INTERNATIONAL STANDARD

ISO/IEC 14496-25

Second edition
2011-05-15

Information technology — Coding of audio-visual objects —

Part 25:
3D Graphics Compression Model

Technologies de l'information — Codage d'objets audiovisuels —

Partie 25: Modèle de compression graphique 3D

Reference number
ISO/IEC 14496-25:2011(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword iv
Introduction vi
1 Scope 1
2 Architecture model 1
2.1 Global view 1
2.2 Layer 1: Textual Data Representation 2
2.3 Layer 2: Binarization 2
2.4 Layer 3: Compression 3
2.5 Interface between Layer 2 and Layer 3 4
3 Decoder model 4
Annex A (informative) Encoding model 6
Annex B (normative) Binarization and Compression of Scene Graph expressed with XMT 7
B.1 Introduction 7
B.2 Scene Graph and Object Graph binarization 7
B.3 3DMCe and SC3DMC Streams 7
B.4 PI, OI, CI Streams 8
B.5 BBA Stream 8
B.6 FAMC Stream 9
Annex C (normative) Binarization and Compression of Scene Graph expressed with COLLADA 10
C.1 Introduction 10
C.2 Scene Graph and Object Graph binarization 10
C.3 3DMCe and SC3DMC Streams 10
C.4 PI, OI, CI Streams 11
C.5 BBA Stream 11
C.6 FAMC Stream 11
Annex D (normative) Binarization and Compression of Scene Graph expressed with X3D 12
D.1 Introduction 12
D.2 Scene Graph and Object Graph binarization 12
D.3 3DMCe and SC3DMC Streams 12
D.4 PI, OI, CI Streams 13

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

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.

ISO/IEC 14496-25 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 14496-25:2009), which has been technically revised.

ISO/IEC 14496 consists of the following parts, under the general title *Information technology — Coding of audio-visual objects*:

— *Part 1: Systems*

— *Part 2: Visual*

— *Part 3: Audio*

— *Part 4: Conformance testing*

— *Part 5: Reference software*

— *Part 6: Delivery Multimedia Integration Framework (DMIF)*

— *Part 7: Optimized reference software for coding of audio-visual objects* [Technical Report]

— *Part 8: Carriage of ISO/IEC 14496 contents over IP networks*

— *Part 9: Reference hardware description* [Technical Report]

— *Part 10: Advanced Video Coding*

— *Part 11: Scene description and application engine*

— *Part 12: ISO base media file format*

— *Part 13: Intellectual Property Management and Protection (IPMP) extensions*

- *Part 14: MP4 file format*
- *Part 15: Advanced Video Coding (AVC) file format*
- *Part 16: Animation Framework eXtension (AFX)*
- *Part 17: Streaming text format*
- *Part 18: Font compression and streaming*
- *Part 19: Synthesized texture stream*
- *Part 20: Lightweight Application Scene Representation (LASeR) and Simple Aggregation Format (SAF)*
- *Part 21: MPEG-J Graphics Framework eXtensions (GFX)*
- *Part 22: Open Font Format*
- *Part 23: Symbolic Music Representation*
- *Part 24: Audio and systems interaction* [Technical Report]
- *Part 25: 3D Graphics Compression Model*
- *Part 26: Audio conformance*
- *Part 27: 3D Graphics conformance*

Introduction

Within ISO/IEC 14496-11 and ISO/IEC 14496-16, several tools for compression of 3D graphics primitives are standardized (for geometry, texture and animation). Such tools are designed to apply on representations information as defined in ISO/IEC 14496-11 (commonly know as BIFS).

In this International Standard, the paradigm changes, making it possible to apply the compression tools for 3D graphics defined in ISO/IEC 14496-11 and ISO/IEC 14496-16 to potentially any representation formats for scene graph and graphics primitives expressed in XML.

Currently, the model is implemented for XMT, COLLADA and X3D.

Information technology — Coding of audio-visual objects —

Part 25:
3D Graphics Compression Model

1 Scope

This part of ISO/IEC 14496 describes a model for connecting 3D graphics compression tools defined in ISO/IEC 14496 to graphics primitives defined in any other standard, specification or recommendation.

The goal of this part of ISO/IEC 14496 is to specify an architectural model able to accommodate

- third-party XML based descriptions of scene graph and graphics primitives with
- (potential) binarization tools and with
- MPEG-4 3D graphics compression tools specified in ISO/IEC 14496-2, ISO/IEC 14496-11 and ISO/IEC 14496-16.

The advantages of such an approach are on the one side the use of powerful compression tools for graphics and on the other side the generality of graphics primitives representation. Hence, compression tools developed in ISO/IEC 14496-2, ISO/IEC 14496-11 and ISO/IEC 14496-16 would be applied not only to the scene graph defined by ISO/IEC 14496-11 but to any scene graph definition. The bitstreams obtained when using the model are MP4 formatted and contain XML (or binarized XML) for the scene graph and binary elementary streams for graphics compression (geometry, texture and animation).

2 Architecture model

2.1 Global view

The architectural model, illustrated in Figure 1, has three layers: Textual Data Representation, Binarization and Compression.

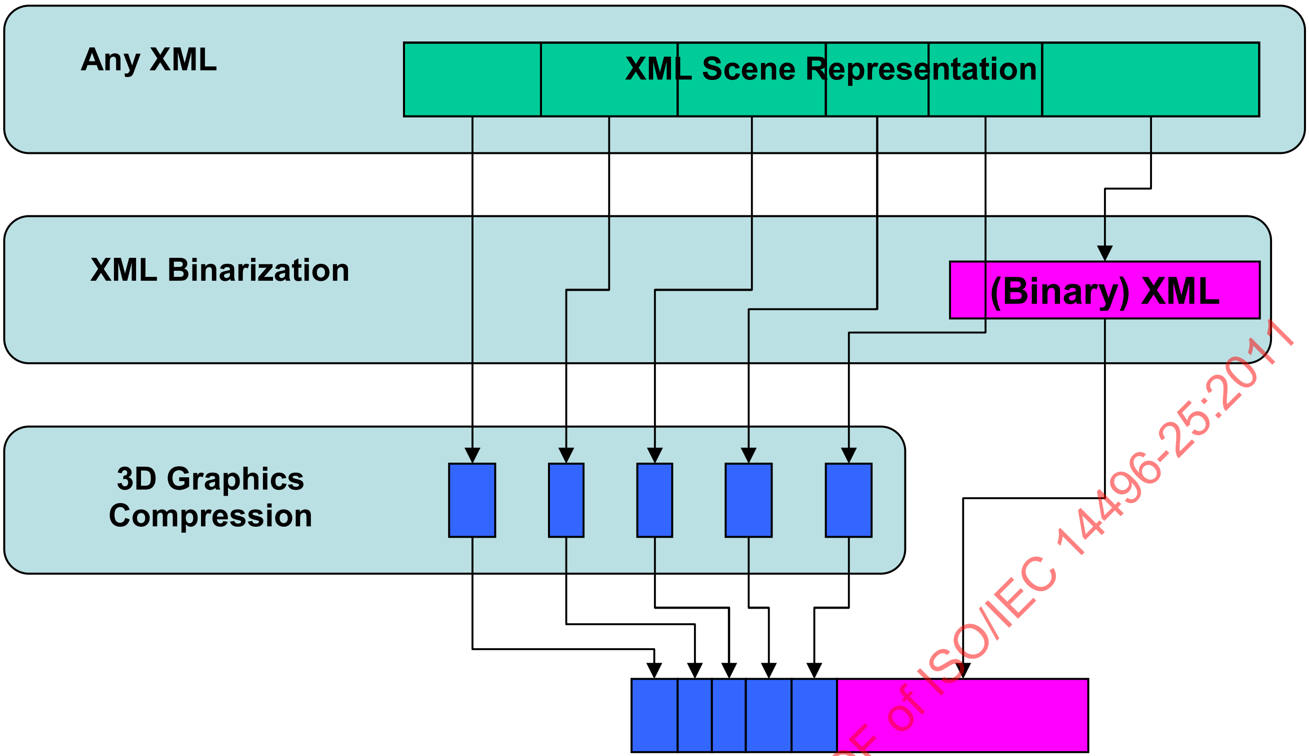

Figure 1 — Layers of the 3D Graphics Compression Model

The three layers of the architecture are defined in the following Subclauses.

2.2 Layer 1: Textual Data Representation

The current model can accommodate any scene graph and graphics primitives' representation formalism. The only requirement on this representation is that it should be expressed in XML. Any XML Schema (specified by MPEG or by external bodies) may be used. The particular case of each specific schema is treated in dedicated annex.

Currently the following XML Schemas are supported:

- XMT (Annex B),
- COLLADA (Annex C) and
- X3D (Annex D).

2.3 Layer 2: Binarization

The Binarization layer is processing the XML data that is not encoded by the MPEG-4 elementary bitstreams encoders (e.g. scene graph elements). This data is encapsulated in the "meta" atom of the MP4 file and it can be textual or binarized by using gzip.

2.4 Layer 3: Compression

2.4.1 General

The Compression layer includes the following MPEG-4 compression tools specified within ISO/IEC 14496-11 and ISO/IEC 14496-16:

AFX object code	Compression Tool	Type	Standard
0x03 / CI	Coordinate Interpolator	Animation	ISO/IEC 14496-11
0x04 / OI	Orientation Interpolator	Animation	ISO/IEC 14496-11
0x05 / PI	Position Interpolator	Animation	ISO/IEC 14496-11
0x07 / BBA	Bone-based Animation	Animation	ISO/IEC 14496-16
0x09 / 3DMCe	3D Mesh Compression Extension	Geometry	ISO/IEC 14496-16
0x0B / FAMC	Frame-based Animated Mesh Compression	Animation	ISO/IEC 14496-16
0x0C / SC3DMC	Scalable Complexity 3D Mesh Compression	Geometry	ISO/IEC 14496-16

For image compression, MPEG-4 natively supports PNG, JPEG and JPEG 2000. An image referred in the XML data structure having a format other than PNG, JPEG or JPEG 2000 should be converted to one of these in order to be encapsulated in the MP4 file format.

Each elementary stream has an associated track. In order to indicate to which element of the initial XML scene a specific track corresponds the following steps are performed:

- a unique identifier is assigned to the element of the initial XML; this is named element_ID
- the following parameters of the MP4 file metabox are updated:
o the data_references table is incremented with a new line having the following form:
§ next_index #ffp (track_ID="anID") - #ffp refers to the File Format Pointer scheme as defined in ISO/IEC 14496-12
o the item_location table is incremented with a new line having the following form:
§ location dref index next_index, offset=0, length=0
o the item_name table is incremented with a new line having the following form:
§ location name=element_ID, mime=video/mp4
Example of a meta_box parameters:
data references
0 (self)
1 #ffp(track_ID=2)
2 #ffp(track_ID=3)
item location
10 dref index 1, offset=0, length=0
11 dref index 2, offset=0, length=0
item name
10 name=mesh_1, mime=video/mp4
11 name=mesh_2, mime=video/mp4
In this example, the element_ID in the original XML file are mesh_1 and mesh_2.

2.4.2 3D-MeshCompression and 3D-Mesh Compression Extension Streams

3DMCe [ISO/IEC 14496-16:2006/Amd.1:2007, 5.9 (3DMC extension tools)] compresses the geometry features present in the IndexedFaceSet node as defined in VRML (ISO/IEC 14772-1:1997) as well as multiple and user defined attributes per vertex. 3DMCe has also the property of keeping the vertex and/or triangle order.

2.4.3 CoordinateInterpolator, OrientationInterpolator and PositionInterpolator Streams

CI, OI and PI [ISO/IEC 14496-11:2005, 8.9 (Interpolator compression)] are compression tools allowing to encode series of arrays consisting in

- (X, Y, Z) x (number of vertices) x (number of key frames) for CI,
- (X, Y, Z) x (1) x (number of key frames) for PI, and
- (X, Y, Z, theta) x (1) x (number of key frames) for OI.

While these tools were initially designed for nodes Coordinate, Position and Orientation Interpolators as defined by VRML, the compression scheme may apply to any time series as defined above.

2.4.4 Bone-Based Animation Stream

BBA [ISO/IEC 14496-16:2006, 5.4 (Compressed Bone-based animation)] is a compression tool for geometric transforms of bones (used in skinning-based animation) and weights (used in morphing animation). These elements are defined in the scene graph and should be uniquely identified by IDs. The BBA stream refers to these IDs and contains the new transforms of bones (expressed as Euler Angles or Quaternion) and the new weights of the morph targets (expressed as floats).

2.4.5 Frame-based Animated Mesh Compression Stream

FAMC (ISO/IEC 14496-16:2006/Amd.2) is a frame-based compression tool for animated vertex attributes (coordinates, normals …). It is applicable to temporal series of vertex attributes and uses the connectivity for efficient prediction.

2.4.6 Scalable Complexity 3D-MeshCompression Stream

SC3DMC compresses the geometry features present in the IndexedFaceSet node as defined in VRML (ISO/IEC 14772-1:1997) as well as multiple and user defined attributes per vertex. SC3DMC contains a rich compression toolbox including several prediction methods and several entropy encoders. This ensures scalability with respect to compression/decompression complexity.

2.5 Interface between Layer 2 and Layer 3

While the relationship between Layer 1 and 2 is straightforward, one being a textual and the other a binary representation of the same information, in order to attach the compression tools (Layer 3) to the scene graph (Layer 2), specific interfaces are defined for each implementation of Layer 1 (for each XML Schema used for scene graph description). These interfaces are normative and are specified in Annexes B, C and D, respectively:

- XMT Schema (standardized in ISO/IEC 14496-11, used for representing textual MPEG-4 scenes)
- COLLADA Schema (proposed by Khronos Group, available at http://www.khronos.org/collada/)
- X3D Schema (standardized by Web3D in ISO/IEC 19775:2004, available at http://www.web3d.org/).

3 Decoder model

The third party scene representations defined by using XML are stored in the 'meta' atom of an MP4 file/stream. The 'meta' atom specifies as well if this content is textual or binary. When binary, the default schema is 'gzip'. Geometry, texture and animation elementary streams (ES) are used in XML as external resources. These ES are stored in the 'moov' atom of the MP4 file/stream as media tracks. After decoding them it is possible to reconstruct the original XML data. The data flow is represented in Figure 2.

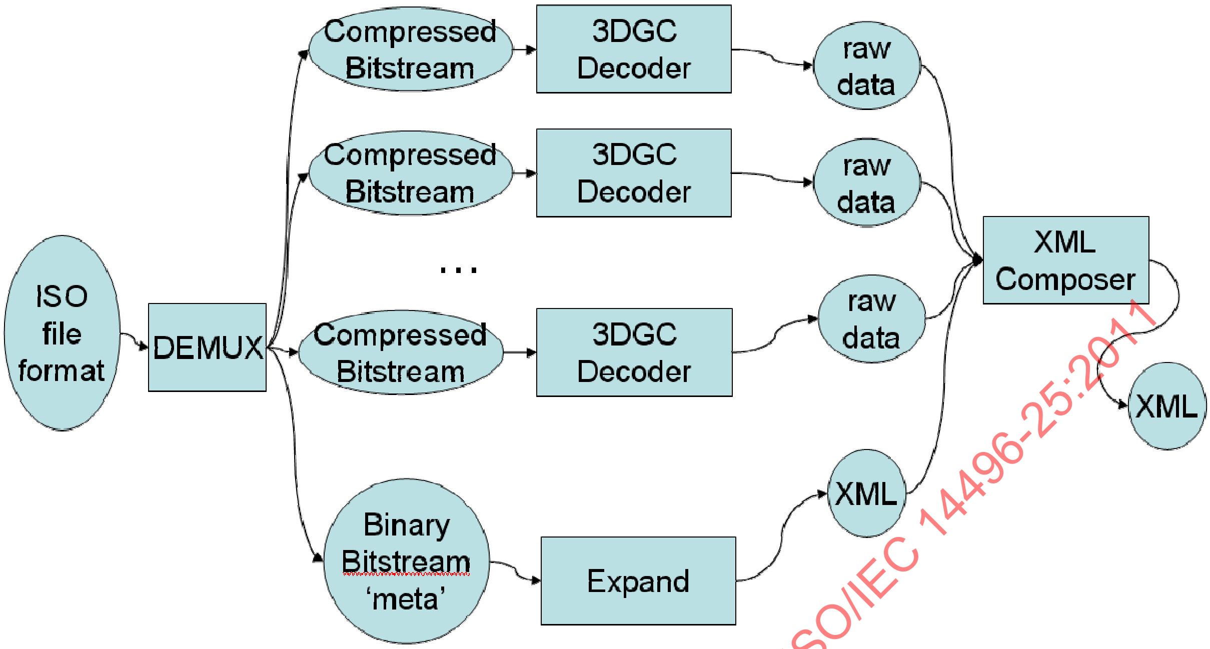

Figure 2 — Decoding path for the 3D Graphics Compression Model

Annex A
(informative)

Encoding model

A potential implementation of an encoder compatible with this part of ISO/IEC 14496 is illustrated in Figure A.1.

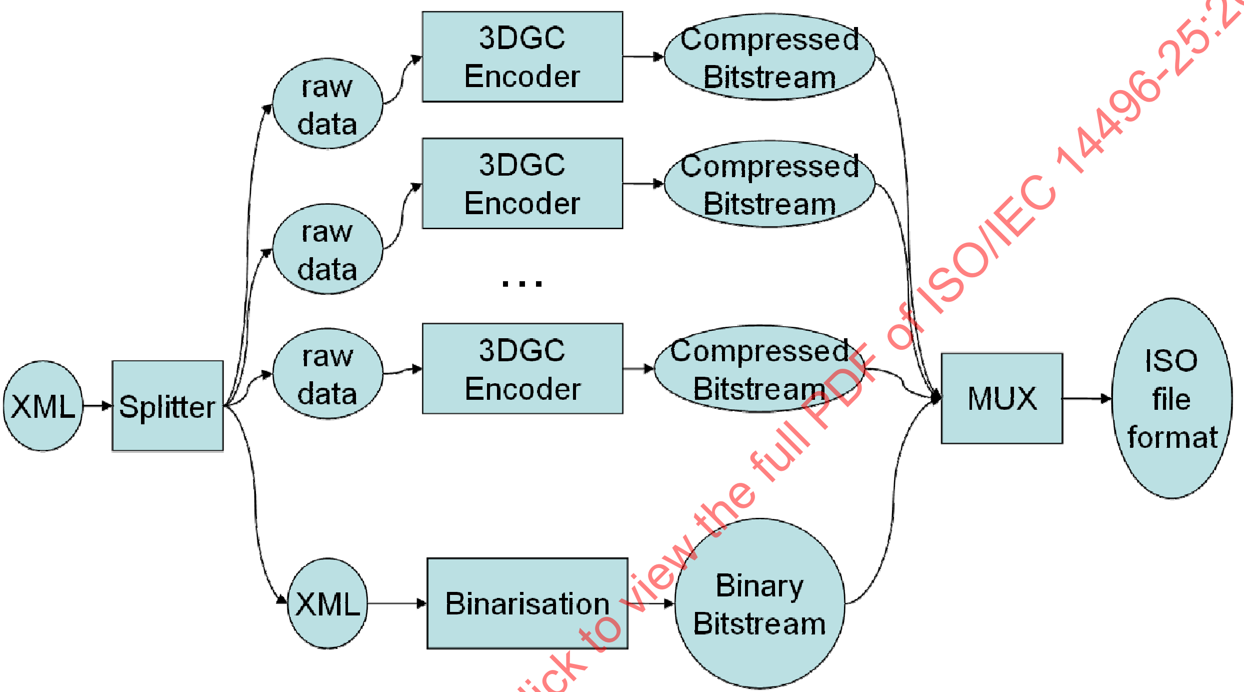

Figure A.1 — Encoding path for the 3D Graphics Compression Model

First, an XML document is split in two sections. The first section contains the graphics primitives that are supported by the MPEG-4 compression tools. The second section contains the other elements of the initial document as well as references to the resources in the first section.

Then, each document is processed individually:

- The first section is processed by the set of MPEG-4 3D graphics encoders introduced in Subclause 1.3 and the output is a collection of elementary streams.
- The second section is optionally processed by the binarization tool in which case the output is one binary stream.

The output of the encoders and the output of the binarization tool are multiplexed in an ISO file format.

Annex B
(normative)

Binarization and Compression of Scene Graph expressed with XMT

B.1 Introduction

This Annex provides the mapping between XMT scene graph nodes and 3D Graphics tools specified in ISO/IEC 14496-16.

B.2 Scene Graph and Object Graph binarization

The binarization tool for XMT content is 'gzip'. The binary buffer is represented into the 'meta' atom.

B.3 3DMCe and SC3DMC Streams

3DMCe and SC3DMC are connected to the IndexedFaceSet element defined as follows:

```
<!-- IndexedFaceSet -->
  <complexType name="IndexedFaceSetType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
      <element name="color" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SFColorNodeType" minOccurs="0" />
        </complexType>
      </element>
      <element name="coord" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SFCoordinateNodeType" minOccurs="0" />
        </complexType>
      </element>
      <element name="normal" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SFNormalNodeType" minOccurs="0" />
        </complexType>
      </element>
      <element name="texCoord" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SFTextureCoordinateNodeType" minOccurs="0" />
        </complexType>
      </element>
    </all>
    <attribute name="ccw" type="xmta:SFBool" use="optional" default="true" />
    <attribute name="colorIndex" type="xmta:MFInt32" use="optional" />
    <attribute name="colorPerVertex" type="xmta:SFBool" use="optional" default="true" />
    <attribute name="convex" type="xmta:SFBool" use="optional" default="true" />
    <attribute name="coordIndex" type="xmta:MFInt32" use="optional" />
    <attribute name="creaseAngle" type="xmta:SFFloat" use="optional" default="0" />
    <attribute name="normalIndex" type="xmta:MFInt32" use="optional" />
    <attribute name="normalPerVertex" type="xmta:SFBool" use="optional" default="true" />
    <attribute name="solid" type="xmta:SFBool" use="optional" default="true" />
    <attribute name="texCoordIndex" type="xmta:MFInt32" use="optional" />
    <attributeGroup ref="xmta:DefUseGroup" />
  </complexType>
```

All attributes are supported by 3DMCe and SC3DMC.

B.4 PI, OI, CI Streams

PI, OI and CI are connected to the PositionInterpolator, OrientationInterpolator and CoordinateInterpolator, elements, respectively, defined as follows:

```
<!-- PositionInterpolator -->
  <complexType name="PositionInterpolatorType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
    </all>
    <attribute name="key" type="xmta:MFFloat" use="optional" />
    <attribute name="keyValue" type="xmta:MFVec3f" use="optional" />
    <attributeGroup ref="xmta:DefUseGroup" />
   </complexType>



  <!-- OrientationInterpolator -->
  <complexType name="OrientationInterpolatorType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
    </all>
    <attribute name="key" type="xmta:MFFloat" use="optional" />
    <attribute name="keyValue" type="xmta:MFRotation" use="optional" />
    <attributeGroup ref="xmta:DefUseGroup" />
  </complexType>



<!-- CoordinateInterpolator -->
  <complexType name="CoordinateInterpolatorType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
    </all>
    <attribute name="key" type="xmta:MFFloat" use="optional" />
    <attribute name="keyValue" type="xmta:MFVec3f" use="optional" />
    <attributeGroup ref="xmta:DefUseGroup" />
  </complexType>
```

All attributes are supported by PI, OI and CI respectively.

B.5 BBA Stream

The BBA stream is connected to the bone transforms defined by SBBone element and weights field defined by MorphShape element.


```
  <!-- SBBone -->
  <complexType name="SBBoneType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
      <element name="children" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SF3DNodeType" minOccurs="0" maxOccurs="unbounded" />
        </complexType>
      </element>
    </all>
    <attribute name="boneID" type="xmta:SFInt32" use="optional" default="0" />
    <attribute name="center" type="xmta:SFVec3f" use="optional" default="0 0 0" />
    <attribute name="endpoint" type="xmta:SFVec3f" use="optional" default="0 0 1" />
    <attribute name="falloff" type="xmta:SFInt32" use="optional" default="1" />
    <attribute name="ikChainPosition" type="xmta:SFInt32" use="optional" default="0" />
    <attribute name="ikPitchLimit" type="xmta:MFFloat" use="optional" />
    <attribute name="ikRollLimit" type="xmta:MFFloat" use="optional" />
    <attribute name="ikTxLimit" type="xmta:MFFloat" use="optional" />
    <attribute name="ikTyLimit" type="xmta:MFFloat" use="optional" />
    <attribute name="ikTzLimit" type="xmta:MFFloat" use="optional" />
    <attribute name="ikYawLimit" type="xmta:MFFloat" use="optional" />
```

```
    <attribute name="rotation" type="xmta:SFRotation" use="optional" default="0 0 1 0" />
    <attribute name="rotationOrder" type="xmta:SFInt32" use="optional" default="0" />
    <attribute name="scale" type="xmta:SFVec3f" use="optional" default="0 0 0" />
    <attribute name="scaleOrientation" type="xmta:SFRotation" use="optional" default="0 0 1 0" />
    <attribute name="sectionInner" type="xmta:MFFloat" use="optional" />
    <attribute name="sectionOuter" type="xmta:MFFloat" use="optional" />
    <attribute name="sectionPosition" type="xmta:MFFloat" use="optional" />
    <attribute name="skinCoordIndex" type="xmta:MFInt32" use="optional" />
    <attribute name="skinCoordWeight" type="xmta:MFFloat" use="optional" />
    <attribute name="translation" type="xmta:SFVec3f" use="optional" default="0 0 0" />
    <attributeGroup ref="xmta:DefUseGroup" />
  </complexType>

<!-- MorphShape -->
  <complexType name="MorphShapeType">
    <all>
      <element ref="xmta:IS" minOccurs="0" />
      <element name="baseShape" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SF3DNodeType" minOccurs="0" />
        </complexType>
      </element>
      <element name="targetShapes" form="qualified" minOccurs="0">
        <complexType>
          <group ref="xmta:SF3DNodeType" minOccurs="0" maxOccurs="unbounded" />
        </complexType>
      </element>
    </all>
    <attribute name="morphID" type="xmta:SFInt32" use="optional" default="0" />
    <attribute name="weights" type="xmta:MFFloat" use="optional" />
    <attributeGroup ref="xmta:DefUseGroup" />
    </complexType>
```

The BBA stream contains updated values of the following elements: `center`, `rotation`, `scale`, `scaleOrientation`, `translation` and `weights`.

B.6 FAMC Stream

FAMC is connected to the IndexedFaceSet element when the latter is animated by using a coordinate interpolator.

Annex C
(normative)

Binarization and Compression of Scene Graph expressed with COLLADA

C.1 Introduction

This Annex provides the mapping between COLLADA scene graph and graphics elements and 3D Graphics tools specified in ISO/IEC 14496-11 and ISO/IEC 14496-16.

C.2 Scene Graph and Object Graph binarization

The binarization tool for COLLADA content is 'gzip'. The binary buffer is represented into the 'meta' atom.

C.3 3DMCe and SC3DMC Streams

In COLLADA, meshes are represented in a MESH element with the following structure:

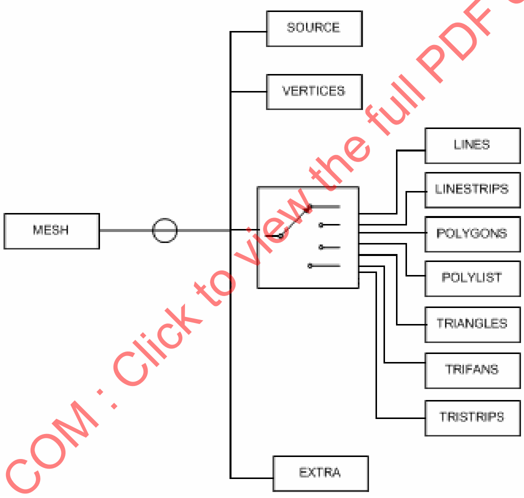

Figure C.1 — COLLADA MESH element

3DMCe and SC3DMC substitute COLLADA MESH element which contains triangles or polygons.

Polylist have to be converted to list of indexed polygons, trifans and tristrips have to be converted to indexed triangles.

3DMCe can be used in order to encode multiple attributes of same type per vertex, such as tangent vectors, binormal vectors or generic vectors (called UVs in COLLADA).

