

GY

中华人民共和国广播电视和网络视听行业标准

GY/T XXX—XXXX

自由视角视频系统技术要求

Technical requirements for free-viewpoint video system

(报批稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

国家广播电视总局 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 系统典型架构及工作流程	2
5.1 系统典型架构	2
5.2 工作流程	3
6 技术要求	3
6.1 拍摄采集子系统	4
6.2 媒体处理子系统	6
6.3 分发子系统	7
6.4 分发网络	7
6.5 终端播放子系统	8
7 自由视角视频数据格式	8
附录 A（资料性） 自由视角视频应用场景示例	11
A.1 大型综艺、晚会节目的应用	11
A.2 体育赛事的应用	11

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

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专利的责任。

本文件由全国广播电影电视标准化技术委员会（SAC/TC 239）归口。

本文件起草单位：国家广播电视总局广播电视规划院、中央广播电视总台、国家广播电视总局监管中心、国家广播电视总局广播电视科学研究院、华为技术有限公司、中兴通讯股份有限公司、北京爱奇艺科技有限公司、湖南芒果无际科技有限公司、浙江广播电视集团、央视频融媒体发展有限公司、湖南快乐阳光互动娱乐传媒有限公司、成都索贝数码科技股份有限公司、优酷信息技术（北京）有限公司、上海宽娱数码科技有限公司、北京百度网讯科技有限公司、北京达佳互联信息技术有限公司、百视通网络电视技术发展有限责任公司、互影科技（北京）有限公司、智令互动（深圳）科技有限公司、咪咕文化科技有限公司、成都华栖云科技有限公司、深圳看到科技有限公司、腾讯科技（深圳）有限公司、清华大学、四川传媒学院。

本文件主要起草人：邓向冬、刘文翰、宁金辉、覃毅力、谭阳、范晓轩、孙文涛、冯海亮、魏娜、孟凡辉、毛连贵、邓治民、欧阳霁、郑红哲、郑磊、胡钰琦、陈亮、张琪浩、赵盛洋、程相昱、侯明强、盛骁杰、汤毅、鹄鹏、张瑞圣、刘松鹏、黄信云、蔡树焕、徐永太、王家福、季向阳、李丹、李小雨、缪川扬、颜士富、于慧霞、王威、潘永杰、孟令康、高杨、郭小傲、郑洛、郑欣、马俊、刘璐、刘杰、曹虹、连晓聪、马建明、李俊敏。

自由视角视频系统技术要求

1 范围

本文件规定了自由视角视频系统的技术要求,包括拍摄采集子系统、媒体处理子系统、分发子系统、分发网络、终端播放子系统的技术要求,以及自由视角视频数据格式要求。

本文件适用于自由视角视频系统的设计、建设、验收、运行和维护,以及自由视角视频内容的拍摄、交换、分发、播放等。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;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 41808—2022 高动态范围电视节目制作和交换图像参数值

GB/T 41809—2022 超高清清晰度电视系统节目制作和交换参数值

GY/T 155—2000 高清晰度电视节目制作及交换用视频参数值

3 术语和定义

下列术语和定义适用于本文件。

3.1

自由视角视频系统 free-viewpoint video system

采用多机位或环绕机位拍摄,获得一定范围内环绕焦点多角度的视频信息,经过制作、传输、分发、播放等环节,可为用户提供特定范围内多视角连续切换,并形成空间位置自由移动视觉效果的一种视频直播/点播系统。

3.2

时间定格空间推移 static space shift effect

对于多路摄像机的同一个帧同步时刻,将所有摄像机的当前帧时刻的拍摄画面按一定顺序组合,实现一种时间静止条件下空间位置自由移动的视觉效果。

3.3

时间推移空间推移 dynamic space shift effect

对于多路摄像机的不同时刻(可等间隔或非等间隔),将所有摄像机的对应帧时刻所拍摄画面按一定顺序组合,实现一种时间流动条件下空间位置自由移动的视觉效果。

4 缩略语

下列缩略语适用于本文件。

DASH 基于HTTP的动态自适应流(Dynamic Adaptive Streaming over HTTP)

- DTS 解码时间戳 (Decoding Time Stamp)
- GoP 图像组 (Group of Pictures)
- HLG 混合对数伽马 (Hybrid Log-Gamma)
- HLS 基于HTTP的自适应码率流媒体传输协议 (HTTP Live Streaming)
- MPD 媒体呈现描述 (Media Presentation Description)
- PQ 感知量化 (Perception Quantization)
- PTS 显示时间戳 (Presentation Time Stamp)
- RTC 实时通信 (Real-Time Communication)
- RTMP 实时消息传输协议 (Real-Time Messaging Protocol)
- RTMPS 经过安全套接层加密的RTMP协议 (Real-Time Messaging Protocol over SecureSocket Layer)
- RTSP 实时流传输协议 (Real-Time Streaming Protocol)
- SDI 数字分量串行接口 (Serial Digital Interface)
- SRT 安全可靠传输协议 (Secure Reliable Transport)
- UDP 用户数据报协议 (User Datagram Protocol)

5 系统典型架构及工作流程

5.1 系统典型架构

自由视角视频系统由拍摄采集子系统、媒体处理子系统、分发子系统、分发网络、终端播放子系统五部分组成，系统典型架构见图1，应用场景示例见附录A，各部分组成及主要功能如下。

- a) 拍摄采集子系统主要包括摄像机组、同步控制与信号传输连接、采集机位固定装置、编码器及相关软件等部分，前端部署摄像机组获取实时拍摄的音视频流，并可根据需要选配编码器，对摄像机组采集的音视频流进行编码。
- b) 媒体处理子系统实现自由视角视频的制作，可通过本地或云端部署。本地部署通过媒体处理服务器、导播制作平台等硬件部署，云端部署方式通过云计算平台部署。
- c) 分发子系统实现自由视角内容的审核和分发。
- d) 分发网络可包括有线网络、无线网络等一种或多种，通过分发网络实现自由视角视频内容传输分发。
- e) 终端播放子系统实现自由视角视频内容在电视、电脑、平板、手机等终端设备上的播放和人机互动操作。

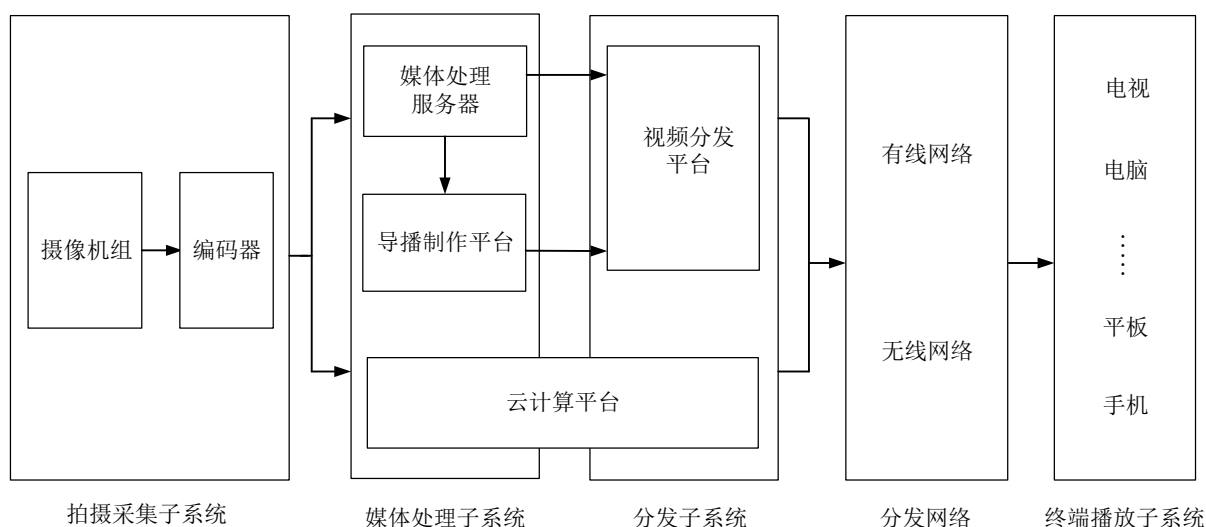


图1 自由视角视频系统典型架构

5.2 工作流程

自由视角视频系统端到端工作流程见图2，具体如下：

- a) 在拍摄采集子系统中，拍摄主体围绕拍摄对象，按业务需求部署摄像机组进行拍摄，同时对多台摄像机进行同步，以获取拍摄现场音视频流，并根据实际需要选配编码器对音视频流统一编码；
- b) 在媒体处理子系统中，音视频流经过前端媒体处理环节完成多路音视频流同步及格式封装转换后进入自由视角内容制作环节，制作生成自由视角直播内容、自由视角点播内容，并封装为第7章规定的格式后传输至分发子系统；
- c) 在分发子系统中，完成对自由视角内容的审核和分发；
- d) 通过分发网络分发至电视、电脑、平板、手机等播放终端；
- e) 通过终端播放子系统实现自由视角视频的加载、解码、渲染、播放和交互等。

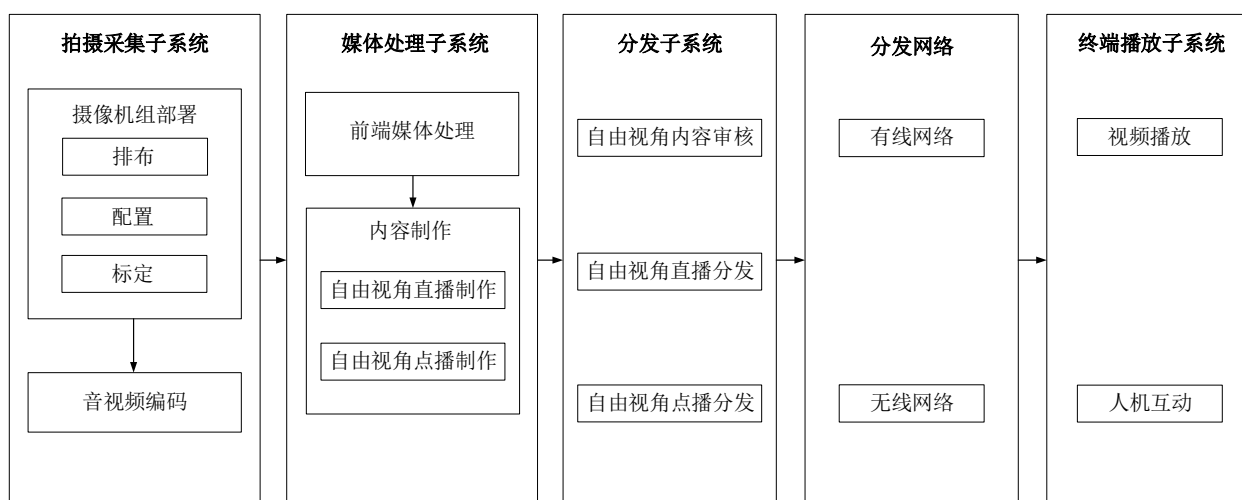


图2 自由视角视频系统工作流程

6 技术要求

6.1 拍摄采集子系统

6.1.1 技术参数

拍摄采集子系统中，单路摄像机技术参数应符合表1的规定。

表1 拍摄采集子系统技术参数

序号	项目	高清	4K/8K 超高清
1	视频编码格式	AVS+或与其相当	AVS2/AVS3 或与其相当
2	媒体协议	RTMP/RTMPS/SRT/UDP/RTSP/HLS/DASH	RTMP/RTMPS/SRT/UDP/RTSP/HLS/DASH
3	视频文件封装格式	FLV/MP4/MOV/MXF 等	FLV/MP4/MOV/MXF 等
4	视频分辨率	宜为 1920×1080	4K 宜为 3840×2160, 8K 宜为 7680×4320
5	视频帧率 (fps)	24/25/30/50/60, 宜为 25/50	宜为 50/60/100/120
6	位深 (bit)	8/10	10/12
7	非线性转换曲线	GY/T 155—2000 或 GB/T 41808—2022 规定的 HLG 或 PQ	GY/T 155—2000 或 GB/T 41808—2022 规定的 HLG 或 PQ
8	色域	GY/T 155—2000 规定的 BT. 709 色域	GB/T 41809—2022 规定的 BT. 2020 色域
9	码率 (Mbps)	单路码率宜为 120	4K 单路码率宜为 500, 8K 单路码率宜为 2000
10	音频采样频率 (kHz)	48	48
11	音频量化比特	≥16	≥16
12	音频声道数	≥2	≥2

6.1.2 摄像机组部署

6.1.2.1 部署形态要求

应支持多种摄像机场地部署形态，包括正圆形、圆弧形、直线形等，以及上述形态的组合。摄像机组部署参数应符合表2的要求。

表2 摄像机组部署参数

序号	项目	要求
1	覆盖角度	正圆形/圆弧形布局：100°~360°；见图 3、图 4 直线形布局：宜≥100°；见图 5
2	拍摄最短距离	正圆形/圆弧形布局：≥3m 直线形布局：≥5m
3	俯仰方式	支持俯拍、仰拍、平拍三种方式
4	相邻摄像机间夹角	宜≤5°

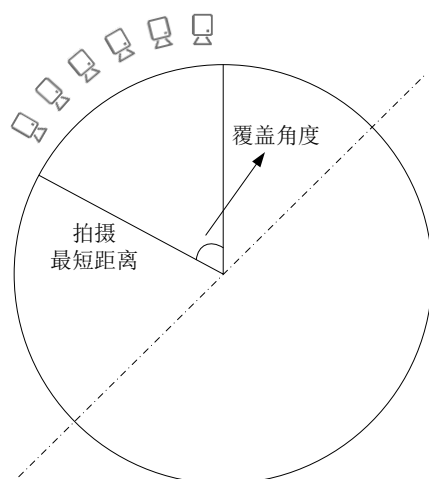


图3 正圆形布局

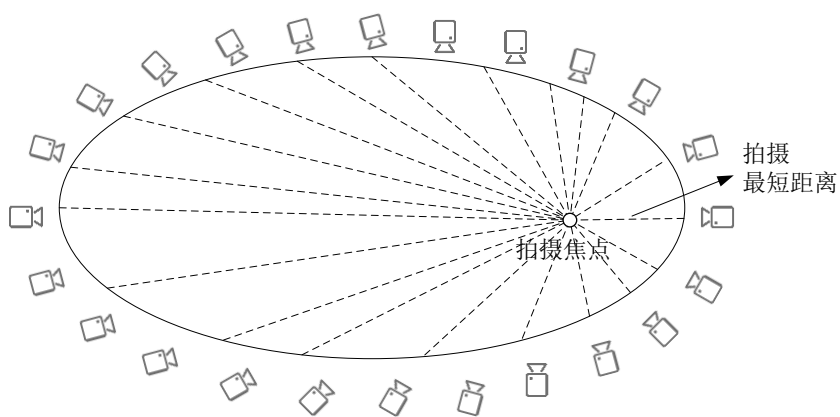


图4 圆弧形布局

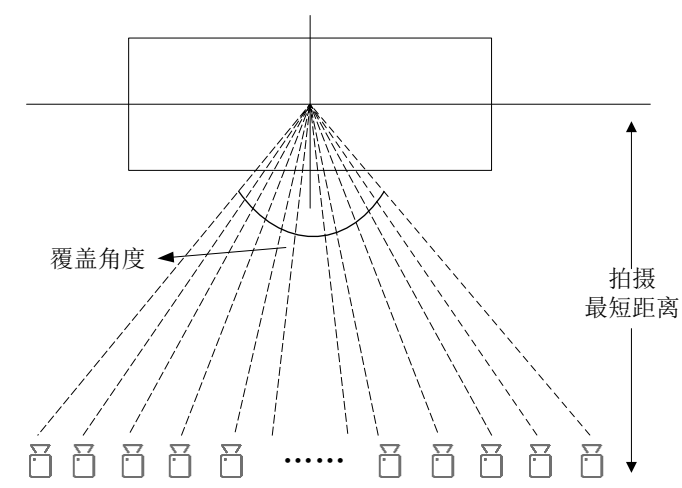


图5 直线形布局

6.1.2.2 摄像机组同步

应支持所有摄像机时间基准对齐，机位间实现帧同步。

6.1.2.3 摄像机组配置

应支持摄像机配置、镜头配置，以保证拍摄焦点和画面呈现的一致性。

摄像机参数包括但不限于感光度、位深、帧率、曝光时间、非线性转换曲线、色域，以及输出格式、编码方式、分辨率等，镜头参数应包括但不限于光圈、焦距等。

6.1.2.4 摄像机组坐标误差标定

应支持通过软件预先计算每个摄像机成像的坐标误差并生成标定文件。预设朝向和实际朝向误差应不大于 3° ，预设三维坐标位置 and 实际三维坐标位置误差应不大于2cm。

6.2 媒体处理子系统

6.2.1 前端媒体处理

应支持所有音视频流进行时间基准对齐，并符合如下要求：

- a) 所有机位间应实现帧同步；
- b) 画面同步不应存在累积误差；
- c) 应为每一路音视频流打时间戳，以保持同步。

6.2.2 自由视角内容制作

6.2.2.1 自由视角编码

应支持多路音视频流转码，支持生成自由视角直播、点播内容，并符合如下要求。

- a) 输入媒体要求：直播可支持RTMP/RTMPS/SRT/UDP/RTSP协议，点播可支持FLV/MP4/MOV/MXF格式。
- b) 输入编码、视频分辨率应符合6.1.1中表1的要求。
- c) 应配置输出的协议、分辨率和码率。
- d) 编码格式为AVS+或与其相当的编码格式时，输出高清视频码率应不小于12Mbps；编码格式为AVS2或与其相当的编码格式时，输出4K超高清视频码率应不小于36Mbps；编码格式为AVS3或与其相当的编码格式时，输出8K超高清视频码率应不小于100Mbps。
- e) 输出内容应包含摄像机总数量、默认视角信息。
- f) 输出内容中应标明视频对应的机位编号。

6.2.2.2 音视频同步

自由视角音视频信息采集制作的过程中，关键时间戳信息可同步嵌入至各机位的拍摄音视频流编码信息中，保持所有机位音视频信息同步。输出音视频流中，视频和音频相对时间误差应不超过 $+25\text{ms} \sim -100\text{ms}$ （正值表示音频超前视频、负值表示音频滞后视频）。

6.2.2.3 精彩瞬间生成

通过自由视角视频系统可生成精彩瞬间，包括时间定格空间推移、时间推移空间推移两种呈现效果。应具备对于同一个节目同时用多路摄像机进行多角度拍摄后，选择某个机位的某个精彩时刻，以及所经历机位的运动轨迹，由系统自动生成一段该精彩时刻的多视角连续视频片段的功能，支持FLV/MP4/MOV/MXF等格式以及SDI或无损IP音视频流输出，视频长度可按需选取。

当自由视角精彩瞬间生成的信号作为一路视频源,传输至节目现场制作系统时,选定角度和时长后,渲染重现视频速度应不低于一倍速,且保证画面观看流畅;精彩瞬间视频信号的输出通常采用SDI接口,精彩瞬间音频可采用现场制作系统音频子系统制作输出的音频信号进行合成。

6.2.2.4 自由视角直播制作

应支持对多路高清或超高清摄像机拍摄内容以音视频流方式进行自由视角内容在线制作,并对多路音视频流进行自由视角转码,音视频流传输协议可为RTMP/RTMPS/SRT/UDP/RTSP/DASH等,编码格式为AVS+或与其相当的编码格式时,输出高清视频码率应不小于12Mbps;编码格式为AVS2或与其相当的编码格式时,输出4K超高清视频码率应不小于36Mbps;编码格式为AVS3或与其相当的编码格式时,输出8K超高清视频码率应不小于100Mbps。

6.2.2.5 自由视角点播制作

应支持对多路高清或超高清摄像机拍摄内容录制并对录制的视频转存文件进行后期制作,视频文件封装格式可为FLV/MP4/MOV/MXF,宜支持对自由视角视频文件的二次编辑等后期制作。

6.3 分发子系统

6.3.1 内容审核

自由视角视频内容审核应对所有视点的视频和音频内容进行审核。

6.3.2 分发预处理

应支持自由视角视频处理功能,完成自由视角直播音视频流和点播内容的聚合、切片、帧同步等预处理,提供自由视角直播、点播、时移和回看等业务。要求如下:

- a) 应支持自由视角视频预处理,并生成初始化文件,标明码流和摄像机、位置、角度、分辨率等对应关系;
- b) 应支持帧同步矫正,对多路入向码流(时间差异不大于5s)进行校验和矫正;
- c) 应支持GoP不大于3的高密度I帧视频输入和处理,提供快速解析视频I帧、降低转码时延的功能;
- d) 应具备断流异常处理功能,当自由视角多路视频流输入存在部分断流,可对异常码流做隔离处理,且不影响其他正常码流处理,不影响整体自由视角业务;
- e) 应具备断流恢复功能,断流视频恢复时,可自动恢复该路视频业务。

6.3.3 内容分发

内容分发技术要求如下:

- a) 应支持自由视角直播通道创建,可接收自由视角多路机位音视频码流,并增加自由视角直播通道标识;
- b) 应支持自由视角点播内容注入,并增加自由视角内容标识;
- c) 应支持自由视角直播、点播、时移和回看业务分发;
- d) 应支持识别客户端手势(或遥控器)操作请求,实时输出对应角度的视频流,输出可采用DASH/RTC分发,支持按视频帧级别服务和加速;
- e) 应支持多种终端(例如电视、机顶盒、平板、手机、VR类终端等)服务请求;
- f) 组网宜支持多节点服务模式,应支持聚合回源,下级节点可通过向上级节点回源获取音视频流,减少回源请求流量并且提高回源效率。

6.4 分发网络

分发网络技术参数要求如下：

- a) 上行带宽：前端视频回传上行带宽资源应高于所有入向视频流所需总带宽（按照80%使用率计算）；
- b) 下行带宽：高清场景宜不小于20Mbps，4K超高清场景宜不小于50Mbps，8K超高清场景宜不小于150Mbps；
- c) 网络时延（分发子系统到用户终端）宜不大于40ms。

6.5 终端播放子系统

6.5.1 视频播放

视频播放技术要求如下：

- a) 应支持视频、音频数据的同步播放，视频和音频误差不超过+90ms~-185ms（正值表示音频超前视频、负值表示音频滞后视频）；
- b) 应及时响应用户的交互信息，响应时间宜不大于60ms；
- c) 应保证多个视角视频流DTS、PTS同步，切换视角时用户无明显画面延迟体验，且根据用户切换视角速度，保证画面流畅无明显卡顿。

6.5.2 人机互动

终端播放子系统宜支持触控滑屏、陀螺仪传感器感应、手势识别、遥控器操控等方式实现人机互动，以便于用户自由切换播放视角，技术要求如下：

- a) 当用户连续操作旋转时，视频帧率宜不小于25fps，且应根据用户切换速度调整视角加载和渲染，保证画面无明显卡顿；
- b) 无论视频在播放或暂停中，用户应均可旋转视角自由观看；
- c) 对于用户旋转视角的操作，响应时间应不大于60ms。

7 自由视角视频数据格式

宜在媒体处理子系统或分发子系统将接入的多路音视频流封装为DASH协议格式，在MPD中扩展增加自由视角视频相关信息，扩展信息在MPD中的位置见图6。

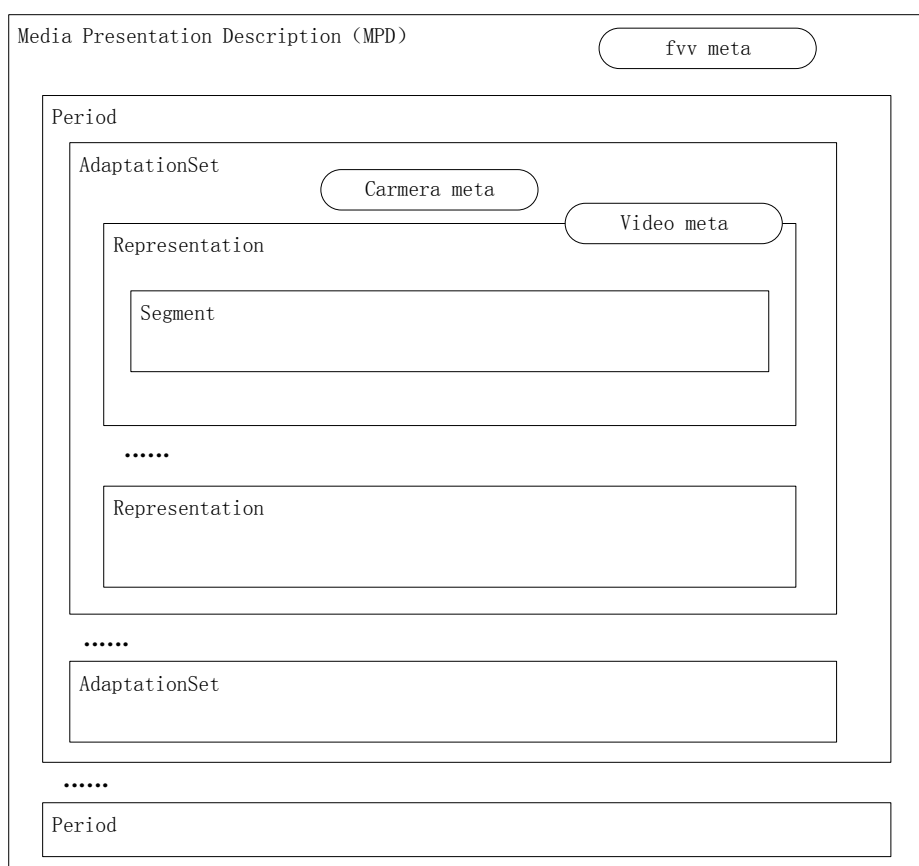


图6 媒体描述呈现示意图

在MPD中增加fvv meta，用于描述视角的数量、角度、默认视角等元数据，fvv meta扩展数据项见表3。

表3 fvv meta 扩展数据项

序号	字段名	说明
1	mediaType	表示流为自由视角流
2	frameCntPerGop	一个 GoP 里面多少帧
3	gopCntPerSeg	一个分片里面多少 GoP

除增加fvv meta部分外，MPD中可包含若干个片段（Period）字段，每个片段（Period）字段相互独立，代表每一个可以独立加载播放的视频分段。片段（Period）字段扩展数据项见表4。

表4 片段（Period）字段扩展数据项

字段名	说明
user	自定义扩展数据，制作片源者

一个片段(Period)字段可包含多个自适应集(AdaptationSet)字段,每个自适应集(AdaptationSet)字段代表一个数据集合，自适应集（AdaptationSet）字段扩展数据项见表5。

表5 自适应集 (AdaptationSet) 字段扩展数据项

序号	字段名	说明
1	schemeIdUri	统一资源标识符
2	contentType	类型: video/audio
3	maxWidth	支持的最大视频宽度, 单位为像素
4	maxHeight	支持的最大视频高度, 单位为像素
5	maxFrameRate	支持的最大视频帧率, 单位为帧每秒
6	colourGamut	色域

一个自适应集 (AdaptationSet) 字段可包含多个表示 (Representation) 字段, 每个表示 (Representation) 字段代表一个独立流。一个表示 (Representation) 字段可包含多个段 (Segment) 字段, 每个段 (Segment) 字段代表一个GoP的信息, 段 (Segment) 字段扩展数据项见表6。

表6 段 (Segment) 字段扩展数据项

序号	字段名	说明
1	baseUrl	获取资源的地址
2	startTime	此段的开始时间, 单位为粒度
3	mediaRange	此段在文件中的索引, 单位为字节

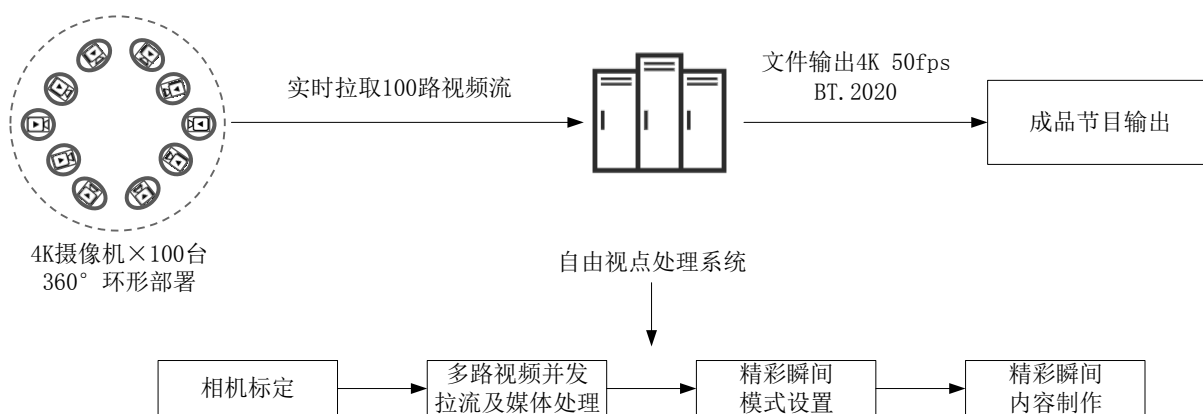
附录 A

(资料性)

自由视角视频应用场景示例

A.1 大型综艺、晚会节目的应用

自由视角视频在大型综艺晚会中已广泛应用,针对大型歌舞、武术、杂技等节目中演员动作难度高、变换幅度大、观赏性强等特点,利用自由视角系统可打破常规拍摄方式局限性,用户可通过手指滑动等方式自由切换特定范围的视角,创造更沉浸的互动视听体验。某大型晚会节目自由视角系统框图见图 A.1,系统部署见图 A.2。



图A.1 某大型晚会节目自由视角系统框图

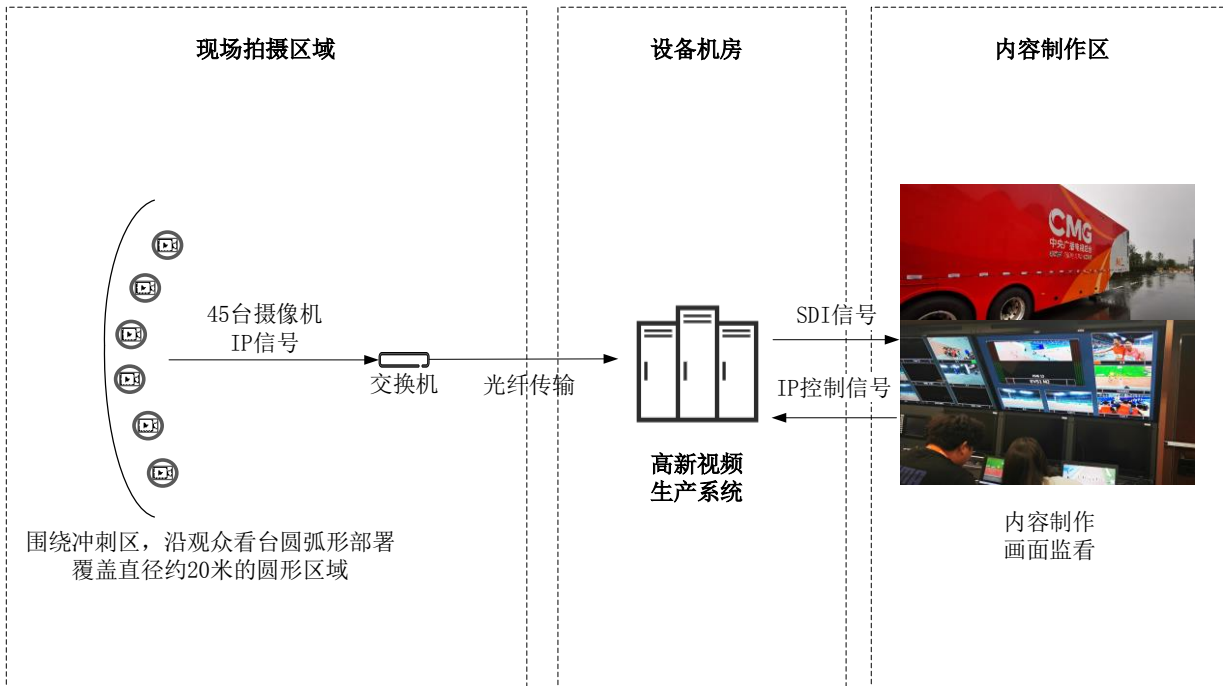


图A.2 某大型晚会节目自由视角系统部署图

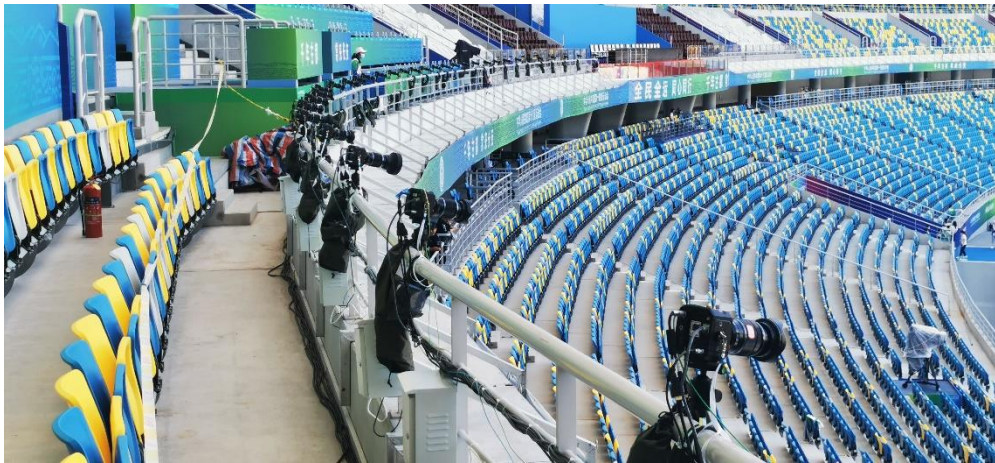
A.2 体育赛事的应用

自由视角视频可广泛应用于短跑、乒乓球、拳击、篮球、体操、棒球等各类体育赛事转播报道中,

用户能够自由切换特定范围的视角，结合时间推移空间推移的呈现效果，通过连续的角度变化观看选手正面、侧面等多角度竞技姿态，使用户获得更好的空间感与沉浸感。某体育赛事自由视角系统框图见图A.3，系统部署见图A.4。



图A.3 某体育赛事自由视角系统框图



图A.4 某体育赛事自由视角系统部署图