

中华人民共和国国家标准

GB/T XXXXX—XXXX

非结构化数据表示规范

Unstructured data expression specification

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上

XXXX—XX—XX 发布

XXXX—XX—XX 实施

中华人民共和国国家质量监督检验检疫总局
中国国家标准化管理委员会 发布

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 非结构化数据 DDL 表示框架	2
5.1 概述	2
5.2 整体框架	2
5.3 基本抽象描述符	3
5.4 基本抽象描述方案	6
5.5 根元素	8
6 数据类型	9
6.1 基本数据类型	9
6.1.1 数值数据类型	9
6.1.2 字符串数据类型	10
6.1.3 日期数据类型	10
6.2 扩展数据类型	10
6.2.1 标量	10
6.2.2 向量	11
6.2.3 矩阵	12
7 基本属性描述	13
7.1 基本元素	13
7.1.1 唯一标识符描述符	13
7.1.2 时间描述方案	14
7.1.3 人物描述符	14
7.1.4 地点描述方案	15
7.2 创建信息描述方案	16
7.3 格式与大小描述方符	16
7.4 数据定位描述符	16
8 语义特征描述	17
8.1 对象描述方案	17
8.2 事件描述方案	17
8.3 概念描述方案	17
8.4 数据编码描述方案	18
8.5 数据语义描述方案	18
9 底层特征描述	19
9.1 图像类型	19

9.1.1 颜色相关描述符	19
9.1.2 纹理特征	26
9.1.3 形状特征	27
9.1.4 局部特征	28
9.2 音频类型	30
9.2.1 标量序列	30
9.2.2 低级音频描述符	32
9.2.3 静音描述符	34
9.2.4 MFCC 特征描述符	35
9.3 视频类型	35
9.3.1 时间域特征描述类型	35
9.3.2 运动特征	37
9.3.3 镜头和关键帧描述符	40
9.4 文本类型	41
9.4.1 关键字序列描述符	41
9.4.2 TF 描述符	42
9.4.3 TF-IDF 描述符	42
10 原始数据描述	42
11 关联描述符	43
11.1 语义特征-底层特征关联描述符	43
11.2 基本属性-原始数据关联描述符	43
11.3 语义特征-原始数据关联描述符	44
11.4 底层特征-原始数据关联描述符	44
附 录 A (规范性附录) Unsigned 类型语法定义	45
附 录 B (资料性附录) 非结构化数据表示 XML Schema 定义	50
B.1 非结构化数据表示 XML Schema 定义	50
B.2 InstanceBasicAttributionDSType	50
B.3 InstanceSemanticDSType	51
B.4 InstanceLowFeatureDSType	51
B.5 InstanceRawDataDSType	55
B.6 InstanceAssociationDType	55

前 言

本标准按照GB/T 1.1-2009给出的规则起草。

本标准由全国信息技术标准化技术委员会（SAC/TC28）提出并归口。

本标准起草单位：

本标准主要起草人：

非结构化数据表示规范

1 范围

本标准规定了文本、音频、视频、图像等非结构化数据的统一数据表示。
本标准适用于非结构化管理系统中非结构化数据的表示。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 18793-2002 信息技术 可扩展置标语言（XML）1.0

ISO/IEC 15938 信息技术 多媒体内容描述接口（Information technology — Multimedia content description interface）

XML Schema 可扩展置标语言模式（XML Schema, W3C Recommendation, 2 May 2001）

3 术语和定义

下列术语和定义适用于本文件。

3.6

特征 feature

数据的特性。

3.7

逆向文件频率 inverse document frequency, IDF

一个词语普遍重要性的度量，由总文件数目除以包含该词语的文件数目，将得到的商取对数获得。

3.8

词频 term frequency, TF

某一给定的词语在文件中出现的次数。

4 缩略语

下列缩略语适用于本标准。

D 描述符（Descriptor）

DDL 描述定义语言（Description Definition Language）

DS 描述方案（Description Scheme）

IDF 逆向文件频率（Inverse Document Frequency）

GPS 全球定位系统（Global Positioning System）

MFCC 梅尔频率倒谱系数（Mel Frequency Cepstrum Coefficient）

SIFT 尺度不变特征转换（Scale-Invariant Feature Transform）

TF 词频 (Term Frequency)

TF-IDF 词频-逆向文件频率 (term frequency - inverse document frequency)

UML 统一建模语言 (Unified Modeling Language)

5 非结构化数据 DDL 表示框架

5.1 概述

本标准通过抽象类定义的方式，用XML Schema来表达非结构化数据DDL表示框架，XML Schema的语法及使用方式见GB/T 18793-2002和XML Schema。

5.2 整体框架

非结构化数据表示的 UML 类图如图 1 所示。

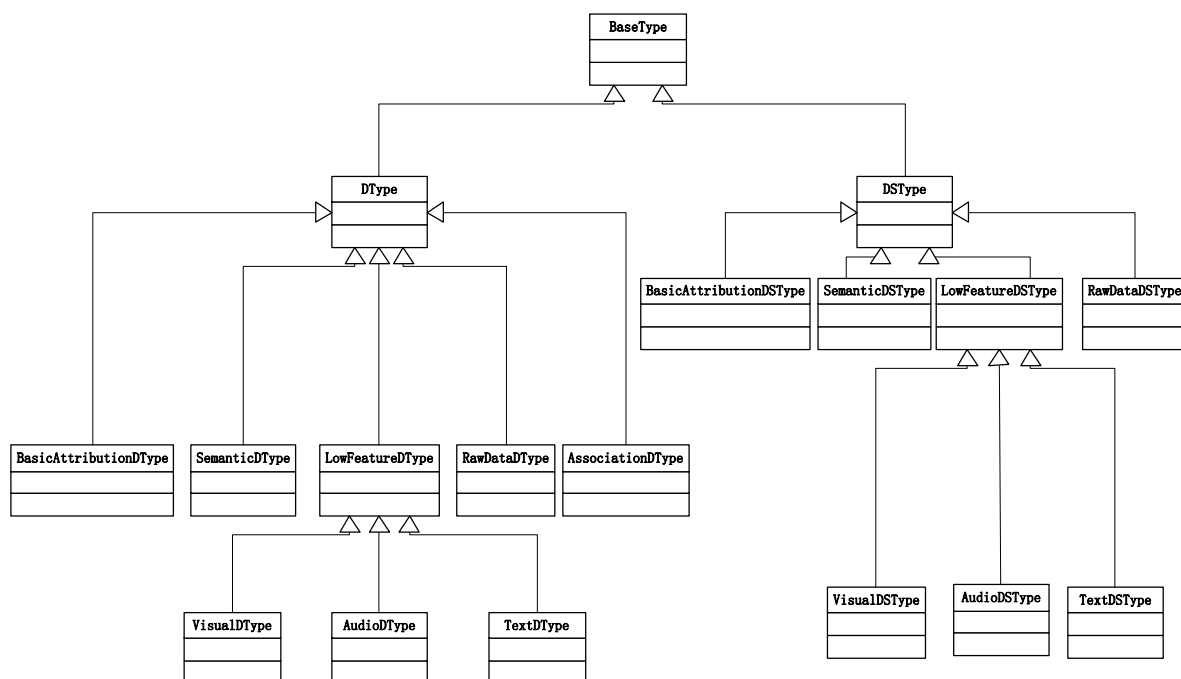


图 1 标准整体框架 UML 类图

本标准定义了 BaseType、DType、DStype、UnstructuredDataType、BasicAttributionDStype、SemanticDStype、LowFeatureDStype、RawDataDStype、AssociationDType、BasicAttributionDType、SemanticDType、LowFeatureDType、RawDataDType、VisualDType、AudioDType、TextDType、VisualDStype、AudioDStype 和 TextDStype 共 18 个抽象类。具体各抽象类的含义如下：

- BaseType 是基准类型；
- DType 是抽象的描述符类型，是所有描述符的超类；
- DStype 是抽象的描述方案类型，是所有描述方案的超类；
- BasicAttributionDStype 是基本属性的描述方案基准类型；
- SemanticDStype 是语义特征的描述方案基准类型；
- LowFeatureDStype 是底层特征的描述方案基准类型；
- RawDataDStype 是原始数据的描述方案基准类型；

- g) AssociatioDType 用于描述基本属性、语义特征、底层特征和原始数据之间的关联关系;
- h) BasicAttributionDType 是基本属性的描述符基准类型;
- i) SemanticDType 是语义特征的描述符基准类型;
- j) LowFeatureDType 是底层特征的描述符基准类型;
- k) RawDataDType 是原始数据的描述符基准类型;
- l) VisualDSType 是图像或视频类数据对其中的图像部分进行操作提取的描述方案基准类型;
- m) AudioDSType 是音频相关特征描述方案基准类型;
- n) TextDSType 是文本相关特征描述方案基准类型;
- o) VisualDType 是图像或视频类数据对其中的图像部分进行操作提取的描述符基准类型;
- p) AudioDType 是音频相关特征描述符基准类型;
- q) TextDType 是文本相关特征描述符基准类型。

其中 BaseType 的语法形式定义如下:

```
<complexType name="BaseType" abstract="true">
  <complexContent>
    <restriction base="anyType"/>
  </complexContent>
</complexType>
```

DType的语法形式定义如下:

```
-<complexType name="DType" abstract="true">
  -<complexContent>
    <extension base="BaseType"/>
  </complexContent>
</complexType>
```

DSType的语法形式定义如下:

```
<complexType name="DSType" abstract="true">
  <complexContent>
    <extension base="BaseType"/>
  </complexContent>
</complexType>
```

5.3 基本抽象描述符

非结构化数据基本抽象描述符的关系如图 2 所示:

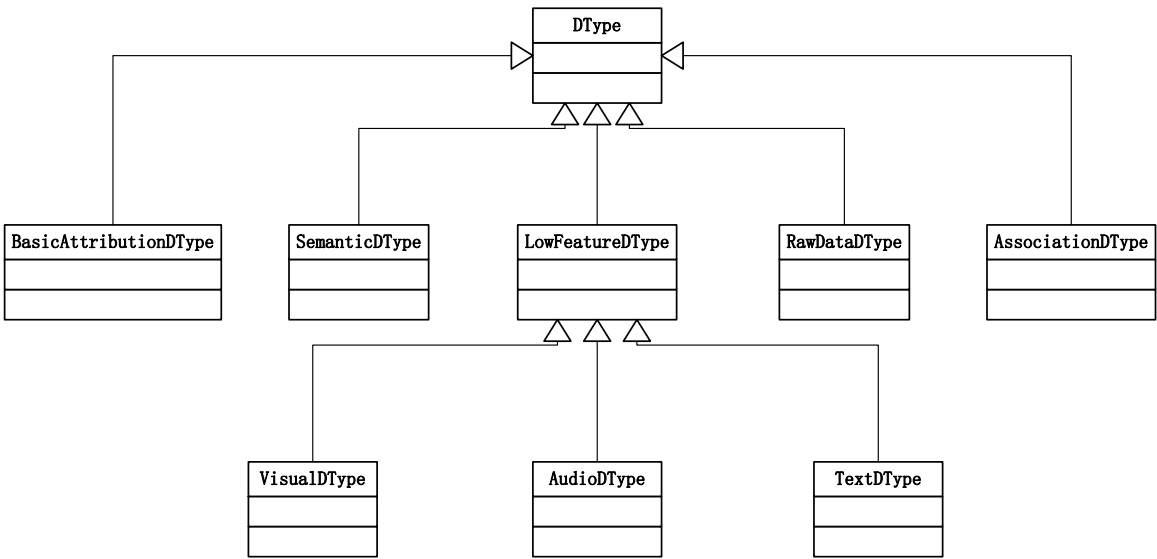


图 2 基本抽象描述符

基本属性描述符、语义特征描述符、底层特征描述符、原始数据描述符和关联描述符这 5 类基本抽象描述符的语法定义如下：

```

-<complexType name="BasicAttributionDType" abstract="true">
  -<complexContent>
    <extension base="DType"/>
  </complexContent>
</complexType>

-<complexType name="SemanticDType" abstract="true">
  -<complexContent>
    -<extension base="DType">
  </complexContent>
</complexType>

<complexType name="LowFeatureDType" abstract="true">
  -<complexContent>
    <extension base="DType"/>
  </complexContent>
</complexType>

<complexType name="RawDataDType" abstract="true">
  -<complexContent>
    <extension base="DType"/>
  </complexContent>
</complexType>

<complexType name="AssociationDType" abstract="true">
  -<complexContent>
    <extension base="DType"/>
  </complexContent>
</complexType>

```

视频、音频、文本的描述符抽象类型的语法定义如下，其中视频描述符主要指基于图像内容进行操作所获得的描述符：

```
-<complexType name="VisualDType" abstract="true">
  -<complexContent>
    <extension base="LowFeatureDType"/>
  </complexContent>
</complexType>

-<complexType name="AudioDType" abstract="true">
  -<complexContent>
    -<extension base="LowFeatureDType">
      <attribute name="channels" type="integerVector" use="optional"/></extension>
    </complexContent>
  </complexType>

<complexType name="TextDType" abstract="true">
  -<complexContent>
    <extension base="LowFeatureDType"/>
  </complexContent>
</complexType>
```

5.4 基本抽象描述方案

非结构化数据基本抽象描述方案的关系如图 3 所示：

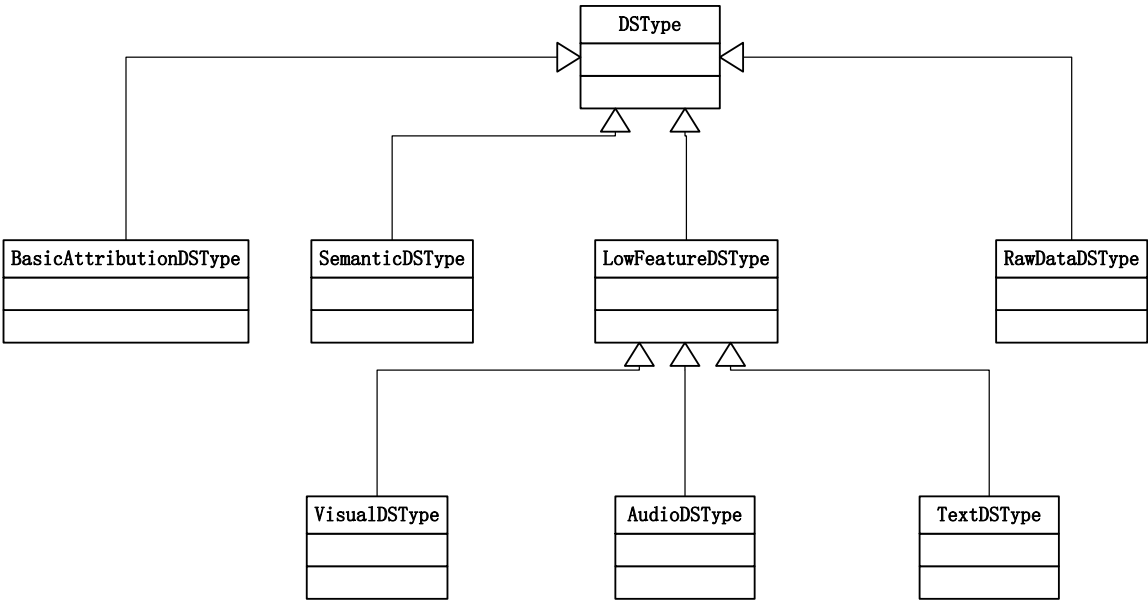


图 3 基本抽象描述方案

基本属性描述方案，语义特征描述方案，底层特征描述方案和原始数据描述方案语法形式定义如下：

```

<!-- 基本属性描述方案定义 -->
-<complexType name="BasicAttributionDSType" abstract="true">
  <complexContent>
    <extension base="DSType"/>
  </complexContent>
</complexType>
<!-- 语义特征描述方案定义 -->
-<complexType name="SemanticDSType" abstract="true">
  <complexContent>
    <extension base="DSType"/>
  </complexContent>
</complexType>
<!-- 底层特征描述方案定义 -->
<complexType name="LowFeatureDSType" abstract="true">
  <complexContent>
    <extension base="DSType"/>
  </complexContent>
</complexType>
<!-- 原始数据描述方案定义 -->
<complexType name="RawDataDSType" abstract="true">
  <complexContent>
    <extension base="DSType"/>
  </complexContent>
</complexType>
<!-- 关联描述符定义 -->
<complexType name="AssociationDType" abstract="true">
  <complexContent>
    <extension base="DType"/>
  </complexContent>
</complexType>

```

视频、音频、文本的描述方案抽象类型的语法定义如下：

```

-<complexType name="VisualDSType" abstract="true">
-  <complexContent>
    <extension base="LowFeatureDSType"/>
  </complexContent>
</complexType>

-<complexType name="AudioDSType" abstract="true">
  -<complexContent>
    -<extension base="LowFeatureDSType">
      <attribute name="channels" type="integerVector" use="optional"/>
    </extension>
  </complexContent>
</complexType>

-<complexType name="TextDSType" abstract="true">
  -<complexContent>
    <extension base="LowFeatureDSType"/>
  </complexContent>
</complexType>

```

5.5 根元素

XML 描述文件根元素从 unstructuredData 节点开始，子元素为 UnstructuredDataType 实例序列，每一个 UnstructuredDataType 实例由其基本属性描述方案、语义属性描述方案、底层特征描述方案以及原始数据描述方案及其 ID 属性构成。其语法形式定义如下：

```

<!-- 根元素 -->
<element name="unstructuredData">
  <complexType>
    <complexContent>
      <extension base="BaseType">
        <!-- 根元素包含多个非结构化数据 -->
        <sequence>
          <element name="datum" type="UnstructuredDataType" minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
      </extension>
    </complexContent>
  </complexType>
</element>
-

```

其中 UnstructuredDataType 包含非结构化数据类型定义，包括基本属性描述方案，语义特征描述方案，底层特征描述方案、原始数据描述方案以及关联关系描述符，其语法形式定义如下：

```
<complexType name=" UnstructuredDataType ">
  <sequence>
    <element name="basicDS" type="BasicAttributionDSType" minOccurs="1" maxOccurs="unbounded"/>
    <element name="semanticDS" type="SemanticDSType" minOccurs="1" maxOccurs="unbounded"/>
    <element name="lowFeatureDS" type="LowFeatureDSType" minOccurs="1" maxOccurs="unbounded"/>
    <element name="rawDataDS" type="RawDataDSType" minOccurs="1" maxOccurs="unbounded"/>
    <element name="association" type="AssociationDType " minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
  <attribute name="id" type="ID" use="required">
</complexType>
</complexType>
```

6 数据类型

6.1 基本数据类型

6.1.1 数值数据类型

数值数据类型如表1所示。

表1 数值数据类型

名称	描述
byte	有正负的 8 位整数
decimal	十进制数
int	有正负的 32 位整数
integer	整数值
long	有正负的 64 位整数
negativeInteger	仅包含负值的整数（..，-2，-1）
nonNegativeInteger	仅包含非负值的整数（0，1，2，..）
nonPositiveInteger	仅包含非正值的整数（..，-2，-1，0）
positiveInteger	仅包含正值的整数（1，2，..）
short	有正负的 16 位整数
unsignedLong	无正负的 64 位整数
unsignedInt	无正负的 32 位整数
unsignedShort	无正负的 16 位整数
unsignedByte	无正负的 8 位整数
boolean	布尔型
double	浮点双精度
float	浮点单精度

hexBinary	二进制块类型
-----------	--------

6.1.2 字符串数据类型

字符串数据类型如表2所示。

表2 字符串数据类型

名称	描述
ID	在 XML 中提交 ID 属性的字符串
IDREF	在 XML 中提交 IDREF 属性的字符串
IDREFS language	包含合法的语言 id 的字符串
Name	包含合法 XML 名称的字符串
NMTOKEN	在 XML 中提交 NMTOKEN 属性的字符串
normalizedString	不包含换行符、回车或制表符的字符串
string	字符串
token	不包含换行符、回车或制表符、开头或结尾空格或者多个连续空格的字符串

6.1.3 日期数据类型

日期数据类型如表3所示。

表3 日期数据类型

名称	描述
date	定义一个日期值：“YYYY-MM-DD”
dateTime	定义一个日期和时间值：“YYYY-MM-DDThh:mm:ss”
duration	定义一个时间间隔：时间间隔使用下面的格式来规定：“PnYnMnDTnHnMnS”，其中 P 表示周期（必需）；nY 表示年数；nM 表示月数；nD 表示天数；T 表示时间部分的起始（若规定小时、分钟和秒，则此选项为必需）；nH 表示小时数；nM 表示分钟数；nS 表示秒数。
gDay	定义日期的一个部分 —— 天（DD）
gMonth	定义日期的一个部分 —— 月（MM）
gMonthDay	定义日期的一个部分 —— 月和天（MM-DD）
gYear	定义日期的一个部分 —— 年（YYYY）
gYearMonth	定义日期的一个部分 —— 年和月（YYYY-MM）
time	定义一个时间值：“hh:mm:ss”

6.2 扩展数据类型

6.2.1 标量

6.2.1.1 Unsigned 类型

语法定义见附录 A。

6.2.1.2 取值区间[0,1]数据类型

语法形式定义如下：


```

<simpleType name="zeroToOneType">
  <restriction base="float">
    <minInclusive value="0.0"/> <maxInclusive value="1.0"/>
  </restriction>
</simpleType>

```

6.2.1.3 取值区间 $[-1, 1]$ 数据类型

语法形式定义如下：

```

<simpleType name="minusOneToOneType">
  <restriction base="float">
    <minInclusive value="-1.0"/> <maxInclusive value="+1.0"/>
  </restriction>
</simpleType>

```

6.2.1.4 非负实数类型

语法形式定义如下：

```

<simpleType name="nonNegativeReal">
  <restriction base="double">
    <minInclusive value="0.0"/>
  </restriction>
</simpleType>

```

6.2.2 向量

6.2.2.1 整型向量类型

语法形式定义如下：

```

-<simpleType name="integerVector">
  <list itemType="integer"/>
</simpleType>

```

6.2.2.2 浮点向量类型

语法形式定义如下：

```
<simpleType name="floatVector">
  <list itemType="float"/>
</simpleType>
```

6.2.2.3 字符串向量类型

语法形式定义如下：

```
<simpleType name="stringVector">
  <list itemType="string"/>
</simpleType>
```

6.2.2.4 元素取值区间[0, 1]向量类型

语法形式定义如下：

```
<simpleType name="zeroToOneVector">
  <list>
    <simpleType>
      <restriction base="float">
        <minInclusive value="0"/>
        <maxInclusive value="1.0"/>
      </restriction>
    </simpleType>
  </list>
</simpleType>
```

6.2.3 矩阵

6.2.3.1 整型矩阵类型

语法形式定义如下：

```
-<complexType name="IntegerMatrixType">
  -<simpleContent>
    -<extension base="integerVector">
      <attribute use="required" ref="reservedDim"/>
    </extension>
  </simpleContent>
</complexType>
```

其中 reservedDim 类型语法形式定义如下，该内置属性类型仅用于矩阵类型的定义使用。

```

<attribute name="reservedDim">
  <simpleType>
    <restriction base="listOfPositiveIntegerForDim">
      <minLength value="1"/>
    </restriction>
  </simpleType>
</attribute>

<simpleType name="listOfPositiveIntegerForDim">
  <list itemType="positiveInteger"/>
</simpleType>

```

6.2.3.2 浮点型矩阵类型

语法形式定义如下：

```

-<complexType name="FloatMatrixType">
  -<simpleContent>
    -<extension base="floatVector">
      <attribute use="required" ref="reservedDim"/>
    </extension>
  </simpleContent>
</complexType>

```

7 基本属性描述

7.1 基本元素

7.1.1 唯一标识符描述符

此描述符可作为多媒体内容描述的唯一标识，用此描述符可唯一区分和定位对应非结构化数据。其语法形式定义如下：

```

<complexType name="UIdentifierDType">
  <element name="IdValue" type="string"/>
  <attribute name="IdOrganization" type="string"/>
  <attribute name="IdName" type="string"/>
</complexType>

```

其中：

UIdentifier：描述多媒体内容唯一标识符的类型；

IdValue：唯一标识符的值；

IdOrganization：指定标识的组织或者机构；

IdName：标识符的名称。

7.1.2 时间描述方案

此描述方案用于规范时间段描述，由时间点描述符（开始时间）和时间段描述符（持续时间）两个元素组成。其语法形式定义如下：

```
<complexType name="TimeDSType">
  <sequence>
    <element name="TimePoint" type="TimePointDType"/>
    <element name="Duration" type="DurationDType"/>
  </sequence>
</complexType>
```

其中，TimePoint描述符用于描述具体某个时间点，其语法格式定义如下：

```
<simpleType name="FractionalTime" base="string" derivedBy="restriction">
  <pattern value="-?(\d+(\-\d{2})(\-\d{2})?)?(T\d{2}(:\d{2}(:\d{2}(\:\d+(\.\d{2})?)?)?)?
    ((\-\|+)\d{2}:\d{2})?(F\d+)?)"/>
</simpleType>
<simpleType name="TimePointDType" base="FractionalTime" derivedBy="restriction">
  <pattern value="-?(\d+(\-\d{2})(\-\d{2})?)?(T\d{2}(:\d{2}(:\d{2}(\:\d+(\.\d{2})?)?)?)?
    ((\-\|+)\d{2}:\d{2})?(F\d+)?)"/>
</simpleType>
```

Duration 描述符用于描述某个具体时长，其语法格式定义如下

```
<simpleType name="FractionalDuration" base="string" derivedBy="restriction">
  <pattern value="-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?(\d{2}f)?)?
    (\-\|+)\d{2}:\d{2}Z)?(\d+F)"/>
</simpleType>
<simpleType name="DurationDType" base="FractionalDuration" derivedBy="restriction">
  <pattern value="-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?)
    (\-\|+)\d{2}:\d{2}Z)?(\d+F)"/>
</simpleType>
```

7.1.3 人物描述符

此描述符主要包括人物ID和姓名两个属性。其语法形式定义如下：

```

<complexType name="PersonDType">
  <attribute name="id" type="ID"/>
  <attribute name="name" type="string"/>
</complexType>
<element name="Person" type="Person"/>

```

其中：

PersonDType：人物描述符类型；

id：一个人物描述实例的标识符；

name：一个人物描述实例的姓名。

7.1.4 地点描述方案

此描述方案用于位置的描述，位置可以是真实的、历史的或是虚构的。该描述方案中前缀为“mpeg7:”的数据类型定义见ISO/IEC 15938，其语法形式定义如下：

```

<complexType name="PlaceDSType">
  <element name="PlaceName" type="mpeg7:TextualDescription" maxOccurs="unbounded"/>
  <element name="PlaceRole" type="mpeg7:ControlledTerm" minOccurs="0"/>
  <element name="Planet" type="mpeg7:ControlledTerm" minOccurs="0"/>
  <element name="GPSCoordinates" type="mpeg7:GPSCoordinates" minOccurs="0"/>
  <element name="Country" type="mpeg7:ISO3166-1CountryCode" minOccurs="0"/>
  <element name="Region" type="mpeg7:ISO3166-2RegionCode" minOccurs="0"/>
  <element name="PostingIdentifier" type="mpeg7:TextualDescription" minOccurs="0"/>
  <element name="AdministrativeUnit" type="string" minOccurs="0"/>
  <element name="PostalAddress" type="mpeg7:PostalAddress" minOccurs="0"/>
  <element name="InternalCoordinates" type="string" minOccurs="0"/>
  <attribute name="id" type="ID" use="optional"/>
</complexType>

```

其中：

PlaceDSType：描述方案描述的对象；

PlaceName：位置的名称；

PlaceRole：位置的作用，例如射击位置、邮寄地址等；

Planet：星球的名称，默认为地球；

GPSCoordinates：GPS系统上的坐标；

Country：国家代码；

Region：地区代码；

PostingIdentifier：邮政编码；

AdimistrativeUnit：行政单位的名称；

PostalAddress：邮政地址；

InternalCoordinates：用于描述内部地址的文本；

id：位置描述实例的标识符。

7.2 创建信息描述方案

此描述方案用于描述内容的创建，包括创建者、创建地点、创建时间和创建描述。其语法形式定义如下：

```
<complexType name="CreationDSType">
  <element name="Creator" type="PersonDType" minOccurs="0" maxOccurs="unbounded"/>
  <element name="CreationLocation" type="PlaceDSType" minOccurs="0"
maxOccurs="unbounded"/>
  <element name="CreationDate" type="TimePointDType" minOccurs="0"
maxOccurs="unbounded"/>
  <element name="CreationDescription" type="string" minOccurs="0" maxOccurs="unbounded"/>
</complexType>
```

其中：

CreationDSType：表示非结构化数据类型内容的创建，包括时间、地点等；

Title：媒体内容的标题，可以是文本或者图像；

Creator：非结构化数据的创作者；

CreationLocation：非结构化数据的创建地点；

CreationDate：非结构化数据的创建日期；

CreationDescription：基于文本注释的文本内容的描述。

7.3 格式与大小描述方符

此描述符用于描述非结构化数据的文件格式以及文件大小。其语法形式定义如下：

```
<complexType name="FormatAndSizeDType">
  <element name="FileFormat" type="string"/>
  <element name="FileSize" type="nonNegativeInteger" minOccurs="0"/>
</complexType>
```

其中：

FileFormat：非结构化数据实例的文件格式；

FileSize：非结构化数据的物理存储大小。

7.4 数据定位描述符

此描述符用于指定引用的媒体数据的位置。其语法形式定义如下：

```
<complexType name="DataLocatorDType">
  <element name="DataURL" type="DataURL" minOccurs="0"/>
</complexType>
<simpleType>
  <element name="DataURL" type="String"/>
</simpleType>
```

其中：

DataLocatorDType:引用的数据位置的描述符类型;
DataURL: 使用URL定位。

8 语义特征描述

8.1 对象描述方案

此描述方案描述可察觉的或抽象的对象。一个可察觉的对象是一个存在的实体，具有时空性；一个抽象对象是对一个可察觉对象的抽象和概括的结果。该描述方案中的mpeg7:ReferenceType类型定义见ISO/IEC 15938，其语法形式定义如下：

```
<complexType name="ObjectDSType">
  <complexContent>
    <choice maxOccurs="unbounded" minOccurs="0">
      <element name="Object" type="ObjectDSType"/>
      <element name="ObjectRef" type="mpeg7:ReferenceType"/>
    </choice>
  </complexContent>
</complexType>
```

8.2 事件描述方案

此描述方案用于描述可察觉的或抽象的事件。一个可察觉的事件与一个或多个对象存在一种动态关系，还与叙述世界的时空性相关；一个抽象事件是对可察觉事件的抽象和概括的结果。该描述方案中的mpeg7:ReferenceType类型定义见ISO/IEC 15938，其语法形式定义如下：

```
<complexType name="EventDSType">
  <complexContent>
    <sequence>
      <choice maxOccurs="unbounded" minOccurs="0">
        <element name="Event" type="EventDSType"/>
        <element name="EventRef" type="mpeg7:ReferenceType"/>
      </choice>
      <element name="Place" maxOccurs="unbounded" type="PlaceDSType" minOccurs="0"/>
      <element name="Time" maxOccurs="unbounded" type="TimeDSType" minOccurs="0"/>
    </sequence>
  </complexContent>
</complexType>
```

8.3 概念描述方案

此描述方案描述不能用普通方法或抽象方法对其对象、事件、时间、位置和状态进行描述的语法实体，它是用一种特性或多种特性群来表述的。其语法形式定义如下：

```

<complexType name="ConceptDSType">
  <complexContent>
    <sequence>
      <element name="Definition" type="string" minOccurs="0"/>
      <element name="Property" maxOccurs="unbounded" type="string" minOccurs="0"/>
      <element name="Label" maxOccurs="unbounded" type="string"/>
    </sequence>
  </complexContent>
</complexType>

```

8.4 数据编码描述方案

此描述方案用于指定数据的编码参数。其语法形式定义如下：

```

<complexType name="DataCodingDSType">
  <element name="Resolution" type="nonNegativeInteger" minOccurs="0"/>
  <element name="FrameWidth" type="nonNegativeInteger" minOccurs="0"/>
  <element name="FrameHeight" type="nonNegativeInteger" minOccurs="0"/>
  <element name="FrameRate" type="float" minOccurs="0"/>
  <element name="AudioSamplingRate" type="float" minOccurs="0"/>
  <element name="CompressionFormat" type="string" minOccurs="0"/>
  <element name="QuantizerCoarseness" type="nonNegativeInteger" minOccurs="0"/>
  <element name="PredictionMode" type="string" minOccurs="0"/>
  <element name="BitRate" type="float" minOccurs="0"/>
  <element name="Bandwidth" type="float" minOccurs="0"/>
  <attribute name="id" type="ID"/>
</complexType>

```

其中：

DataCoding：表示数据编码参数的描述方案；

Resolution：每帧的分辨率；

FrameWidth：视频图像帧的宽度，以像素为单位；

FrameHeight：视频图像帧的高度，以像素为单位；

FrameRate：帧速率，以赫兹为单位；

AudioSamplingRate：音频的采样率，以赫兹为单位；

CompressionFormat：视频采用的压缩格式；

QuantizerCoarseness：采样的准确性；

PredictionMode：视频内容的预测编码所采用的算法类型；

BitRate：视频编码的同步传输所需的位速率；

Bandwidth：视频编码覆盖的赫兹带宽；

id：数据编码参数描述方案实例的标识符。

8.5 数据语义描述方案

此方案主要用于存储数据的描述。其语法形式定义如下：

```
<element name="DataInformation" type="DataInformation"/>
  <complexType name="DataInformation">
    <element name="DataIdentification" type="DataIdentification"/>
    <element name="DataProfile" type="DataProfile" maxOccurs="unbounded"/>
    <attribute name="id" type="ID"/>
  </complexType>
<complexType name="DataProfile">
  <element name="DataFormat" type="string"/>
  <element name="DataCoding" type="DataCodingDSType" minOccurs="0" maxOccurs="unbounded"/>
  <element name="Title" type="String"/>
  <element name="Theme" type="String"/>
  <attribute name="id" type="ID"/>
</complexType>
```

- 其中：
- DataInformation：数据信息包含一个或多个数据内容描述；
 - DataIdentification：数据实体的标识；
 - DataProfile：被描述实体的数据内容描述, 该描述主要包括数据格式，数据编码，标题，主题等；
 - id：数据信息描述实例的标识符。

DataIdentification类型的语法形式定义如下：

```
<!-- Definition the DataIdentification DS -->
<complexType name="DataIdentification">
  <element name="Identifier" type="UIIdentifier" minOccurs="0"/>
  <element name="AudioDomain" type="ControlledTerm" minOccurs="0" maxOccurs="unbounded"/>
  <element name="VisualDomain" type="ControlledTerm" minOccurs="0" maxOccurs="unbounded"/>
  <element name="ImageDomain" type="ControlledTerm" minOccurs="0" maxOccurs="unbounded"/>
  <attribute name="id" type="ID"/>
</complexType>
```

- 其中：
- DataIdentification：表示数据内容标识的描述方案；
 - AudioDomain：有关音频的来源、采集或使用的信息；
 - VisualDomain：有关视频的来源、采集或使用的信息；
 - ImageDomain：有关图像的来源、采集或使用的信息。

9 底层特征描述

9.1 图像类型

9.1.1 颜色相关描述符

9.1.1.1 颜色空间类型

颜色空间可由单独的 3x3 的整型矩阵或内置的颜色空间来表示，内置的颜色空间包括：

- a) RGB 类型：RGB 颜色空间；
- b) YCbCr 类型：用 RGB 的一个线性变换来表示：

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

$$Cb = -0.169 * R - 0.331 * G + 0.500 * B$$

$$Cr = 0.500 * R - 0.419 * G - 0.081 * B$$
- c) HSV 类型：用 RGB 的一个非线性变换来表示，由表示颜色的亮度的值（Value）、表示颜色丰富程度的饱和度（Saturation）和表示占主导的频谱颜色的色调（Hue，用 0 度到 360 度的角度来表示）所组成。它们的值由规一化的 RGB 值导出（范围从 0 到 1），变换方法见 ISO/IEC 15938。
- d) HMMD 类型：定义为非线性的、源于 RGB 颜色空间的可逆变换。在 HMMD 颜色空间中有以下五个特性：
 - 1) Hue：与 HSV 中的定义一样；
 - 2) Max：显示有多黑的颜色，给出阴影或黑色的程度；
 - 3) Min：显示有多白的颜色，给出浅色或白色的程度；
 - 4) Diff：显示包含了多少的灰度，并且与纯色有多接近，给出色调和色的程度，转换公式为 $Diff = Max - Min$ ；
 - 5) Sum：仿真颜色的亮度，其转换公式为 $Sum = (Max + Min) / 2$ 。
 Max、Min 和 Hue 的转换公式和 HSV 颜色空间中 Min、Max 和 Hue 的转换公式相同。
- e) LinearMatrix 类型：用一个来自于参照颜色空间成分的线性矩阵变换来表示：

$$C1 = ColorTransMat[0][0] * R + ColorTransMat[0][1] * G + ColorTransMat[0][2] * B$$

$$C2 = ColorTransMat[1][0] * R + ColorTransMat[1][1] * G + ColorTransMat[1][2] * B$$

$$C3 = ColorTransMat[2][0] * R + ColorTransMat[2][1] * G + ColorTransMat[2][2] * B$$
 其中 ColorTransMat 为一 3x3 线性矩阵，其每个元素取值范围为 $[-1, 1]$ 。
- f) Monochrome 类型：用 RGB 的一个线性变换来表示：

$$Y = 0.299 * R + 0.587 * G + 0.114 * B$$

颜色空间的语法形式定义如下：

```

<complexType name="ColorSpaceType">
  <complexContent>
    <extension base="VisualIDType">
      <choice>
        <element name="LinearMatrix">
          <complexType base="IntegerMatrix" derivedBy="restriction">
            <minInclusive value="0"/>
            <maxInclusive value="65535"/>
            <attribute name="sizes" use="fixed" value="3 3"/>
          </complexType>
        </element>
        <attribute name="type" type="colorSpaceType"/>
      </choice>
    </extension>
  </complexContent>
</complexType>

<simpleType name="colorSpaceType" base="string">
  <enumeration value="RGB"/>
  <enumeration value="YCbCr"/>
  <enumeration value="HSV"/>
  <enumeration value="HMMD"/>
  <enumeration value="LinearMatrix"/>
  <enumeration value="Monochrome"/>
</simpleType>

```

9.1.1.2 主导颜色描述符

主导颜色描述符主要用于定义图像/区域中的 1~8 种主要颜色，包括定义了其色彩空间，量化方式，主导颜色的值，色彩值的偏差等。其语法形式定义如下：

```

-<complexType name="DominantColorType" final="#all">
-<complexContent>
  -<extension base="VisualIDType">
    -<sequence>
      <element name="ColorSpace" type="ColorSpaceType" minOccurs="0"/>
      <element name="ColorQuantization" type="ColorQuantizationType" minOccurs="0"/>
      <element name="SpatialCoherency" type="unsigned5"/>
      -<element name="Value" maxOccurs="8">
        -<complexType>
          -<sequence>
            <element name="Percentage" type="unsigned5"/> -<element name="Index">
              -<simpleType>
                -<restriction>
                  -<simpleType>
                    <list itemType="unsigned12"/></simpleType>
                    <length value="3"/>
                  </restriction>
                </simpleType>
              </element>
            -<element name="ColorVariance" minOccurs="0">
              -<simpleType>
                -<restriction>
                  -<simpleType>
                    <list itemType="unsigned1"/></simpleType>
                    <length value="3"/>
                  </restriction>
                </simpleType>
              </element>
            </sequence>
          </complexType>
        </element>
      </sequence>
    </extension>
  </complexContent>
</complexType>

```

其中 ColorQuantizationType 类型的语法形式定义如下：

```

-<complexType name="ColorQuantizationType" final="#all">
-<sequence maxOccurs="3">
  -<element name="Component">
    -<simpleType>
      -<restriction base="string">
        <enumeration value="R"/>
        <enumeration value="G"/>
        <enumeration value="B"/>
        <enumeration value="Y"/>
        <enumeration value="Cb"/>
        <enumeration value="Cr"/>
        <enumeration value="H"/>
        <enumeration value="S"/>
        <enumeration value="V"/>
        <enumeration value="Max"/>
        <enumeration value="Min"/>
        <enumeration value="Diff"/>
        <enumeration value="Sum"/>
      </restriction>
    </simpleType>
  </element>
  <element name="NumOfBins" type="unsigned12"/>
</sequence>
</complexType>

```

9.1.1.3 色彩内容描述符

色彩内容描述符主要包括由 harr 变换得到的色彩直方图、系数数量以及每个系数的位数，其语法形式定义如下：

```

<complexType name="ScalableColorType">
  <complexContent>
    <extension base="VisualDType">
      <element name="Coefficients" type="integerVectorType"/>
      <attribute name="numberOfCoefficients" type="unsigned3"/>
      <attribute name="numberOfBitplanesDiscarded" type="unsigned3"/>
    </extension>
  </complexContent>
</complexType>

```

9.1.1.4 颜色结构描述符

颜色结构描述符语法形式定义如下：

```
-<complexType name="ColorStructureType" final="#all">
  -<complexContent>
    -<extension base="VisualDType">
      -<sequence>
        -<element name="Values">
          -<simpleType>
            -<restriction>
              -<simpleType>
                <list itemType="unsigned8"/>
              </simpleType>
              <minLength value="1"/>
              <maxLength value="256"/>
            </restriction>
          </simpleType>
        </element>
      </sequence>
      <attribute name="colorQuant" type="unsigned3" use="required"/> </extension>
    </complexContent>
  </complexType>
```

9.1.1.5 颜色布局描述符

颜色布局描述符的语法形式定义如下：

```

<complexType name="ColorLayoutType">
  <complexContent>
    <extension base="VisualDType">
      <element name="YCcoeff">
        <complexType>
          <element name="YDCCcoeff" type="unsigned6"/>
          <element name="YACCcoeff">
            <simpleType base="unsigned5" derivedBy=list"/>
            <maxLength value="63"/>
          </simpleType>
        </complexType>
      </element>
      <element name="CbCcoeff">
        <complexType>
          <element name="CbDCCcoeff" type="unsigned6"/>
          <element name="CbACCcoeff">
            <simpleType base="unsigned5" derivedBy=list"/>
            <maxLength value="63"/>
          </simpleType>
        </complexType>
      </element>
      <element name="CrCcoeff">
        <complexType>
          <element name="CrDCCcoeff" type="unsigned6"/>
          <element name="CrACCcoeff">
            <simpleType base="unsigned5" derivedBy=list"/>
            <maxLength value="63"/>
          </simpleType>
        </complexType>
      </element>
      <attribute name="numOfYCcoeff" type="numberOfCoeffType" use="default" value="6"/>
      <attribute name="numOfCCcoeff" type="numberOfCoeffType" use="default" value="3"/>
    </extension>
  </complexContent>
</complexType>

```

其中 numberOfCoeffType 的定义如下:

```

<simpleType name="numberOfCoeffType" base="positiveInteger">
  <enumeration value="1"/>
  <enumeration value="3"/>
  <enumeration value="6"/>
  <enumeration value="10"/>
  <enumeration value="15"/>
  <enumeration value="21"/>
  <enumeration value="28"/>
  <enumeration value="64"/>
</simpleType>

```

9.1.2 纹理特征

9.1.2.1 均匀纹理特征

均匀纹理特征主要用于表示图像中的纹理方面的特征，其语法形式定义如下：

```

-<complexType name="HomogeneousTextureType">
  <complexContent>
    <extension base="VisualDType">
      <element name="Average" type="unsigned8"/>
      <element name="StandardDeviation" type="unsigned8"/>
      <element name="Energy">
        <simpleType base="unsigned8" derivedBy="list">
          <length value="30"/>
        </simpleType>
      </element>
      <element name="EnergyDeviation" minOccurs="0" maxOccurs="1">
        <simpleType base="unsigned8" derivedBy="list">
          <length value="30"/>
        </simpleType>
      </element>
    </extension>
  </complexContent>
</complexType>

```

9.1.2.2 边缘直方图描述符

边缘直方图描述符主要用于图像中的边缘信息，其语法形式定义如下：


```

-<complexType name="EdgeHistogramType">
  <complexContent>
    <extension base="VisualDType">
      <element name="BinCounts">
        <simpleType base="unsigned3" derivedBy="list">
          <length value="80"/>
        </simpleType>
      </element>
    </extension>
  </complexContent>
</complexType>

```

9.1.3 形状特征

9.1.3.1 区域形状描述符

区域形状描述符的语法形式定义如下：

```

<complexType name="RegionShapeType">
  <complexContent>
    <extension base="VisualDType">
      <element name="ArtDE">
        <simpleType base="unsigned4" derivedBy="list">
          <length value="35"/>
        </simpleType>
      </element>
    </extension>
  </complexContent>
</complexType>

```

9.1.3.2 轮廓形状描述符

轮廓形状描述符的语法形式定义如下：

```

<complexType name="ContourShapeType">
<complexContent>
  <extension base="VisualDType">
    <element name="GlobalCurvatureVector">
      <simpleType base="unsigned6" derivedBy="list">
        <length value="2"/>
      </simpleType>
    </element>
    <element name="PrototypeCurvatureVector" minOccurs="0"
      maxOccurs="1">
      <simpleType base="unsigned6" derivedBy="list">
        <length value="2"/>
      </simpleType>
    </element>
    <element name="HighestPeak" type="unsigned7"/>
    <element name="Peak" maxOccurs="63">
      <complexType>
        <element name="xpeak" type="unsigned6"/>
        <element name="ypeak" type="unsigned3"/>
      </complexType>
    </element>
    <attribute name="numberOfPeaks" type="unsigned6"/>
  </extension>
</complexContent>
</complexType>

```

9.1.4 局部特征

9.1.4.1 局部区域描述符

局部区域描述符的语法形式定义如下：

```

-<complexType name="RegionLocatorType" final="#all">
  -<sequence>
    -<element name="CoordRef" minOccurs="0">
      -<complexType>
        <attribute name="ref" type="IDREF" use="required"/>
        <attribute name="spatialRef" type="boolean" use="optional" default="false"/>
      </complexType>
    </element>
    -<element name="Box" maxOccurs="unbounded" minOccurs="0">
      -<complexType>
        -<simpleContent>
          -<extension base="BoxListType">
            <attribute name="unlocatedRegion" type="boolean" use="optional" default="false"/>
          </extension>
        </simpleContent>
      </complexType>
    </element>
    -<element name="Polygon" maxOccurs="unbounded" minOccurs="0">
      <complexType>
        <sequence>
          <element name="Coords" type="IntegerMatrixType"/>
        </sequence>
        <attribute name="unlocatedRegion" type="boolean" use="optional" default="false"/>
      </complexType>
    </element>
  </sequence>
</complexType>

```

其中 BoxList 的语法形式定义如下：

```

-<complexType name="BoxListType">
  -<simpleContent>
    <restriction base="IntegerMatrixType"/>
  </simpleContent>
</complexType>

```

9.1.4.2 SIFT 特征描述符

SIFT 特征描述具有尺度不变性，可在图像中检测出关键点，是一种局部特征描述。SIFT 特征描述符的语法形式定义如下：

```

<complexType name=" SIFTDType ">
<complexContent>
  <extension base="VisualDType"
    <sequence>
      <element name="position" minOccurs="0" maxOccurs="1">
        <simpleType >
          <element name="x" type="Integer"/>
          <element name="y" type="Integer"/>
        </simpleType>
      </element >
      <element name="mainDirection" type="float" minOccurs="0" maxOccurs="1"/>
      <element name="feature" minOccurs="1" maxOccurs="1">
        <complexType >
          <element name="descript" type="floatVector"/>
          <attribute name="dimCount" type=" Integer " />
        </complexType>
      </element >
    </sequence>
  </extension>
</complexContent>
</complexType>

```

其中：

SIFTDType： 定义了 sift 描述符的类型；

Position： 定义了 sift 特征所在的位置点，有 x 和 y 两个元素，可选；

MainDirection： 定义了 sift 特征点的主方向；

Feature： 定义了具体的特征描述值，由于提取 sift 特征的方式不同，导致 sift 特征描述值有 128 或 64 位等差别，因此采用可变向量的描述形式，用 dimCount 来记录维度。

9.2 音频类型

9.2.1 标量序列

9.2.1.1 标量序列基类型

标量序列基类型的语法形式定义如下：

```

<complexType name="ScalableSeriesType" abstract="true">
  <sequence>
    <element name="Scaling" minOccurs="0" maxOccurs="unbounded">
      <complexType>
        <attribute name="ratio" type="positiveInteger" use="required"/>
        <attribute name="numOfElements" type="positiveInteger" use="required"/>
      </complexType>
    </element>
  </sequence>
  <attribute name="totalNumOfSamples" type="positiveInteger" use="required"/>
</complexType>

```

9.2.1.2 标量序列类型

标量序列类型的语法形式定义如下：

```

<complexType name="SeriesOfScalarType">
  <complexContent>
    <extension base="ScalableSeriesType">
      <sequence>
        <element name="Raw" type="floatVector" minOccurs="0"/>
        <element name="Min" type="floatVector" minOccurs="0"/>
        <element name="Max" type="floatVector" minOccurs="0"/>
        <element name="Mean" type="floatVector" minOccurs="0"/>
        <element name="Random" type="floatVector" minOccurs="0"/>
        <element name="First" type="floatVector" minOccurs="0"/>
        <element name="Last" type="floatVector" minOccurs="0"/>
        <element name="Variance" type="floatVector" minOccurs="0"/>
        <element name="Weight" type="floatVector" minOccurs="0"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

```

9.2.1.3 向量序列类型

向量序列类型的语法形式定义如下：

```

<complexType name="SeriesOfVectorType">
  <complexContent>
    <extension base="ScalableSeriesType">
      <sequence>
        <element name="Raw" type="FloatMatrixType" minOccurs="0"/>
        <element name="Min" type="FloatMatrixType" minOccurs="0"/>
        <element name="Max" type="FloatMatrixType" minOccurs="0"/>
        <element name="Mean" type="FloatMatrixType" minOccurs="0"/>
        <element name="Random" type="FloatMatrixType" minOccurs="0"/>
        <element name="First" type="FloatMatrixType" minOccurs="0"/>
        <element name="Last" type="FloatMatrixType" minOccurs="0"/>
        <element name="Variance" type="FloatMatrixType" minOccurs="0"/>
        <element name="Covariance" type="FloatMatrixType" minOccurs="0"/>
        <element name="VarianceSummed" type="floatVector" minOccurs="0"/>
        <element name="MaxSqDist" type="floatVector" minOccurs="0"/>
        <element name="Weight" type="floatVector" minOccurs="0"/>
      </sequence>
      <attribute name="vectorSize" type="positiveInteger" default="1"/>
    </extension>
  </complexContent>
</complexType>

```

9.2.2 低级音频描述符

9.2.2.1 音频 LLD 标量

音频 LLD 标量的语法形式定义如下

```

-<complexType name="AudioLLDScalarType" abstract="true">
  -<complexContent>
    -<extension base="AudioDType">
      -<choice>
        <element name="Scalar" type="float"/>
        -<element name="SeriesOfScalar" maxOccurs="unbounded">
          -<complexType>
            -<complexContent>
              -<extension base="SeriesOfScalarType">
                <attribute name="hopSize" type="mediaDurationType" default="PT10N1000F"/>
              </extension>
            </complexContent>
          </complexType>
        </element>
      </choice>
      <attribute name="confidence" type="zeroToOneType" use="optional"/>
    </extension>
  </complexContent>
</complexType>

```

其中 mediaDurationType 类型的语法形式定义如下：

```

-<simpleType name="mediaDurationType">
  -<restriction base="basicDurationType">
    <pattern value="\-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?)(\d+F)?" />
  </restriction>
</simpleType>

```

9.2.2.2 音频 LLD 向量

音频 LLD 向量的语法形式定义如下：

```

-<complexType name="AudioLLDVectorType" abstract="true">
-<complexContent>
  -<extension base="AudioDType">
    -<choice>
      <element name="Vector" type="floatVector"/>
      -<element name="SeriesOfVector" maxOccurs="unbounded">
        -<complexType>
          -<complexContent>
            -<extension base="SeriesOfVectorType">
              <attribute name="hopSize" type="mediaDurationType" default="PT10N1000F"/>
            </extension>
          </complexContent>
        </complexType>
      </element>
    </choice>
  </extension>
</complexContent>
</complexType>

```

9.2.2.3 音频波谱

音频波谱的语法形式定义如下：

```

-<complexType name="AudioWaveformType">
  -<complexContent>
    -<extension base="AudioLLDScalarType">
      <attribute name="minRange" type="float" use="optional"/>
      <attribute name="maxRange" type="float" use="optional"/>
    </extension>
  </complexContent>
</complexType>

```

9.2.3 静音描述符

静音描述符的语法形式定义如下：


```

<complexType name="SilenceDType">
  <complexContent>
    <extension base="AudioDType">
      <attribute name="confidence" type="zeroToOneType" default="1.0"/>
      <attribute name="minDurationRef" type="anyURI" use="optional"/>
    </extension>
  </complexContent>
</complexType>

```

9.2.4 MFCC 特征描述符

MFCC 特征描述符的语法形式定义如下：

```

<complexType name="MFCCDType ">
  <complexContent>
    <extension base="AudioDType"
      <element name="feature" type="floatVector"/>
      <attribute name="dimCount" type="nonNegativeInteger" />
    </extension>
  </complexContent>
</complexType>

```

9.3 视频类型

9.3.1 时间域特征描述类型

媒体时间域特征类型的语法形式定义如下：

```

-<complexType name="MediaTimeType">
  -<sequence>
    -<choice>
      <element name="MediaTimePoint" type="mediaTimePointType"/>
      <element name="MediaRelTimePoint" type="MediaRelTimePointType"/>
      <element name="MediaRelIncrTimePoint" type="MediaRelIncrTimePointType"/>
    </choice>
    -<choice minOccurs="0">
      <element name="MediaDuration" type="mediaDurationType"/>
      <element name="MediaIncrDuration" type="MediaIncrDurationType"/>
    </choice>
  </sequence>
</complexType>

```

其中 mediaTimePointType 类型的语法形式定义如下:

```

-<simpleType name="mediaTimePointType">
  -<restriction base="basicTimePointType">
    <pattern value="(-?\d+(-\d{2}(-\d{2}))?)?(T\d{2}(:\d{2}(:\d{2}(:\d+))?)?)?(F\d+)?"/>
  </restriction>
</simpleType>

```

mediaDurationType 类型的语法形式定义如下:

```

-<simpleType name="mediaDurationType">
  -<restriction base="basicDurationType">
    <pattern value="-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?)(\d+F)?"/>
  </restriction>
</simpleType>

```

MediaIncrDurationType 类型的语法形式定义如下

```

-<complexType name="MediaIncrDurationType">
  -<simpleContent>
    -<extension base="integer">
      <attribute name="mediaTimeUnit" type="mediaDurationType" use="optional"/>
    </extension>
  </simpleContent>
</complexType>

```

MediaRelTimePointType 类型的语法形式定义如下：

```
-<complexType name="MediaRelTimePointType">
-<simpleContent>
  -<extension base="mediaTimeOffsetType">
    <attribute name="mediaTimeBase" type="XPathRefType" use="optional"/>
  </extension>
</simpleContent>
</complexType>
-<simpleType name="mediaTimeOffsetType">
-<restriction base="basicDurationType">
  <pattern value="\-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?)(\d+F)?"/>
</restriction>
</simpleType>
```

xPathRefType 类型的语法形式定义如下：

```
<!-- Definition of XPathRef datatype -->
-<simpleType name="XPathRefType">
  -<restriction base="XPathType">
    <pattern value="/?((((child::)?((i\c*?)?(\i\c*)(\[\d+\])))\.|\(\.\.)))*((((child::)?((i\c*?)?(\i\c*)(\[\d+\])))\.|\.)(\(((attribute::|@)((i\c*?)?(\i\c*|\*)))))"/>
  </restriction>
</simpleType>
<!-- Definition of XPath datatype -->
-<simpleType name="XPathType">
  <restriction base="token"/>
</simpleType>
```

MediaRelIncrTimePointType 的类型定义如下：

```
-<complexType name="MediaRelIncrTimePointType">
-<simpleContent>
  -<extension base="integer">
    <attribute name="mediaTimeUnit" type="mediaDurationType" use="optional"/>
    <attribute name="mediaTimeBase" type="XPathRefType" use="optional"/>
  </extension>
</simpleContent>
</complexType>
```

9.3.2 运动特征

9.3.2.1 镜头运动描述符

镜头运动描述符语法形式定义如下：

```
<complexType name="CameraMotionType">
  <complexContent>
    <extension base="VisualIDType">
      <sequence minOccurs="0" maxOccurs="unbounded">
        <element name="Segment" type="SegmentedCameraMotionType"/>
      </sequence>
      <attribute name="numberOfSegments" type="unsigned" use="required"/>
      <attribute name="mode" type="boolean" use="required"/>
    </extension>
  </complexContent>
</complexType>
<complexType name="SegmentedCameraMotionType">
  <element name="Time" type="MediaTimeType"/>
  <element name="FractionalPresence" type="FractionalPresenceType"/>
  <element name="AmountOfMotion" type="AmountOfMotionType"/>
  <element name="FOE_FOC_HorizontalPosition" type="float"/>
  <element name="FOE_FOC_VerticalPosition" type="float"/>
</complexType>
```

其中 AmountOfMotionType 的语法形式定义如下：

```
<complexType name="AmountOfMotionType">
  <element name="TRACK_LEFT" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="TRACK_RIGHT" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="BOOM_DOWN" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="BOOM_UP" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="DOLLY_FOWARD" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="DOLLY_BACKWARD" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="PAN_LEFT" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="PAN_RIGHT" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="TILT_DOWN" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="TILT_UP" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="ROLL_CLOCKWISE" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="ROLL_ANTICLOCKWISE" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="ZOOM_IN" type="unsigned11" minOccurs="0" maxOccurs="1"/>
  <element name="ZOOM_OUT" type="unsigned11" minOccurs="0" maxOccurs="1"/>
</complexType>
```

FractionalPresenceType 的语法形式定义如下：

```
</complexType name="FractionalPresenceType">
  <element name="TRACK_LEFT" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="TRACK_RIGHT" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="BOOM_DOWN" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="BOOM_UP" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="DOLLY_FOWARD" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="DOLLY_BACKWARD" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="PAN_LEFT" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="PAN_RIGHT" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="TILT_DOWN" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="TILT_UP" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="ROLL_CLOCKWISE" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="ROLL_ANTICLOCKWISE" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="ZOOM_IN" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="ZOOM_OUT" type="unsigned7" minOccurs="0" maxOccurs="1"/>
  <element name="FIXED" type="unsigned7" minOccurs="0" maxOccurs="1"/>
</complexType>
```

9.3.2.2 运动活动性描述符

运动活动性描述符语法形式定义如下：

```

<complexType name="MotionActivityType">
  <complexContent>
    <extension base="VisualIDType">
      <element name="Intensity" type="intensityType"/>
      <element name="DominantDirection" type="unsigned3" minOccurs="0" maxOccurs="1"/>
      <element name="SpatialParameters" minOccurs="0" maxOccurs="1">
        <complexType>
          <element name="Nsr" type="unsigned6"/>
          <element name="Nmr" type="unsigned5"/>
          <element name="Nlr" type="unsigned5"/>
        </complexType>
      </element>
      <element name="SpaLocNumber" type="SpaLocNumberType" minOccurs="0" maxOccurs="1"/>
      <element name="SpatialLocalizationParameters" minOccurs="0" maxOccurs="1">
        <simpleType base="unsigned3" derivedBy="list"/>
      </element>
      <element name="TemporalParameters" minOccurs="0" maxOccurs="1">
        <simpleType base="unsigned6" derivedBy="list">
          <length value="5"/>
        </simpleType>
      </element>
    </extension>
  </complexContent>
</complexType>

```

其中 intensityType 和 SpaLocNumberType 的语法形式定义如下：

```

<simpleType name="intensityType" base="nonNegativeInteger">
  <minInclusive value="1"/>
  <maxInclusive value="5"/>
</simpleType>
<simpleType name="SpaLocNumberType" base="positiveInteger">
  <enumeration value="4"/>
  <enumeration value="16"/>
  <enumeration value="64"/>
  <enumeration value="256"/>
</simpleType>

```

9.3.3 镜头和关键帧描述符

9.3.3.1 镜头描述符

镜头描述符语法形式定义如下：

```
<complexType name="ShotDType ">
<complexContent>
  <extension base="VisualDType"
    <sequence>
      <element name="startTimePoint" type="mediaTimePointType"/>
      <element name="duration" type="mediaDurationType"/>
    </sequence>
  </extension>
</complexContent>
</complexType>
```

其中：

ShotDType：定义了shot描述符的类型

startTimePoint：表示镜头的开始时间点；

duration：表示镜头的时长。

9.3.3.2 关键帧描述符

关键帧描述符的语法形式定义如下：

```
<complexType name="KeyframeDType ">
<complexContent>
  <extension base="VisualDType"
    <sequence>
      <element name="timePoint" type="mediaTimePointType"/>
      <element name="keyframeContent" type=" hexBinary"/>
    </sequence>
  </extension>
</complexContent>
</complexType>
```

其中：

KeyframeDType：定义了shot描述符的类型

timePoint：表示该关键帧在视频中的时间位置；

keyframeContent：表示该关键帧的内容，具体为图像内容，用二进制文件存储；

9.4 文本类型

9.4.1 关键字序列描述符

关键字序列描述符的语法形式定义如下：

```
<complexType name="keyWordListType">
  <complexContent>
    <extension base="TextDType">
      <element name="keyWordList" type="stringVector"/>
      <attribute name="dimCount" type="nonNegativeInteger" />
    </extension>
  </complexContent>
</complexType>
```

9.4.2 TF 描述符

TF 描述符的语法形式定义如下：

```
<complexType name="termFrequencyType">
  <complexContent>
    <extension base="TextDType">
      <element name="frequencyList" type="zeroToOneVector"/>
      <attribute name="dimCount" type="nonNegativeInteger" />
    </extension>
  </complexContent>
</complexType>
```

9.4.3 TF-IDF 描述符

TF-IDF 描述符的语法形式定义如下：

```
<complexType name="TFIDFType">
  <complexContent>
    <extension base="TextDType">
      <element name="tfidfList" type="floatVector"/>
      <attribute name="dimCount" type="nonNegativeInteger" />
    </extension>
  </complexContent>
</complexType>
```

10 原始数据描述

定位原始数据的描述符的语法形式定义如下：


```

<complexType name="InsRawDataDType">
  <complexContent>
    <extension base="RawDataDType"/>
    <sequence minOccurs="0" maxOccurs="unbounded">
      <!-- DATA_File_id 为数据文件的存放路径 -->
      <element name="DATA_File_path" type="String"/>
    </sequence>
    <attribute name="fileCount" type="nonNegativeInteger"/>
  </complexContent>
</complexType>

```

11 关联描述符

11.1 语义特征-底层特征关联描述符

该关联描述符主要用于描述语义特征描述符与底层特征描述符存在关联关系。其语法形式定义如下：

```

<complexType name="Semantic-LowfeatureAssociation">
  <complexContent>
    <extension base="AssociationDType">
      <sequence>
        <element name="semantic" type="SemanticDType" minOccurs="1" maxOccurs="unbounded"/>
        <element name="lowfeature" type="LowFeatureDType" minOccurs="1" maxOccurs="unbounded"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

```

SemanticDType 和 LowFeatureDType 均为抽象类，继承自这两个类的可实例化的类型均可刻画这种关联关系。

11.2 基本属性-原始数据关联描述符

该关联描述符主要用于描述语义特征描述符与底层特征描述符存在关联关系。其语法形式定义如下：

```

<complexType name=" BasicAttribution- RawDataAssociation">
<complexContent>
  <extension base="AssociationDType">
    <sequence>
      <element name="basicAttibution" type=" BasicAttributionDType " minOccurs="1" maxOccurs="
unbounded "/>
      <element name="rawData" type=" RawDataDType " minOccurs="1" maxOccurs=" unbounded"/>
    </sequence>
  </extension>
</complexContent>
</complexType>

```

11.3 语义特征-原始数据关联描述符

该关联描述符主要用于描述语义特征描述符与底层特征描述符存在关联关系。其语法形式定义如下：

```

<complexType name=" Semantic-RawDataAssociation">
<complexContent>
  <extension base="AssociationDType">
    <sequence>
      <element name="semantic" type=" SemanticDType " minOccurs="1" maxOccurs=" unbounded "/>
      <element name="rawData" type=" RawDataDType " minOccurs="1" maxOccurs=" unbounded"/>
    </sequence>
  </extension>
</complexContent>
</complexType>

```

11.4 底层特征-原始数据关联描述符

该关联描述符主要用于描述语义特征描述符与底层特征描述符存在关联关系。其语法形式定义如下：

```

<complexType name=" LowFeature-RawDataAssociation">
<complexContent>
  <extension base="AssociationDType">
    <sequence>
      <element name="lowFeature" type=" LowFeatureDType " minOccurs="1" maxOccurs=" unbounded
"/>
      <element name="rawData" type=" RawDataDType " minOccurs="1" maxOccurs=" unbounded"/>
    </sequence>
  </extension>
</complexContent>
</complexType>

```

附 录 A
(规范性附录)
Unsigned 类型语法定义

```
-<simpleType name="unsigned1">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="1"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned2">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="3"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned3">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="7"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned4">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="15"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned5">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="31"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned6">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="63"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned7">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="127"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned8">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="255"/>
  </restriction>
</simpleType>
```

```

-<simpleType name="unsigned9">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="511"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned10">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="1023"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned11">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="2047"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned12">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="4095"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned13">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="8191"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned14">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="16383"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned15">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="32767"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned16">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="65535"/>
  </restriction>
</simpleType>

```

```
-<simpleType name="unsigned17">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="131071"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned18">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="262143"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned19">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="524287"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned20">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="1048575"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned21">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="2097151"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned22">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="4194303"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned23">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="8388607"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned24">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="16777215"/>
  </restriction>
</simpleType>
```

```

-<simpleType name="unsigned25">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="33554431"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned26">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="67108863"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned27">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="134217727"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned28">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="268435455"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned29">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="536870911"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned30">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="1073741823"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned31">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="2147483647"/>
  </restriction>
</simpleType>
-<simpleType name="unsigned32">
  -<restriction base="nonNegativeInteger">
    <minInclusive value="0"/> <maxInclusive value="4294967295"/>
  </restriction>
</simpleType>

```

附 录 B
(资料性附录)
非结构化数据表示 XML Schema 定义

B.1 非结构化数据表示XML Schema定义

非结构化数据表示的XML Schema定义如下：

```
<complexType name=" UnstructuredDataType ">
  <sequence>
    <element name="basicDS" type="InstanceBasicAttributionDSType" minOccurs="1"
maxOccurs="unbounded"/>
    <element name="semanticDS" type="InstanceSemanticDSType" minOccurs="1"
maxOccurs="unbounded"/>
    <element name="lowFeatureDS" type="InstanceLowFeatureDSType" minOccurs="1"
maxOccurs="unbounded"/>
    <element name="rawDataDS" type="InstanceRawDataDSType" minOccurs="1"
maxOccurs="unbounded"/>
    <element name="association" type="InstanceAssociationDType " minOccurs="0"
maxOccurs="unbounded"/>
  </sequence>
  <attribute name="id" type="ID" use="required">
</complexType>
```

其中 InstanceBasicAttributionDSType、InstanceSemanticDSType、InstanceLowFeatureDSType、InstanceRawDataDSType 和 InstanceAssociationDType 的定义见 B.2 至 B.6。

B.2 InstanceBasicAttributionDSType

InstanceBasicAttributionDSType的语法形式定义如下：


```

<complexType name="InstanceBasicAttributionDSType">
<complexContent>
  <extension base="BasicAttributionDSType">
    <sequence>
      <element name="creationInfo" type="CreationDSType"/>
      <element name="fileFormat" type="string"/>
      <element name="fileSize" type="nonNegativeInteger" minOccurs="0"/>
    </sequence>
  </extension>
</complexContent>
</complexType>

```

其中CreationDSType定义见9.2节。

B.3 InstanceSemanticDSType

InstanceSemanticDSType的语法形式定义如下：

```

<complexType name="InstanceSemanticDSType">
<complexContent>
  <extension base="SemanticDSType">
    <sequence>
      <element name="DataCoding" type="DataCodingDSType" minOccurs="0"
maxOccurs="unbounded"/>
      <element name="Title" type="String"/>
      <element name="Theme" type="String"/>
      <element name="Keyword" type="String"/>
      <attribute name="id" type="ID"/>
    </sequence>
  </extension>
</complexContent>
</complexType>

```

其中DataCodingDSType见10.4节定义。

B.4 InstanceLowFeatureDSType

InstanceLowFeatureDSType的语法形式定义如下：

```

-<complexType name="InstanceLowFeatureDSType">
-<complexContent>
-<extension base="LowFeatureDSType">
  <sequence>
    <element name="imageFeature" type="StillRegionFeatureType" minOccurs="0"/>
    <element name="audioFeature" type="AudioSignatureType" minOccurs="0"/>
    <element name="videoFeature" type="VideoSegmentType" minOccurs="0"/>
    <element name="textFeature" type="simpleTextDescription" minOccurs="0"/>
  </sequence>
</extension>
</complexContent>
</complexType>

```

其中 StillRegionFeatureType 给出一个图像的描述方案，该描述方案包含有主导颜色描述符、色彩内容描述符、颜色结构描述符、颜色布局描述符、边缘直方图描述符等描述符，该方案从各个方面来描述图像的底层特征，其 XML Schema 的语法形式定义如下：

```

-<complexType name="StillRegionFeatureType">
-<complexContent>
-<extension base="VisualDSType">
  <sequence>
    <element name="DominantColor" type="DominantColorType" minOccurs="0"/>
    <element name="ScalableColor" type="ScalableColorType" minOccurs="0"/>
    <element name="ColorStructure" type="ColorStructureType" minOccurs="0"/>
    <element name="ColorLayout" type="ColorLayoutType" minOccurs="0"/>
    <element name="Edge" type="EdgeHistogramType" minOccurs="0"/>
    <element name="HomogeneousPattern" type="HomogeneousTextureType" minOccurs="0"/>
    <element name="ShapeMask" type="RegionShapeType" minOccurs="0"/>
    <element name="Contour" type="ContourShapeType" minOccurs="0"/>
    <element name="Sift" type="SIFTDType" minOccurs="0"/>
  </sequence>
</extension>
</complexContent>
</complexType>

```

AudioSignatureType 给出了用于音频实例检索的音频指纹的描述方案，其语法格式定义如下：

```

<complexType name="AudioSignatureType">
  <complexContent>
    <extension base="AudioDSType">
      <sequence>
        <element name="Flatness" type="AudioSpectrumFlatnessType"/>
        <element name="MFCC" type="MFCCDType"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

<complexType name="AudioSpectrumFlatnessType">
  <complexContent>
    <extension base="AudioLLDVectorType">
      <attribute name="loEdge" type="float" default="250"/>
      <attribute name="hiEdge" type="float" default="16000"/>
    </extension>
  </complexContent>
</complexType>

```

VideoSegmentType给出了视频片段特征描述方案，其语法形式定义如下：

```

<complexType name="VideoSegmentType">
  <sequence>
    <element name="MediaTime" type="MediaTimeType" minOccurs="0"/>
    <choice maxOccurs="unbounded" minOccurs="0">
      <element name="VisualDescriptor" type="VisualDType"/>
      <element name="VisualDescriptionScheme" type="VisualDSType"/>
      <element name="VisualTimeSeriesDescriptor" type="VisualTimeSeriesType"/>
      <element name="GofGopFeature" type="GofGopFeatureType"/>
    </choice>
  </sequence>
</complexType>

-<complexType name="GofGopFeatureType">
  <sequence>
    <element name="Descriptor" type="VisualDType"/>
  </sequence>
  <attribute name="aggregation" use="optional">
    <simpleType>
      <restriction base="string">
        <enumeration value="Average"/>
        <enumeration value="Median"/>
        <enumeration value="SplitMerge"/>
      </restriction>
    </simpleType>
  </attribute>
</complexType>

-<complexType name="VisualTimeSeriesType" abstract="true">
  <sequence>
    <element name="TimeIncr" type="mediaDurationType"/>
  </sequence>
  <attribute name="offset" type="mediaDurationType" use="optional" default="PT0S"/>
</complexType>

```

simpleTextDescription给出了文本描述方案，该描述方案主要包含关键词序列，词频以及TF-IDF分值，用该描述方案可进行简单的文本相似性比较。其语法形式定义如下：

```

<complexType name="simpleTextDescription">
  <complexContent>
    <extension base="TextDSType">
      <sequence>
        <element name="keyWordList" type="keyWordListType" minOccurs="0" maxOccurs="1"/>
        <element name="termFrequency" type="termFrequencyType" minOccurs="0" maxOccurs="1"/>
        <element name="tfidf" type="TFIDFType" minOccurs="0" maxOccurs="1"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

```

B.5 InstanceRawDataDSType

InstanceRawDataDSType的语法形式定义如下：

```

<complexType name="InstanceRawDataDSType">
  <complexContent>
    <extension base="RawDataDType"/>
    <element name="rawData" type="InsRawDataDType"/>
  </extension>
</complexContent>
</complexType>

<complexType name="InsRawDataDType">
  <complexContent>
    <extension base="RawDataDType"/>
    <sequence minOccurs="0" maxOccurs="unbounded">
      <!-- DATA_File_id 为数据文件的存放路径 -->
      <element name="DATA_File_path" type="String"/>
    </sequence>
    <attribute name="fileCount" type="nonNegativeInteger"/>
  </complexContent>
</complexType>

```

B.6 InstanceAssociationDType

InstanceAssociationDType的语法形式定义如下：

```

<complexType name="InstanceAssociationDType">
  <complexContent>
    <extension base="AssociationDType">
      <sequence>
        <element name="semantic" type="patternType " minOccurs="1" maxOccurs="unbounded "/>
        <element name="lowfeature" type="instanceFeatureDType " minOccurs="1"
maxOccurs="unbounded"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

<simpleType name="patternType">
  -<complexContent>
    <extension base="SemanticDType"/>
    <element name="pattern" type="string"/>
  </extension>
  </complexContent>
</simpleType>

<simpleType name="instanceFeatureDType">
  -<complexContent>
    <extension base="LowFeatureDType"/>
    <sequence minOccurs="0" maxOccurs="unbounded">
      <element name="color" type="DominantColorType"/>
    </sequence>
  </extension>
  </complexContent>
</simpleType>

```
