



# 中华人民共和国国家标准

GB/T 34952—2017

---

## 多媒体数据语义描述要求

Multimedia data semantic description requirement

2017-11-01 发布

2018-05-01 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 .....	I
1 范围 .....	1
2 术语和定义 .....	1
3 概述 .....	1
4 文本 .....	2
4.1 标题特征 .....	2
4.2 版权管理 .....	2
4.3 统计特征 .....	2
4.4 词性特征 .....	2
4.5 句法特征 .....	3
4.6 文档类型 .....	3
5 图形 .....	3
5.1 标题特征 .....	3
5.2 版权管理 .....	3
5.3 几何特征 .....	3
6 图像 .....	4
6.1 标题特征 .....	4
6.2 版权管理 .....	4
6.3 情感特征 .....	4
6.4 行为特征 .....	4
6.5 场景特征 .....	4
6.6 空间关系特征 .....	4
6.7 对象特征 .....	4
6.8 底层特征 .....	5
7 音频 .....	5
7.1 标题特征 .....	5
7.2 版权管理 .....	5
7.3 音频内容 .....	5
7.4 音频类别 .....	5
7.5 底层特征 .....	6
8 视频 .....	6
8.1 标题特征 .....	6
8.2 版权管理 .....	6
8.3 视频内容 .....	6
8.4 视频类别 .....	7
8.5 底层特征 .....	7

## 前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本标准由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本标准起草单位：中国电子技术标准化研究院、北京信息科技大学、北京华宇软件股份有限公司、杭州海康威视数字技术股份有限公司。

本标准主要起草人：侯霞、卫凤林、张展新、高冀鹏、赵世范。

# 多媒体数据语义描述要求

## 1 范围

本标准规定了文本、图形、图像、音频和视频五类多媒体数据的语义描述要求。  
本标准适用于多媒体数据语义分析、标识和多媒体数据应用系统的数据字典设计。

## 2 术语和定义

下列术语和定义适用于本文件。

### 2.1

#### **多媒体 multimedia**

综合表现文本、图形、图像、音频和视频的信息组合。

### 2.2

#### **文本 text**

以字符、符号、字、短语、段落、句子、表格或者其他字符排列形式出现的数据,旨在表达一个意义,其解释主要以读者对某种自然语言或人工语言的了解为基础。例:打印在纸上或显示在屏幕上的业务信件。

[GB/T 5271.1—2000,定义 01.01.03]

### 2.3

#### **图形 graphics**

用来表示一个变量相对于其他变量变化情况的线条图,或用以代替文字说明一个概念和思想的图解和表格。

### 2.4

#### **图像 image**

物体通过感光材料、光电传感处理设备和计算机而重显的影像。

### 2.5

#### **音频 audio**

一种数字化动态媒体形态,用于描述声音及其顺序,给人以时序的感觉,并可进行数字人工合成。

### 2.6

#### **视频 video**

一种数字化动态媒体形态,用于描述运动图像,进行高速信息传送或显示瞬间的相互关系。

## 3 概述

本标准规定了文本、图形、图像、音频和视频五类多媒体数据的语义描述要求,其中:

- a) 文本数据包括标题特征、版权管理、统计特征、词性特征、句法特征和文档类型;
- b) 图形数据包括标题特征、版权管理和几何特征;
- c) 图像语义要求包括标题特征、版权管理、情感特征、行为特征、场景特征、空间关系特征、对象特征和底层特征;
- d) 音频语义要求标题特征、版权管理、音频内容、音频类别和底层特征;

- e) 视频语义要求包括标题特征、版权管理、视频内容、视频类别和底层特征。

## 4 文本

### 4.1 标题特征

描述内容应包括：

- a) 标题；
- b) 主题；
- c) 创建者；
- d) 作者；
- e) 最后修改者；
- f) 摘要；
- g) 创建日期；
- h) 编辑次数；
- i) 编辑时间；
- j) 创建应用程序；
- k) 分类；
- l) 文档模板；
- m) 关键字集；
- n) 用户自定义元数据集；
- o) 单位名称。

### 4.2 版权管理

描述内容应包括：

- a) 作者；
- b) 出版日期；
- c) 非原创使用者时的授权信息。

### 4.3 统计特征

描述内容应包括：

- a) 字数；
- b) 词数；
- c) 句数；
- d) 英文字符数；
- e) 中文字符数；
- f) 行数；
- g) 段落数；
- h) 对象数；
- i) 页数；
- j) 词频；
- k) 字节数；
- l) 字频。

### 4.4 词性特征

描述内容可包括：

- a) 文本中出现的名词；
- b) 文本中出现的动词；
- c) 文本中出现的形容词；
- d) 文本中出现的数量词；
- e) 文本中出现的虚词。

#### 4.5 句法特征

描述内容可包括：

- a) 文本中出现的句子的句法；  
注：通常句法包括：主语、谓语、宾语、定语、状语和补语等。
- b) 文本中出现的词语属性。  
注：通常词语属性包括：
  - 1) 上位词：在词义的层次关系中位于高层的词；
  - 2) 下位词：在词义的层次关系中位于低层的词；
  - 3) 形近词：字形和另外的词相似但笔画/拼写不尽相同的词；
  - 4) 首选词：一组同义词中，用户偏爱选用的词；
  - 5) 相关词：通过词义关联关系，和另一个词密切关联的词；
  - 6) 同音词：词形不同但发音相同或相近的词；
  - 7) 同义词：词形不同但意思相同或相近的词；
  - 8) 族首词：在词义的层次关系中位于最高层的词。

#### 4.6 文档类型

描述内容应包括：

- a) 流式文档类型；
- b) 版式文档类型。

### 5 图形

#### 5.1 标题特征

描述内容应包括：

- a) 名称；
- b) 制作人；
- c) 制作公司；
- d) 制作工具；
- e) 用途；
- f) 制作时间；
- g) 主体颜色。

#### 5.2 版权管理

描述内容应包括：

- a) 作者；
- b) 出版日期；
- c) 非原创使用者时的授权信息。

#### 5.3 几何特征

描述内容应包括：

- a) 图元：
  - 1) 类型；  
注：图元类型一般包括点、线、多边形、面、曲线、循环线、复合曲线、曲线多边形、多点等。
  - 2) 位置；
  - 3) 层次。
- b) 轮廓：
  - 1) 轮廓线型；
  - 2) 轮廓颜色。
- c) 填充：
  - 1) 图案；
  - 2) 颜色与渐变。
- d) 阴影；
- e) 文本。

## 6 图像

### 6.1 标题特征

描述内容应包括：

- a) 名称；
- b) 制作人；
- c) 制作公司；
- d) 制作时间；
- e) 制作工具。

### 6.2 版权管理

描述内容应包括：

- a) 作者；
- b) 出版日期；
- c) 非原创使用者时的授权信息。

### 6.3 情感特征

应描述图像给人带来的主观情绪。

注：主观情绪一般包括：喜悦、愤怒、悲伤、恐惧、嫉妒、惭愧、羞耻、自豪、激动、振奋和迷茫等。

### 6.4 行为特征

应描述图像所表示的行为和动作。

### 6.5 场景特征

应描述图像所表示的场景。

### 6.6 空间关系特征

应描述图像中对象间的空间关系。

### 6.7 对象特征

应描述图像中出现的对象，如：人脸。

## 6.8 底层特征

描述内容应包括：

- a) 高度；
- b) 宽度；
- c) 编码；
- d) 图像特定颜色(颜色直方图)；
- e) 颜色模型；
- f) 纹理；
- g) 形状；
- h) 结构布局；
- i) 构图显著性等特征及其组合,如:红色的圆形；
- j) 上下文特征；
- k) 文件大小；
- l) 图片格式；
- m) 文本特征。

## 7 音频

### 7.1 标题特征

描述内容应包括：

- a) 标题；
- b) 专辑名；
- c) 压缩标准；
- d) 时长；
- e) 制作时间；
- f) 文件大小。

### 7.2 版权管理

描述内容应包括：

- a) 作者；
- b) 出版日期；
- c) 非原创使用者时的授权信息。

### 7.3 音频内容

描述内容应包括：

- a) 音色；
- b) 旋律；
- c) 流派；
- d) 情感；
- e) 对音频的描述。

### 7.4 音频类别

描述内容应包括：



- a) 静音；
- b) 和谐背景音；
- c) 对话；
- d) 环境背景音。

## 7.5 底层特征

描述内容应包括：

- a) 频谱信息；
- b) 采样频率；
- c) 采样精度；
- d) 声道数；
- e) 音强；
- f) 比特率(压缩)；
- g) 码率模式；
- h) 编码(压缩)格式。

## 8 视频

### 8.1 标题特征

描述内容应包括：

- a) 字幕；
- b) 时长；
- c) 片段；
- d) 原作；
- e) 制作人；
- f) 制作公司；
- g) 制作时间；
- h) 制作工具；
- i) 参与制作的全体成员；
- j) 片头曲或主题曲；
- k) 片尾曲；
- l) 版本类型；
- m) 阶段。

### 8.2 版权管理

描述内容应包括：

- a) 作者；
- b) 出版日期；
- c) 非原创使用者时的授权信息。

### 8.3 视频内容

描述内容应包括：

- a) 视频序列；

注：视频序列的属性一般包括场景的个数和持续时间。

## b) 视频场景；

注：视频场景的属性一般包括标题、持续时间、镜头数目、开始镜头、结束镜头等。

## c) 镜头；

注 1：镜头的属性一般包括持续时间、开始帧号、结束帧号、代表帧集合、特征空间向量等。

注 2：镜头切换方式一般包括突变和渐变两种。

## d) 帧；

注：帧的属性一般包括直方图、轮廓图等。

#### 8.4 视频类别

描述内容应包括：

- a) 卡通；
- b) 广告；
- c) 运动；
- d) 综艺节目；
- e) 纪录片；
- f) 新闻。

#### 8.5 底层特征

描述内容应包括：

- a) 拍摄方式；
  - b) 拍摄角度；
  - c) 拍摄地点；
  - d) 关键帧数量；
  - e) 关键帧间隔；
  - f) 关键帧的颜色；
  - g) 关键帧的纹理；
  - h) 关键帧的形状；
  - i) 帧宽度；
  - j) 帧高度；
  - k) 数据速率；
  - l) 总比特率；
  - m) 帧速率；
  - n) 比特率(压缩)；
  - o) 码率模式；
  - p) 视频时长；
  - q) 色彩空间；
  - r) 位深度图像；
  - s) 编码(压缩)格式。
-