

ITU-R BS.1548-8建议书

(05/2025)

BS系列：广播业务（声音）

数字广播音频编码系统的用户需求

前言

无线电通信部门的作用是确保所有无线电通信业务，包括卫星业务，合理、公平、有效和经济地使用无线电频谱，并开展没有频率范围限制的研究，在此基础上通过建议书。

无线电通信部门制定规章制度和政策的职能由世界和区域无线电通信大会以及无线电通信全会完成，并得到各研究组的支持。

知识产权政策（IPR）

国际电联无线电通信部门（ITU-R）的 IPR 政策述于 ITU-R 第 1 号决议所参引的《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 <https://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 建议书系列	
（可同时在以下网址获得： https://www.itu.int/publ/R-REC/zh ）	
系列	标题
BO	卫星传输
BR	用于制作、存档和播放的记录；用于电视的胶片
BS	广播业务（声音）
BT	广播业务（电视）
F	固定业务
M	移动、无线电测定、业余无线电以及相关卫星业务
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	卫星固定和固定业务系统之间频率共用和协调
SM	频谱管理
SNG	卫星新闻采集
TF	时间信号和标准频率发射
V	词汇和相关课题

注：本ITU-R建议书英文版已按ITU-R第1号决议规定的程序批准。

电子出版物
2025年，日内瓦

© 国际电联 2025

版权所有。未经国际电联书面许可，不得以任何手段翻印本出版物的任何部分。

ITU-R BS.1548-8建议书

数字广播音频编码系统的用户需求

(ITU-R第19-1/6号课题)

(2001-2002-2006-2012-2013-2017-01/2019-10/2019年)

范围

本建议书详细说明了与包括电视在内的声音广播中音频信源编码系统的使用有关的要求。本建议书涵盖了接收、分发和发射的应用。

关键词

音频、音频编码、广播、数字、广播、声音、电视、编解码器、高级音响系统、基于声道的声音信号、基于对象的声音信号、音频相关元数据、音频定义模型 (ADM)

国际电联无线电通信全会，

考虑到

- a) 电视和声音广播的音响系统所要求的基本音频和立体声图像的质量可能是最高级的，通常是无法与原始资料辨别的；
- b) 一些发射应用所要求的音频质量应当等效于或优于FM模拟广播业务的良好接收质量；
- c) 应该仔细考虑涉及诸如接收和分发链路这样的节目连接线路的互操作性和网络操作；
- d) 应该仔细考虑与现有的诸如矩阵环绕解码器和分离多声道解码器这样的消费者多声道音频设备的互操作性；
- e) 当在现有的广播业务中引入一个多声道音响系统时，必须考虑与现有接收机的兼容性以保持业务；
- f) 更普遍的，考虑到这类系统的许多应用，应该明确规定所有技术、质量和操作方面的要求；
- g) 音频编码系统的性能是普遍地依赖于系统运行时的配置（比特速率、预矩阵化的使用，合成编码的使用等）；
- h) 因此，广播公司需要有必要的信息来设置被推荐系统的所有可用到的编码参数；
- i) 引入与类似性能特性不兼容的系统是非常不希望的；
- j) 还没有开始业务的那些广播公司应该能够选择最适合他们应用且最具有成本效益的系统，

注意到

- a) 有或没有伴图的多声道音响系统是ITU-R BS.775建议书的议题；

- b) ITU-R BS.2051建议书规定的、用于高级音响系统的扬声器布局和声道配置；
- c) ITU-R BS.2076、ITU-R BS.2094、ITU-R BS.2125和ITU-R BS.2168建议书中的音频定义模型（ADM）规定了用于制作的音频相关元数据格式；
- d) ITU-R BS.2127建议书规定了ADM的渲染算法；
- e) 用于数字广播的音频编码是ITU-R BS.1196建议书的主题；
- f) ITU-R BS.1283建议书为主观评估声音质量提供了ITU-R建议书的指南，

建议

- 1 用于接收和分发应用的数字电视和声音广播的音频编码系统应该满足附件1中列出的要求；
- 2 用于发射应用的数字电视和声音广播的音频编码系统应该满足附件2中列出的要求；
- 3 附件3中列出的音频质量的种类应当支配建议1和2中的音频质量和应用。

注1 – 接收和分发应用中已经被证明满足质量和其他要求的系统的信息被包含在附件1的附录1中。

注2 – 发射应用中已经被证明满足质量和其他要求的系统的信息被包含在附件2的附录1中。

目录

页码

知识产权政策（IPR）	ii
附件1 – 接收和分发的要求	4
1 业务要求	4
1.1 声道配置	4
1.2 声道的灵活分配	5
1.3 辅助数据	5
1.4 高级音响系统的音频相关元数据	5
2 性能要求	6
2.1 音频质量	6
2.2 编码时延	7
2.3 误码适应力	7
2.4 恢复时间	7
3 功能和操作要求	7
3.1 比特速率和编码方案	7

3.2	合成编码	8
附件1的附录1 – （资料性） 有关已经被验证满足接收和分发的质量及其他用户需求的编码系统的信息		8
附件2 – 发射要求.....		10
1	业务要求	10
1.1	声道配置	10
1.2	音频业务	11
1.3	声道的灵活分配	12
1.4	辅助数据	12
1.5	高级音响系统的音频相关元数据	12
2	性能要求	12
2.1	音频质量	12
2.2	编码时延	14
2.3	误码适应力	14
2.4	恢复时间	14
3	多声道系统的功能要求和操作要求	14
3.1	与单声道/立体声系统的兼容性（ITU-R BS.775建议书）	14
3.2	比特速率	15
3.3	解码器复杂度	15
附件2的附录1 – （资料性） 有关已经被验证满足发射的质量及其他用户需求的编码系统的信息		15
附件3 – 广播应用的音频质量等级.....		23
附件4（资料性） – 与ITU-R BS.1548-7建议书相比的修改摘要.....		23

附件1

接收和分发的要求

对于接收和分发应用，数字电视和声音广播的音频编码系统应该满足下面列出的要求。

1 业务要求

1.1 声道配置

对于音频业务，应该根据应用需求支持下列声道配置中的至少一项。

1.1.1 ITU-R BS.775建议书规定的声道配置

表1

声道数	声道配置	声道分配
1 声道	1/0	单声道
2 声道	2/0	左、右
3 声道	3/0 2/1	左、右、中 左、右/环绕
4 声道	3/1 2/2	左、右、中/环绕 左、右/环绕左、环绕右
5 声道	3/2	左、右、中间/环绕左、环绕右
注：对于声道配置“a/b”“a”和“b”，分别指明正面声道和背面声道的数量		

另外，对于接收，可能有必要传递用与上面列出的那些格式不一样的其他格式（比如，3/4）生成的节目，这样，编码系统应该允许采用其他的高质量声道。

1.1.2 ITU-R BS.2051建议书规定的、用于基于声道的高级音响系统的声道配置

表2

音响系统 标签	声道数	声道配置	LFE声道数	声道分配
系统C	8	2+5+0 (2/0+3/2+0)	1	左上方正面、右上方正面+左、右、中/左环绕、右环绕。LFE
系统D	10	4+5+0 (2/2+3/2+0)	1	左上方正面、右上方正面/左上方背面、右上方背面+左、右、中/左环绕、右环绕。LFE
系统E	11	4+5+1 (2/2+3/2+1/0)	1	左上方正面、右上方正面/左上方背面、右上方背面+左、右、中/左环绕、右环绕+中下方正面。LFE

表2（结束）

音响系统 标签	声道数	声道配置	LFE声道号	声道分配
系统F	12	3+7+0 (2/1+3/2/2+0)	2	左高、右高/中高+左、右、中/左侧、右侧/左后、右后。左LFE、右LFE
系统G	14	4+9+0 (2/2+5/2/2+0)	1	左上方正面、右上方正面/左上方后面、右上方后面+左、右、中、左屏幕、右屏幕/左侧环绕、右侧环绕/左背面环绕、右背面环绕。LFE
系统H	24	9+10+3 (3/3/3+5/2/3+0)	2	上方正面左、上方正面右、上方正面中/上方侧面左、上方侧面右、上方中/上方后面左、上方后面右、上方后面中+正面左、正面右、正面左中、正面右中、正面中/侧面左、侧面右/后面左、后面右、后面中+下方正面左、下方正面右、下方正面中。LFE-1、LFE-2
系统I	8	0+7+0 (0+3/2/2+0)	1	左、右、中/左侧环绕、右侧环绕/左背面环绕、右背面环绕。LFE
系统J	12	4+7+0 (2/2+3/2/2+0)	1	左上方正面、右上方正面/左上方后面、右上方后面+左、右、中/左侧面环绕、右侧面环绕/左背面环绕、右背面环绕。LFE

注：对于声道配置“a/b/c+a/b/c+a/b/c”，第一、第二和第三“a/b/c”部分分别指明上、中、下各层面的音频声道数。“a”“b”和“c”分别指明正面、侧面和背面的声道数。当侧面的声道数为0时，“a/b/c”可以写成“a/c”。当层面的音频声道数为0时，“a/b/c”可以写成“0”。

对于接收，可能有必要传递用与上面列出的那些格式不一样的其他格式生成的节目；这样，编码系统应该允许采用其他的高质量声道。

1.2 声道的灵活分配

比特流应该提供用作声音配置的信令和控制的识别数据。在播送系统中，必须要能在§ 1.1 中列出的声道配置之间动态地切换。

1.3 辅助数据

音频编码系统应该提供传输辅助数据的可能性。辅助数据能够传送各种类型的信息，包括动态范围控制、响度控制、用户数据及将对传递给消费者的最终音频进行编码的发射编码器所需要的任何元数据。

1.4 高级音响系统的音频相关元数据

音频编码系统应根据需要为高级音响系统提供音频相关元数据的传输可能性。

2 性能要求

2.1 音频质量

2.1.1 基本音频质量

对于大多数音频节目素材的类型，在一个参考接收/分发级联（5个接收编解码器和3个分发编解码器顺序工作）后再生的音频的质量应该与原始资料在主观上无法区分。采用描述于 ITU-R BS.1116 建议书 – 包含多声道音响系统的音频系统中用于主观评估小损伤的方法 – 中的隐藏参考的三激励双盲测试，对位于参考收听位置的收听者，这要求在 5 级损伤评定等级中平均分通常大于 4.5。最劣额定项目应该不低于 4 级。

注1 – 对于一个编解码器和项目，与单个平均分有关的置信区间（误差条）给出了真实分数可能会以某种程度的确定性（通常为95%）落于其中的规定的平均分的上下范围。一个编解码器和项目的真实分数可能会差到位于规定分值的置信区间的下限。为了对级联编解码器的预期性能作一个有意义的评估，与单个编解码器报告的平均分关联的置信区间必须近似等于或低于被比较分值之间的差。

注2 – 与发射编解码器的基本音频质量相比，当与发射编解码器前后放置时，接收/分发级联不应该引起质量方面的明显下降。精确的规范有待进一步研究。

注3 – 与ITU-R BS.1387建议书一致，接收/分发的客观音频质量参数可以在稍后并入。

注4 – 被称为“基本音频质量”的主观音频质量属性在ITU-R BS.1116建议书中进行描述。

注5 – 如果接收/分发编解码器支持，则聆听测试中的扬声器配置应与节目制作中使用的扬声器配置相同。

注6 – 在分发中，多对象常联合编码（类似于基于声道的内容）以提高编码效率。此外，在广播应用中，通常不打算单独回放多个对象。

2.1.2 量化精度

分发所要求的精度应该至少为 18 比特，而对于接收，20 比特或更多是更好的。

2.1.3 采样频率

与 ITU-R BS.646 建议书 – 广播演播室中数字音频信号的信源编码相一致，采样频率应该是 48 kHz。

2.1.4 带宽

主音频声道：20-20 000 Hz。

LFE声道：15-120 Hz。

2.1.5 加重

音频编码系统应该是无加重的。

2.1.6 级联能力

所要求的级联能力依赖于表 3 所给的应用：

表3

分发	3个编解码器级联
接收	5个编解码器级联

这些数字取自于先前为评估双声道声音广播系统所做的实验（参见 ITU-R BS.1196 建议书），并且可能不能代表实际的无线电和电视广播运行情况。需要更多的信息来更好地说明这个方面。

2.1.7 后处理能力

需要的后处理能力很大程度上依赖于应用。对于分发，交叉衰减可以与动态范围控制一起应用。

2.2 编码时延

一个节目中所有声道的编码时延必须是相同的。考虑到所需的编码性能（即比特速率缩减的量），编码时延应该尽可能的小。在电视伴音的情形中，音频的延迟必须与视频的延迟匹配。音频编码器产生的编码后音频帧（接入单元）准确对应到匹配视频帧的时间周期是所希望的。

2.3 误码适应力

在音频比特流中必须提供一种机制以允许解码器识别残余的声道误差并采用适当的屏蔽方法。

2.4 恢复时间

恢复时间应该尽可能短。在应用音频接入单元（AAU）的情况下，恢复时间应在几个 AAU 之内，更为希望的是，在接收到第一个无差错的 AAU 后应该恢复音频。

3 功能和操作要求

3.1 比特速率和编码方案

对于分发和接收链路，在级联高达 5 级的情况下，ITU-R BS.1196 建议书推荐采用如国际标准化组织/国际电工委员会（ISO/IEC）IS 11172-3 中规定的、每声道 180 kbit/s 或更高比特速率的 MPEG1 层 II，如 ISO/IEC 14496-3 中规定的、每声道 144 kbit/s 或更高比特速率的 MPEG-4 AAC，以及 ISO/IEC 23008-3 中规定的、每声道 144 kbit/s 或更高比特速率的 MPEG-H 3D 音频。另外，在级联高达 5 级的情况下，还包括 ETSI TS 103 190-1 v1.3.1 和 ETSI TS 103 1901-2 v1.3.1 中规定的、每声道 128 kbit/s 或更高比特速率的 AC-4。出于若干原因，系统可能会以不同的比特速率来应用，或者可能采用其他的系统。

可能包括下面几个原因：

- 支持可能会在编码产生块之间插入的信号处理的额外的编码余量（在ITU-R BS.1196建议书的开发过程中这并没有得到测试或验证）；
- 在分发和接收链路中获得更低的比特速率；
- 获得更高的质量；
- 与伴随的视频信号同步和转换的适宜性。

注1 – 当高级音响系统中带有音频对象的声音节目要在没有用户交互的情况下播放时，对于应该同时播放的所有音频对象的叠加，可以确认音频质量。

注2 – 当通过可单独回放某些音频对象的用户互动回放高级音响系统中带有音频对象的声音节目时，应分别确认每个这些音频对象的音频质量。此类单个音频对象所需比特率加上其余音频内容的必要比特率之和可视为包含音频对象的声音节目所需比特率的上限。

注3 – 如果限制互动性来增加/减少某些音频对象的电平，则必须通过在整个混音中用这些音频对象的最大和最小电平设置来确认质量。

注4 – 如果场景中有许多对象，则每个对象所需的比特率都会降低，因为听者能够感知的数据总量是有限的。

3.2 合成编码

双声道或多声道节目素材经常包含一些声道间的统计相关性。合成编码是减少声道间不相干性或冗余性从而提高编码效率的一种有效方法。一些编码系统使用感性准则通过在人耳分辨声源方向能力较差的频率区域上联合两个或多个声道来消除声道间不相干的部分。这项技术的缺点是在随后的阶段中不能够正确地重新配置通常在原始声道中存在的声音信息。对于接收和许多分发应用来说，不应该采用此种合成编码方案。

附件1的 附录1 (资料性)

有关已经被验证满足接收和分发的质量及 其他用户需求的编码系统的信息

表4的左边一栏中列出了附件1中规定的要求。右边一栏中给出了满足这些要求的特定编解码器的能力。预计未来对本建议书的修订将包括有关其他编解码器的额外信息。

表4

附件1的要求列表	编解码器: Dolby E [参考文献1]	MPEG-4 AAC	AC-4	MPEG-H 3D音频
1.1.1 ITU-R BS.775规定的声道配置	满足, [参考文献1, p. 6]	满足	满足	满足
1.1.2 ITU-R BS.2051规定的、基于声道的高级声音系统的声道配置 (默认支持)	N/A	系 统 C 、 H、I	系 统 C 、 D 、 G 、 I、J	系 统 C 、 D、F 至J
1.2 声道的灵活分配	满足, [参考文献1, p. 15]	满足	满足	满足
1.3 辅助数据	满足, [参考文献1, p. 14]	满足	满足	满足
1.4 高级音响系统的音频相关元数据	N/A	N/A	满足	满足
2.1.1 基本音频质量	满足, [参考文献2]	满足	满足	满足
2.1.2 量化精度	满足, [参考文献1, p. 5]	满足	满足	满足
2.1.3 采样频率	满足, [参考文献1, p. 5]	满足	满足	满足
2.1.4 带宽	满足, [参考文献1, p. 9]	满足	满足	满足
2.1.5 加重	满足, [参考文献1]	满足	满足	满足
2.1.6 级联能力	满足, [参考文献2]	满足	满足	满足
2.1.7 后处理能力	未经验证	满足	满足	满足
2.2 编码时延	满足 ⁽¹⁾ , [参考文献1, p. 7]	满足	满足	满足
2.3 误码适应力	满足, [参考文献1, p. 15]	满足	满足	满足
2.4 恢复时间	满足, [参考文献1, p. 15]	满足	满足	满足
3.1 比特速率和编码方案	满足 ⁽²⁾ , [参考文献1, p. 6]	满足	满足	满足
3.2 合成编码	满足, [参考文献1]	满足	满足	满足

⁽¹⁾ 为了便于电视声音的操作, 编码或解码延迟与相应的视频帧速率 (1/24、1/25、1/30秒) 相同。访问单元对应于视频帧。

⁽²⁾ 比特速率/信道为250 kbit/s, 以便获得§ 3.1下第一、第三和第四点中指出的优点。

参考文献

- [1] FIELDER, L. D., LYMAN, S. B., VERNON, S. and TODD, C. C. [September 1999] *Professional audio coder optimized for use with video*. 107th AES Convention, New York, NY, United States of America.
- [2] GRANT, D., DAVIDSON, G. and FIELDER, L. [21-24 September 2001] *Subjective evaluation of an audio distribution coding system*. 111th AES Convention, New York, NY, United States of America.

附件2

发射要求

对于发射应用，数字电视和声音广播的音频编码系统应该满足下面列出的要求。

1 业务要求

1.1 声道配置

对于音频业务，根据应用需求至少应支持下列声道配置中的一项。

1.1.1 ITU-R BS.775建议书规定的声道配置

表5

声道数	声道配置	声道分配
1 声道	1/0	单声道
2 声道	2/0	左、右
3 声道	3/0 2/1	左、右、中 左、右/环绕
4 声道	3/1 2/2	左、右、中/环绕 左、右/环绕左、环绕右
5 声道	3/2	左、右、中/环绕左、环绕右

注：对于声道配置“a/b”“a”和“b”，分别指明正面声道和背面声道的数量

1.1.2 ITU-R BS.2051规定的、基于声道的高级音响系统的声道配置

表6

音响系统 标签	声道数	声道配置	LFE声道数	声道分配
系统C	8	2+5+0 (2/0+3/2+0)	1	左上方正面、右上方正面+左、右、中/左环绕、右环绕。LFE
系统D	10	4+5+0 (2/2+3/2+0)	1	左上方正面、右上方正面/左上方背面、右上方背面+左、右、中/左环绕、右环绕。LFE
系统E	11	4+5+1 (2/2+3/2+1/0)	1	左上方正面、右上方正面/左上方背面、右上方背面+左、右、中/左环绕、右环绕+中下方正面。LFE
系统F	12	3+7+0 (2/1+3/2/2+0)	2	左高、右高/中高+左、右、中/左侧、右侧/左后、右后。左LFE、右LFE

表6（结束）

音响系统 标签	声道数	声道配置	LFE声道数	声道分配
系统G	14	4+9+0 (2/2+5/2/2+0)	1	左上方正面、右上方正面/左上方后面、右上方后面+左、右、中、左屏幕、右屏幕/左侧环绕、右侧环绕/左背面环绕、右背面环绕。LFE
系统H	24	9+10+3 (3/3/3+5/2/3+3/0)	2	上方正面左、上方正面右、上方正面中/上方侧面左、上方侧面右、上方中/上方后面左、上方后面右、上方后面中+正面左、正面右、正面左中、正面右中、正面中/侧面左、侧面右/后面左、后面右、后面中+下方正面左、下方正面右、下方正面中。LFE-1、LFE-2
系统I	8	0+7+0 (0+3/2/2+0)	1	左、右、中/左侧环绕、右侧环绕/左背面环绕、右背面环绕。LFE
系统J	12	4+7+0 (2/2+3/2/2+0)	1	左上方正面、右上方正面/左上方后面、右上方后面+左、右、中/左侧面环绕、右侧面环绕/左背面环绕、右背面环绕。LFE

注：对于声道配置“a/b/c+a/b/c+a/b/c”，第一、第二和第三“a/b/c”部分分别指明上、中、下各层面的音频声道数。“a”“b”和“c”分别指明正面、侧面和背面的声道数。当侧面的声道数为0时，“b”可以写成“a/c”。当层面的音频声道数为0时，“a/b/c”可以写成“0”

1.2 音频业务

与主要的音频业务一起，根据应用的需要，还能够提供下列相关联的音频业务：

- 多语言的业务 – 包括一个或多个独立的声道，用于分发一个具有使用一种或多种语言的解说词的节目；
- 听力和视力障碍者的音频业务 – 视力障碍者的业务通常包含对图片内容的声音描述，而听力障碍者的业务应该包括清晰的对话，不伴有或伴有较低等级的音乐和特殊效果以提高语言的可理解性；
- 辅助数据 – 传送各种类型的信息，包括：动态范围控制、响度控制和用户数据（ITU-R BS.775建议书）。

各种业务可以被分组为：

- 主业务（主业务的每个声道都分配给同一个节目，包括可选的LFE声道）。
- 扩展业务，可以是：
 - 独立业务（对于附加节目，是与主业务节目无关的，例如解说词，或包括两个或更多个声道的其他业务；可以按照§ 1.1中的表来选择声道配置）。
 - 可选业务（对于试图替代主业务声道中的一个或多个声道的节目，例如多语言的、听力障碍的）。
 - 附加业务（包含有将被添加到主业务声道中的声道，例如解说词，或者用于像3D TV这样高级音响系统的附加声道）。

由于任何传输系统都应当包括一个能够执行多路复用操作的系统层，因此，采用单个比特流传送上面列出的所有音频业务不是所需要的。

1.3 声道的灵活分配

比特流应该提供用作声音配置的信令和控制的识别数据。播送系统必须提供在§ 1.1 中列出的任何声道配置之间动态切换的能力。

1.4 辅助数据

音频编码系统应该提供传输辅助数据的可能性。辅助数据能够传送各种类型的信息，包括动态范围控制、响度控制和用户数据。

1.5 高级音响系统的音频相关元数据

音频编码系统应根据需要提供传输高级音频系统的音频相关元数据的可能性。

2 性能要求

2.1 音频质量

对于附件 3 中所示的发射应用，假设了两类音频质量。这些是高质量（“CD 质量”）发射和中等质量发射。

音频质量由几个参数来表征，特别是音频编码方法、采样速率和比特速率。满足所要求的音频质量所需要的比特速率是由音频编码方法和采样速率所主导的。

2.1.1 基本音频质量

2.1.1.1 高质量发射

广播公司典型地具有折衷音频质量和音频所用比特速率的能力。理想情况下，对于大多数音频节目素材类型，解码后再生的音频的质量在主观上是类似于原始信号的。采用 ITU-R BS.1116 建议书中描述的隐藏参考的三激励双盲测试，这要求在参考收听位置上平均值要一直高于 ITU-R BS.1116 建议书 5 级损伤评定等级中的 4 级。实际上，商业需求有时会导致以低于达到这个质量水平所必需的比特速率进行工作。然而，系统应该提供给广播公司操作在这个质量水平上的选择权。

注1 – 与ITU-R BS.1387建议书一致，接收/分发的客观音频质量参数可以在稍后并入。

注2 – 如果接收/分发编解码器支持，则聆听测试中的扬声器配置应与节目制作中使用的扬声器配置相同。

注3 – 多个对象通常联合编码（类似于基于声道的内容）以提高编码效率。此外，在广播应用中，通常不打算单独回放多个对象。

注4 – 音频对象也可用于多种语言。带有对话的音频对象所需的比特率可能取决于语言。

2.1.1.2 中等质量发射

在一些发射应用中，可能需要低于“CD 质量”但等于或优于 FM 或 AM 模拟广播业务良好接收时的音频质量。使用 ITU-R BS.1534 建议书中描述的 MUSHRA 方法，可能需要对应于“优秀”和“良好”等级的平均分。可能也会使用测试中用作支撑的未处理音频信号的低通滤波版本，因为这些代表了现有模拟声音广播系统的音频质量。

2.1.2 立体声质量

在双声道立体音响或多声道配置的情形中，原始素材的声音画面的质量应该被保留。对于包含一个中间声道的配置（3/0、3/1、3/2），在大于常规双声道立体声所提供的收听区域上，正面声音画面的方向稳定性应该保持在合理的限度内。对于包括环绕声道的配置（2/1、2/2、3/1、3/2），空间真实性（氛围）的感觉应该比常规双声道立体声（ITU-R BS.775 建议书）所提供的有显著地增强。

2.1.3 量化精度

要求的精度应该至少是 16 比特。

2.1.4 采样频率

2.1.4.1 高质量发射

与 ITU-R BS.646 建议书相一致，采样频率应该为 48 kHz。

2.1.4.2 中等质量发射

当不要求“CD 质量”时，应该允许采用低于 48 kHz 的采样频率。与 ITU-R BS.1196 建议书相一致，采样频率应该是 32 kHz 或者 48 kHz。进一步考虑到通过使用一个降低的采样频率可以改善甚低比特速率的可察觉到的音频质量，并且 MPEG-2 音频允许采用更低的采样频率，即半采样频率（16、22.05 和 24 kHz）及四分之一采样频率（8、11.025 和 12 kHz），更低的采样频率对于中等质量的发射可能是合适的。

2.1.5 带宽

2.1.5.1 高质量发射

主音频声道：20-20 000 Hz。

LFE声道：15-120 Hz。

2.1.5.2 中等质量发射

带宽依赖于采样频率。

2.1.6 加重

音频编码系统不应当采用加重。

2.1.7 后处理能力

需要的后处理能力很大程度上取决于应用。对于发射链路，可以限制于均衡和动态范围调整（例如，把节目素材的动态范围匹配到收听环境的动态范围）。

2.2 编码时延

一个节目中所有声道的编码时延必须是相同的。在电视伴音的情况中，音频的时延必须与视频的时延匹配。

2.3 误码适应力

在音频比特流中必须提供一种机制以允许解码器识别残余的声道误码并采用适当的屏蔽方法。

2.4 恢复时间

恢复时间应该尽可能短。对于提供音频接入单元（AAU）的系统，恢复时间应该在几个 AAU 之内，理想情况是在单个 AAU 内。

3 多声道系统的功能要求和操作要求

3.1 与单声道/立体声系统的兼容性（ITU-R BS.775建议书）

3.1.1 向下兼容性

一个多声道比特流格式必须是能够被不同复杂度的各类解码器解码的。根据用户的再现能力，解码器中必须要能够用比发射声道数要少的一定数量的声道来安排一次呈现，并且除了立体声或多声道局部效应损失之外不会带来其他的损伤。

已经确认两种方法可以提供与低接收机复杂度的向下兼容性。第一种方法需要采用矩阵处理。低成本接收机只需要 A-和 B-声道，这与 2/0 系统的情形一样，即系统不使用后向兼容性矩阵。第二种方法适用于离散的 3/2 传递系统。被传递的信号使用方程式来进行数字合成，这要能够提供需要的信号数。在低比特速率信源编码信号的情况中，3/2 信号的向下混合可以在解码过程的合成部分（大部分复杂性位于这里）之前进行。

3.1.2 后向兼容性

这个要求适用于现有的单声道/立体声应用必须被升级到多声道音频但必须维持对现有接收机的业务的情形。在已经采用单声道或者立体声的系统中，多声道低比特速率编码的后向兼容性意味着解码器应该正确地解码基本的立体声信息，由来自所有源声道的音频信息恰当的向下混合所组成。为了满足这个需求，或者可以采用同时联播方法，或者可以采用矩阵化方法。

同时联播方法

一种方法是继续提供现有的单声道/立体声业务，并增加新的 3/2 声道业务。这种方法被称为同时联播操作。此方法的优点是可以在将来的某个时间点停止现有的单声道/立体声业务，并且 2/0 和 3/2 节目混合可以被相互独立地优化。

矩阵化方法

为了产生需要的音频声道数，另一种方法是采用兼容矩阵，它是对发射声道中传送的信号进行线性组合来实现的。矩阵方程可以用来提供与现有接收机的兼容性。在这种情况下，现有的左和右发射声道用于传送兼容的 A 和 B 矩阵信号。另外的发射声道用来传送 T、Q₁ 和 Q₂ 矩阵信号。这种方法的优点可能是在增加新业务时需要较少的额外的数据容量。

3.1.3 前向兼容性

对于必须与单声道/立体声系统共存的新的多声道系统的应用，可能要求解码器能够解码单声道/立体声音频比特流。

3.2 比特速率

ITU-R BS.1196 建议书详细说明了高质量发射应用中立体声信号所需的比特速率。这样，在不需要后向兼容（见§ 3.1.2）的情况下，2.5 倍的比特速率（即 $5/2 \times 144$ kbit/s 到 $5/2 \times 256$ kbit/s）可以被认为是 5 声道主业务的一个上限。就像合成编码技术可以提供额外的编码增益，对于§ 2.1 中定义的音频质量，采用新的多声道编码系统应该可以达到比特速率的明显降低。

注1 – 多个对象通常联合编码（类似于基于声道的内容）以提高编码效率。当高级音响系统中带有音频对象的声音节目要在没有用户交互的情况下回放时，可以确认同时回放的所有音频对象的音频质量。

注2 – 当通过可单独回放某些音频对象的用户互动回放高级音响系统中带有音频对象的声音节目时，应分别确认每个这些音频对象的音频质量。此类单个音频对象所需比特率加上其余音频内容的必要比特率之和可视为包含音频对象的声音节目所需比特率的上限。

注3 – 如果限制互动性来增加/减少某些音频对象的电平，则必须通过在整个混音中用这些音频对象的最大和最小电平设置来确认质量。

注4 – 如果场景中有许多对象，则每个对象所需的比特率都会降低，因为听者能够感知的数据总量是有限的。

注5 – 一些编码算法动态地在音频对象/声道之间分配比特率，并利用对象间/声道间的冗余来提高编码效率。原则上，此类编码器根据编码器输入端音频目标/声道之间的电平关系在重放期间保持不变的假设来分配比特率。然而，在基于对象的音频中，某些音频对象可以仅使用对话增强或使对话静音来收听，在这种情况下，可能会出现掩蔽消失并可能感知到编码损伤。

3.3 解码器复杂度

音频节目的解码器应该具有不过度高的复杂性，以便于解码器成本可以保持在较低的水平。在从包含 N 个声道的一个音频节目再生出较少的声道数 M 的情形中，解码器复杂度应该小于完整的 N 声道解码器的复杂度。

附件2的 附录1 (资料性)

有关已经被验证满足发射的质量及其他 用户需求的编码系统的信息

表7和8的左边一栏中分别列出了附件2中规定的高质量发射和中等质量发射的要求，仅用于基于声道的信号。其余各列（目前有4列）给出了满足这些要求的特定编解码器的能力。预计未来对此建议书的修订将包括有关其他编解码器的额外信息。

表9包含了对由基于声道的和基于对象的信号组成的配置的要求。对于后者，必要的比特率取决于内容和与用户有关的因素。

表7

高质量发射（仅包含基于声道信号的配置）

附件2的要求列表	AAC LC 配置文件 ⁽³⁾	带MPEG环绕的 AAC LC	AC-3/E- AC-3	MPEG-2 层II	AC-4 ⁽⁶⁾	MPEG-H LC和 BL 配置文件 ⁽¹⁰⁾	DTS-UHD ⁽⁹⁾
1.1.1 ITU-R BS.775规定的声道配置	满足	满足	满足	满足	满足	满足	满足
1.1.2 ITU-R BS.2051规定的、基于声道的高级音响系统的声道配置（默认支持）	系统C、H、I	系统C、H、I	N/A	N/A	系统C、D、G到J	系统C、D、F到J	系统C至 J
1.2 音频业务	满足	满足	满足	满足	满足	满足	满足
1.3 声道的灵活分配	满足	满足	满足	满足	满足	满足	满足
1.4 辅助数据	满足	满足	满足	满足	满足	满足	满足
1.5 高级音响系统的音频相关元数据	N/A	N/A	N/A	N/A	满足	满足	满足
2.1.1 基本音频质量	在每2声道 144 kbit/s时 满足[1]	在每5声道 384 kbit/s时满足 ⁽⁴⁾	在每2声道 192 kbit/s时 满足[1]	在每2声道 256 kbit/s时 满足[1]	在每声道 48 kbit/s[11]、每2 声道96 kbit/s、每5 声道192 kbit/s和每 11.1声道288 kbit/s （系统J）时满足 ⁽⁷⁾	在每声道 48 kbit/s[11]、每 22.2声道 768 kbit/s（系统 H）时满足[8] [10]	分别在每2、5和 11声道128、 192、288 kbit/s 时满足 ⁽⁸⁾
2.1.2 空间音频质量	满足	满足	满足	满足	满足	满足	满足
2.1.3 量化精度	满足	满足	满足	满足	满足	满足	满足
2.1.4 采样频率	满足	满足	满足	满足	满足	满足	满足
2.1.5 带宽	满足	满足	满足	满足	满足	满足	满足
2.1.6 加重	满足	满足	满足	满足	满足	满足	满足

表7（结束）

附件2的要求列表	AAC LC 配置文件 ⁽³⁾	带MPEG环绕的 AAC LC	AC-3/E- AC-3	MPEG-2 层II	AC-4 ⁽⁶⁾	MPEG-H LC和 BL 配置文件 ⁽¹⁰⁾	DTS-UHD ⁽⁹⁾
2.1.7 后处理	未证实	未证实	未证实	未证实	未证实	未证实	未证实
2.2 编码时延	满足 ⁽¹⁾	满足 ⁽¹⁾	满足 ⁽¹⁾	满足 ⁽¹⁾	满足	满足	满足
2.3 误码适应力	满足	满足	满足	满足 ⁽²⁾	满足	满足	满足
2.4 恢复时间	满足	满足	满足	满足	满足	满足	满足
3.1.1 向下兼容性	满足	满足	满足	满足	满足	满足	满足
3.1.2 后向兼容性	同时联播方法满足	通过设计或同时联播方法满足 ⁽⁵⁾	同时联播方法满足	矩阵化方法满足	同时联播方法满足	同时联播方法满足	同时联播方法满足
3.1.3 前向兼容性	双解码器满足	满足	双解码器满足	满足	双解码器满足	双解码器满足	同时联播方法满足
3.2 比特速率	满足	满足	满足	满足	满足	满足	满足
3.3 解码器复杂度	满足	满足	满足	满足	满足	满足	满足

⁽¹⁾ 固有的编码时延足够小，应用可以容易地匹配视频和音频时延。

⁽²⁾ 层II基本流中提供了某些误码适应力，其他的适应力典型地由应用提供。

⁽³⁾ AAC LC包含在扩展的HE AAC、HE AAC v2和HE AAC中。因此，所有这些AAC版本同时也满足了附件2的要求列表。

⁽⁴⁾ 384 kbit/s用于传统立体声AAC解码器作为2/0下降混合解码的多声道比特流。

⁽⁵⁾ 如前2声道服务使用AAC编码，通过设计满足该要求。如初始2声道使用其他编解码技术，则通过同时联播方法满足该要求。

⁽⁶⁾ AC-4核是由ETSI TS 103 190-1 v1.1.1（2015-06）定义的，这里使用的提供增强比特流的ETSI TS 103 190-2 v1.2.1（2015-09）对此进行标准参考。

⁽⁷⁾ 比特速率是根据支持方内部测试结果确定的。

⁽⁸⁾ 比特率基于尚未公布的第三方主观测试结果。

⁽⁹⁾ UHD DTS在ETSI TS 103 491中定义。

⁽¹⁰⁾ LC配置文件=低复杂度配置文件，BL配置文件 = 基线配置文件

表8

中等质量发射（仅包含基于声道信号的配置）

附件2 的要求列表	HE-AAC	带MPEG环绕的 HE-AAC	HE-AAC v2	扩展的HE-AAC	AC-4	MPEG-H LC和BL 配置文件	DTS-UHD
1.1.1 ITU-R BS.775规定的声道配置	满足	满足	满足	满足	满足	满足	满足
1.1.2 ITU-R BS.2051规定的、基于声道的高级音响系统的声道配置（默认支持）	系统C、H、I	系统C、H、I	系统C、H、I	系统C、H、I	系统C、D、G到J	系统C、D、F到J	系统C至J
1.2 音频业务	满足	满足	满足	满足	满足	满足	满足
1.3 声道的灵活分配	满足	满足	满足	满足	满足	满足	满足
1.4 辅助数据	满足	满足	满足	满足	满足	满足	满足
1.5 高级音响系统的音频相关元数据	N/A	N/A	N/A	N/A	满足	满足	满足
2.1.1 基本音频质量	在每2声道 48 kbit/s时 满足（优秀） [2] [4]； 在每2声道 32 kbit/s时满足 （良好）[2] [4]； 在每1声道 24 kbit/s 时满足 （良好）[3]	在每5声道 64 kbit/s时 满足（良好）[7]	在每2声道 24 kbit/s 时 满足（良好）[2]	在每2声道 16 kbit/s时 满足（良好） [5]； 在每1声道 12 kbit/s时 满足（良好）[5]	在每2声道48 kbit/s时 满足（优秀）[9]； 在每5.1声道128 kbit/s 时满足（优秀）[9]； 满足 在每11.1声道 256 kbit/s（系统J） 时满足（优秀）	在每2声道48 kbit/s 时满足（优秀） [8]；在每5.1声道 128 kbit/s时满足 （系统B）（优 秀）[8]	分别在每2、5 和11声道64、 144、 192 kbit/s时满 足（优秀）
2.1.2 空间音频质量	满足	满足	满足	满足	满足	满足	满足
2.1.3 量化精度	满足	满足	满足	满足	满足	满足	满足
2.1.4 采样频率	满足	满足	满足	满足	满足	满足	满足
2.1.5 带宽	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2.1.6 加重	满足	满足	满足	满足	满足	满足	满足
2.1.7 后处理	未验证	未验证	未验证	未验证	未证实	未证实	未证实

表8（结束）

附件2 的要求列表	HE-AAC	带MPEG环绕的 HE-AAC	HE-AAC v2	扩展的HE-AAC	AC-4	MPEG-H LC和BL 配置文件	DTS-UHD
2.2 编码时延	满足 ⁽¹⁾	满足 ⁽¹⁾	满足 ⁽¹⁾	满足 ⁽¹⁾	满足	满足	满足
2.3 误码适应力	满足	满足	满足	满足	满足	满足	满足
2.4 恢复时间	满足	满足	满足	满足	满足	满足	满足
3.1.1 向下兼容性	满足	满足	满足	满足	满足	满足	满足
3.1.2 后向兼容性	同时联播 方法满足	（从设计上）满 足	同时联播 方法满足	同时联播 方法满足	同时联播 方法满足	同时联播 方法满足	同时联播 方法满足
3.1.3 前向兼容性	双解码器满足	满足	双解码器 满足	双解码器 满足	双解码器 满足	双解码器 满足	双解码器 满足
3.2 比特速率	满足	满足	满足	满足	满足	满足	满足
3.3 解码器复杂度	满足	满足	满足	满足	满足	满足	满足

N/A: 不适用。

注 – ITU-R BS.1534建议书对“优秀”和“良好”的属性做出了定义。

⁽¹⁾ 固有的编码时延足够小，应用可以容易地匹配视频和音频时延。

表9提供的比特率说明见第3.2节。

表 9
发射（基于声道的和基于对象的信号的配置）

附件2 的要求列表	AC-4	MPEG-H LC和BL配置文件	DTS-UHD
1.1.1 ITU-R BS.775规定的声道配置	满足	满足	满足
1.1.2 ITU-R BS.2051规定的、基于声道的高级音响系统的声道配置（默认支持）	系统C、D、G至J	系统C、D、F至J	系统C至J
1.2 音频业务	满足	满足	满足
1.3 声道的灵活分配	满足	满足	满足
1.4 辅助数据	满足	满足	满足
1.5 高级音响系统的音频相关元数据	满足	满足	满足
2.1.1 基本音频质量（见第3.2节）	满足 配置的比特率可以根据表7中提供的声道比特率计算，加入对象（通常是对话），使每个对象的整体比特率增加48至64 kbit/s[11]。	满足 配置的比特率可以从表7中提供的声道比特率计算，通过添加对象将每个对象的总比特率增加48至64 kbit/s[11]。测试的具体配置可以在[8]和[10]中找到。	满足 可以根据表7中的声道比特率计算配置的比特率，附加对象将每个对象的总比特率增加64至96 kbit/s，由提议者估计 ⁽¹⁾ 。
2.1.2 空间音频质量	满足	满足	满足
2.1.3 量化精度	满足	满足	满足
2.1.4 采样频率	满足	满足	满足
2.1.5 带宽	满足	满足	满足
2.1.6 加重	满足	满足	满足
2.1.7 后处理	未证实	未证实	未证实
2.2 编码时延	满足	满足	满足

表 9（结束）

附件2 的要求列表	AC-4	MPEG-H LC和BL配置文件	DTS-UHD
2.3 误码适应力	满足	满足	满足
2.4 恢复时间	满足	满足	满足
3.1.1 向下兼容性	满足	满足	满足
3.1.2 后向兼容性	通过联播方式实现	通过联播方式实现	通过联播方式实现
3.1.3 前向兼容性	双解码器满足	双解码器满足	双解码器满足
3.2 比特速率	满足	满足	满足
3.3 解码器复杂度	满足	满足	满足

⁽¹⁾ 对象比特率由提议者根据表7中的双信道比特率来估计。

由于上述原因，为AdvSS中的对象指定最小比特率是不可行的。由第三方测试的内容规范和比特率给出了指导原则：

- 对于MPEG-H：验证报告[8]和[10]给出了比特率场景的示例。两份报告都是MPEG的公开输出文件，见：[https://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w16584_3D_Audio_Verification_Test_Report\).docx](https://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w16584_3D_Audio_Verification_Test_Report).docx) 和https://www.mpeg.org/wp-content/uploads/mpeg_meetings/131_OnLine/w19407.zip。对象是每份报告中测试1（UHD广播，总比特率768 kbit/s）和测试2（HD广播，测试总比特率512 kbit/s、384 kbit/s和256 kbit/s）中一些项目内容的一部分。每项详情列于[8]附件1。
- 对于AC-4：对于对象（特别是对话），可使用表7中提供的单声道比特率。因此，具有声道和附加对象的配置的比特率可以由表7中的声道配置的比特率加上每个附加对象的48至64 kbit/s来确定。
- 对于DTS-UHD：对于对象，可以使用单声道比特率，根据表7中提供的配置比特率计算得出。因此，在声道配置中增加对象将使每增加一个对象的比特率增加64至96 kbit/s。

参考文献

- [1] GRANT, D., DAVIDSON, G. and FIELDER, L. (21-24 September 2001) *Subjective evaluation of an audio distribution coding system*. 111th AES Convention, New York, NY, United States of America.
- [2] ISO/IEC JTC 1/SC 29/WG 11 N6009 (October 2003) *Report on the Verification Tests of MPEG-4 High Efficiency AAC*.
- [3] ISO/IEC JTC 1/SC 29/WG 11 N7137 (April 2005) *Listening test report on MPEG-4 High Efficiency AAC v2*.
- [4] KOMORI, T., SUGIMOTO, T., and KUROZUMI, K. (2005) *AAC + SBR Audio coding quality used for the mobile digital terrestrial broadcasting*. Proc. Spring meeting of the Acoustical Society of Japan.
- [5] ISO/IEC JTC 1/SC 29/WG11 N12232 [July 2011] *USAC Verification Test Report*.
- [6] HERRE, J., et al., (May 2007) *MPEG Surround – The ISO/MPEG Standard for Efficient and Compatible Multi-Channel Audio Coding*. 122nd AES Convention, Vienna, Austria.
- [7] Rödén, J., et al., (October 2007) *A study of the MPEG Surround quality versus bit-rate curve*. 123rd AES convention, New York, NY, United States of America.
- [8] ISO/IEC JTC1/SC29/WG11 N16584 (January 2017) *MPEG-H 3D Audio Verification Test Report*.
- [9] Riedmiller J., et al., (March 2017) *Delivering Scalable Audio Experiences using AC-4*, *IEEE Transactions on Broadcasting*, Vol. 63, No. 1.
- [10] ISO/IEC JTC1/SC29/WG11 N19497 (July 2020) *MPEG-H 3D Audio Baseline Profile Verification Test Report*.
- [11] Japanese Ministry of Internal Affairs and Communications, the Broadcasting System Subcommittee, (February 2022) *Comparative study of audio coding methods for advanced terrestrial digital broadcasting*, (in Japanese). https://www.soumu.go.jp/main_content/000795468.pdf

附件3

广播应用的音频质量等级

为广播应用设立了下列三个音频质量等级。

表10

等级	音频质量	应用
(1)	甚高质量，有足够的质量余量允许级联（串联）和后处理	接收、分发、制作，以及后制作
(2)	主观透明质量，足够满足最高质量广播	高质量（“CD质量”）发射
(3)	相当于或优于良好的FM业务质量，或者相当于或优于良好的AM业务质量	中等质量发射

附件4
(资料性)

与ITU-R BS.1548-7建议书相比的修改摘要

修订8的修改如下。

- 在附件1新的第1.4节和表4以及附件2新的第1.5节以及表7和8中增加了高级音响系统的音频相关元数据的要求。
- 在附件1第2.1.1节和附件2第2.1.1.1节中增加了基于对象的音频的音频质量注释。
- 在附件1第3.1节和附件2第3.2节中增加了有关基于对象的音频比特率的注释。
- 增加了新的表9，显示了满足高级音响系统要求的系统和配置。
