

团 体 标 准

T/CAIACN 013—2024

低延迟低复杂度高清音频 编解码蓝牙传输适配规范

Technology Specification for transmission Low Latency
Low Complexity High Resolution Audio Codec via Bluetooth

2024 - 10 - 18 发布

2024 - 12 - 18 实施

中国电子音响行业协会 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.....	2
5 技术要求.....	3
5.1 系统技术要求.....	3
5.2 源端技术要求.....	3
5.3 宿端技术要求.....	4
6 L2HC 蓝牙传输能力配置.....	5
6.1 蓝牙音频能力协商与建链过程.....	5
6.2 编解码器规格信息.....	6
6.3 蓝牙音频能力协商与建链过程.....	9
6.4 L2HC 位流数据.....	9
6.5 边信息解码.....	10
6.6 位流格式.....	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。

中国电子音响行业协会（China Audio Industry Association，简称CAIA）自1983年成立以来就以“服务企业，献策政府”为宗旨。是我国最早成立的跨地区、跨部门、跨系统，具有社团法人资格的全国性社会团体（国家一级行业协会）。

组织开展电子音响领域国际、国内标准化活动，制定中国电子音响行业协会团体标准（以下简称：中音协团标），满足行业需要，推动行业标准化工作，是中国电子音响行业协会的重要工作。协会的所有会员，均有权利提出制、修订中音协团标的建议并参与有关工作。

中音协团标按《中国电子音响行业协会团体标准建设管理办法》进行制定和管理。

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。

在本标准实施过程中，如发现需要修改或补充之处，请将意见和有关资料报送中国电子音响行业协会，以便修订时参考。

本文件由中国电子音响行业协会、中国电子技术标准化研究院、华为终端有限公司提出。

本文件由中国电子音响行业协会归口。

本文件起草单位：中国电子技术标准化研究院、华为终端有限公司、东莞市漫步者科技有限公司、深圳市云动创想科技有限公司、深圳市冠旭电子股份有限公司、瀛通通讯股份有限公司、深圳瑞利声学技术股份有限公司、深圳市盛佳丽电子有限公司、浙江声研科技有限公司、上海市浦东新区先进音视频技术协会。

本文件主要起草人：董桂官、张帅超、肖锴、曾泽雄、曾德钧、彭久高、涂长青、李德涛、盛成龙、陈荣、卫文港、唐铖、张波、黄晖、张洋、姚赞。

本文件为首次制定。

低延迟低复杂度高清音频编解码蓝牙传输适配规范

1 范围

本文件规定了低延迟低复杂度高清音频编解码作为第三方供应商编解码（Vendor Codec）适配蓝牙无线信道和传输协议，实现无线音频能力协商和传输的过程。

本文件适用于无线音频、实时通信、广播流媒体、网络电视、虚拟现实和增强现实、视频监控等领域。

2 规范性引用文件

GB/T 9002—2017 音频、视频和视听设备及系统词汇

GB/T 5271.1 信息技术 词汇 第1部分：基本术语（GB/T 5271.1—2000 eqv ISO/IEC 2382-1: 1993）

GB/T 5271.4 信息技术 词汇 第4部分：数据的组织（GB/T 5271.4—2000 eqv ISO/IEC 2382-4: 1987）

GB/T 5271.9 信息技术 词汇 第9部分：数据通信（GB/T 5271.9—2001 eqv ISO/IEC 2382-9: 1995）

T/CAIACN 009—2023 低延迟低复杂度高清音频编解码技术规范

Bluetooth SIG Specification. Bluetooth Core Specification Version 5.2 or later.

Bluetooth SIG Specification. Advanced Audio Distribution Profile 1.4 or later.

Bluetooth SIG Specification. Audio/Video Distribution Transport Protocol Version 1.3 or later.

Bluetooth SIG Document. Assigned Numbers.

ITU-R BS.1387 Method for objective measurements of perceived audio quality

3 术语和定义

GB/T 9002—2017、GB/T 5271.1、GB/T 5271.4 和 GB/T 5271.9 界定的以及下列术语和定义适用于本文件。

3.1

编码 coding

读入音频采样值并产生一个符合本文件的有效位流的过程。

3.2

解码 decoding

在本文件中定义的一种数据处理，即读入编码位流并输出音频采样值的过程。

3.3

编码位流 coded bitstream

音频信号的编码表示。

3.4

边信息 side information

位流中控制解码的必要信息。

3.5

声道 channel

用于传送到单个扬声器或其他重放设备的一组有序音频样本集合。

3.6

双声道立体声 stereo audio

一种音频格式，该格式下，使用两个声道承载有一定相位关系的音频信号，通常通过位于听音者前方的两个对称的扬声器或使用耳机重放，带给听音者更宽的声场感觉。

3.7

源端 source

产生或发送音频信号的设备。

3.8

宿端 sink

接受并处理音频信号的设备。

3.9

恒定比特流 constant bitstream

在同一码率挡位下传输码流包长不变。

3.10

恒定码率 constant bitrate

编码码率不随传输干扰环境变化而切换的模式。

3.11

自适应码率 adaptive bitrate

系统根据传输干扰变化自适应调整码率挡位的模式。

4 缩略语

下列缩略语适用于本文件。

L2HC 低延迟低复杂度高清音频编解码 (Low-latency Low-complexity High Resolution Audio Codec)

SRC: 信号源 (Signal Source)

SNK: 信号宿 (Signal Sink)

Tx: 无线信号发射机 (Transmitter)

Rx: 无线信号接收机 (Receiver)

A2DP: 高级音频分发协议 (Advanced Audio Distribution Profile)

LE Audio: 低功耗音频 (Low Energy Audio)

AVDTP: 音视频分发传输协议 (Audio/Video Distribution Transport Protocol)

HFP: 蓝牙免提协议 (Hands Free Profile)

THD+N: 总谐波失真加噪声 (Total Harmonic Distortion plus Noise)

DAC: 数字模拟转换 (Digital to Analog Conversion)

AMP: 放大器 (Amplifier)

Codec: 编解码器 (Coder-decoder)

PCM: 脉冲编码调制 (Pulse-code Modulation)

DRM: 数字版权管理 (Digital Rights Management)

MTU: 最大传输单元 (Maximum Transmission Unit)

ODG: 客观差异等级 (Objective Difference Grade)

CBR 恒定码率 (Constant Bitrate)

ABR 自适应码率 (Adaptive Bitrate)

5 技术要求

5.1 系统技术要求

L2HC 蓝牙传输系统如图 1 所示。

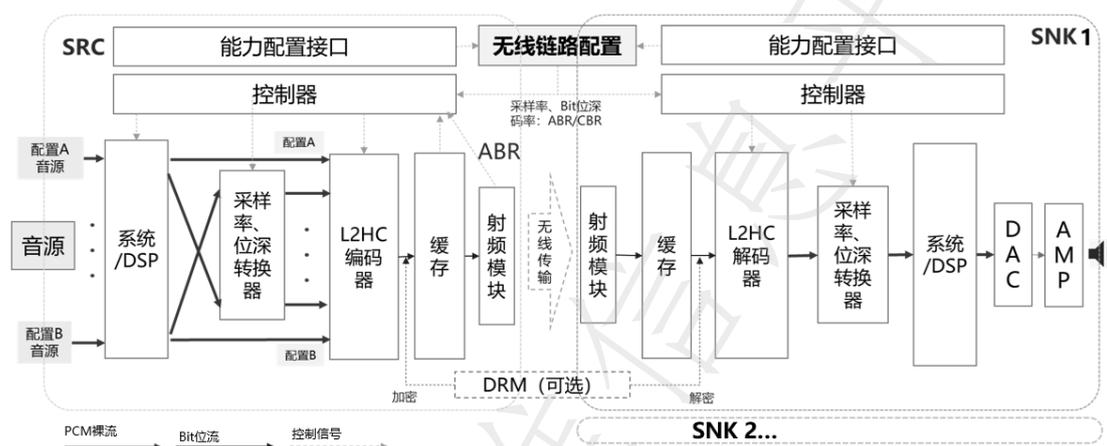


图1 L2HC 蓝牙传输系统音频编解码器框架

5.2 源端技术要求

L2HC 蓝牙传输适配源端技术要求如表 1 所示。

表1 L2HC 蓝牙传输适配源端技术要求

序号	项目	功能要求	
1	L2HC 编码	应具备将 PCM 音频裸流编码 L2HC 音频流的功能，位流表示方式应符合 T/CAIACN 009-2023 标准要求	
2	蓝牙技术要求	蓝牙核心传输协议	应支持 5.2 版本或以上的蓝牙核心传输协议（Bluetooth Core Specification）
3		AVDTP	应支持 1.3 版本或以上的蓝牙音视频分发传输协议（Audio/Video Distribution Transport Protocol）
4		A2DP	应支持 1.3.2 版本或以上的高级音频分发协议（Advanced Audio Distribution Profile）
5		LE Audio	宜支持蓝牙低功耗音频（LE Audio）
6		声道数	应支持单声道、双声道立体声；宜支持双声道
7	采样率	应支持 96kHz、48kHz，宜支持 88.2kHz、44.1kHz、32kHz	
8	采样位深	应支持 32bit、24bit、16bit	

表1 L2HC 蓝牙传输适配源端技术要求（续）

序号	项目	功能要求
9	传输码率	传输双声道立体声时： 应支持 960kbps、640kbps、480kbps、320kbps、256kbps、192kbps； 宜支持 1920kbps、1600kbps、1280kbps；可支持 128kbps（不建议） 传输单声道时： 应支持 480kbps、320kbps、256kbps、192kbps、128kbps、96kbps； 宜支持 960kbps、640kbps；可支持 64kbps（不建议）
10	帧长	应支持 10ms 普通帧长模式和 5ms 短帧长模式
11	码率切换模式	应支持 ABR 模式，宜支持 CBR 模式
12	采样率位深转换	应支持将输入 PCM 音频裸流的采样率和位深转换到蓝牙传输协议配置的采样率和位深规格
13	数字版权管理	可支持数字版权管理（DRM）
14	编码器声音质量	最低传输速率时：ODG 应 ≥ -0.5 640kbps 时传输双声道立体声：ODG 应 ≥ 0.0 960kbps 时传输双声道立体声：THD+N $\leq -120\text{dB}$

5.3 宿端技术要求

L2HC 蓝牙传输适配宿端技术要求如表 2 所示。

表2 L2HC 蓝牙传输适配宿端技术要求

序号	项目	功能要求
1	L2HC 解码	应具备正确识别 T/CAIACN 009-2023 标准要求的 L2HC 音频位流和语法，并将 L2HC 音频流解码为 PCM 音频裸流的功能
2	蓝牙技术要求	蓝牙核心传输协议 应支持 5.2 版本或以上的蓝牙核心传输协议（Bluetooth Core Specification）
3		AVDTP 应支持 1.3 版本或以上的蓝牙音视频分发传输协议（Audio/Video Distribution Transport Protocol）
4		A2DP 应支持 1.3.2 版本或以上的高级音频分发协议（Advanced Audio Distribution Profile）
5		LE Audio 宜支持蓝牙低功耗音频（LE Audio）
6	声道数	具备回放双声道及以上能力的音频产品应支持双声道立体声；宜支持单声道 只支回放持单声道的音频产品应支持单声道
7	采样率	应支持 96kHz、48kHz，宜支持 88.2kHz、44.1kHz、32kHz
8	采样位深	应支持 32bit、24bit、16bit
9	传输码率	传输双声道立体声时： 应支持 960kbps、640kbps、480kbps、320kbps、256kbps、192kbps； 宜支持 1920kbps、1600kbps、1280kbps；可支持 128kbps（不建议） 传输单声道时： 应支持 480kbps、320kbps、256kbps、192kbps、128kbps、96kbps； 宜支持 960kbps、640kbps；可支持 64kbps（不建议）
10	帧长	应支持 10ms 普通帧长模式，宜支持 5ms 短帧长模式

表 2 L2HC 蓝牙传输适配宿端技术要求（续）

序号	项目	功能要求
11	码率切换模式	应支持 ABR 模式，宜支持 CBR 模式
12	电性能要求	总谐波失真加噪声（THD+N） $\leq 0.5\%$ （1 kHz）
13	声性能要求	声频率响应应符合国标和行标要求

6 L2HC 蓝牙传输能力配置

6.1 蓝牙音频能力协商与建链过程

L2HC 作为供应商特定编解码（Vendor Specific Codec）适配蓝牙 A2DP（v1.4 4.7 章节）规范传输时，蓝牙音频能力协商与建链过程如图 2 所示。

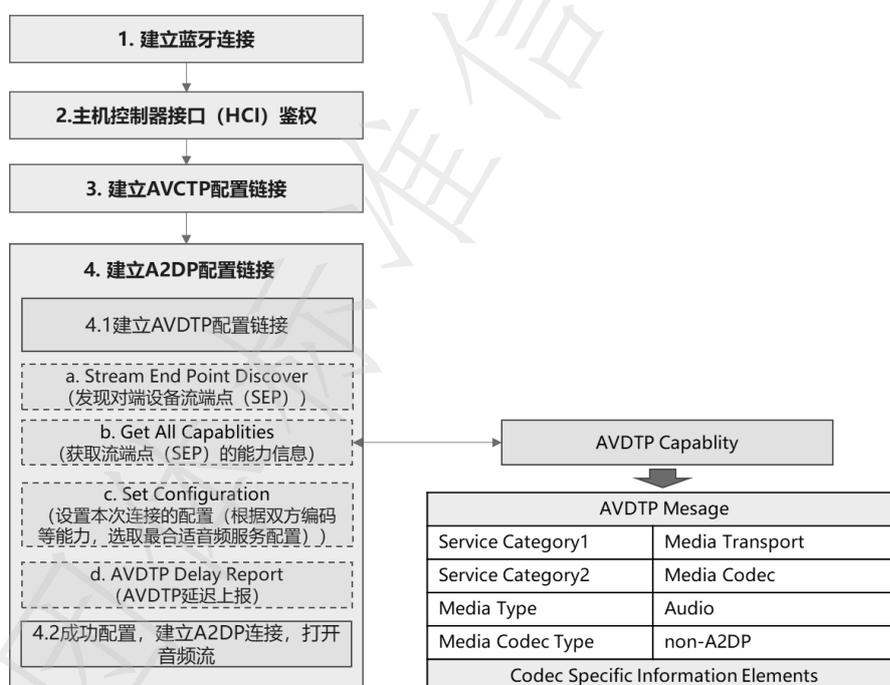


图2 L2HC 蓝牙音频能力协商与建链过程

- a) SRC 端发现 SNK 端的建立蓝牙连接；
- b) SRC 端与 SNK 端进行主机控制接口（HCI）鉴权；
- c) SRC 端与 SNK 端建立 AVCTP 配置链接；
- d) SRC 端与 SNK 端建立 A2DP 配置链接：
 - 1) SRC 端与 SNK 端建立 AVDTP 配置链接：
 - SRC 端发现对端设备流端点（SEP）；
 - SRC 端获取流端点（SEP）的能力信息，SNK 端通过 AVDTP 信息反馈 AVDTP 能力（AVDTP Capability）；
 - 设置本次连接的配置（根据双方编码等能力，选取最合适音频服务配置）；
 - AVDTP 延迟上报；
 - 2) 成功配置，建立 A2DP 连接，打开 A2DP 音频流。

6.2 编解码器规格信息

L2HC 作为供应商特定编解码（Vendor Specific Codec）适配蓝牙 A2DP（v1.4 4.7 章节）规范传输时，编解码器规格信息元素（Codec Specific Information Elements）的语义结构如表 3 所示。

表3 编解码器规格信息元素（Codec Specific Information Elements）语义结构

	MSB							LSB
Bit	7	6	5	4	3	2	1	0
Octet								
0	Vendor ID: (0x0000) 0x 0CCF 注: Vendor ID: 32bit, 高16位为0, 低16位为在蓝牙SIG Assigned Numbers Document中注册的Company ID							
1								
2								
3								
4	Codec ID: 0x CA01							
5								
6	Version Capability				Reserved	<u>Resolution</u> 32bits	<u>Resolution</u> 24bits	<u>Resolution</u> 16bits
7	Reserved	<u>Sample rate</u> 192kHz	<u>Sample rate</u> 176.4kHz	<u>Sample rate</u> 96kHz	Sample rate 88.2kHz	<u>Sample rate</u> 48kHz	Sample rate 44.1kHz	Sample rate 32kHz
8	Reserved	<u>Bitrate</u> 1920kbps	<u>Bitrate</u> 1600kbps	<u>Bitrate</u> 1280kbps	<u>Bitrate</u> 960kbps	<u>Bitrate</u> 640kbps	<u>Bitrate</u> 480kbps	<u>Bitrate</u> 320kbps
9	<u>Bitrate</u> 256kbps	<u>Bitrate</u> 192kbps	Bitrate 128kbps	Bitrate 96kbps	<u>Bitrate</u> 64kbps	Reserved	Frame Duration	
10	Frame Duration			Chanel Number			Reserved	Reserved
11	Reserved	Reserved	Reserved	Reserved	Reserved	Reserved	Reserved	Reserved

<u>双声道必选</u>	仅用于单声道	<u>可选</u>
0 为不支持, 1 为支持		

6.2.1 第三方供应商 ID（Vendor ID）

根据 A2DP 规范：应使用 32 位第三方供应商 ID（Vendor ID），高 16 位为 0，低 16 位为在蓝牙 SIG Assigned Numbers Document 中注册的 Company ID。L2HC 编解码器对应的 Company ID 信息如表 4 所示。

表4 L2HC Vendor ID 信息定义

Value	Name
0x0CCF	Shenzhen CESI Information Technology Co., Ltd.

因此 L2HC 编解码器对应 Vendor ID 为 0x0000CCF。下表为 Vendor ID 参数配置：

Octet 0	Octet 1	Octet 2	Octet 3
0xCF	0x0C	0x00	0x00

6.2.2 第三方供应商编解码器 ID (Codec ID)

L2HC 对应的第三方供应商编解码器 ID(Codec ID)为 0xCA01。Codec ID 参数配置如表 5 所示。

表5 L2HC Codec ID 信息定义

Octet 4	Octet 5
0x01	0xCA

6.2.3 版本信息

L2HC 编解码器版本信息定义如表 6 所示。

表6 L2HC 版本信息定义

位置	版本信息
Octet6; b4~b7	0000

6.2.4 采样位深

采样位深支持16位定点，24位定点和32位定点/浮点，编码器和解码器位深可以不同，具体配置如表7所示。

表7 L2HC 采样位深信息定义

位置	采样位深	SRC端支持	SNK端支持
Octet6; b0	16bits 定点	双声道必选	双声道必选
Octet6; b1	24bits 定点	双声道必选	双声道必选
Octet6; b2	32bits 定点/浮点	双声道必选	双声道必选

6.2.5 采样率

输入/输出的音频信号采样率支持44.1kHz、48kHz、88.2kHz和96kHz，具体配置如表8所示。

表8 L2HC 采样率信息定义

位置	采样率	SRC端支持	SNK端支持
Octet7; b0	32kHz	可选	可选
Octet7; b1	44.1kHz	可选	可选
Octet7; b2	48kHz	双声道必选	双声道必选
Octet7; b3	88.2kHz	可选	可选
Octet7; b4	96kHz	双声道必选	双声道必选
Octet7; b5	176.4kHz	可选	可选
Octet7; b6	192kHz	可选	可选

6.2.6 传输速率

L2HC编码模式为恒定比特流 (Constant Bitstream) 模式，单声道编码码率范围为64kbps到960kbps，立体声编码码率范围为128kbps到1920kbps，具体配置如表9所示。

表9 L2HC 传输速率信息定义

位置	传输速率	SRC端支持	SNK端支持
Octet9; b3	64kbps	可选	可选
Octet9; b4	96kbps	仅用于单声道	仅用于单声道
Octet9; b5	128kbps	可选	可选
Octet9; b6	192kbps	双声道必选	双声道必选
Octet9; b7	256kbps	双声道必选	双声道必选
Octet8; b0	320kbps	双声道必选	双声道必选
Octet8; b1	480kbps	双声道必选	双声道必选
Octet8; b2	640kbps	双声道必选	双声道必选
Octet8; b3	960kbps	双声道必选	双声道必选
Octet8; b4	1280kbps	可选	可选
Octet8; b5	1600kbps	可选	可选
Octet8; b6	1920kbps	可选	可选

6.2.7 帧长模式

L2HC编解码器支持10ms标准帧长模式和5ms短帧长模式，具体配置如表10、表11和表12所示。

表10 L2HC 帧长信息定义

位置	帧长	SRC端支持	SNK端支持
Octet9; b1	10ms 模式	必选	C1
Octet9; b0	7.5ms模式	- (待定)	- (待定)
Octet10; b7	5ms模式	必选	C1
Octet10; b6	Reserved		

C1: 至少支持一种模式

表11 10ms 标准帧长模式下 L2HC 采样点数与帧长、采样率对应关系

采样率	算法时延	帧内采样点数
96kHz	算法时延为12.5ms，其中帧长为10ms，前向附加时延为2.5ms	960点
88.2kHz	算法时延为13.6ms，其中帧长为10.88ms，前向附加时延为2.72ms	960点
48kHz	算法时延为12.5ms，其中帧长为10ms，前向附加时延为2.5ms	480点
44.1kHz	算法时延为13.6ms，其中帧长为10.88ms，前向附加时延为2.72ms	480点

表12 5ms 短帧长模式下 L2HC 采样点数与帧长、采样率对应关系

采样率	算法时延	帧内采样点数
96kHz	算法时延为7.5ms，其中帧长为5ms，前向附加时延为2.5ms	480点
88.2kHz	算法时延为8.16ms，其中帧长为5.44ms，前向附加时延为2.72ms	480点
48kHz	算法时延为7.5ms，其中帧长为5ms，前向附加时延为2.5ms	240点
44.1kHz	算法时延为8.16ms，其中帧长为5.44ms，前向附加时延为2.72ms	240点

6.2.8 声道数

L2HC编解码器包括单声道编解码、立体声（双声道）编解码。（待定，当前默认为双声道），具体配置如表13所示。

表13 L2HC 声道数信息定义

位置	声道数	SRC端支持	SNK端支持
Octet10; b5		Reserved	Reserved
Octet10; b4	四声道	Reserved	Reserved
Octet10; b3	双声道	必选	C1, 支持立体声设备必选
Octet10; b2	单声道	可选	C1
C1: 至少支持一种模式			

6.3 蓝牙音频能力协商与建链过程

按照 6.1 节要求完成蓝牙音频能力协商与 A2DP 音频流建链后，L2HC 蓝牙音频流传输过程如图 3 所示。

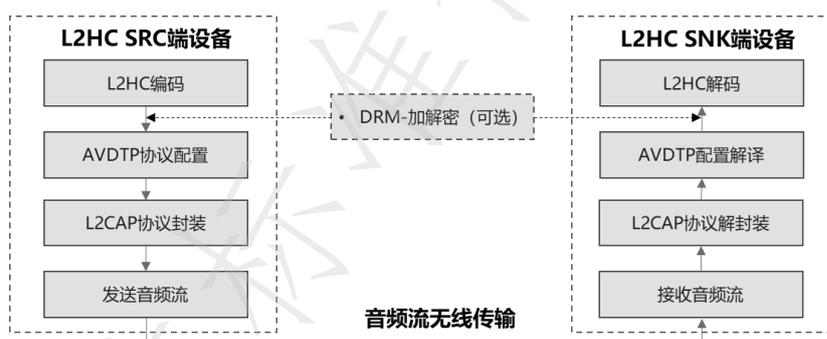


图3 L2HC 蓝牙音频流传输过程

6.4 L2HC 位流数据

6.4.1 语法

L2HC 位流数据语法如表 14 所示。

表14 L2hcDecodeRawData()语法

L2hcDecodeRawData()语法	比特数	助记符
L2hcDecodeRawData(){	—	—
L2hcHeaderUnpack()	—	—
switch(codecFormat) {	—	—
case 0x0: L2hcCommonDec(chNumDec=chNum)	—	—
case 0x1: L2hcHighCostDec()	—	—
}	—	—
PcmCopyChannel()	—	—
}	—	—

包头信息解码如表 15 所示。

表15 L2hcHeaderUnpack()语法

L2hcHeaderUnpack()语法	比特数	助记符
L2hcHeaderUnpack(){	—	—
codecType	2	uimsbf
sampleRate	2	uimsbf
chNum	1	uimsbf
chNum = chNum + 1	—	—
frameLength	2	uimsbf
}	—	—

6.5 边信息解码

6.5.1 语法

边信息解码语法如表 16 所示。

表16 DecodeCoreSideBits() 语法

DecodeCoreSideBits() 语法	比特数	助记符
DecodeCoreSideBits(chNumDec) {	—	—
lowBrFlag	1	uimsbf
if (chNumDec == 2) {	—	—
msFlag	1	uimsbf
}else{	—	—
msFlag = 0	—	—
}	—	—
dr	5	uimsbf
drQuarter	3	uimsbf
sflD	3	uimsbf
if (lowBrFlag) {	—	—
bandNum	4	uimsbf
bandNum += 16	—	—
}else{	—	—
bandNum = 32	—	—
}	—	—
diffFlag	1	uimsbf
for (ch = 0; ch < chNumDec; ch++) {	—	—
DecodeScaleFactor()	—	—
}	—	—
for (ch = 0; ch < chNumDec; ch++) {	—	—
hufTupId_4tuple[ch]	2	uimsbf

表 16 DecodeCoreSideBits() 语法 (续)

DecodeCoreSideBits()语法	比特数	助记符
hufTupId_2tuple[ch]	2	uimsbf
hufTupId_1tuple[ch]	2	uimsbf
}	—	—
}	—	—

6.6 位流格式

6.6.1 概述

编码器处理单声道信号时，按照单声道位流格式进行打包。位流结构如表 17 所示。

单声道位流格式包含三部分。第一部分包头位流如表 18 所示。第二部分边信息位流如表 19 所示。其中子带包络(sf)编码比特数不固定，记为 nBit_sf。第三部分 Payload 位流如表 20 所示。其中量化的 MDCT 系数使用 Huffman 编码，所用比特数不固定，记为 nBits_Payload。

表 19 中 bandNum 仅在低码率模式 (lowBrFlag=1) 下传递。

表 17 单声道位流格式

Header	Side Info	Payload
--------	-----------	---------

表 18 包头位流格式

Header				
字段	Codec Type	Sample Rate	Channel Number	Frame Length
比特数	2	2	1	2

表 19 边信息位流格式

Side Info				
字段	lowBrFlag	dr	drQuater	sfId
比特数	1	5	3	3
Side Info				
字段	bandNum	diffFlag	sf	hufTupID
比特数	4	1	nBits_sf	6

表 20 Payload 位流格式

Payload		
字段	mdctQ	mdctRes
比特数	nBits_Payload	

编码器处理立体声信号时，涉及两种位流打包格式：立体声交织位流和立体声非交织位流，其位流格式如表 21、表 22、表 23 和表 24 所示。

当 codecFormat 等于 0 时，使用立体声交织位流，位流结构如表 21 所示。立体声交织位流格式包含三部分，第一部分为包头位流，复用表 18，第二部分交织格式的边信息位流如表 22 所示，第三部分交织格式的 Payload 位流如表 23 示。表 22 中 bandNum 仅在低码率模式（lowBrFlag=1）下传递。

当 codecFormat 等于 1 时，使用立体声非交织位流，位流结构如表 24 所示。立体声非交织位流格式其包含五部分，第一部分包头位流复用表 18，第二部分左声道的边信息位流复用表 19，第三部分左声道的 Payload 位流复用表 20，第四部分右声道的边信息位流复用表 19，第五部分右声道的 Payload 位流复用表 20。

表21 立体声交织位流格式

Header	Side Info(common)	Payload(L,R)
--------	-------------------	--------------

表22 立体声交织格式边信息位流

Side Info(common)					
字段	lowBrFlag	msFlag	dr	drQuarter	sfId
比特数	1	1	5	3	3
Side Info(common)					
字段	bandNum	diffFlag	sf	hufTupID	
比特数	4	1	nBits_sf(L, R)	12	

表23 立体声交织格式 Payload 位流

Payload(L,R)		
字段	mdctQ(L,R)	mdctRes(L,R)
比特数	nBits_Payload(L, R)	

表24 立体声非交织位流格式

Header	Side Info(L)	Payload(L)	Side Info(R)	Payload(R)
--------	--------------	------------	--------------	------------