

AoIP原理与实践

Principles and Practice of AoIP

钱岳林 ◎ 主 编
邓纶晖 ◎ 副主编

GY/T 304 Q-LAN AES67
GY/T 322 Dante Ravenna NMOS WheatNet-IP
LiveWire+ **Audio-over-IP** IEEE 1588
SMPTE ST 2110 CobraNet
EtherSound AES70 AVB

图书在版编目(CIP)数据

AoIP 原理与实践 / 钱岳林主编. --北京 : 中国传媒大学出版社, 2020.10
ISBN 978-7-5657-2735-1

I. ①A… II. ①钱… III. ①数字音频技术 IV. ①TN912.2

中国版本图书馆 CIP 数据核字(2020)第 126744 号

AoIP 原理与实践

AoIP YUANLI YU SHIJIAN

主 编 钱岳林
副 主 编 邓纶晖
策 划 陈 默
责任编辑 王 硕
特约编辑 陈 默
封面设计 蒋华挺
封面制作 郭 琳
责任印制 李志鹏

出版发行 中国传媒大学出版社
社 址 北京市朝阳区定福庄东街 1 号 邮 编 100024
电 话 86-10-65450528 65450532 传 真 65779405
网 址 <http://cucp.cuc.edu.cn>
经 销 全国新华书店

印 刷 艺堂印刷(天津)有限公司
开 本 787mm×1092mm 1/16
印 张 22.5
字 数 493 千字
版 次 2020 年 10 月第 1 版
印 次 2020 年 10 月第 1 次印刷

书 号 ISBN 978-7-5657-2735-1/TN·2735 定 价 98.00 元

本社法律顾问:北京李伟斌律师事务所 郭建平

版权所有 翻印必究 印装错误 负责调换

本书编委会

主 编 钱岳林

副主编 邓纶晖

编 著 (按姓氏笔画排序)

兆 翦 陈 武 罗 攀

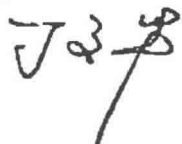
单雪松 郭 晓 谢凌云

序言

数字化技术深刻改变着我们的生活方式。信息技术的发展,使得信息无处不在,随时随地都可以获得,大数据、人工智能等新技术的快速出现和应用,将使经济发展方式和社会管理模式产生巨大的变革,科学与技术成为社会发展的巨大推动力量。

作为大众传播的主要载体的广播电视行业,经过近20年的发展,已经完成了由模拟向数字化的转变。一方面,国家三网融合的大趋势要求广播电视的传输向IP化发展;另一方面,在超高清电视领域,IP化传输是解决基带信号传输的手段和方式。国内的一些电台、电视台在演播室、播控中心和传输的IP化方面也做了许多尝试,行业对统一的技术标准的需求变得更加迫切。

钱岳林总工程师带领的团队,在完成了AES67和AES70标准的国内版本译制工作后,集中力量完成了本书,为在广播电视行业和演艺行业普及AoIP技术提供了参考,是目前国内唯一的关于AoIP技术的专著。相信此书的出版,对国内的工程技术人员了解世界上AoIP的发展状况、掌握AoIP相关的技术标准、推动该技术在国内的教学和应用,将起到积极的推进作用。



中国工程院院士
2019.7

前 言

AoIP 技术是随着互联网技术的发展而兴起的利用 IP 封装传输音频的技术,特点是低延时、无压缩,随之而来的是互联网技术中的发现、同步与组播技术等。

2017 年 1 月,由中央人民广播电台牵头,国内广电界科研院所、设备生产厂家、用户等参与的参照 AES67-2015 起草的 GY/T 304—2016《高性能流化音频在 IP 网络上的互操作性规范》发布,几乎同时,美国电影电视工程师协会将 AES67 纳入 SMPTE 的标准簇,命名为 SMPTE ST 2110-30,这意味着这个标准不仅仅是一个音频行业协会的标准,它同时还被高清和超高清电视行业采纳。

2017 年 5 月,在杭州的一次讨论 AoIP 应用的小范围会议上,我有两点突出的感受:一是不仅仅广电行业需要标准,演艺扩声等多行业也需要;二是我们目前的大学教育,学科面比较窄,学音频和声学的不懂计算机和网络,学计算机和网络的不懂音频。当时,我和中国传媒大学及一些厂家的同行不谋而合,认为有必要写一本书,打通音频与计算机网络的界限,为培养复合型应用人才而努力。

我们查阅了相关的资料,找到的已出版的相关书籍只有 Steve Church 和 SkipPizzi 合著的 *Audio over IP—Building Pro AoIP System with Livewire*,它主要涉及公司产品的网络设计,且在 2009 年出版。这更坚定了我们写好这本书的信心,我们希望写出一本能满足国内专业音频、演艺扩声、设备研制生产、大学教育、专业人员进修的书籍,能够比较全面地介绍 AoIP 的发展历程、技术特点、对网络的要求等。

在撰写本书的过程中,我们同时完成了 AES70 的翻译工作,于 2019 年 4 月发布(GY/T 322.1—2019;322.2—2019;322.3—2019)。AES67 的主要作用是满足不同厂家设备的互联互通,而 AES70 则对控制架构、类结构和用于 TCP/IP 网络的协议进行了规范,是一个比较完整的标准体系,这部分内容在本书中也有详细介绍,以保证本书的权威和完整。

经过一年多的努力,本书终于成稿了。它是集体智慧的结晶,其中中国传媒大学副研究员谢凌云编写了第 1 章,中国传媒大学高级工程师郭晓编写了第 2 章和第 5 章的 5.1、5.2、5.3、5.4、5.5、5.6 节,中央人民广播电台高级工程师罗攀编写了第 4 章,AIMS 大中华区代表、QSC 亚太区商务经理兆翦编写了第 3 章的 3.1、3.2、3.3 节和第 5 章的 5.7 节、第 7 章的 7.1、7.4 节以及第 8 章,苏州福川科技有限公司技术总监、高级

工程师陈武编写了第3章的3.4节、第6章和第7章的7.2、7.3节,杭州联汇科技股份有限公司广电技术副总监、高级工程师单雪松撰写了第5章的5.8、5.9节和第6章的6.3节中的部分内容,中央电视台转播部高级工程师周磊编写了第6章的6.5节。本书主编——中央人民广播电台总工程师钱岳林负责本书的规划统筹、全篇校订,中国传媒大学副研究员邓纶晖作为副主编承担了部分统稿及部分修订工作。

中国工程院院士丁文华为本书撰写了序言,苏州福川科技有限公司和杭州联汇科技股份有限公司为本书的出版提供了大力支持,在此一并表示感谢。

希望本书的出版能对电台和电视中心的IP化建设、演艺界的IP化应用等起到积极的规范和推动作用,同时,对高校复合型人才的培养、广播电视中心技术人员的再教育和进修具有积极的参考价值。

钱岳林

2019.7.5

目 录

1	数字音频基础	1
1.1	概述 / 1	
1.2	采样 / 1	
1.3	量化 / 4	
1.4	编码 / 6	
1.5	传输 / 6	
1.6	抖动 / 8	
1.7	同步 / 9	
2	网络技术基础	12
2.1	概述 / 12	
2.2	物理层 / 14	
2.3	数据链路层 / 17	
2.4	网络层 / 21	
2.5	传输层 / 31	
2.6	应用层 / 36	
3	AoIP 技术发展和常见体系	42
3.1	AoIP 的定义与适用范围 / 42	
3.2	AoIP 技术演进 / 44	
3.3	相关技术系统的特点 / 50	
3.4	AoIP 相关标准体系 / 60	
4	AoIP 工作原理	74
4.1	概述 / 74	
4.2	AoIP 同步和媒体时钟 / 75	
4.3	高性能实时音频流的传输 / 100	
4.4	AoIP 设备及音频流管理 / 116	

5 AoIP 系统设计	140
5.1 系统架构 /	140
5.2 设计原则 /	142
5.3 网络架构与设计 /	144
5.4 VLAN 配置 /	150
5.5 DSCP 系统配置 /	153
5.6 静态路由配置 /	156
5.7 组播路由配置 /	162
5.8 同步系统配置 /	167
5.9 AoIP 互通操作配置 /	175
6 AoIP 在广播电视中的应用	182
6.1 AoIP 在广播制播系统中的应用 /	182
6.2 AoIP 在广播总控系统中的应用 /	204
6.3 AoIP 在广播监测系统中的应用 /	229
6.4 广播远程传输系统 /	248
6.5 AoIP 在央视 4K UHD/HDR IP 转播车(B-5)系统中的应用 /	253
7 AoIP 在专业音频系统中的应用	259
7.1 大型公共广播 /	259
7.2 演艺系统 /	273
7.3 会议系统 /	281
7.4 影院系统 /	300
8 AoIP 各协议研读	305
8.1 CobraNet 协议研读 /	305
8.2 EtherSound 协议研读 /	309
8.3 AVB/TSN 技术研读 /	313
8.4 Livewire+AES67 协议研读 /	316
8.5 Dante 协议研读 /	321
8.6 Ravenna 协议研读 /	328
8.7 Q-LAN 协议研读 /	334
8.8 WheatNet-IP 协议研读 /	341
参考文献 /	344

1 数字音频基础

1.1 概述

随着信息和通信技术(Information and Communication Technology, ICT)的不断进步,音频技术也逐渐由模拟系统过渡到了数字系统。相对于模拟音频系统而言,数字音频系统在信号动态范围、失真、信噪比等方面有着更为广阔的提高空间,因而各种数字音频设备逐步取代了原有的模拟音频设备,成为制作、存储、传输和播放各类音频的主流设备。

从模拟信号到数字信号,要经过采样、量化、编码、存储和传输这一套数字化流程,如图 1-1 所示。本章将围绕这个处理流程,介绍必要的数字音频基础知识,对音频信号的模数转换(采样和量化)、存储(以数字音频格式)、编码和传输(基于 AES3)等方面一一进行介绍,让读者对数字音频技术及其应用有一个基本的认识。

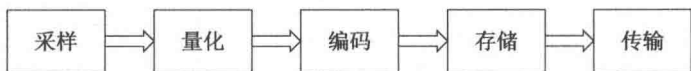


图 1-1 模拟信号数字化流程

1.2 采样

数字系统中,将连续的模拟信号转换成离散的数字信号的过程被称为采样。这个过程实质上就是将振幅随时间连续变化的模拟信号,按照一定的时间间隔取出信号样本,得到一个在时间上离散的数字信号序列。采样一般由专门的模拟数字转换设备完成。每秒钟采集到的样本数,即采样的密度,又被称为采样频率或者抽样频率;两次采样之间的时间间隔则被称为采样周期,它和采样频率互为倒数。如图 1-2 所示,不同的采样频率,意味着对模拟信号的还原精度不同。从通常意义上来说,采样频率越高,

对信号的重构还原度越好。

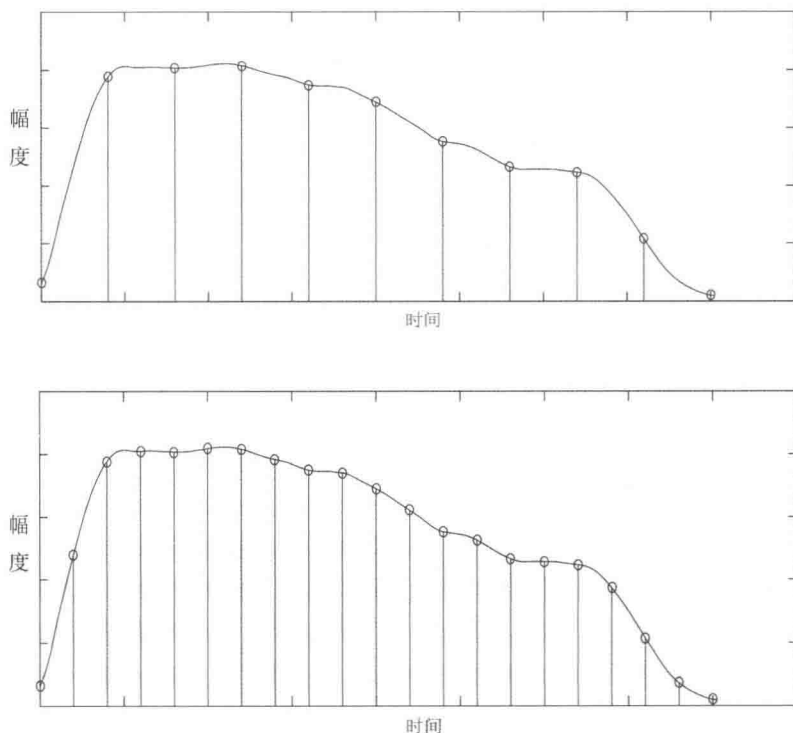


图 1-2 对连续模拟信号进行不同频率采样

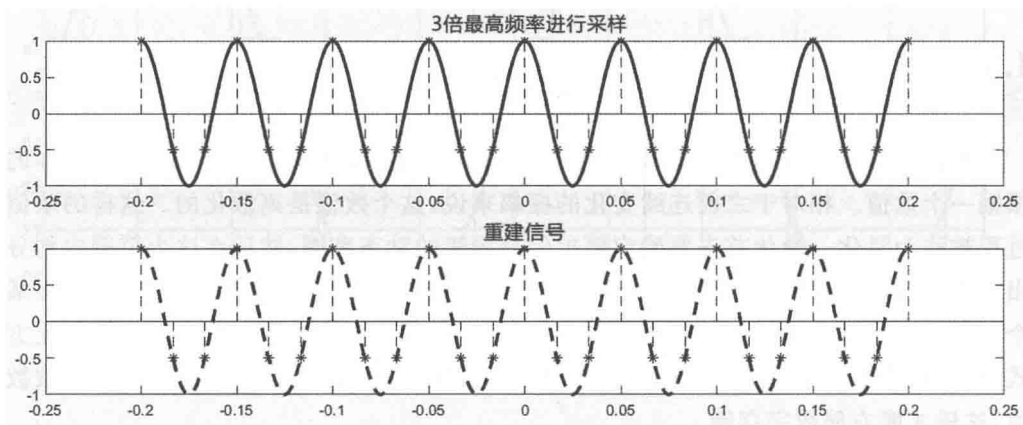
如果要尽可能完美地重构一段模拟信号,使其不会产生失真,采样频率就必须满足一定的要求。1928年美国工程师奈奎斯特提出了著名的“奈奎斯特采样定理”,对这个要求进行了阐述,即采样频率大于模拟信号中最高频率的2倍时,采样之后的数字信号才会完整保留模拟信号中的信息。具体到实际的信号时,不同的信号对应的奈奎斯特采样频率又会有所不同,但本质上都是以人的感知为依据的。

以音频信号为例,公认的人类能听到的最高频率是20kHz,如果要让采样后的数字信号保留人耳能感知的所有信息,那么根据奈奎斯特采样定理,采样频率大于40kHz即可满足要求。所以数字音频系统的采样频率大多是40kHz~50kHz,例如CD的标准采样频率是44.1kHz,主流的声卡还会支持48kHz,这已经可以保留可听范围内的全部信息了。如果是语音信号,其常用频带范围是300Hz~3400Hz,对应的理想数字采样频率一般是8kHz。为满足高保真音质的需求,采样频率要求较高,支持44.1kHz/48kHz/96kHz/192kHz。

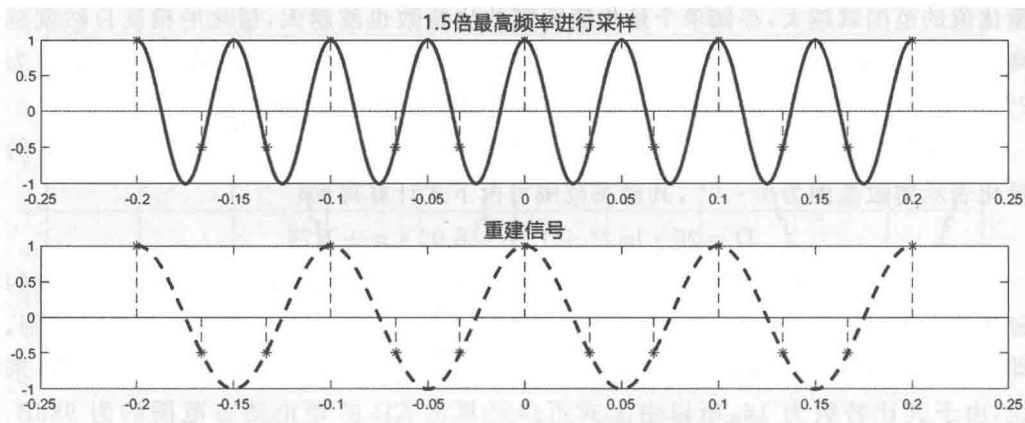
有时会有这样的质疑:既然超出了人类听觉范围,那么过高的采样频率会不会是冗余和浪费的?事实上,大多数人只是对20kHz以上的声音不那么敏感而已,并不是完全无法感知,而且20kHz以上的高次谐波对于很多声音(尤其是乐器的声音)来说是比较重要的泛音成分,缺少这部分谐波会导致声音听起来缺少自然感和丰满感(与

现场听到的真实声音相比)。同时,较高的采样频率也能为后期音频处理带来的失真留下改善的空间。所以,目前专业录音领域内支持的最高采样频率已经达到了192kHz。

如果采样频率 F_s 低于模拟信号最高频率的2倍,信号中大于 $F_s/2$ 的频率成分并不会在采样后的数字信号中消失,而是会对称映射到 $F_s/2$ 以下的频带中,和原有的 $F_s/2$ 以下的频率成分混合叠加在一起,这就是频率混叠现象,它带来的干扰即混叠干扰,会导致信号重建时发生错误,无法正确恢复原始信号的频率。如图1-3(a)所示,图中上半部分是20Hz的连续信号,对其进行60Hz采样频率的采样,采样点在图中以虚线表示。由图中下半部分的重建信号可以看出,对20Hz的原始频率的恢复很准确。但如果采样频率低于信号最高频率的2倍,例如采样频率仅为30Hz,那么如图1-3(b)



(a) 符合采样定理的3倍最高频率采样和重建



(b) 会导致错误的1.5倍最高频率采样和重建

图1-3 不同的采样频率对信号重建的影响示意图

所示,就会出现错误:重建信号和原始信号不一致。为了消除这种干扰,最直接有效的方法就是提高采样频率。然而实际的数字信号处理系统会受到实时性能和存储容量的限制,没办法处理大量的数据,达不到理想的采样频率。所以,提高采样频率以避免混叠干扰的方法在解决实际问题的時候还有很大的局限性。这时,一般会采用另一种预处理的方法,即放置一个抗混叠低通滤波器,滤掉模拟信号中高于 $F_s/2$ 的频率成分,这样也可以消除混叠干扰。

经过一系列处理后,如果要把模数转换后的数字信号重放出来,还得经过数模转换。这个过程正好和采样相反,需要把时间上不连续的采样值之间的空隙给填补上,得到与原来连续变化的模拟信号差不多的结果。数模转换一般采用插值滤波器的方法,通过专门的数模转换电路来实现。

1.3 量化

对模拟信号进行离散化的采样后,要完成信号数字化的过程,还需要赋予样本的振幅一个数值。相对于之前连续变化的振幅来说,这个数值是离散化的。这样的赋值过程被称为量化。量化首先要确定模拟信号振幅的动态范围,然后在这个范围内划分出若干个区间,每个区间对应一个量值。输入的模拟信号采样的振幅大小如果落在某个区间内,则该振幅量化的结果是这个区间对应的那个量值。用这种多对一的量化方式,就能将原本连续变化、有近乎无限取值可能的信号振幅转变为数量有限的离散数值,这样才能方便数字存储。

量化数值(又称量化电平)在数字系统中需要用一个二进制数来存储,这个二进制数的长度(也就是需要用到的比特数)和量化区间的多少有关。划分的量化区间越多,量化值的范围就越大,存储单个量化值所需的比特数也就越大,量化的精度自然就越高。例如,在 CD 制品中,一般采用 16bit 进行量化采样,这就意味着量化值的数目为 $2^{16} = 65,536$ 。有时也将量化比特数称为量化深度(Bits Depth)。

从量化深度也可以推导出数字信号的理论动态范围。以定点处理为例, n bit 的量化表示整数范围为 $1 \sim 2^n$,其动态范围可由下式计算得到:

$$D = 20 * \lg 2^n + 1.76 = 6.02 * n + 1.76$$

其中, D 表示动态范围。这个公式来源于 1948 年贝尔实验室的 Bennett 发表的经典论文 *Spectra of Quantized Signals*^①。而上述公式可以进一步简化为 6dB 规则,即:每增加 1bit 的量化深度,数字系统的动态范围将增加约 6dB。对于常见的 CD 来说,由于其比特数为 16,所以由上式可以计算出 CD 的理论动态范围约为 98dB。AoIP 可支持的量化深度有 16bit/24bit/32bit,这就意味着相应的动态范围最高可以到 194dB。

① BERNETT W R. Spectra of quantized signals[J]. Bell system technical journal,1948(27):446-472.

实际的模拟样本值和量化以后的数值是有差别的,这种伴随量化而产生的误差被称为量化误差,又叫量化噪声。量化精度越低,量化误差越大,信噪比越低。图 1-4 给出了一个对简单的连续正弦信号进行不同精度量化的例子,其中,量化比特数分别为 2(量化值有 $2^2=4$ 个)和 4(量化值有 $2^4=16$ 个)。从图 1-4 中明显可以看出,2bit 这样粗糙的量化会带来极大的量化误差,而 4bit 的量化就能对图示的正弦信号给出精度较高的量化结果,其产生的量化误差也要小很多。

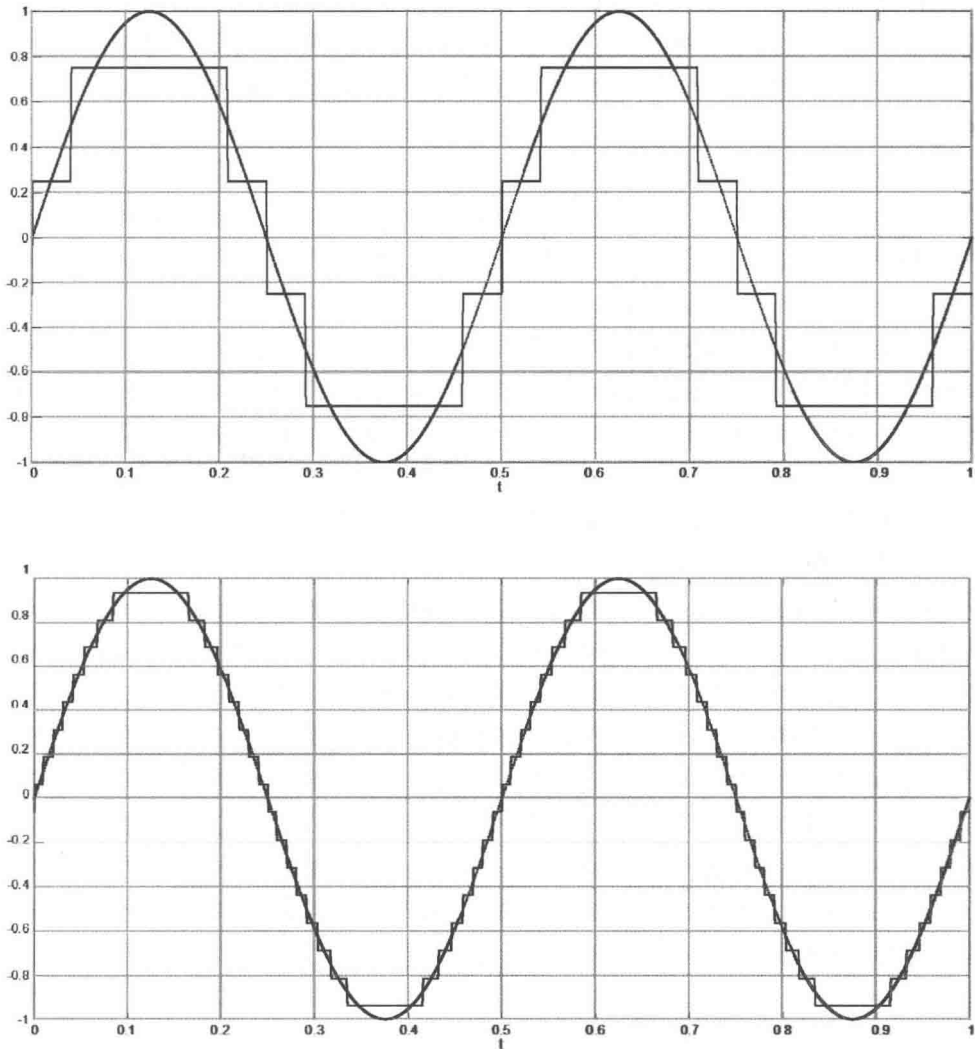


图 1-4 比特数为 2(上图)和 4(下图)的量化示例

相邻的量化数值之间的距离被称为量化距离。如果各个量化区间的大小相等且均匀分布在整个模拟信号振幅的取值范围内,那么量化距离也是相等的。这时候的量化就是均匀量化,又被称为线性编码。图 1-4 中给出的两种量化都是均匀量化。除此

以外,还有一类量化方法。这类方法会根据需要,在信号振幅的不同范围内划分大小不同的量化区间。例如,语音信号一般都是小幅度的信号,这时可以在中低幅值范围划分较多较小的量化区间,而在高幅值范围划分较少较大的量化区间,这样可以保证量化后的数字信号能够更精确地保留重要的中低幅值语音信息。这类量化方法被称为非均匀量化或非线性量化。目前常见的非均匀量化有 A 律和 μ 律,它们都是采用对数曲线来设定量化区间的,广泛应用于模拟电话数字化领域。

如果小幅度信号采用非均匀量化,信噪比会较高,而如果大幅度信号采用非均匀量化,信噪比则会较低。这对语音信号来说也许是合适的,但是对动态范围很大、大幅度信号很常见的音乐信号而言就不见得了,所以 AoIP 支持的是高比特数的线性均匀量化。尽管这会带来传输效率上的损失,但是它保证了 AoIP 更看重的高质量、宽频带特性。

1.4 编码

经过采样和量化的数字信号的数值转换为给定字长的二进制格式的过程就是编码,正式的名称是脉冲调制编码(Pulse Code Modulation, PCM)。这个概念是英国工程师 Alec Reeves 于 1937 年提出来的。根据线性和非线性量化的不同,具体的 PCM 方法也会不同,不过通常我们说的 PCM 都是指线性 PCM,这也是在 AoIP 上传输的音频数据编码方式,属于无压缩的编码方式。

PCM 数据在计算机上以音频文件的形式存储。从数据信息损失与否的角度来看,数字音频信号的存储文件格式有无损和有损两种。我们常见的 WAV 文件就是一种无损格式,通常用来保存一些没有压缩过的音频数据。类似的无损格式还有 APE 和 FLAC。APE 本质上是一种无损音频压缩技术,WAV 文件被压缩为 APE 文件后,还可以还原为 WAV 格式,音频数据也和之前相同,没有损失。APE 的压缩率在 50% 左右。FLAC 的全称是 Free Lossless Audio Codec,是开放源码的无损压缩格式,压缩方式类似于我们常用的 RAR 和 ZIP 等一般文件压缩软件,但它专门针对 PCM 音频数据的特点设计了算法,压缩率也在 50% 左右。有损压缩的格式以常见的 MP3、AAC 和 OGG 为主。其中,MP3 和 AAC 都是 MPEG 标准的一部分,OGG 则是开放源码的免费音频压缩格式。

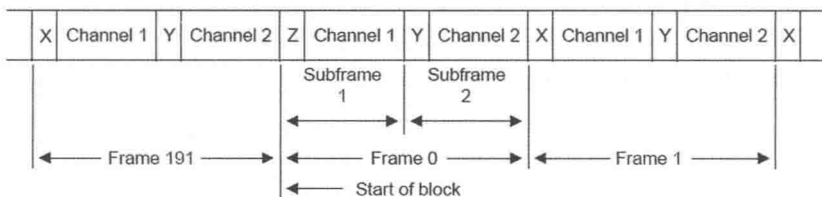
1.5 传输

音频数字信号的传输并不是简单地通过通信网络的发送和接收实现的,还需要考虑不同设备之间的同步。尤其是数字系统中有非常重要的时钟概念,要通过系统时钟来确定各个设备的数字信号的传输速率和帧起始点。如果没有统一的音频信号传输

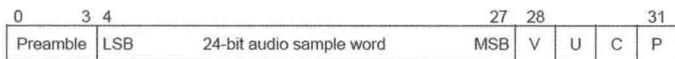
和接口标准,那么不管是信号内容编码还是时钟同步,都会产生错误。在这样的背景下,音频工程协会(Audio Engineering Society, AES)和欧洲广播联盟(European Broadcasting Union, EBU)于1985年提出了AES3标准,这一标准其后几经修订和更新,以适应音频设备不断发展的要求。目前最新的标准版本是AES3-2009(r2014),其全称为*AES Standard for Digital Audio Engineering—Serial Transmission Format for Two-channel Linearly Represented Digital Audio Data*。我国与其相对应的广播电影电视标准为GY/T 158—2000,即《演播室数字音频信号接口》标准。

AES3是传输和接收数字音频信号的数字设备接口协议,它通过单根绞合线对来传输数字音频数据,是一种串行位传输协议。它支持平衡差分连接,一般使用XLR(卡农)接口及专用的数字音频屏蔽电缆,也支持单端非平衡连接,一般使用BNC接头的数字音频同轴电缆。这两种传输接口所传输的数据帧结构是一致的。通常AES3传输的都是周期采样和线性量化的PCM音频数据,它们符合AES3帧结构标准。

AES3的数据帧结构如图1-5所示,其中图1-5(a)表示的是标准帧结构,图1-5(b)表示的是标准帧中的子帧结构。AES3标准数据帧提供两个通道的音频数据,它们以串行的方式排列传输,最高可以达到24bit量化。除此以外,数据帧还提供了传输控制的方法和状态信息的表示方法,并具备了一定的误码检测能力。从图1-5(b)可知,一个AES数据帧包含两个子帧,分别对应两个通道的音频数据。每个子帧里面除了音频数据以外,在帧的开头还会有四位标识数据,即Preamble,又可以称其为前置同步字。这样的标识数据有三种类型:X类型的标识数据,表示在标识数据后跟的是通道1的音频数据;Y类型的标识数据,表示在标识数据后跟的是通道2的音频数

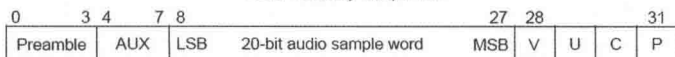


(a)标准帧结构



(a)

V Validity bit
U User data bit
C Channel status bit
P Parity bit
AUX Auxiliary sample bits



(b)

(b)子帧结构

图 1-5 AES3 标准帧结构和子帧结构

[资料来源: AES3-2009(r2014)标准]

据;Z 类型的标识数据,表示在标识数据后跟的是一个通道 1 的音频数据,同时也表示一个新的状态数据块的开始。在 Preamble 之后有 4 位辅助数据,即 AUX。这部分数据的功能主要就是为将样本分辨率从 20 位扩展到 24 位。

AES3 无须均衡即可在长达 100 米的距离上传输数据,如果均衡,则可以传输得更远。它的时钟信息是由传输端控制的,它的三个标准采样频率是 32kHz、44.1kHz、48kHz,当然许多接口能够在其他的采样频率上工作。AES3 提供“专业”和“消费”两种模式。两者最大的不同在于信道状态位格式。专业模式的状态位格式里包括数字信道的源和目的地址、日期时间码、采样点数、字节长度和其他信息。消费模式包括的东西比较少,但有拷贝保护信息。另外,AES3 标准提供“用户数据”,它的位流里包含用户说明(例如厂商说明等)。

之后,为了拓展多通道音频数据传输的应用范围,AES 在 AES3 标准的基础上,又于 1991 年开始制定面向多通道传输的标准 AES10,提出了多通道音频数字串行接口(serial Multichannel Audio Digital Interface, MADI)。该标准目前最新的版本是 AES10-2008 (r2019),全称为 *AES Recommended Practice for Digital Audio Engineering—Serial Multichannel Audio Digital Interface (MADI)*。我国与之相对应的广播电影电视标准是 GY/T 187—2002,即《多通路音频数字串行接口》标准。

AES10 采用 75Ω 同轴电缆或光纤来进行线性音频数据的串行传输,可支持 32、56 和 64 个通道,采样频率从 32kHz 到 96kHz。单通道分辨率最高可以到 24bit,实际传输速率最高可到 100Mb/s。AES10 的每一帧包含 32、56 或 64 个通道,每个通道块的数据格式包含了 32 个比特,头 4 个比特是 MADI 通道标识,之后的 28 个比特则和 AES3 的数据帧格式兼容。总体来说,AES10 定义了专业音频设备上的接口标准,打造了数字调音台和音频矩阵这类多通道系统之间的传输网络,目前已经应用到了大规模的数字音频系统和多通道数字设备互联中。

1.6 抖动

理想情况下,以 1MHz 为例,其理想的脉冲信号的持续时间应该是 $1\mu\text{s}$,每 500ns 有一个跳变沿。实际通信系统中,这种信号并不存在,其周期长度会有一定变化,从而导致下一个跳变沿的到来时间不确定。这种实际脉冲跳变时间点和理想脉冲跳变时间点之间的不确定差异就是抖动,如图 1-6 所示。

抖动是信号在时间上的变化,它从本质上描述了信号周期与理想值相比偏差了多少,一般使用真实时钟信号抖动的时间来衡量,即抖动幅度,也就是实际跳变相对于理想跳变的时间改变量,用纳秒或者单位间隔(Unit Interval, UI)来表述。1 个 UI 是数字信号状态变化的最小间隔,等同于 1 个比特的传送时间长度,其和采样频率大小有关。由于 AES3 使用双相位标志编码,传输时钟信号为 2 倍的比特率,故 UI 等同于 1 个比特的音频原始数据传送时间的一半。AES3 标准给出了对应不同采样频率信号

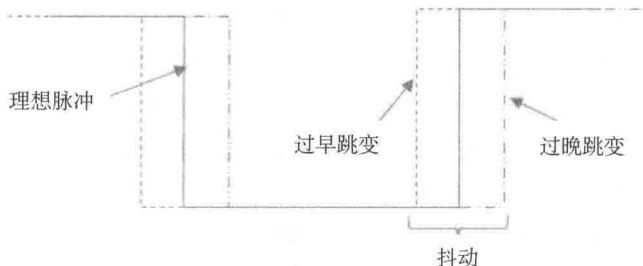


图 1-6 抖动概念示例

的 UI 值,其中 32kHz、44.1kHz、48kHz 信号对应的单个 UI 值分别是 244.14ns、177.15ns、162.76ns。

传统数字音频传输过程中的抖动通常都是发生在数字信号从一个设备传输到另一个设备的时候,又被称为接口抖动。接口抖动源于输出设备本身固有的抖动特性,也会由传输线路中电缆的阻抗变化所引发。因此,一个信号由输出设备输出后,会带有其固有的抖动,然后经过传输链路上的每一个设备和每条电缆时,又会纳入这些设备的固有抖动以及电缆引发的抖动,形成所谓的传输抖动增益,最后到达接收设备。AES3 标准规定设备输出的固有抖动须低于 0.025UI,而传输时带来的正弦波抖动增益要小于 2dB。

对于接收设备而言,也存在着一个和抖动相关的重要技术指标,即抖动容限。因为随着抖动电平的增加,接收设备对传输信号的解码就会出现错误,导致解码失败,所以要定义一个接收设备产生错误前可容忍的最大抖动电平,这就是抖动容限。抖动容限的单位也用 UI 表示,它也和频率有关:抖动频率较低的时候,抖动容限较大。在 AES3 标准中,8kHz 以上的抖动频率对应 0.25UI 的抖动容限,200Hz 以下的抖动频率对应 10UI 的抖动容限。

抖动是数字信号传输中非常重要的问题,它能导致信号失真,但无法被消除,尤其是设备本身固有的抖动。而对于传输链路中的抖动干扰,可以通过电路设计改进和屏蔽等措施来减轻或者消除。

1.7 同步

在一个设备众多的音频传输系统中,保证这些设备的同步,是一个基本要求,否则传输信号可能不会被正确地接收。同步的方式有好几种,在模拟时代最常用的是时间码同步方式,也就是在一个系统中确定一个统一的时间码,所有的设备都以该时间码为基准,从相同的时间点开始操作。这种同步方式也被应用在如今的数字音频系统中,被称为“字时钟”(Word Clock)。

字时钟是一种采样频率的方波参考信号,用于在音频系统的不同设备中产生或转