: an **absolute time mode** where the PUs' PTS are specified according

to a same time origin and a **relative time mode** where they are defined relatively to the PTS of the previous PU.

The properties related to time information are specified as follows:

— `timescale` is one of {`undefined`, an integer value}. It specifies the time scale, i.e., the number of ticks per second. This property is inherited and its default value is `undefined`.

— `ptsDelta` is one of {`undefined`, an integer value}. It specifies the relative interval in time ticks after the preceding anchor element. The default value for the processing time stamp of the first anchor element is 0. This property is inherited and its default value is undefined.

— `absTimeScheme` is one of {`undefined`, a string value}: It specifies the absolute time scheme used. This property is inherited and its default value is undefined.

— `absTime` is one of {`undefined`, a string value}: It specifies the absolute time of the anchor element. Its syntax and semantics are specified according to the time scheme used (`absTimeScheme` property). This property is not inherited and its default value is `undefined`.

— `pts` is one of {`undefined`, an integer value}. It specifies the absolute time of the anchor element as the number of ticks since the origin. This property is not inherited and its default value is `undefined`.

5.2.4 Fragmenter

The fragmenter can generate process units (PUs) using an XML document with XSI attributes described in the previous section.

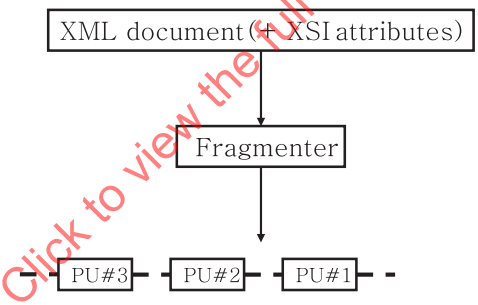


Figure 3 — Normative processing related to XML streaming instruction

The fragmenter parses the XML document in a depth-first order. XML streaming instructions properties are computed by fragmenter. An element with the `anchorElement` property set to true indicates an anchor element and a new process unit. The process unit then comprises connected elements according to the `puMode` property of the anchor element (Figure 3).

5.2.5 Fragmenting SEM example

The following instance is a SEM instance with XSI attributes.

```
 ?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:ct="urn:mpeg:mpeg-v:2010:01-CT-NS"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <DescriptionMetadata>
      <ClassificationSchemeAlias href="urn:mpeg:mpeg-v:01-SI-ColorCS-NS"
alias="COLOR"/>
   </DescriptionMetadata>

   <Declarations>
      <Parameter xsi:type="ColorCorrectionParameterType" id="ccpt001">
         <ToneReproductionCurves>
            <ct:DAC_Value>0</ct:DAC_Value>
            <ct:RGB_Value>0 0 0</ct:RGB_Value>
            <ct:DAC_Value>255</ct:DAC_Value>
            <ct:RGB_Value>255 255 255</ct:RGB_Value>
      </ToneReproductionCurves>
         <ConversionLUT>
            <ct:RGB2XYZ_LUT mpeg7:dim="3 3">
               86.60 67.60 38.0
               46.00 137.0 16.5
               2.365 19.41 203.9
            </ct:RGB2XYZ_LUT>
            <ct:RGBScalar_Max>
               1.0 1.0 1.0
            </ct:RGBScalar_Max>
            <ct:Offset_Value>
               0.2150 0.2050 0.4250
            </ct:Offset_Value>
            <ct:Gain_Offset_Gamma mpeg7:dim="3 3">
               1.0228 -0.0228 1.6222
               1.0242 -0.0242 1.5624
```



```
                1.0220 -0.0220 1.6180
            </ct:Gain_Offset_Gamma>
            <ct:InverseLUT mpeg7:dim="3 3">
               0.0155 -0.0073 -0.0023
               -0.0052 0.0099 0.0002
               0.0003 -0.0009 0.0049
            </ct:InverseLUT>
         </ConversionLUT>
         <ColorTemperature>
            <ct:xy_Value x="0.1" y="0.8"/>
            <ct:Y_Value>100</ct:Y_Value>
         </ColorTemperature>
         <InputDeviceColorGamut>
            <ct:IDCG_Type>BARCO</ct:IDCG_Type>
            <ct:IDCG_Value mpeg7:dim="3 2">
               0.2835 0.6043
               0.1509 0.0624
               0.6244 0.3410
            </ct:IDCG_Value>
         </InputDeviceColorGamut>
         <IlluminanceOfSurround>500</IlluminanceOfSurround>
      </Parameter>
   </Declarations>


      <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescendants"
si:anchorElement="true" activate="true" color=":COLOR:white" intensity-value="50"
intensity-range="0.0 100.0" duration="5000" si:pts="1000"/>
      <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescendants"
si:anchorElement="true" activate="true" color=":COLOR:red" duration="7000"
intensity-value="30" intensity-range="0.0 100.0" si:pts="6000"/>


      <GroupOfEffects si:puMode="ancestorsDescendants" si:anchorElement="true"
id="explosion_effect" si:pts="20000">
         <Effect xsi:type="sev:FlashType" intensity-range="0.0 100.0" intensity-
value="65.0" color=":COLOR:white" duration="5000" activate="true" si:ptsDelta="0"/>
         <Effect xsi:type="sev:TemperatureType" intensity-range="0.0 100.0"
intensity-value="40.0" duration="5000" activate="true" si:ptsDelta="0"/>
         <Effect xsi:type="sev:WindType" intensity-range="0.0 100.0" intensity-
value="50.0" duration="5000" activate="true" si:ptsDelta="0"/>
         <Effect xsi:type="sev:VibrationType" intensity-range="0.0 100.0"
intensity-value="60.0" duration="5000" activate="true" si:ptsDelta="0"/>
      </GroupOfEffects>
</SEM>
```

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:ct="urn:mpeg:mpeg-v:2010:01-CT-NS"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <DescriptionMetadata>
      <ClassificationSchemeAlias href="urn:mpeg:mpeg-v:01-SI-ColorCS-NS"
alias="COLOR"/>
   </DescriptionMetadata>

   <Declarations>
      <Parameter xsi:type="ColorCorrectionParameterType" id="ccpt001">
         . . .
      </Parameter>
   </Declarations>

   <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescendants"
si:anchorElement="true" activate="true" color=":COLOR:white" intensity-value="50"
intensity-range="0.0 100.0" duration="5000" si:pts="1000"/>
   <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescendants"
si:anchorElement="true" activate="true" color=":COLOR:red" duration="7000"
intensity-value="30" intensity-range="0.0 100.0" si:pts="6000"/>
   <GroupOfEffects si:puMode="ancestorsDescendants" si:anchorElement="true"
id="explosion_effect" si:pts="20000">
      <Effect xsi:type="sev:FlashType" intensity-range="0.0 100.0" intensity-
value="65.0" color=":COLOR:white" duration="5000" activate="true" si:ptsDelta="0"/>
      <Effect xsi:type="sev:TemperatureType" intensity-range="0.0 100.0" intensity-
value="40.0" duration="5000" activate="true" si:ptsDelta="0"/>
       <Effect xsi:type="sev:WindType" intensity-range="0.0 100.0" intensity-
value="50.0" duration="5000" activate="true" si:ptsDelta="0"/>
       <Effect xsi:type="sev:VibrationType" intensity-range="0.0 100.0" intensity-
value="60.0" duration="5000" activate="true" si:ptsDelta="0"/>
    </GroupOfEffects>
</SEM>
```

The XSI attributes of the `SEM` root element, `puMode` and `timeScale`, are inherited to its child anchor elements. There are four anchor elements in the instance. The `timeScale` from the `SEM` root element and `pts` become the time information of each processing unit.

The following instances are the examples of processing units generated by the fragmenter. Because the `puMode` of the `SEM` root element is `sequential`, the first processing unit includes following elements until a next element is flagged as an anchor element. Therefore, the first processing unit contains the `SEM` root, the `DescriptionMetadata`, and the `Declarations`. Because the `puMode` of the second, third, and forth `enchorElement` are `ancestorsDescedents`, it overides the `puMode` of the `SEM` root element.

```
 /* PU #1 */
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <DescriptionMetadata>
     <ClassificationSchemeAlias href="urn:mpeg:mpeg-v:01-SI-ColorCS-NS" alias="COLOR"/>
   </DescriptionMetadata>

   <Declarations>
      <Parameter xsi:type="ColorCorrectionParameterType" id="ccpt001">
         <ToneReproductionCurves>
            <ct:DAC_Value>0</ct:DAC_Value>
            <ct:RGB_Value>0 0 0</ct:RGB_Value>
            <ct:DAC_Value>255</ct:DAC_Value>
            <ct:RGB_Value>255 255 255</ct:RGB_Value>
      </ToneReproductionCurves>
         <ConversionLUT>
            <ct:RGB2XYZ_LUT mpeg7:dim="3 3">
               86.60 67.60 38.0
               46.00 137.0 16.5
               2.365 19.41 203.9
            </ct:RGB2XYZ_LUT>
            <ct:RGBScalar_Max>
               1.0 1.0 1.0
            </ct:RGBScalar_Max>
            <ct:Offset_Value>
               0.2150 0.2050 0.4250
            </ct:Offset_Value>
            <ct:Gain_Offset_Gamma mpeg7:dim="3 3">
               1.0228 -0.0228 1.6222
               1.0242 -0.0242 1.5624
               1.0220 -0.0220 1.6180
```

```
            </ct:Gain_Offset_Gamma>
            <ct:InverseLUT mpeg7:dim="3 3">
               0.0155 -0.0073 -0.0023
               -0.0052 0.0099 0.0002
               0.0003 -0.0009 0.0049
            </ct:InverseLUT>
         </ConversionLUT>
         <ColorTemperature>
            <ct:xy_Value x="0.1" y="0.8"/>
            <ct:Y_Value>100</ct:Y_Value>
         </ColorTemperature>
         <InputDeviceColorGamut>
            <ct:IDCG_Type>BARCO</ct:IDCG_Type>
            <ct:IDCG_Value mpeg7:dim="3 2">
               0.2835 0.6043
               0.1509 0.0624
               0.6244 0.3410
            </ct:IDCG_Value>
         </InputDeviceColorGamut>
         <IlluminanceOfSurround>500</IlluminanceOfSurround>
      </Parameter>
   </Declarations>
</SEM>
```

```
/* PU #2 */
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescendants"
si:anchorElement="true" activate="true" color=":COLOR:white" intensity-
value="50" intensity-range="0.0 100.0" duration="5000" si:pts="1000"/>
</SEM>
```

```
/* PU #3 */
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <Effect xsi:type="sev:LightType" si:puMode="ancestorsDescenda
nts" si:anchorElement="true" activate="true" color=":COLOR:red"
duration="7000" intensity-value="30" intensity-range="0.0 100.0"
si:pts="6000"/>
</SEM>
```

```
/* PU #4 */
<SEM
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xmlns="urn:mpeg:mpeg-v:2010:01-SEDL-NS"
xmlns:sev="urn:mpeg:mpeg-v:2010:01-SEV-NS"
xmlns:mpeg7="urn:mpeg:mpeg7:schema:2004"
xmlns:si="urn:mpeg:mpeg21:2003:01-DIA-XSI-NS"
xsi:schemaLocation="urn:mpeg:mpeg-v:2010:01-SEV-NS MPEG-V-SEV.xsd"
si:puMode="sequential"
si:anchorElement="true"
si:timeScale="1000">
   <GroupOfEffects si:puMode="ancestorsDescendants"
si:anchorElement="true" id="explosion_effect" si:pts="20000">
      <Effect xsi:type="sev:FlashType" intensity-range="0.0 100.0"
intensity-value="65.0" color=":COLOR:white" duration="5000"
activate="true" si:ptsDelta="0"/>
      <Effect xsi:type="sev:TemperatureType" intensity-range="0.0 100.0"
intensity-value="40.0" duration="5000" activate="true" si:ptsDelta="0"/>
      <Effect xsi:type="sev:WindType" intensity-range="0.0 100.0"
intensity-value="50.0" duration="5000" activate="true" si:ptsDelta="0"/>
      <Effect xsi:type="sev:VibrationType" intensity-range="0.0 100.0"
intensity-value="60.0" duration="5000" activate="true" si:ptsDelta="0"/>
   </GroupOfEffects>
</SEM>
```

5.3 Encoding process units and packaging multiple sensorial media application format

The encoding method of the PUs is MPEG-V binary representation.

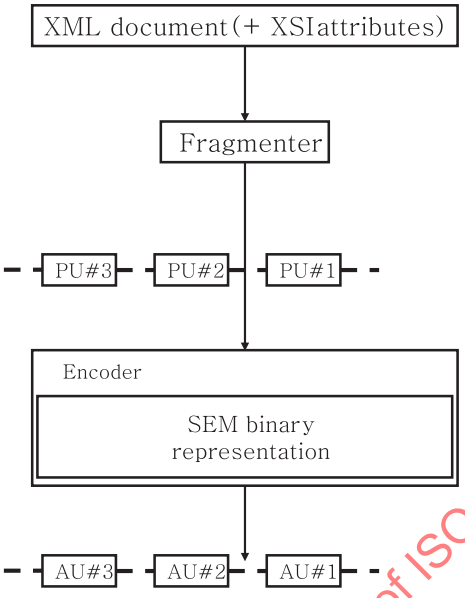


Figure 4 — Access unit (AU) encoder

Figure 4 shows the encoding process of access units using process units. Every processing unit becomes the well-formed XML documents so that any binary encoding using the well-formed documents can again generate valid binary documents.

6 Components of multiple sensorial media application format

6.1 Supported components

Table 3 shows a brief summary of the supported components of the multiple sensorial media application format which conists of the ISO/IEC Standards.

The multiple sensorial media application format includes ISO/IEC 14496-2 Simple Profile at Level 3 and ISO/IEC 14496-10 Baseline Profile at Level 1.3 for visual, ISO/IEC 14496-3 AAC and HE-AAC Profile for audio, and ISO/IEC 23005-3 SEM for multiple sensorial effect. For this specification, ISO/IEC 14496-12 ISO base media file format is used for a base file format structure.

Table 3 — Supported components of multiple sensorial media application format

Type	Component name	Specification	Standard
File format	ISO base media file format	ISO/IEC 14496-12	ISO/IEC Standards
Visual	MPEG-4 Video	ISO/IEC 14496-2 Simple Profile Level 3	
	MPEG-4 AVC	ISO/IEC 14496-10 Baseline Profile Level 1.3	
Audio	MPEG-4 Audio AAC	ISO/IEC 14496-3	
	MPEG-4 Audio HE-AAC	ISO/IEC 14496-3	
Data	MPEG-V Sensory Effect Meta-data	ISO/IEC 23005-3	

ISO/IEC 14496-12 ISO base media file format is a flexible, extensible format which contains timed media information in order to facilitate interchange, management, editing, and presentation of the media as

being shown in Figure 5. The ISO base media file format is a base format for the multiple sensorial media application format.

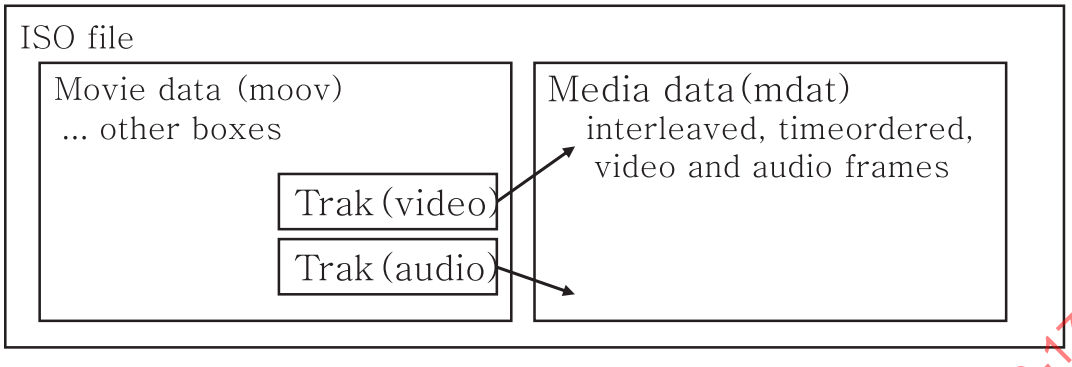


Figure 5 — Example of a simple ISO base media file format

7 MSMAF file structures

7.1 Table for boxes

The multiple sensorial media application format contains various boxes based on the ISO base media file format.

The normative file structure consists of `'ftyp'`, `'moov'` and `'mdat'` boxes. Mandatory boxes are marked with an asterisk (*).

The `'ftyp'` box indicates the type of the file format which complies to the structure defined for the multi sensorial AF. Thus, an application should be able to play multiple sensorial media application format files when it supports the brands of `'ftyp'` box field. A detailed description of the brands of multiple sensorial media application format is provided in 8.1.

The `'moov'` box contains one or more tracks for multiple sensorial video sequences, a track for SEM, and also contain tracks for audio.

The `'trak'` boxes contain temporal and spatial information of the media data. For multi sensorial application format, each track contains its associated `'mdia'` box.

The `'mdia'` box contains fragment information of the multiple sensorial contents in the track.

The `'mdat'` box contains the media data which are described in the `'trak'` boxes.

Table 4 briefly shows the structure of the boxes and their descriptions.

Table 4 — Boxes of multiple sensorial media application format

*	ftyp						file type and compatibility
	pdin						Progressive download Information
*	moov						container for all the metadata
*		mvhd					movie header, overall declarations
*		trak					container for an individual track or stream
*			tkhd				track header, overall information about the track
			edts				edit list container
				elst			an edit list
*			mdia				container for the media information in a track
*				mdhd			media header, overall information about the media