

跨区域视频监控联网共享技术规范 第 5 部分：视音频编解码技术要求 及媒体插件管理

Technical specifications for inter-domain video surveillance networking and sharing
Part5: Video/Audio CODEC technical specifications and media plugin manage

2011 - 06 - 27 发布

2011 - 07 - 27 实施

前 言

DB33/T 629《跨区域视频监控联网共享技术规范》分为六个部分：

- 第1部分：总则；
- 第2部分：联网单元；
- 第3部分：设备描述与控制协议；
- 第4部分：用户及设备编码管理；
- 第5部分：视音频编解码技术要求及媒体插件管理；
- 第6部分：设备运行监测管理。

本部分为DB33/T 629的第5部分，依据GB/T 1.1—2009给出的规则起草。

本部分由浙江省公安厅提出并归口。

本部分起草单位：浙江省公安科技研究所、杭州海康威视数字技术股份有限公司、浙江立元通信技术有限公司。

本部分参加起草单位：杭州华三通信技术有限公司、浙江广信智能建筑研究院有限公司、浙江大华技术股份有限公司、浙江方圆检测集团股份有限公司。

本部分主要起草人：蒋乐中、赵问道、夏海元、马伯康、焦庆春、徐力、赵己绩、李旻、杨军喜、李伟强。

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。

跨区域视频监控联网共享技术规范

第 5 部分：视音频编解码技术要求及媒体插件管理

1 范围

DB33/T 629的本部分规定了跨区域视频监控联网共享建设中联网单元传输与控制协议消息体的技术要求。

本部分适用于视频监控系统的跨区域联网共享建设，及相关系统和软硬件产品的设计与开发。

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，仅所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。

ISO/IEC 13818-1	Specifies the system layer of the coding.
ITU-T G. 711	A ITU-T Standard for audio companding
ITU-T G. 722A	ITU-T standard 7 kHz wideband speech codec operating at 48, 56 and 64 kbit/s.
ITU-T G. 723	A ITU-T standard speech codec
ITU-T G. 729	An audio data compression algorithm for voice that compresses digital voice in packets of 10 milliseconds duration.
ITU-T H. 264 05/2003	Advanced video coding for generic audiovisual services.
AMR-NBA	patented audio data compression scheme optimized for speech coding.
MPEG-2 Layer 2	A standard for "the generic coding of moving pictures and associated audio information.
MPEG-4	Is a collection of methods defining compression of audio and visual (AV) digital data.
3GPP	3rd Generation Partnership Project.
RFC 3984	RTP Payload Format for H.264 Video.

3 术语和定义

下列术语和定义是用于本部分

3.1

解码插件

用于在客户端显示压缩图像时使用的解码的插件。

3.2

高清

清晰度标准格式在720p及以上的清晰度。

3.3

标清

清晰度标准格式在720p以下（不包括720p）的清晰度。

4 缩略语

CBR Constant Bit Rate, 定码率

VBR Variable Bit Rate, 变码率

CAVLC Context-based Adaptive Variable Length Coding, 上下文自适应变长编码

CABAC Context-based Adaptive Binary Arithmetic Coding, 上下文自适应二进制算术编码

Slice 条带, 由连续的整数个宏块以扫描顺序构成的图像区域

3GPP The 3rd Generation Partnership Project

5 音频编解码

音频编解码应符合以下要求:

- a) 视频监控前端设备单元所采用的音频编码器应支持以下标准中的一种:
 - 1) G.711A
 - 2) G.722.1
 - 3) G.723.1
 - 4) G.729
 - 5) AMR-NB
 - 6) MPEG-2 Layer 2
- b) 视频监控平台所采用的音频解码器应支持以上六种。
- c) 语音对讲的音频解码器应采用 ITU-T G.711A 标准。

6 视频编解码

6.1 编解码器要求

编解码器应符合以下要求:

- a) 视频监控设备单元的视频编解码器应支持 ISO/IEC 14496 - 10, 即 ITU-T H.264 压缩标准, 简称 H.264/AVC。视频码流的语法内容应依据 ISO/IEC 14496-10 标准, 应至少支持 5.1 定义的 H.264/AVC Constrained Baseline Profile 要求, 宜支持 5.1 定义的 Main Profile 要求, 宜支持 5.1 定义的 High Profile 要求, 不应包含非法私有数据格式。
- b) 视频监控前端设备单元的视频解码器应支持自身编码器支持的最高档次和级别。
- c) 视频监控平台所采用的解码器 (3GPP 除外) 支持的档次和级别不应低于相应编码器支持的最高档次和级别。
- d) 编码器应支持双码流编码模式, 即主码流和辅码流。主码流的视频分辨率应至少达到 4CIF, 辅码流的视频分辨率应支持 CIF、QCIF 或 QVGA。主码流带宽至少为 128kbps~4Mbps, 辅码流

带宽至少为 64kbps~1Mbps。主码流和辅码流均应支持 CBR 和 VBR 两种码率控制方式，CBR 码率波动不应超过 15%，VBR 波动幅度可以不作要求。

6.2 编码码流要求

编码码流的要求如下：

- a) H.264/AVC 编码器应支持 H.264/AVC Constrained Baseline Profile，包括如下选项：
 - 1) 支持 H.264/AVC 基本语法格式；
 - 2) 支持 I Slices, P Slices，其中 P Slice 只支持 1 个参考图像（参考图像可为 I Slice 或者 P Slice）；
 - 3) 支持 CAVLC 自适应变长编码（CAVLC Entropy Coding）；
 - 4) 支持 Loop Filter 环路滤波（In-Loop Deblocking Filter）；
 - 5) 支持整像素、1/2 像素运动搜索和 1/4 像素运动搜索。
- b) H.264/AVC 编码器宜支持 H.264/AVC Main Profile，包括如下选项：
 - 1) 支持 Interlace 编码格式；
 - 2) 支持 B Slice 编码，B Slice 仅使用 2 个参考图像，B Slice 本身不作参考；
 - 3) 支持 CABAC 自适应算术编码。
- c) H.264/AVC 编码器宜支持 H26 High Profile，编码器生成的 High Profile 的码流应符合 H.264/AVC 标准规定。
- d) 对于标清及以下应用，编码 Level 不应超过 3.0；对于高清应用，Level 宜不超过 4.2，部分超高清应用 Level 不应超过 5.1。

6.3 编码数据格式

为了保证码流解析的效率，对编码器产生的码流有如下格式要求：

- a) 应采用单帧/单场一个 Slice 的编码方式（3GPP 除外），当使用多 Slice 的编码方式时，比特流中应当出现 AUD 语法元素进行划界；
- b) 比特流中应当在每个 I 帧之前都出现相应的 SPS 和 PPS。

7 封装格式要求

7.1 传输封装格式

H.264 标准的视频主码流和辅码流实时传输应符合 RFC3984 规范，音频采用相应 RFC 标准的 RTP 封装及 RTSP 协议，辅码流应支持 3GPP 标准。

7.2 存储封装格式

主码流和辅码流采用 ISO/IEC 13818-1 定义的 PS (Program Stream) 流或者 TS (Transport Stream) 流作为存储封装格式。PS 流和 TS 流所使用的 stream_type 定义见表 1。

表1 流类型定义

码流标准	stream_type 定义
G. 711A	0x90
G. 722. 1	0x92
G. 723. 1	0x93
G. 729	0x99
AMR-NB	0x9A
MPEG-2 Layer 1/2/3	0x04
H. 264/AVC	0x1b

8 测试规范

8.1 总则

通过本节给出方法进行测试，确认编码设备输出码流是否符合本附录规范的要求。若某编码设备输出码流没有超出本规范的定义，而某解码器无法正常解码，则认为该解码器无法满足兼容性要求。

8.2 测试条件

测试使用如下工具：

- a) 录像工具，要求能够正确无误将录像存为无封装格式裸码流，建议由设备供应厂家提供；
- b) 标准解码软件，要求公开源代码，通过标准组织的一致性测试，建议使用 VLC 1.1.0(及以上版本)、JM 10.0（及以上版本）；
- c) 码流分析软件，要求通过标准组织的一致性测试，记录或者显示码流语法元素的内容，建议使用 JM10.0（及以上版本）、ElecCard Stream Eye Tools；
- d) 网络协议分析软件，要求能够抓取、分析网络协议 RTP 包数，建议使用 WireShark。

8.3 测试步骤

测试应按以下步骤进行：

- a) 打开待测试设备，产生视频码流；
- b) 采用录像工具，对设备产生的视频码流进行录像，录像时间不小于 120 秒，存为录像文件；
- c) 采用网络协议分析软件抓取设备产生的数据，抓包时间不小于 120 秒，存为抓包文件；
- d) 采用标准解码软件对录像文件进行解码播放；
- e) 采用码流分析软件对录像文件进行码流分析；
- f) 采用网络协议分析软件对抓包文件的 RTP 层数据进行分析。

8.4 预期结果

- a) 标准解码软件能够正确解码存储文件，且解码出来的图像播放正常；
- b) 码流分析软件检查码流的结构和配置满足本部分第 4、5、6 章的要求；
- c) 网络协议分析软件检查 RTP 层数据满足本部分第 6 章要求。

9 插件管理

9.1 DecoderTag 定义

- a) 插件管理的目的是为了解决不同厂家设备视频编码格式不一致导致的客户端视频解码问题;
- b) 解码插件以设备来区分,推送设备信息的联网单元需要提供 DecoderTag 标签及对应的解码插件库,单个联网单元提供商可以有多个 DecoderTag 及相对应的解码插件库;解码插件接口定义见附录 B;
- c) 联网单元在推送设备信息时在报文中添加 DecoderTag 字段告知接收设备信息的联网单元该设备的解码插件标记。联网单元通过 DecoderTag 找到对应的解码插件库,从而正确解码;
- d) 客户端宜通过配置文件的形式来读取解码插件 DecoderTag 及相对应的解码插件库、解码插件接口名,以达到在不编译程序的情况下支持新设备新插件。

9.2 解码库命名

解码库统一命名。后缀为 _Decoder.dll,前缀为 DecoderTag 名称;H.264 标准码流解码插件命名定为 H264_Decoder.dll;MPEG-4 标准码流解码插件命名定为 MPEG4_Decoder.dll;标准码流解码插件定义见附录 A。

9.3 码流传输要求

解码插件只接收符合 RTP (RFC3550) 格式打包的数据。RTP 包中的净荷数据 (除去 RTP 头及 RTP 扩展头部分) 应是设备厂商提供的解码库能够解码的数据,不带有平台厂商的私有数据。

附 录 A
(规范性附录)
标准码流解码插件名称定义

表A. 1规定了标准码流解码插件名称定义。

表A. 1 标准码流解码插件名称定义

标准码流名称	解码插件标签	解码库命名	说明
MPEG-4	MPEG4	MPEG4_Decoder.dll	
H. 264	H264	H264_Decoder.dll	

附 录 B
(规范性附录)
解码插件接口定义

表B.1 消息回调指针

名称	内容
函数指针名称	typedef int (__stdcall *CBF_Msg)(int iSessionId, int iMsgType, int iMsgCode, char* szMsg, void* pUsrData);
描述	消息回调函数指针
输入参数	iSessionId: preview_start 或 playback_start 成功的返回值 iMsgType: 消息回调类型, 需要详细定义 iMsgCode: 消息码, 需要详细定义 szMsg: 消息描述 pUsrData: 用户自定义数据, preview_start 或 playback_start 传入的 pUsrData 字段。
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	

表B.2 数据回调指针

名称	内容
函数指针名称	typedef int (__stdcall* CBF_Stream)(int iSessionId, int iStreamType, char* pData, int iDataLen, void* pUsrData);
描述	数据回调函数指针
输入参数	iSessionId: preview_start 或 playback_start 成功的返回值 iSessionId: preview_start 或 playback_start 成功的返回值 iStreamType: 消息回调类型, 0 为头数据, 1 为视频数据, 2 为音频数据。 pData: 数据指针, 去掉 RTP 头及扩展头数据 iDataLen: 数据长度 pUsrData: 用户自定义数据, preview_start 或 playback_start 传入的 pUsrData 字段
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	

表B.3 初始化解码库

名称	内容
函数名	int __stdcall media_init()
函数描述	初始化插件接口
输入参数	
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	

表B.4 销毁解码库

名称	内容
函数名	void __stdcall media_terminate ()
函数描述	销毁插件接口
输入参数	
输出参数	无
返回值	
调用的函数	
其他	

表B.5 停止所有流

名称	内容
函数名	int __stdcall stop_all()
函数描述	全部停止实时预览
输入参数	
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	

表B.6 设置消息回调

名称	内容
函数名	int __stdcall set_msgcbf(CBF_Msg pCbf)
函数描述	设置消息回调
输入参数	pCbf:消息回调函数
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	

表B.7 设置取流回调

名称	内容
函数名	int __stdcall set_streamcbf(CBF_Stream pCbf);
函数描述	设置取流回调，这个函数必须在 preview_start 和 playback_start 前调用，否则在 preview_start 和 playback_start 使用时无效。
输入参数	pCbf:取流回调函数
输出参数	无
返回值	0 表示成功,-1 表示失败，get_lasterror() 获取错误码
调用的函数	
其他	

表B.8 开始实时流

名称	内容
函数名	int __stdcall preview_start(void* wnd, unsigned short port, void* pUserData)
函数描述	开始预览 回调函数设置有效时，需要回调视音频流，无效时，不回调视音频流。 Wnd 有效时，需要在窗口中播放视频。回调和播放可以同时有效。
输入参数	wnd: 播放窗口句柄，如果 wnd 为 NULL，表示不解码只回调码流。 port: 接收数据的端口号 pUserData: 用户自定义数据
输出参数	无
返回值	>=0 表示成功,-1 表示失败，get_lasterror() 获取错误码
调用的函数	
其他	可以多次调用，释放时必须次数相同。

表B.9 停止实时流

名称	内容
函数名	int __stdcall preview_stop(int sessionid);
函数描述	停止实时预览
输入参数	pConf:级联客户端配置信息
输出参数	无
返回值	0 表示成功,-1 表示失败，get_lasterror() 获取错误码
调用的函数	
其他	

表B.10 开始回放

名称	内容
函数名	int __stdcall playback_start(void* wnd, unsigned short port, const char* sdp, char* filename, void* pUserData);
函数描述	开始回放，wnd 有效时，必须在窗口中播放视频。如果调用过 set_streamcbf 设置取流回调有效，必须在回调函数中回调视频数据。播放和回调可以同时支持。

表 B. 10 开始回放（续）

名称	内容
输入参数	wnd: 播放窗口句柄 port: 接收数据的端口号 sdp: rtsp DESCRIBE 响应时带回来的 SDP 信息 filename: 保存的录像文件名 pUserData: 用户自定义数据
输出参数	无
返回值	
调用的函数	
其他	

表B. 11 停止回放

名称	内容
函数名	void __stdcall playback_stop(int sessionid);
函数描述	停止回放
输入参数	sessionid: play_back 成功的返回值。
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	可以多次调用, 释放时必须次数相同。

表B. 12 抓图

名称	内容
函数名	int __stdcall picture_capture(int sessoinid, const char* picturename, int iPicMode);
函数描述	抓图, 抓图保存格式为默认为 JPEG。
输入参数	sessionid: preview_start 或 playback_start 成功的返回值。 picturename: 抓图保存的文件名 iPicMode: 0 表示 JPEG, 1 表示 BMP
输出参数	无
返回值	0 表示成功, -1 表示失败, get_lasterror() 获取错误码
调用的函数	
其他	

表B. 13 设置音量

名称	内容
函数名	int __stdcall get_lasterror();
函数描述	获取错误码，错误码定义应与本规范第三部分附录 B 一致。
输入参数	
输出参数	
返回值	返回错误号
调用的函数	
其他	