

国 际 电 信 联 盟

ITU-R
国际电联无线电通信部门

ITU-R BT.2133-0建议书
(10/2019)

**基于IP的广播系统中
高级沉浸式视听内容的传送**

**BT系列
广播业务
(电视)**



国际电信联盟

前言

无线电通信部门的职责是确保卫星业务等所有无线电电信业务合理、平等、有效、经济地使用无线电频谱，不受频率范围限制地开展研究并在此基础上通过建议书。

无线电通信部门的规则和政策职能由世界或区域无线电通信大会以及无线电通信全会在研究组的支持下履行。

知识产权政策 (IPR)

ITU-R的IPR政策述于ITU-R第1号决议中所参引的《ITU-T/ITU-R/ISO/IEC的通用专利政策》。专利持有人用于提交专利声明和许可声明的表格可从<http://www.itu.int/ITU-R/go/patents/zh>获得，在此处也可获取《ITU-T/ITU-R/ISO/IEC的通用专利政策实施指南》和ITU-R专利信息数据库。

ITU-R系列建议书

(也可在线查询<http://www.itu.int/publ/R-REC/zh>)

系列	标题
BO	卫星传送
BR	用于制作、存档和播出的录制；电视电影
BS	广播业务（声音）
BT	广播业务（电视）
F	固定业务
M	移动、无线电定位、业余和相关卫星业务
P	无线电波传播
RA	射电天文
RS	遥感系统
S	卫星固定业务
SA	空间应用和气象
SF	卫星固定业务和固定业务系统间的频率共用和协调
SM	频谱管理
SNG	卫星新闻采集
TF	时间信号和频率标准发射
V	词汇和相关问题

说明：该ITU-R建议书的英文版本根据ITU-R第1号决议详述的程序予以批准。

电子出版
2020年，日内瓦

© 国际电联 2020

版权所有。未经国际电联书面许可，不得以任何手段复制本出版物的任何部分。

ITU-R BT.2133-0建议书

基于IP的广播系统中
高级沉浸式视听（AIAV）内容的传送

(2019年)

范围

该建议书提供了使用ISO/IEC 23090-2（全向媒体格式，亦称为OMAF），以便在IP广播系统中传输高级沉浸式音视频（AIAV）内容的指南。

关键词

沉浸式媒体、虚拟现实、360°视频、传输、基于IP的广播

国际电联无线电通信全会

考虑到

- a) 虚拟现实、360°视频及其他沉浸式媒体技术已引起内容提供商、受众和相关消费技术供应商的注意；
- b) ITU-R BT.2123建议书规定了AIAV系统的参数值，这些参数值包括用于广播节目制作和国际节目交换的360°视频；
- c) ISO/IEC 23090-2“全向媒体格式（OMAF）”规范了ISO/IEC 23008-1中规定的MPEG媒体传输（MMT）的全向媒体封装、信令和流，以及ISO/IEC 23009-1中规定的基于HTTP的MPEG动态自适应流（DASH）；
- d) 基于IP的广播系统可以通过扩展支持OMAF传输AIAV内容；
- e) 有关在基于IP的广播系统中使用OMAF的指南，将有助于广播公司开发和部署这些系统，

进一步考虑到

- a) ITU-R BT.1869建议书规范了广播系统中IP分组数据包的复用方案；
- b) ITU-R BT.2074-1建议书定义了使用MMT并基于IP的广播系统所需的业务配置、媒体传输协议及信令信息，

建议

当AIAV内容使用ISO/IEC 23090-2所规定的全向媒体格式（OMAF）标准，在基于IP的广播系统中传输时，应参考附件1中的指南。

注1 - 附件2概述了ISO/IEC 23090-2中规定的OMAF标准，以供参考。

注2 - 附件3给出了基于IP的广播系统中AIAV内容可能的层结构，以供参考。

附件1

基于IP的广播系统 使用ISO/IEC 23090-2的指南

OMAF的特征演示文件定义了针对任意数量媒体类型轨道或项目的要求和约束。应使用独立于OMAF检视端口的基线演示特征。该特征满足支持3D音频、全向和3D视频的基本要求。同时支持普通视频和立体视频。

视频轨道应符合基于HEVC且独立于检视端口的OMAF视频特征，该特征支持最高为360°的普通视频和立体视频。轨道的比特流符合HEVC主层10的特征（主层，级别5.1）。

音轨应该符合OMAF 3D音频基线特征，该特征使用MPEG-H 3D音频编解码器将通道、对象和高阶环境音（HOA）作为3D音频来提供支持。MPEG-H 3D音频为用户的观看方向指定了一个标准接口，并允许将音频场景以低复杂度、低延迟的方式呈现到用户希望的任何方向。轨道的比特流符合MPEG-H 3D音频的低复杂度（LC）特征，其级别为1、2和3。

可以通过传输定时文本为全向视频提供字幕和隐藏式字幕，且这方面应符合W3C TTML 互联网媒体字幕和字幕规范1.1（IMSC1.1）或网络视频文本轨道格式（WebVTT）的规定。定时文本位置可以是固定的，即其中定时文本不随用户的观看方向移动而移动，或者是总是可见的，即不管用户的观看方向如何，定时文本总对用户可见。

初始观看方向可作为定时元数据之一传输。此数据指示初始观看方向，以使内容作者能够为用户指出其推荐的特定观看方向。

附件2

（参考性附件）

ISO/IEC 23090-2规定的OMAF标准概述

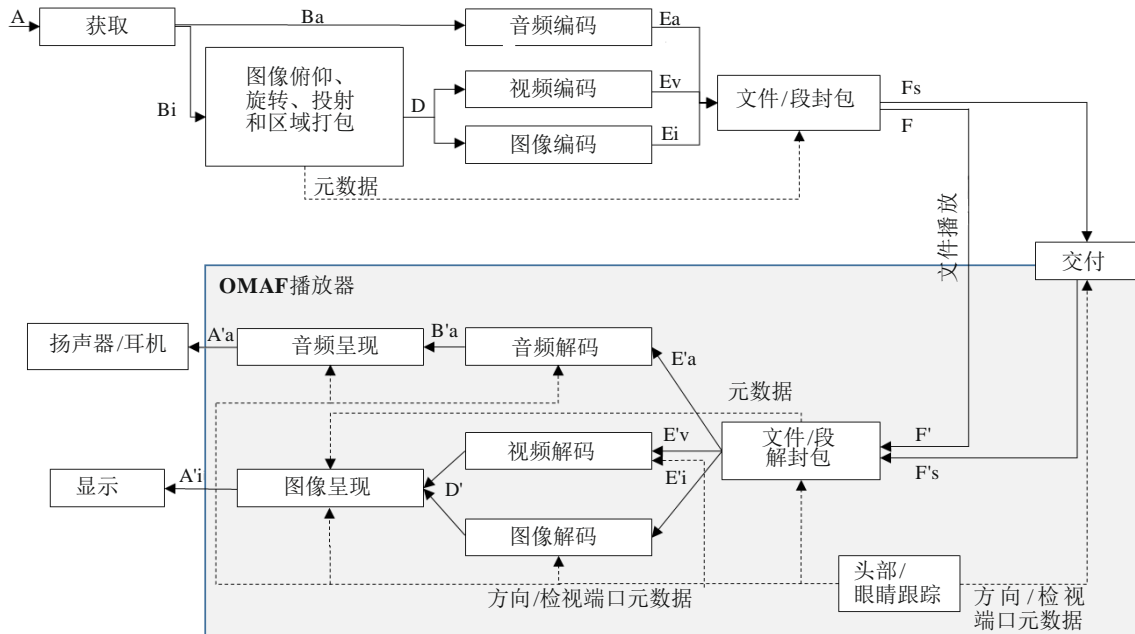
ISO/IEC 23090-2（全向媒体格式，OMAF）规定了一种支持全向媒体应用的媒体格式，侧重于360°视频、图像和音频以及相关的定时文本。此标准还规定了包括区域打包法在内的各种技术。区域打包法可利用从用户设备头部/眼睛跟踪获得的信息来提高图像质量。OMAF的第一个版本支持3DoF¹ VR/360内容。图1显示了带有投影视频的全向媒体应用程序的典型内容流程。

¹ 三向自由：围绕X、Y和Z轴的三向旋转和无限制运动（分别为俯仰、偏转和滚动）。一个典型的用例是用户坐在椅子上自由地向任何方向看。

图 1

含有投影视频的全向媒体内容流过程

(本图摘自 ISO/IEC 23090-2)



A: 真实世界的音像情景
 B: 多传感器捕获的视频或音频
 D/D': 投影/打包视频
 E/E': 编码视频或音频比特流
 F/F': ISO 基础媒体文件格式

BT.2133-01

图1中的文件/段封装过程包括将元数据封装到文件或分段中。元数据示例有助于呈现解码打包图片的投影和区域打包信息。

根据上述流程，OMAF具有以下规定：

- 用于等矩形和立方体地图全向投影格式的坐标系和方程、从局部坐标轴到全局坐标轴的转换以及矩形区域打包；
- 针对全向媒体和球体区域定时元数据的ISO基本媒体文件格式扩展；
- 全向媒体向MMT和MPEG-DASH的扩展；
- 媒体特征，例如基于HEVC且独立于检视端口的视频特征、3D音频基线特征、HEVC图像特征和IMSC1.1定时文本特征；
- 基于某些OMAF媒体特征的演示，例如独立于检视端口的基线演示特征和依赖于检视端口的基线演示特征。

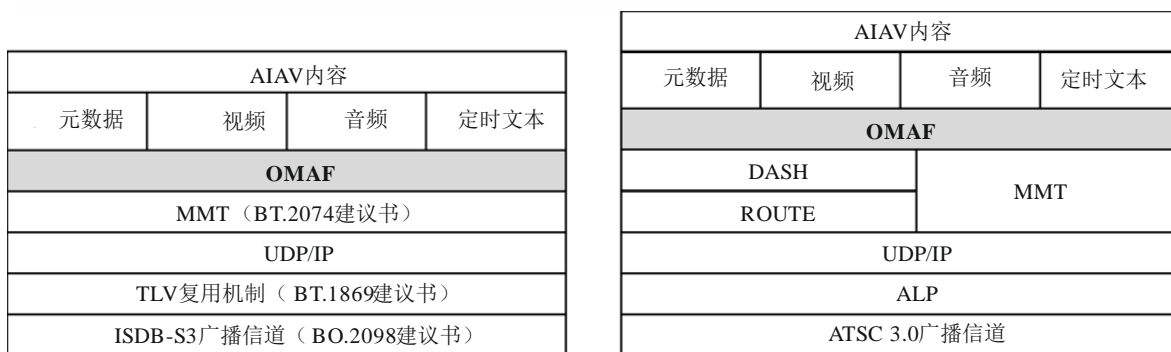
附件3 (参考性附件)

广播系统中传输AIAV内容的可能层结构

OMAF规范规定OMAF内容的交付应使用MMT或MPEG-DASH。封装于OMAF的AIAV内容可以在采用MMT或MPEG-DASH，在基于IP的广播系统中传送。图2显示了ISDB-S3和ATSC 3.0中AIAV内容的可能层结构。在这些情况下，360°全方位视角的视频通过单向广播频道同时传送给大量用户。

图 2

广播系统中传输AIAV内容的可能层结构



* ROUTE: 通过单向传输进行实时对象交付

* ALP: ATSC 3.0链路层协议

BT.2133-02