

ITU-R BS.1115-1 建议书^{*,**}

低比特率音频编码

(ITU-R 19/6 号研究课题)

(1994-2004)

国际电联无线电通信全会，

考虑到

- a) 通过诸如激光唱片 (CD)、磁带机和广播，用户可得到高质量的数字音频；
- b) 数字声音广播系统的基础包含在 ITU-R BS.774 和 ITU-R BO.789 建议书中；
- c) 数字广播中音频编码系统的用户要求在 ITU-R BS.1548 建议书中规定；
- d) 数字声音广播中要求的基本音频和立体声声像质量与 CD 的质量实际上难以区分；
- e) 在数字声音广播系统信号源的节目连接（馈送、分配和解说链路）中应用的低比特率编码系统的性能，必须做到能向数字声音广播发射机传输最高质量的音频信号；
- f) 在馈送和分配链路场合下，由于信号处理和级联余量以及（馈送链路场合下）过载余量或“峰值储备”的要求，这意味着其性能水平会超过 CD 的性能；
- g) 在解说链路场合下，这意味着其性能水平能向听众传输极好质量的话音，但是，当此类链路载送音乐节目内容时，要预想到会有减低的性能水平；
- h) 低比特率编码通常会带来几十毫秒的延时，对于诸如播出提示发令准备之类的操作实践来说，可能有一定影响；
- j) 国际电联无线电通信部门已经测试了高质量音频的低比特率编码，在许多应用场合中获得满意的性能（参见附录 6 和 ITU-R BS.1548 建议书）；
- k) 常规的客观测量方法（例如，测量信噪比和失真度）不再足以评定应用低比特率音频编码方式的系统质量；ITU-R BS.1387 建议书中给出的客观感知测量方法可能最终补充或取代常规的客观评定方法，并有可能最终补充主观评价方法。

* 根据 ITU-R 第 44 号决议，无线电通信第 6 研究组于 2003 年对本建议书做了编辑性修正。

** 应提请电信标准化第 9 和 15 研究组、国际标准化组织/国际电工委员会 (ISO/IEC) (JTC 1/SC 29/WG 11)、欧洲广播联盟 (EBU) 和音频工程协会 (AES) 注意本建议书。

建议

1 对于下面的建议 4 至 6 中列出的数字声音广播应用来说,应采用 ISO/IEC 11172-3 (1993) (见注 1) 和 ISO/IEC 13818-7 (2003) 或根据 ITU-R BS.1196 建议书附件 2 的 AC-3 等标准内说明的低比特率音频编码 (这些标准的简要说明在附录 1、3 和 4 内给出);

注 1 — ISO/IEC 11172-3 编解码器又往往是指 ISO/IEC 13818-3, 因为 ISO/IEC 13818-3 的编解码器中包括 ISO/IEC 11172-3 作为参考。

2 所有应用场合下, 取样频率应为 48 kHz 或 32 kHz;

3 低比特率音频编码器的输入信号应是无预加重的, 编码器中也不应施加预加重;

4 对于数字声音广播的发射 (见注 1), 应采用下面的编解码参数:

— ISO/IEC 11172-3 中的层 II 编码, 单声信号比特率为 128 kbit/s, 立体声信号为 256 kbit/s (也即 2×128 kbit/s), 立体声信号的左、右分量独立地编码 (见注 2);

— ISO/IEC 13818-7 编码中, 立体声信号的比特率至少 144 kbit/s,

— AC-3 编码中, 立体声信号的比特率至少为 192 kbit/s;

注 1 — ITU-R BS.1548 建议书的附件 2 的附录 1 中给出了关于编码系统的信息, 它表明能满足用于发射的质量要求和其他用户要求。

注 2 — 可以预期, 广播商希望在发射的信号内有若干容量供附属数据应用。不过, 它们应知道, 降低音频比特率一般会影响音频质量。

5 对于分配和馈送链路, 应采用 ISO/IEC 11172-3 中的层 II 编码, 每路音频信号 (也即每个单声信号或独立编码的立体声信号中的每个分量信号) 的比特率至少为 180 kbit/s, 附属数据除外 (参见附录 2 和注 2), 但是, 在不存在进一步级联处理的单条分配链路的特殊场合下 (见注 3), 每路音频信号可采用至少 120 kbit/s 的比特率。

注 1 — ITU-R BS.1548 建议书的附件 1 的附录 1 中给出了关于编码系统的信息, 它表明能满足用于馈送和分配的质量要求和其他用户要求。

注 2 — 应用本建议书时应考虑到, 某些电信网当前不容许 64 kbit/s 信道的无限制应用。

注 3 — 当例如是通过单条分配链路对模拟发射机传输时, 或者将录音室内编码到最后发射标准上的信号传送给发射机发射时, 不存在任何进一步的信号解码和编码处理。

6 对于解说链路, 应采用 ISO/IEC 11172-3 中的层 III 编码, 附属数据除外的单声信号的比特率至少为 60 kbit/s (参见注 1), 附属数据除外的、应用联合立体声编码的立体声信号, 比特率至少为 120 kbit/s (见附录 2 和注 2)。

注 1 — 1993 年, 使用头戴耳机进行收听以评价 60 kbit/s 层 III 的单声解说编解码器性能测试中获得的结果指明, 它不能满足语音信号的需求质量。在解说编解码器能达到性能要求之前, 本注释在本建议书中将保留。

注 2 — 应用本建议书时应考虑到，某些电信网当前不容许 64 kbit/s 信道的无限制应用。

7 在编码器和解码器的实施中，延时应做到最小化。

注 1 — ISO/IEC 11172-3 (MPEG-1 音频) 和 ISO/IEC 13818-7 (MPEG-2 AAC) 标准的电子版，可从下列网址上得到：<http://www.iso.org/itu>。

附录 1

ISO/IEC 11172-3 编码方法说明 在最高约为 1.5 Mbit/s 的编码率上用于数字存储媒质的 运动图像和相关音频的编码 — 第 3 部分：音频

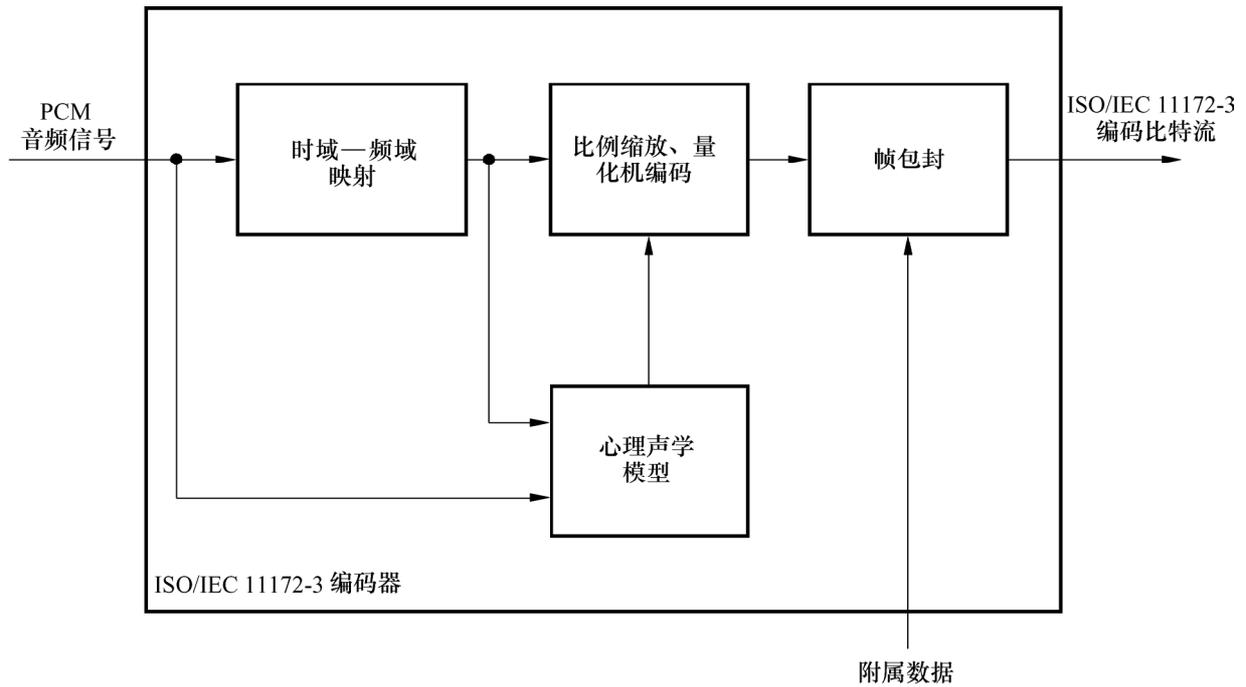
1 编码

编码器对数字音频信号进行处理，产生压缩比特流。编码器算法并无标准化的规定，可采用各种编码手段，诸如听觉掩蔽门限估计、量化和比例缩放（注 1）。编码器输出必须使符合于本建议书规定的解码器能解码给出适合于预定应用的音频信号。

注 1 — 遵循 ISO/IEC 11172-3 (1993) 附体 C 和 D 中技术说明的编码器，能满足最低的性能标准。

下面，说明如图 1 中所示的典型编码器。输入的音频样本传送到编码器上，由时域到频域的映射形成经滤波和亚取样的输入音频流信号。映射后的样本可以是子带样本（如层 I 和层 III 中，参见下述），或者是变换后的子带样本（如层 III 中）。应用与音频信号的时域到频域映射并行运算的快速傅里叶变换，由心理声学模型产生一个数据集以控制量化和编码。取决于实际编码器的实施，这些数据是不相同的。一种可能的方法是应用掩模门限的估计来控制量化器。由“比例缩放、量化和编码”框从映射的输入样本中产生一个编码符号集合。另外，该框的传递函数取决于编码系统的实施。“帧包封”框使来自其他框的输出数据（例如，比特分配数据、比例因子和编码的子带样本）中用于所选定层的实际比特流组装一起，并当需要时在附属数据字段内加上其他信息（例如，误码纠错）。

图 1
典型编码器方框图



1115-01

2 分层

依据应用场合，可以采用具有增大复杂性和增高性能的不同编码系统层级。

层 I：层 I 中包含：将数字音频输入映射入基本的 32 个子带；固定的分段使数据格式化成为数据块；用心理声学模型确定自适应的比例分配以及利用数据块的压扩和格式化进行量化。一个层 I 帧表征每信道 384 个样本。

层 II：层 II 中对比特分配、比例因子和样本提供附加的编码。一个层 II 帧表征每信道 $3 \times 384 = 1\,152$ 个样本。

层 III：层 III 中基于混合滤波器组（带有可变长度修正离散余弦变换的 32 子带滤波器组）引入增大的频率分辨率。它增加了非均匀量化器、自适应分段和量化估熵编码等处理。一个层 III 帧表征每信道 1 152 个样本。

任一层可以有四种不同的模式：

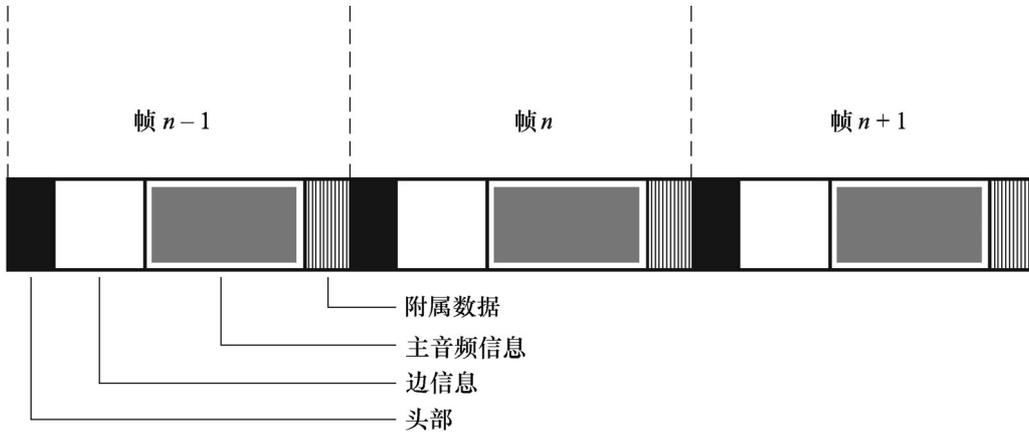
- 单声道；
- 双声道（两个独立的音频信号在一个比特流内编码，例如双语种场合）；
- 立体声（立体声对的左、右信号在一个比特流内编码）；以及，
- 联合立体声（立体声对的左、右信号在一个比特流内编码，带有开发的立体声不相关性处理和冗余处理）。可利用联合立体声模式以使在低比特率上提高音频质量和/或降低立体声信号的比特率。

3 编码比特流格式

ISO/IEC 11172-3 层 II 比特流的概况如图 2 所示，层 III 比特流的概况如图 3 所示。编码比特流内连续帧构成。依据层级，一个帧内包括有如下字段：

图 2

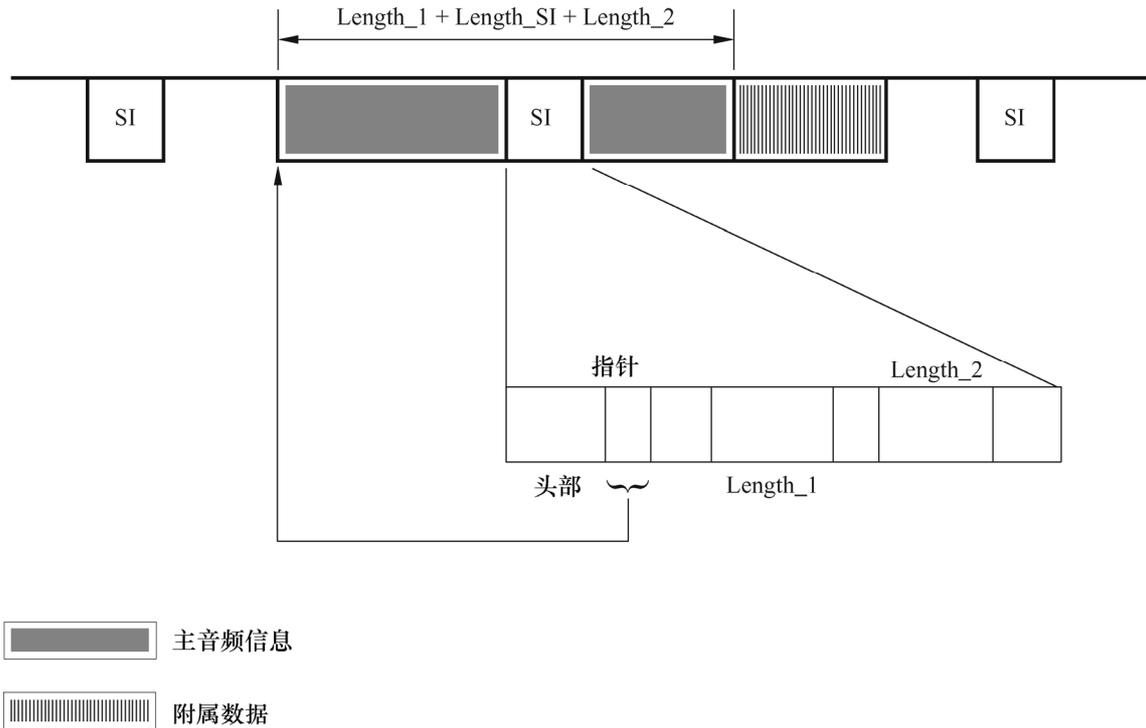
ISO/IEC 11172-3 层 II 比特流格式



层 II:

- 头部: 比特流内包含同步和状态信息的部分
- 边信息: 比特流内包含比特分配和比例因子信息的部分
- 主音频信息: 比特流内包含编码的子带样本的部分
- 附属数据: 比特流内包含用户可定义数据的部分

图 3
ISO/IEC 11172-3 层III比特流格式



层 III:

- 边信息(SI): 比特流中包含首标、指针、length_1 和 length_2 及比例因子信息等的部分；
- 头部: 比特流中包含同步和状态信息的部分；
- 指针: 指明主音频信息的开始；
- Length_1: 主音频信息第一部分的长度；
- Length_2: 主音频信息第二部分的长度；
- 主音频信息: 比特流中包含编码音频的部分；
- 附属数据: 比特流中包含用户可定义数据的部分。

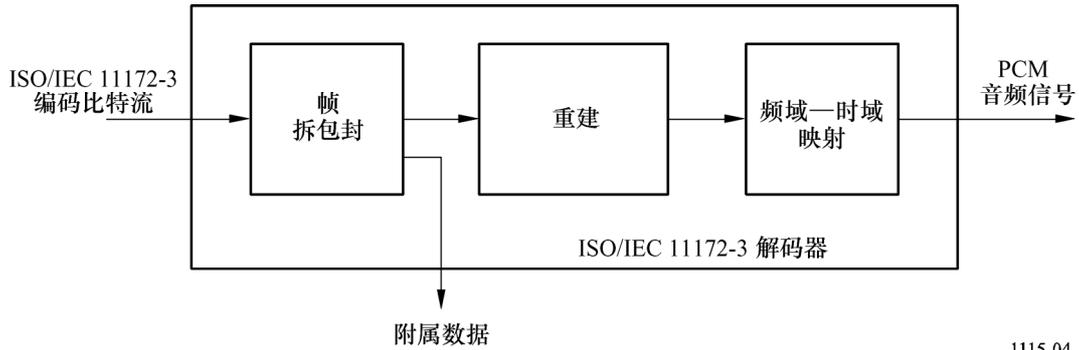
1115-03

4 解码

解码器接收在 ISO/IEC 11172-3 中定义的句法内的编码音频比特流，解码出数据元素，并应用数据信息产生数字音频输出。

如图 4 所示，编码的音频比特流传送入解码器。如果在编码器内加入有误差校验，则比特流拆包封和解码处理时相应实施误差检测。对比特流拆包封以恢复各个信息段，诸如音频帧音标、比特分配、比例因子、映射样本和可选的附属数据等信息。重建处理中，重建映射样本集合中的量化值。频域到时域的映射使那些映射样本变换回线性 PCM 音频样本。

图 4
解码器方框图



1115-04

附录 2

在特定比特率上的 ISO/IEC 11172-3 编解码器应用指南

本附录预定提供某些应用指南，它们对于不是明确地工作于 ISO/IEC 标准内比特率表（表 1）中所列值的特定比特率上的 ISO/IEC 11172-3 编解码器，可能是必需的。考虑到传输通道的可用比特率和该建议书中给出的最小合适比特率时，可能必须选择这些在表 1 中没有的特定比特率。

从 ITU-T J.52 建议书内关于“每路单声信号应用 1、2 或 3 个 64 kbit/s 信道（以及每路立体声最多 6 个 64 kbit/s 信道）的高质量声音节目信号数字传输”的内容中，可得到更完整的说明。

ISO/IEC 标准内明确地列出下列比特率值：

表 1

层 II (kbit/s)	层 III (kbit/s)
32	32
48	40
56	48
64	56
80	64
96	80
112	96
128	112
160	128
192	160
224	192
256	224
320	256
384	320

表 1 内的数字指明的是不论何种工作模式下的每一音频节目的总比特率。有一种附加的格式即自由格式，它可以应用于在表 1 内没有的固定比特率上，对于层 II 最大为 384 kbit/s，对于层 III 最大为 320 kbit/s。

ISO/IEC 11172-3 中，层 II 或层 III 音频帧总是对应于每信道 1152 个音频 PCM 输入样本。一帧的时间长度在 48 kHz 取样频率下为 24 ms，32 kHz 取样频率下为 36 ms。一帧总是由整数个字节构成，字节数的计算是对帧的时长乘以比特率再除以 8。比特率和取样频率在音频帧始端的头部内指明。

借助于下面三种方法可采用在表 1 中未列出的音频比特率进行工作。

1 自由格式

自由格式情况必须在头部内指明。此种场合下，解码器在启动时计量相继的同步字之间的距离，然后，像在规定的比特率场合下那样应用惯性同步程序。可以按上面给出的方法计算以字节为单位的帧长。编码器必须顾及该帧长。

例如，如果在 48 kHz 取样频率下希望立体声节目的比特率为 240 kbit/s，则帧长将为 $0.024 \times 240\,000 / 8 = 720$ 字节。如果希望的比特率值造成一帧内的字节数不是整数，则借助于采用数据填充能实现需求的比特率。这方面的详细说明在 ISO/IEC 11172-3 中第 2.4.2.3 节内给出。

本方法可应用于层 II 和层 III 两者上。

2 应用附属数据字段（适合于层 II）

可以选择表 1 中列出的（高于或等于所需比特率）的一个比特率。通过约束比特率分配，编码器可设置成保留一定量的比特用于附属数据，使得实际音频比特率等于所需比特率。在层 II 中，附属数据位于每个音频帧的最末端。

传输前，附属数据可以去除或者予以忽略。如果去除这些附属数据，则在解码前必须重新插入相应数目的填充比特。

例如，如果在 48 kHz 取样频率下希望立体声节目的比特率为 240 kbit/s，则可选择 256 kbit/s 的比特率，并按照 ISO/IEC 标准在帧头中予以指明。编码器必须设置成每帧保留 48 字节用于附属数据，这对应于 16 kbit/s 的附属数据传输能力。

3 动态比特率切换（适合于层 III）

层 III 中，可应用动态比特率切换来得到表 1 内不存在的比特率。例如，如果要求 120 kbit/s 的比特率，则层 III 编码器可以工作在 ISO/IEC 标准中明确规定的 128 kbit/s 和 112 kbit/s 比特率两种模式下，以交替序列选择它们。工作于此种模式上的编码器必须设置于从那些交替比特率的平均值中导出的一个比特分配数上。层 III 的比特存储技术能保持一种恒定比特率供编码处理应用。此种方法中，同步字之间以比特为单位的距离在两个不同值之间交替。

附录 3

ISO/IEC 13818-7 编码方法说明
运动图像和相关音频信息的通用编码 — 第 7 部分：先进音频编码（AAC）

参见 ISO/IEC 13818-7。

1 引言

ISO/IEC 13818-7 说明了被称为 MPEG-2 先进音频编码（AAC）的 MPEG-2 音频非后向兼容标准，它是一种与要求 MPEG-1 后向兼容相比能达到更高质量的多声道标准。

为了能够在所需存储器和处理能力以及音频质量之间做到兼顾，AAC 系统由三种档型组成：

— 主型

在任一给定的数据率上，主型提供最高的音频质量。可应用除增益控制之外的所有工具以给出高的音频质量。与 LC 型相比，它需要的存储器和处理能力更高。主型解码器能解码 LC 型编码的比特流。

— 低复杂度（LC）型

LC 型所需的处理能力和存储器小于主型，但质量性能可保持较高。LC 型中没有预测器和增益控制工具，但具有阶次有限的时间噪声成形（TNS）。

— 取样频率可缩放（SSR）型

SSR 型中借助增益控制工具能提供取样频率可缩放的信号。它能选择频带进行解码，因而解码器需要较少的硬件。例如，在 48 kHz 取样频率下为了只解码最低的频带，解码器能够以最小的解码复杂度重现 6 kHz 带宽的音频信号。

AAC 系统能支持如表 2 中所示的从 8 到 96 kHz 取样频率范围内的 12 种类型，并可高达 48 路声道。表 3 示明默认的声道配置，它们包括单声道、双声道、5 声道（3 个前方声道/2 个背后声道）和 5 声道加低音效果（LFE）声道（带宽 < 200 Hz）的 5.1 声道等。除默认配置外，还可能在每个位置（前方、侧面和背后）处规定扬声器的数目，以做到灵活的多声道扬声器安排。还支持缩混能力，用户能将多声道音频信号缩混到指定缩混系数的两声道上。所以，能够控制仅应用两声道的重放装置的声音质量。

表 2
支持的取样频率

取样频率 (Hz)
96 000
88 200
64 000
48 000
44 100
32 000
24 000
22 050
16 000
12 000
11 025
8 000

表 3
默认声道配置

扬声器数	接收指令中列出的音频句法元素	映射的扬声器默认元素
1	single_channel_element	前置中央扬声器
2	channel_pair_element	前置左、右扬声器
3	single_channel_element()	前置中央扬声器
	channel_pair_element()	前置左、右扬声器
4	single_channel_element()	前置中央扬声器
	channel_pair_element(),	前置左、右扬声器
	single_channel_element()	后置环绕扬声器
5	single_channel_element()	前置中央扬声器
	channel_pair_element()	前置左、右扬声器
	channel_pair_element()	后置左环绕、右环绕扬声器
5+1	single_channel_element()	前置中央扬声器
	channel_pair_element()	前置左、右扬声器
	channel_pair_element()	后置左环绕、右环绕扬声器
	Lfe_element()	低音效果扬声器
7+1	single_channel_element()	前置中央扬声器
	channel_pair_element(),	前置左、右扬声器
	channel_pair_element()	前置外侧左、右扬声器
	channel_pair_element()	后置左环绕、右环绕扬声器
	lfe_element()	低音效果扬声器

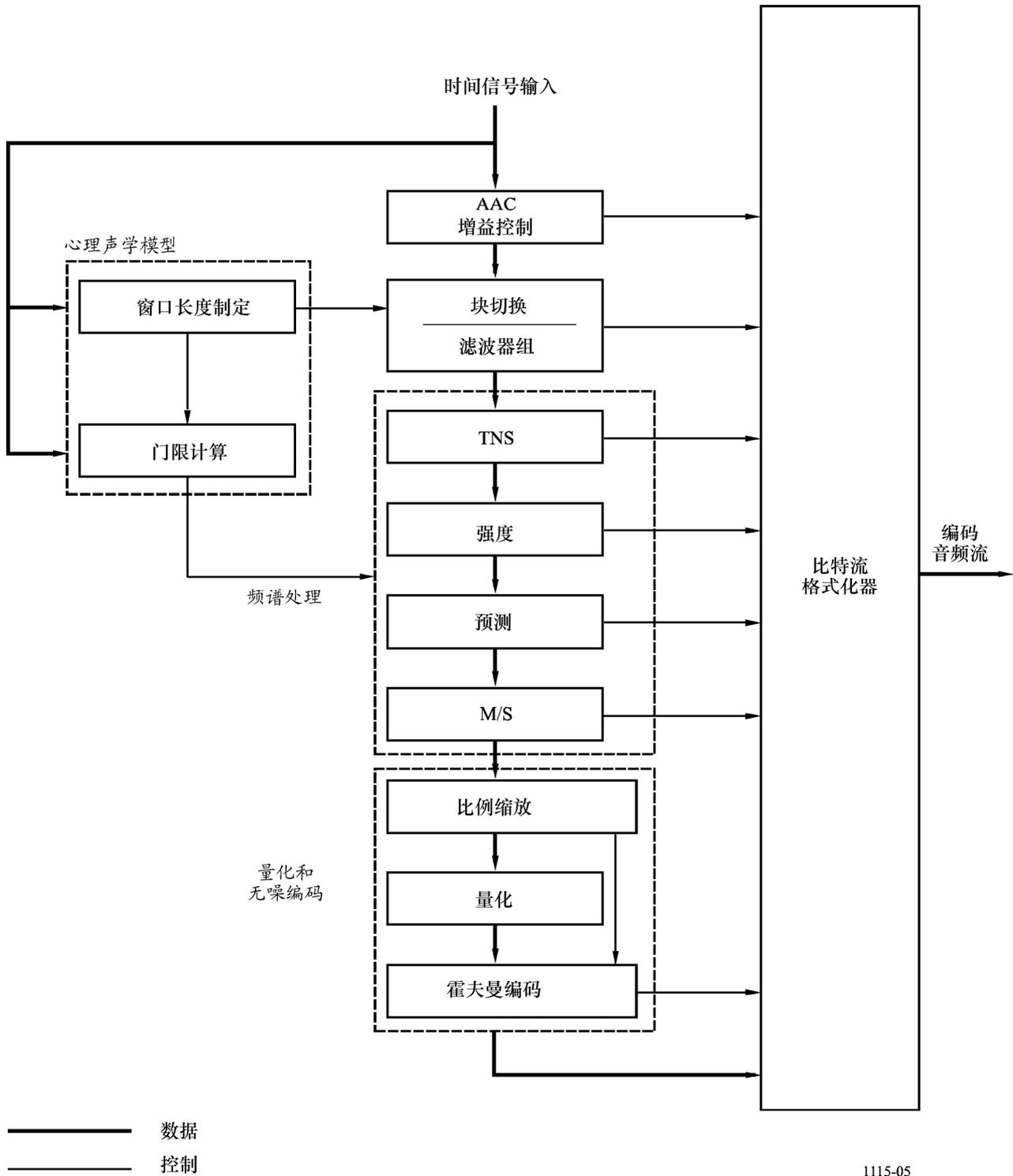
2 编码

MPEG-2 AAC 编码器的基本结构如图 5 所示。AAC 系统由下面的编码工具构成：

- 增益控制：增益控制中，将输入信号分割成 4 个等间隔的频带。增益控制应用于 SSR 型中。
- 滤波器组：修正的离散余弦变换（MDCT）滤波器组将输入信号分解成亚取样的频谱分量，48 kHz 取样时，频率分辨率为 23 Hz，时间分辨率为 21.3 ms（128 个频谱分量），或者频率分辨率为 187 Hz、时间分辨率为 2.6 ms（1024 个频谱分量）。窗口形状在两个可选的窗口形状之间选择。
- 时间噪声成形（TNS）：经滤波器组解析之后，实施 TNS 运算。TNS 技术能使编码器对量化噪声在时间上做出精细结构控制。
- 中间/边侧（M/S）立体声编码和强度立体声编码：对于多声道音频信号，可应用强度立体声编码和 M/S 立体声编码。强度立体声编码中，只传输声能包络以减少传输方向信息。M/S 立体声编码中，采用传输归一化的和信号 M（M 代表中间）及差信号 S（S 代表边侧）以取代传输原来的左与右信号。
- 预测：为降低静止信号的冗余，在相继帧的亚取样频谱分量之间实施时域预测。
- 量化和无噪编码：在量化工具中，应用步级大小为 1.5 dB 的非均匀量化器。对于量化的频谱，不同的比例因子和方向信息，实施霍夫曼编码。
- 比特流格式化器：最后，应用比特流格式化器来复用比特流，其中包括量化和编码的频谱系数以及来自每一工具的某些附加信息。
- 心理声学模型：应用心理声学模型计算输入信号内当前的掩模门限。可采用类似于 ISO/IEC 11172-3 中心理声学模型 2 的心理声学模型。在量化处理中，应用从掩模门限和输入信号电平内导出的信号—掩模比，以使得可闻的量化噪声最小，并附加地借以选择适当的编码工具。

图 5

MPEG-2 AAC 编码器方框图



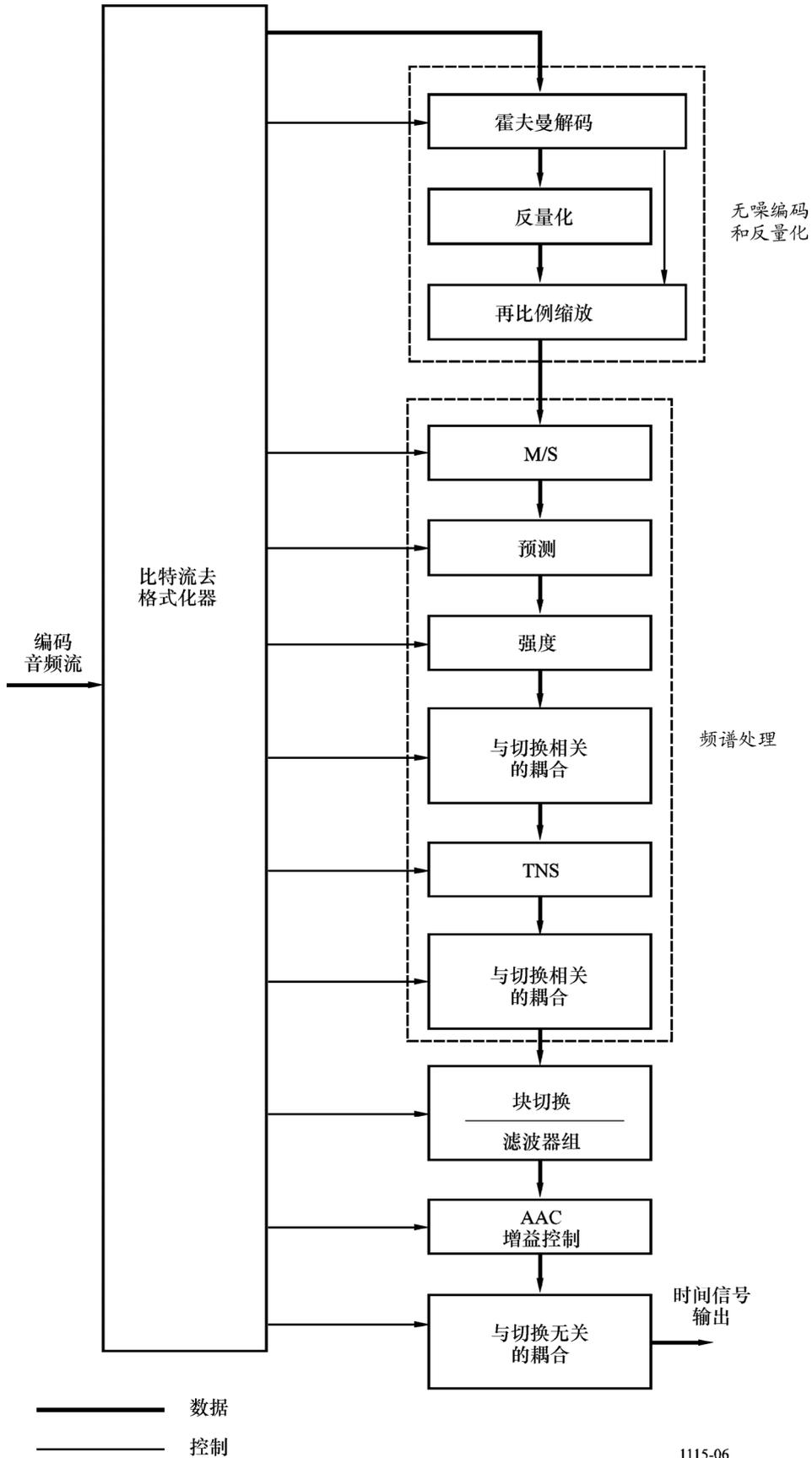
3 解码

MPEG-2 AAC 解码器的基本结构如图 6 所示。解码处理基本上是编码处理的逆过程。

解码器的功能是寻找出比特流中对量化音频频谱的描述，解码其量化值和其他重建信息，重建量化的频谱，借助于比特流中凡是起作用的工具对重建的频谱进行处理，以便得到输入比特流中描述的实际信号频谱，并在最后将频域内频谱变换到时域信号上，处理中可以用或不用可选的增益控制工具。在初始频谱重建和比例缩放的频谱重建之后，有许多供选用的工具可修正一个或多个频谱，以便提供更高效率的编码结果。对于工作于频域内的每一个可选工具，保留其“旁通”选项，在省略频谱运算的所有场合下，可使输入端上的频谱不做修正地直接通过可选工具。

图 6

MPEG-2 AAC 解码器方框图



附录 4

AC-3 编码方法说明

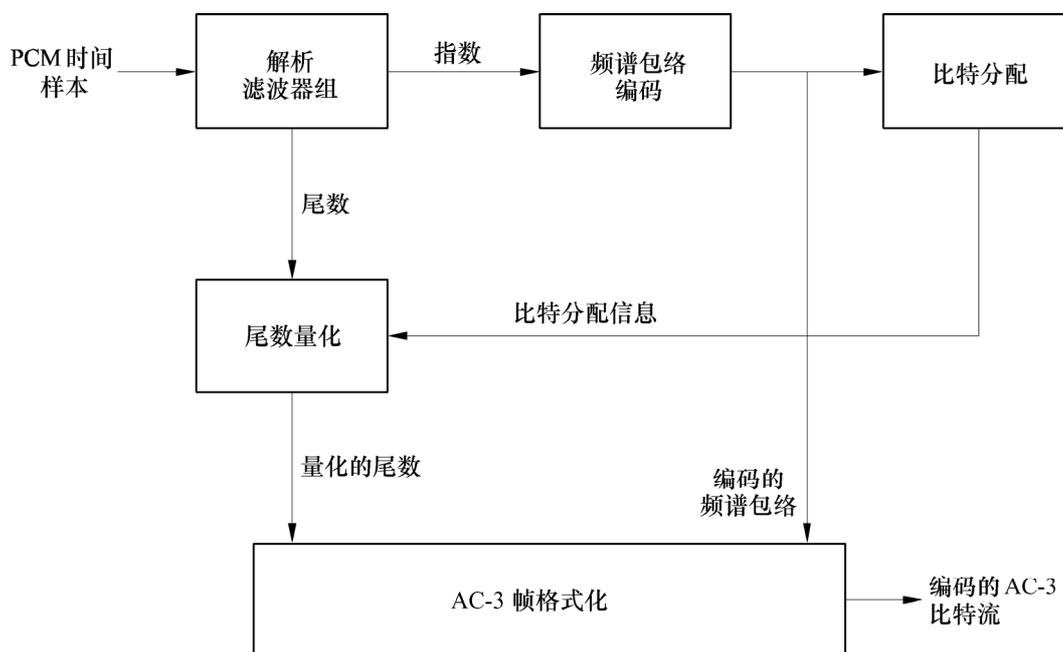
参见 ITU-R BS.1196 建议书的附件 2 和 ATSC 标准 A/53B 的附件 B。

1 编码

AC-3 数字压缩算法能够将来自 PCM 音频信号的 1 至 5.1 声道的音频源编码成数据率范围为 32 kbit/s 至 640 kbit/s 的串行比特流。借助于对音频信号的频域表示做出粗略量化，AC-3 算法能给出高的编码效率（输入比特率与输出比特率之比）。AC-3 编码处理方框图如图 7 所示。编码处理的第一步是将来自 PCM 时间样本序列的音频信号变换成频率系数的数据块序列，由解析滤波器组完成这一变换。对 512 个时间样本的重叠数据块乘以一个时间窗口，并变换到频域上。由于数据块重叠，每个 PCM 输入样本在两个相继的变换块中得到表示。于是，频域内的表示将除以 2，因此，每个数据块包含 256 个频率系数。各个频率系数均以二进制指数标记法表示成二进制的指数和一个尾数。指数集合编码成信号频谱的粗略表示，它被称为频谱包络。频谱包络供核心比特分配程序应用，它决定以多少比特位来对每一个别的尾数进行编码。将 6 个音频数据块（1 536 个音频样本）的频谱包络和粗略量化的尾数格式化成 AC-3 帧。AC-3 比特流是由 AC-3 帧构成的一个序列。

图 7

AC-3 编码器



实际的 AC-3 编码器要比图 7 中示出的复杂，实际编码器中还包括上面未示出的下列功能：

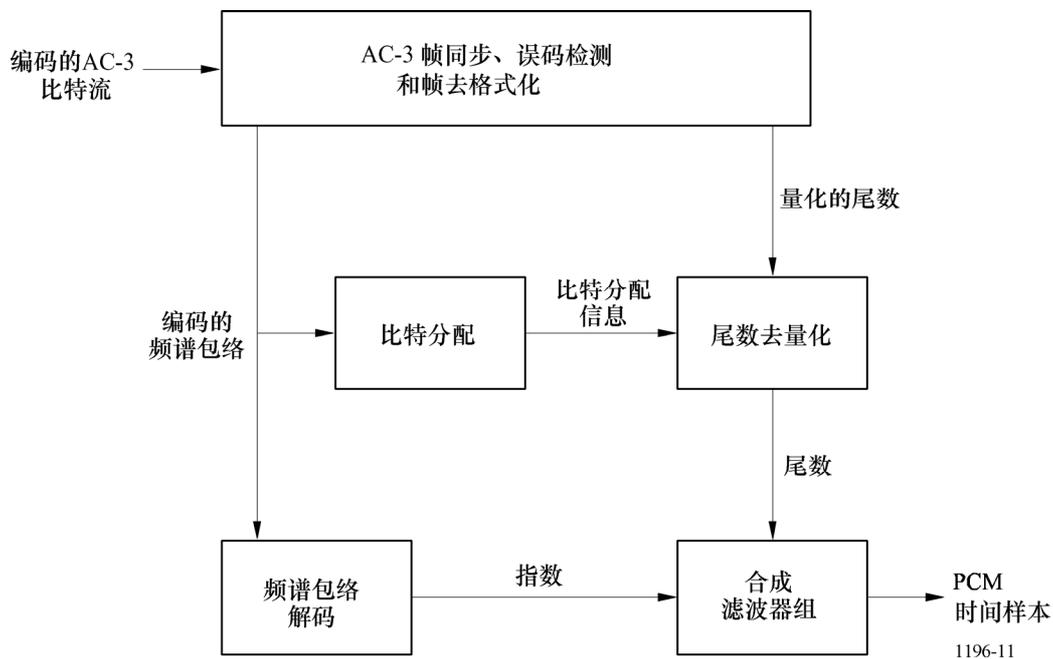
- 附上其中包含当同步和解码已编码 AC-3 比特流时所需信息（比特率、取样频率和编码声道数等）的帧头；
- 插入误码检测码，以使解码器能验明接收的数据帧无误码；
- 解析滤波器组的频谱分辨率可动态地改变，以便能更好地适配每个音频块的时间/频率特性；
- 频谱包络可以用可变的时间/频率分辨率进行编码；
- 用于实现较复杂的比特分配和修正核心比特分配程序的参数，借以得到更优化的比特分配；
- 在高频范围内各声道可以耦合在一起，以便在较低比特率上运行时能达到较高的编码增益；
- 双声道模式中，可以选择性地实施矩阵处理，以便提供附加的编码增益，并在用矩阵环绕声解码器解码双声道信号的情况下能得到改善的结果。

2 解码

解码处理基本上是编码处理的逆过程。图 8 中示出的 AC-3 解码器必须与编码比特流同步，校验误码以及对各种类型的数据诸如编码的频谱包络和量化的尾数实施去格式化。运行比特分配程序，并应用其结果进行拆包封和尾数去量化。解码频谱包络以得到指数。将指数和尾数变换回时域内，产生解码的 PCM 时间信号。

图 8

AC-3 解码器



实际的 AC-3 解码器要比图 8 中示出的复杂，实际解码器中还包括上面未示出的下列功能：

- 在检测到数据差错的场合下，可采用误码隐匿或误码抹去处理；
- 对于它们的高音频内容耦合在一起的各声道，必须进行去耦合。
- 对于已实施再矩阵处理的声道，必须进行解矩阵（在双声道模式中）；
- 编码处理中已使得解析滤波器组的频谱分辨率动态改变时，解码时合成滤波器组的频谱分辨率必须以同样的方式动态地改变。

附 录 5

质量感知的客观方法

已经在实验上应用质量感知的客观方法来测量编解码器性能，某些公布的结果表明，它们与主观评价测试结果有良好的关联性。

在国际电联无线电通信部门内，实验上应用的有下面的方法：

- 噪声—掩模比法（NMR）；
- 感知估值模型法（PERCEVAL）；
- 感知音频质量度量法（PAQM）；
- 质量感知客观模型法（POM 620）。

ITU-R BS.1387 建议书中规定了一种感知音频质量的客观测量方法。

附 录 6

1992-1993 年期间在测试实施中建议的 低比特率音频编解码器选择方法

对于各种单声道和常规的双声道应用实施了主观评价测试，并将所得结果与质量要求进行比较。在多声道编码系统方面，开展着进一步的研究。

在发射、分配和馈送三种应用场合下，经编解码器处理而在解码后重现的基本音频质量，必须相当于激光唱片 CD 的质量，也即对于大多数音频节目素材类型来说，主观上难以区分它们的质量差异。应该保持立体声声像的质量。

对于最苛刻的节目素材的发射，编解码器必须做到，音频质量下降是“可感知但不讨厌”的（4分等级）。（在进行主观评价测试时，采用5级损伤评分制（ITU-R BS.562建议书）。）于1992年进行的主观评价测试中，下面的音频系统能满足这些要求：

2×128 kbit/s 的独立声道（参见注1）：

- ISO/IEC 11172-3 层II；
- ISO/IEC 11172-3 层III；
- 杜比 AC-2。

192 kbit/s 的联合立体声：

- ISO/IEC 11172-3 层II；
- ISO/IEC 11172-3 层III。

对于分配来说，信号通过每个独立声道为120 kbit/s的三套编解码器级联处理后，评分等级与原来的声源信号相比不应低0.5分以上（注1）。为此，通过在线性PCM域内降低0.1 dB增益，以对二进制代码再确定其总数。ISO/IEC 11172-3的层II编解码器是能满足要求的惟一系统，所以，作为本建议书的基础提供该场合应用。

注1 — 在此120 kbit/s比特率上，对于大多数严格的条项来说，这类编解码器的评分为4.5分或更高些。当测试方法和编解码器继续改进时，将来的测试会给出不同的评分。

对于馈送来说，信号通过每个独立声道为180 kbit/s的五套编解码器级联处理后，主观感知应与原来的声源信号难以区分质量。实施级联时，应用18比特分辨率的线性PCM。为此，通过在线性PCM域内降低0.1 dB增益，对二进制代码再确定其总数。ISO/IEC 11172-3的层II编解码器是能够满足要求的惟一系统，所以，作为本建议书的基础提供该场合应用。

对于解说链路来说，解码后重现的话音信号的基本音频质量应相当于14比特线性PCM原始的音频质量。实际测试中，采用16比特线性PCM格式，感知的损伤程度与激光唱片的音乐序列播放相比较是可容忍的。1992年，只采用扬声器进行的测试中，ISO/IEC 11172-3层III的单一套编解码器能满意地实现话音重现，而按一般水平，层III编解码器最适合于音乐信号。关于语音信号，在单声（60 kbit/s）和立体声（120 kbit/s 联合立体声）两种模式中，评分等级总高于4.0分（4.0分对应于能感知损伤但不讨厌）。所以，层III编解码器作为本建议书的基础提供该场合应用。在1993年随后的测试中，采用头戴耳机收听以评定60 kbit/s层III单声解说编解码器的性能，得到的结果指出，评分等级低于4.0分，也即它不满足话音信号所需的质量。需要进一步对60 kbit/s进行研究。

复杂性评估表明，ISO/IEC 11172-3层II解码器和杜比AC-2解码器被证明是复杂性最低的。应从下列方面考虑解码器：

- 解码器复杂程度低；
- 对分配和馈送两种应用具有通用性；
- 灵活适应于编码器的改进；
- 可在两种比特率上测试编解码器。

对于在 2×128 kbit/s 上的发射应用, ISO/IEC 11172-3 层 II 格式及其解码器被选择作为本建议书的基础。于 1992 年进行的测试中, ISO/IEC 11172-3 层 II 在 192 kbit/s 上只是勉强地满足要求, 于是决定, 在该比特率上需要改进, 以便确实地满足基本的音频质量要求 (如同在 ISO/IEC 11172-3 层 III 上所做的那样)。1993 年进行的进一步测试, 并不显出在该时段内有改善。需要进一步研究在 192 kbit/s 上开发联合立体声模式, 以便再考虑将该比特率应用于此。征求各主管部门对此课题提出文稿。

1993 年, 对完整的广播链路进行了网络验证测试, 网络中包括 5 套级联的 180 kbit/s 馈送用编解码器、3 套级联的 120 kbit/s 分配用编解码器和一套发射用编解码器。这些测试的结果表明, 在此种链路的极端情况下, 不能满足基本的音频质量要求。

还进行了 8 套级联的 180 kbit/s 编解码器测试, 对于保持可接受的质量来说, 这证明是一种能够令人满意的配置。
