

ICS 33.040.40

M 32

YD

中华人民共和国通信行业标准

YD/T 2883-2015

电信网视频监控系统 视音频编解码技术要求

Visual surveillance system over telecommunication networks
——Video and audio CODEC general technical requirements

2015-07-14 发布

2015-10-01 实施

中华人民共和国工业和信息化部 发布

目 次

前 言	II
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
3.1 术语和定义	1
3.2 缩略语	2
4 总体性原则	2
5 音频编解码技术要求	2
6 视频编解码技术要求	3
6.1 概述	3
6.2 基于MPEG-4的视频编解码技术要求	3
6.3 基于H.264的视频编解码技术要求	4
6.4 基于M-JPEG的视频编解码技术要求	5
7 其他技术要求	5
附录A（规范性附录）参数	6

前 言

本标准是“电信网视频监控系统”系列标准之一，该系列标准的预计结构及名称如下：

— YD/T 2455 《电信网视频监控系统》

- 第1部分：业务需求；
- 第2部分：总体技术要求；
- 第3部分：业务平台测试方法；
- 第4部分：客户端设备技术要求；
- 第5部分：客户端设备测试方法；
- 第6部分：服务质量；
- 第7部分：安全要求。

— YD/T 2456 《电信网视频监控系统 智能分析及传感器叠加应用架构和总体技术要求》；

— YD/T 2883 《电信网视频监控系统 视音频编解码技术要求》；

— YD/T 2882 《电信网视频监控系统 终端设备网管技术要求》；

— YD/T 2881 《电信网视频监控系统 跨平台监控的统一客户端技术要求》；

— 《电信网视频监控系统 无线视频监控业务应用平台》；

— 《电信网视频监控系统 移动视频监控系统的客户端单元》。

本标准按照GB/T 1.1-2009给出的规则起草。

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。

本标准由中国通信标准化协会提出并归口。

本标准起草单位：中国电信集团公司。

本标准主要起草人：冯 明、曹 宁、梁笃国、张艳霞、张琳姝、冯传滨、周 杰、崔瑞琳、任广杰。

电信网视频监控系统

视音频编解码技术要求

1 范围

本标准规定了电信网络视频监控系统的视音频编解码技术要求。

本标准适用于国内电信网视频监控业务及与视频监控编解码相关设备。

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/IEC14496（所有部分） 信息技术 音视频对象编码（Information technology-Coding of audio-visual objects）

IETF RFC2435 JPEG-压缩图象的RTP有效载荷格式（RTP Payload Format for JPEG-compressed Video）

IETF RFC3016 MPEG-4 音频/视频流的RTP负载格式（RTP Payload Format for MPEG-4 Audio/Visual Streams）

IETF RFC3640 MPEG-4载流线传输的RTP有效载荷格式（RTP Payload Format for Transport of MPEG-4 Elementary Streams）

IETF RFC3984 流量工程(TE)的媒介系统到媒介系统(IS-IS)扩展（Intermediate System to Intermediate System (IS-IS) Extensions for Traffic Engineering (TE)）

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本文件。

3.1.1

视音频存储 Video/Audio Storage

音频视频信息记录到载体上的相关的行为和要求，主要包括存储方式、音视频存储格式、存储载体及存储策略。

3.1.2

图像质量 Picture Quality

能够为观察者分辨的光学图像质量，它通常包括像素数量、分辨率、色彩还原性、图像轮廓还原性（灰度级）和信噪比，但主要表现为信噪比。

3.1.3

档次 Profile

根据视觉对象定义了对应码流语法的子集。

3.1.4

等级 Level

根据特定档次，定义了对应采用的限制参数集合。一个档次可包含一个或多个等级。

3.2 缩略语

下列缩略语适用于本文件。

1080P	1080 Progressive	一种高清影像传输格式，在逐行扫描下达到1920×1080的图像分辨率
4CIF		一种影像传输格式，其图像分辨率为CIF的4倍
720P	720 Progressive	一种高清影像传输格式，在逐行扫描下达到1280×720的图像分辨率
ASP	Advanced Real Time Simple Profile	高级实时简单档次
AVS		中国数字音视频编码标准
CIF	Common Intermediate Format	一种常用的影像传输格式，其图像分辨率为（PAL：352×288，NTSC：352×240）
D1		一种影像传输格式，其图像分辨率为PAL：720×576，NTSC：720×480
NTSC	National Television Systems Committee	全国电视系统委员会制式（一种视频制式）
PAL	Phase Alternating Line	逐行倒相制式（一种视频制式）
QCIF	Quarter Common Intermediate Format	一种影像传输格式，其图像分辨率为CIF的1/4
QVGA	Quarter Video Graphics Array	一种小型显示器的显示分辨率，其分辨率为VGA的1/4
SP	Simple Profile	简单档次
VGA	Video Graphics Array	一种电脑显示器早期常用的显示分辨率（640×480）

4 总体性原则

网络视频监控系统中视频编解码器宜采用 MPEG-4、H.264、M-JPEG 视频编解码，也可采用 AVS 编码。

注：当在 AVS 标准和产业成熟时可以采用 AVS 视频编解码。

视频编码器应至少支持其中一种标准。

视频解码器应该具有能力交互，应具有自适应（H.264、MPEG-4、M-JPEG）解码的能力。

5 音频编解码技术要求

音频比特流宜采用标准的RTP进行打包，音频载荷数据应为整数个音频编码帧。

网络视频监控系统中音频编解码器应支持G.722.1音频编解码标准，宜采用G.711/G.723.1/G.729/AMR音频编解码标准；一般解码器应支持G.722.1或者 G.711，手机解码器应支持AMR（其他为可选）；其他音频编解码标准为可选项，具体要求如下：

- G.711，它是采用 A 率或 μ 率量化的 LOG-PCM 的编码。
- G.722.1，它是一种采用重叠调制变换 MLT 的编码算法，可提供宽频带音频。其采样率为 16kHz，支持 24kbit/s、32kbit/s 及扩展支持 16kbit/s 等输出码率。

— AMR, 它是 3GPP 中系统规范采用的自适应多码率的语音编码算法, 编码每帧间应该动态自适应的输出 4.75~12.2 kbit/s 八种码率模式。编码应支持 VAD/CNG 模式, 编码码流应符合 AMR 测试序列兼容性测试。

6 视频编解码技术要求

6.1 概述

本标准规定的视频编解码技术要求, 涉及M-JPEG、MPEG-4和H.264视频编码标准, 主要从编解码的档次和级别、工具选项、码流语法的规定、传输包的负载格式以及比特流和解码器的一致性测试等方面进行描述和要求。

采用MPEG-4编码格式、H.264编码格式的视频流应满足本标准的规定, 不应包含私有数据格式。

6.2 基于 MPEG-4 的视频编解码技术要求

6.2.1 MPEG-4 的档次和级别

MPEG-4的档次和级别提供了ISO/IEC14496中句法和语义的定义子集, 以便于解码器(指软/硬解码)能对指定的比特流进行正确解码。MPEG-4的一个档次是对整个比特流语法限制的子集, 一个级别则是对该比特流参数的限制集合。

采用MPEG-4标准的视频编码器应为MPEG-4简单档次 (Simple Profile) 或高级简单档次 (Advance Simple Profile) 的编码器。基本系统至少应支持ISO/IEC14496-2中规定的简单档次的级别L3: 即MPEG-4 SP@L3。系统最高可支持到高级简单档次的级别L5: 即MPEG-4 ASP@L5。

采用MPEG-4标准的视频解码器支持的档次和级别不应低于编码器支持的最高档次和级别 (MPEG-4 ASP@L5)。

6.2.2 MPEG-4 工具

简单档次的工具包括:

- I-VOP, 帧内编码的矩形视频对象平面;
- P-VOP, 帧间编码的矩形视频对象平面;
- AC/DC Prediction, AC/DC 预测;
- 4-MV, 每个宏块可以有 4 个运动矢量;
- Unrestricted MV, 不受限制的运动矢量。

高级简单档次的工具包括:

- B-VOP, 双向帧间编码的矩形视频对象平面;
- Interlace, 隔行编码。

MPEG-4视频编码器应支持简单档次或高级简单档次的部分或全部工具, 视频解码器(指软/硬解码)应支持高级档次的全部工具。

6.2.3 MPEG-4 的码流语法

为实现系统视频流的互通, 采用MPEG-4标准的视频码流语法应符合ISO/IEC 14496-2中的规定。

6.2.4 MPEG-4 RTP 包的载荷格式

MPEG-4的RTP载荷格式应符合IETF RFC3016和IETF RFC3640的规定, 其中RTP有效载荷是MPEG-4的原始视频流。

6.2.5 MPEG-4 的一致性测试

MPEG-4的一致性比特流是指实现了ISO/IEC 14496-2在通用语法中定义的所有限制的比特流, 包括ISO/IEC 14496-2中第9部分定义的关于档次和级别的限制。

MPEG-4的一致性比特流应满足如下测试: 当MPEG-4的比特流被VLC(xvid插件)解码器解码时, 参考软件解码器也不应出现任何由比特流引起的错误或不一致的消息报告。这种测试不适用于由于传输而产生错误的比特流。

MPEG-4的比特流一致性测试的附加测试见ISO/IEC 14496-4, 同时需要通过中立的第三方硬件测试仪器的测试。

6.3 基于 H.264 的视频编解码技术要求

6.3.1 H.264 的档次和级别

H.264的档次和级别规定了对比特流的限制要求, 因此也限制了解码比特流的能力。档次和级别还可用于指导不同解码器间的互操作性。

每个档次都规定了一个算法特征和限制的子集, 任何遵守某个档次的解码器都应该支持与其相应的子集。

每个级别都规定了一组对标准中语法元素所采用的各种参数值的限制。所有档次都使用一组相同的级别定义, 但是每个档次的具体实现可以支持不同的级别。

视频编码比特流应支持H.264视频标准的基本档次(Baseline Profile), 级别(Level)应至少支持到Level 1.3, 宜扩展支持到主档次(Main Profile)级别Level 3。

视频解码器支持的档次和级别应不低于编码器支持的最高档次和级别, 至少应支持到H.264视频标准主档次(Main Profile)的Level 3。

6.3.2 H.264 基本档次的选项和工具

H.264编码算法有多种选项和工具。

本标准使用的基本档次支持的选项和工具主要有:

- I片和P片(Slice) (P片限定二个参考帧);
- 基于内容自适应的变长编码 CAVLC;
- 去块效应滤波器。

注: 以下工具不做为强制要求: FMO、ASO、Redundant slice。

本标准使用的主档次支持的选项和工具主要有:

- B片(Slice) (B片限定二个参考帧, 前后各一个参考帧, B片不作为参考帧);
- 隔行编码(Interlace PAFF);
- 1/4 像素运动补偿。

注: 以下工具不做为强制要求: MBAFF、CABAC、Weighted Prediction。

6.3.3 H.264 RTP 包的载荷格式

H.264的RTP载荷格式应符合IETF RFC3016与IETF RFC3984的规定, 其中RTP的有效载荷是标准的H.264 NAL。H.264 3GPP 包的载荷格式(可选)。

在移动网络环境下, 编解码应该优选支持H.264-3GPP 码流封装格式(H.264 NAL), 具体要求见IETF RFC 3984。

6.3.4 H.264 的一致性测试

若一个视频编码器产生的比特流是满足ITU-T H.264|ISO/IEC 14496-10标准的一致性比特流, 则称该编码器与该标准相一致。

比特流一致性描述详见ITU-T H.264 |ISO/IEC 14496-10附录C中C.3的规定。

与ITU-T H.264|ISO/IEC 14496-10一致的比特流应通过下面的标准测试。

比特流应能被ITU-T H.264.2|ISO/IEC 14496-5中给出的参考软件解码器解码。当被参考软件解码器解码时, 比特流不应出现任何错误, 参考软件解码器也不应出现任何由比特流引起的错误或不一致的消息报告。这种测试不适用于由于传输而产生错误的比特流。

6.4 基于 M-JPEG 的视频编解码技术要求

6.4.1 解码侧技术要求

对 M-JPEG 解码侧应支持的技术要求如下:

- 支持 JPEG Baseline 编码, 即 Sequential DCT 编码;
- 支持 3 张 JPEG 量化表, 其中 2 张为协议推荐的缺省表;
- 支持 4 张 Huffman 码表, 其中包括 2 张 AC 码和 2 张 DC 码表;
- 支持 semi-planar YUV4:2:0 图像;
- MCU 采用 interleave 顺序组织。

6.4.2 基于 M-JPEG 的 RTP 包的载荷格式:

基于 M-JPEG 的 RTP 包的载荷格式应符合 IETF RFC2435 中的规定, 其中 RTP 有效载荷类型是 JPEG Baseline 编码。

7 其他技术要求

存储的图像数据应保证具有 CIF (352×288) 格式的图像分辨率, 重要目标和报警图像宜具有高于 4CIF (704×576) 格式的图像分辨率。

网络视频监控主要支持的分辨率:

- QCIF (PAL: 176×144, NTSC: 176×120);
- CIF (PAL: 352×288, NTSC: 352×240);
- 4CIF (PAL: 704×576, NTSC: 704×480);
- D1 (PAL: 720×576, NTSC: 720×480);
- VGA (PAL: 640×480, NTSC: 640×400);
- QVGA (PAL: 320×240, NTSC: 320×200);
- 720P (1280×720);
- 1080P (1920×1080)。

附 录 A
(规范性附录)
参 数

A.1 MPEG-4 相关

A.1.1 级别的参数限制

在ISO/IEC 14496-2中, MPEG-4的简单档次支持L0、L1、L2、L3四个级别, 每个级别的参数限制如表A.1所示。MPEG-4 L4a、L5级别的参数限制在表A.2中给出。

表A.1 MPEG-4 (Level 0~Level 3) 级别的参数限制

Level	L3	L2	L1	L0
Typical Visual Session Size	CIF	CIF	QCIF	QCIF
Max number of objects	4	4	4	1
Maximum number objects per type	4×Simple	4×Simple	4×Simple	1×Simple
Max unique Quant Tables	1	1	1	1
Max. VMV buffer size (MB units)	792	792	198	198
Max VCV buffer size (MB)	396	396	99	99
VCV decoder rate (MB/s)	11880	5940	1485	1485
VCV Boundary MB decoder rate (MB/s)	Not Applicable	Not Applicable	Not Applicable	Not Applicable
Max total VBV buffer size (units of 16384 bits)	40	40	10	10
Max vol vbv buffer size (units of 16384 bits)	40	40	10	10
Max. video packet length (bits)	8192	4096	2048	2048
Max sprite size (MB units)	Not Applicable	Not Applicable	Not Applicable	Not Applicable
Wavelet restrictions	Not Applicable	Not Applicable	Not Applicable	Not Applicable
Max bitrate (kbit/s)	384	128	64	64
Max. enhancement layers per object	Not Applicable	Not Applicable	Not Applicable	Not Applicable

表A.2 MPEG-4 L4a、L5 级别的参数限制

Level	L5	L4a
Typical Visual Session Size	720×576	640×480
Max number of objects	4	4
Maximum number objects per type	4×Simple	4×Simple
Max unique Quant Tables	1	1
Max. VMV buffer size (MB units)	3240	2400
Max VCV buffer size (MB)	1620	1200
VCV decoder rate (MB/s)	40500	36000
VCV Boundary MB decoder rate (MB/s)	Not Applicable	Not Applicable
Max total VBV buffer size (units of 16384 bits)	112	80
Max vol vbv buffer size (units of 16384 bits)	112	80
Max. video packet length (bits)	16384	16384
Max sprite size (MB units)	Not Applicable	Not Applicable
Wavelet restrictions	Not Applicable	Not Applicable
Max bitrate (kbit/s)	8000	4000
Max. enhancement layers per object	Not Applicable	Not Applicable

A.2 H.264 相关

A.2.1 H.264级别的限制

H.264级别 (Level 1~ Level 3) 的限制如表A.3所示, 表中“—”表示未做相应的限制。

表A.3 H.264 级别 (Level 1~ Level 3) 的限制

Level number	Max macroblock processing rate MaxMBPS (MB/s)	Max frame sizeMax FS (MBs)	Max decoded picture buffer sizeMax DPB (1024 bytes for 4:2:0)	Max video bit rate Max BR (1000 bit/s or 1200 bit/s)	Max CPB size MaxCPB (1000 bits or 1200 bits)	Min compression ratio MinCR	Max number of motion vectors per two consecutive MBs Max MvsPer2Mb
1	1 485	99	148.5	64	175	2	—
1.1	3 000	396	337.5	192	500	2	—
1.2	6 000	396	891.0	384	1 000	2	—
1.3	11 880	396	891.0	768	2 000	2	—
2	11 880	396	891.0	2 000	2 000	2	—
2.1	19 800	792	1 782.0	4 000	4 000	2	—
2.2	20 250	1 620	3 037.5	4 000	4 000	2	—
3	40 500	1 620	3 037.5	10 000	10 000	2	32

A.2.2 H.264基本档次各级别的参数限制

H.264基本档次各级别的参数限制如表A.4所示, 表中的“—”表示未做相应的限制。

表A.4 H.264 基本档次各级别的参数限制

Level number	MaxSubMbRectSize
1	576
1.1	576
1.2	576
1.3	576
2	576
2.1	576
2.2	576
3	576

A.2.3 H.264各级别的最大帧率限制

当帧的大小在176×144到720×576间变化时, H.264各级别 (Level) 的最大帧率限制如表A.5、表A.6所示, 表中的“—”表示未做相应的限制。

表A.5 H.264 各级别的最大帧率限制(1)

Level					1	1b	1.1	1.2	1.3
Max frame size (macroblocks):					99	99	396	396	396
Max macroblocks/second:					1 485	1 485	3 000	6 000	11 880
Max frame size (samples):					25 344	25 344	101 376	101 376	101 376
Max samples/second:					380 160	380 160	768 000	1 536 000	3 041 280
Format	Luma Width	Luma Height	MBs Total	Luma Samples					

QCIF	176	144	99	25 344	15.0	15.0	30.3	60.6	120.0
------	-----	-----	----	--------	------	------	------	------	-------

表 A.1 (续)

Level					1	1b	1.1	1.2	1.3
QVGA	320	240	300	76 800	—	—	10.0	20.0	39.6
525 SIF	352	240	330	84 480	—	—	9.1	18.2	36.0
CIF	352	288	396	101 376	—	—	7.6	15.2	30.0
525 HHR	352	480	660	168 960	—	—	—	—	—
625 HHR	352	576	792	202 752	—	—	—	—	—
VGA	640	480	1 200	307 200	—	—	—	—	—
525 4SIF	704	480	1 320	337 920	—	—	—	—	—
525 SD	720	480	1 350	345 600	—	—	—	—	—
4CIF	704	576	1 584	405 504	—	—	—	—	—
625 SD	720	576	1 620	414 720	—	—	—	—	—

表A.6 H.264 各级别的最大帧率限制(2)

Level					2	2.1	2.2	3
Max frame size (macroblocks):					396	792	1 620	1 620
Max macroblocks/second:					11 880	19 800	20 250	40 500
Max frame size (samples):					101 376	202 752	414 720	414 720
Max samples/second:					3 041 280	5 068 800	5 184 000	10 368 000
Format	Luma Width	Luma Height	MBs Total	Luma Samples				
QCIF	176	144	99	25 344	120.0	172.0	172.0	172.0
QVGA	320	240	300	76 800	39.6	66.0	67.5	135.0
525 SIF	352	240	330	84 480	36.0	60.0	61.4	122.7
CIF	352	288	396	101 376	30.0	50.0	51.1	102.3
525 HHR	352	480	660	168 960	—	30.0	30.7	61.4
625 HHR	352	576	792	202 752	—	25.0	25.6	51.1
VGA	640	480	1 200	307 200	—	—	16.9	33.8
525 4SIF	704	480	1 320	337 920	—	—	15.3	30.7
525 SD	720	480	1 350	345 600	—	—	15.0	30.0
4CIF	704	576	1 584	405 504	—	—	12.8	25.6
625SD	720	576	1620	414 720			12.5	25

中华人民共和国
通信行业标准
电信网视频监控系统
视音频编解码技术要求
YD/T 2883-2015

*

人民邮电出版社出版发行
北京市丰台区成寿寺路11号邮电出版大厦
邮政编码: 100164
北京康利胶印厂印刷
版权所有 不得翻印

*

开本: 880×1230 1/16 2016年2月第1版
印张: 1 2016年2月北京第1次印刷
字数: 21千字

15115·794

定价: 10元

本书如有印装质量问题, 请与本社联系 电话: (010)81055492