

数字音响系统(一)

沈 赓

(中国科学院声学研究所 北京 100080)

1993年4月3日收到

数字录音机和激光唱片极大地改进了放音质量。九十年代初数字盒式磁带录音机和微型光学唱片的推出,使数字音响能够进入家庭放声领域。超大规模集成电路和数字信号处理技术的进展,不但使录音放音性能达到完美程度,而且也降低了数字音响设备的价格,数字音响替代模拟音响的前景已经明确,此外也突破了唱片只能放音不能录音的传统概念。本文按照家用数字音响的发展首先讨论数字音响系统的一般特点,然后分别介绍 CD 和 MD 等数字唱片放声以及 DAT 和 DCC 等数字磁带放声。最后谈了数字音响系统的前景。

一、引言

二十世纪五十年代实现了密纹唱片和七十年代采用了降噪系统制成了高质量盒式磁带录音机;接着开发了立体声放声,音质有了很大提高;但是录音方式没有改变,都是直接记录音频信号的波形,重放时用唱针扫描槽纹或者用放音磁头来拾取信号。对于唱片放声,在音质上由于唱针针尖和拾音器芯座问题引起的噪声、畸变、串扰、动态范围、抖晃率,经过多年努力改进,已经接近能够解决的极限。对于磁带放声,则采用了伺服、同步、定位、降噪和磁带制造工艺改进等高科技,使模拟式录音机在机械、电和磁特性等方面也达到了很高的技术指标。在这种情况下,高保真放声的进一步发展方向是数字化。所谓数字音响系统(简称数字音响)是指音频信号的数字化,在产生、变换、记录和传输音频信号过程中采用数字编码的音频信号并采用数字技术来处理,而在放声时还可利用数字技术实现根据程序规定的一些自动调节和控制功能。脉码调制(PCM)方法首先使用于电话卫星通信。数字信号的传输和记录过程比连续波形更好些,被量化了的采样值可以经编码构成一个新的信号,进一步采用数字技术进行

处理、记录和传输,它所带来的误差要比原信号直接处理少得多。并且还可以用处理波形信息来替代波形本身或它的某些特征,而经处理后的信息仍然可以用来重新恢复所期望的波形。实验表明,用 PCM 技术获得高质量的记录和重放音质是无容怀疑的。1969 年日本 NHK 研究所在一种四磁头视频录象机基础上,用两英寸磁带,采样频率为 47.25kHz,并使用 13bit 的线性编码开发了 PCM 录音机;稍后日本索尼公司作了改进,将它商品化。随后发展的数字音响系统广泛采用各种编码技术,它们的特点是用计算技术,由超大规模集成电路(VLSI)构成音响设备。1970 年前后,国际上开始研制数字式录音放音设备,便进入数字音频设备的实用化发展阶段。数字音响系统具有失真小、噪声低、动态范围大,多次复制不会下降质量等优点而受重视。但是由于价格昂贵,初期仅在广播、录音专业领域内使用。直到 1982 年开发了激光唱片(CD),1987 年开发了数字式盒带录音机(DAT),达到可接受的家用的价格而商品化,使家用音响系统的性能接近于模拟式专业音响系统的水平。其中 CD 发展顺利,并不断采用新技术提高性能,逐步降低价格。DAT 则由于价格昂贵,磁带商以版权无保障为由不予合作等原因发展较慢;直到 1989 年制定了“连

续复制管理系统 (SCMC)” 协定后才得到解决。90 年代初几乎同时开发了数字式微型盒带机 (DCC) 和微型光学唱片 (MD), 不但具有高性能, 而且价格适中, 因此用数字录音替代模拟录音的前景已经明确。当前数字音响的发展, 由于 VLSI 和数字信号处理 (DSP) 的进展, 计算技术和高密度记录和重放技术的应用, 使数字音响朝着集成化、自动化、智能化和全固态化方向发展。

二、数字音响系统的特点

1. 脉码调制系统的工作原理

对模拟信号经采样和量化并用二进制码表示的过程称为脉码调制 (PCM)。PCM 系统用固定的时间进行采样, 获得的采样值经量化用二进制码表示。由于音频信号有负值, 因此需要采用能表示负数的二进制码。初看起来, 采样和量化好象只取信号的一部分而不是象模拟系统那样处理整个电压波形; 并且用近似的量化值去代替。这时, 不可避免地会产生量化误差和舍入误差。但是, 如果我们用很高的采样频率并且用很小的分级去量化; 那么获得的一组量化取值可以恢复与原信号完全相同的波形。甚至这些离散采样值的集合要比连续波形更好些, 因为已经量化的采样值可以经编码而构成一个新的信号, 它可以用数字技术来处理、记录、传输并且还可以重新恢复成模拟信号。这些过程所带来的误差要比对原信号直接处理而引起的误差小得多。并且可以用处理波形信息来代替处理波形本身或它的某些特性。

通常用模数 (A/D) 转换器通过采样、量化和编码把一个模拟信号数字化。采样过程是确定观察信号的瞬时值, 通常采用等时间间隔的采样方式, 若模拟信号的幅度被周期地采样, 用 T 表示时间间隔, 则采样频率 f_s 为时间间隔的倒数,

$$f_s = \frac{1}{T} \quad (1)$$

图 1 示音频信号的采样, 其中 (a) 是模拟信号 $f(t)$ 的波形, (b) 是等时间间隔的采样脉

冲 $u(t)$, t 是脉冲宽度, T 是采样时间间隔, (c) 是采样后信号 $g(t)$, 其包络类似于模拟信号的波形。应该注意, 采样点太密会产生相关的和大量多余的数据, 从而不必要地增加需要处理的数据。如果采样点太少, 就会在原始数据中引起低频和低频分量的混淆。若信号中最高频率分量的频率为 f_h ; 根据采样定理, 采样

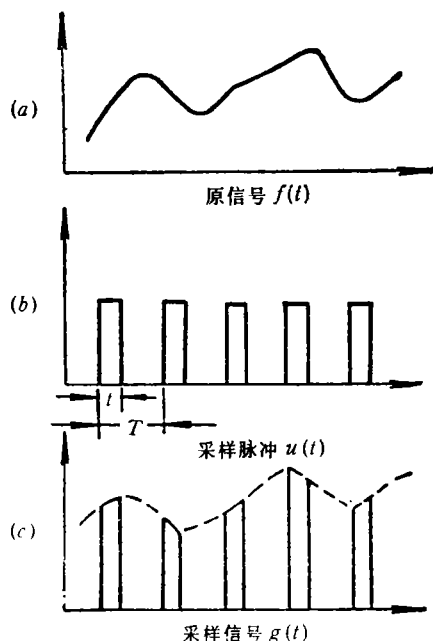


图 1 音频信号的采样

频率 f_s 至少应是 f_h 的两倍。对于高频分量的频率达 15—20kHz 声频范围内的高保真信号, 采样频率应该大于 30—40kHz。

模拟信号经采样后需进一步量化。量化是把连续变化的采样值转换为一组已选定的数值, 即用一组指定电平近似地表示连续的无限个电平。指定的电平数越多, 这种近似越精确, 如果量化电平数量很少就会出现失真。通常在量化过程中会产生近似误差称为量化噪声, 即模拟信号的一个特定值转换成一个稍有差异的数值, 这种幅度误差是随机的, 其效果就象是噪声。另外大于某个数值的信号会被限幅为该数值, 由于这种现象产生的噪声称过载噪声。

编码是对每个选定的量化电平给一个只有两个状态 (即“0”和“1”) 的二进制码, 这两个数

字的各种组合可以表示不同数值。通常把模拟信号量化为一组电平,每一电平用十进制数表示,然后再把这些十进制数表示为二进制码。因此量化可以把复杂信号用电码信号传输,而这种码最基本的信号由“0”和“1”组成。 n 比特(bit) 二进制码能够区别 2^n 个状态的信息,一个给定二进制码的比特数限定了能够表示的最大的二进制数,也限制了二进制码的分辨能力。 n 比特表示最大二进制数是 2^{n-1} ,而分辨力限制为 2^n 。因此比特分辨力为 $1/256$ 。在高保真放声中,至少要用 13bit,其分辨率才能满足数字音响的要求。相应的电平数为 8192 个,因为信号有正值和负值,所以只有 4096 个离散值。

数字信号可以通过解码来恢复模拟信号。每一个采样脉冲可以看作是一个窄的矩形脉冲。当脉冲通过一个理想低通滤波器时,采样得到的脉冲串将产生一系列响应,其中每一个响应都被延时相当于采样间隔的时间。这些响应的代数和构成的合成波形与被采样的模拟信号波形相同,如图 2 所示。图中给出了输入模拟信号的波形变化和解码后的合成波形。解码是在数模(D/A)转换器中完成的。解码后信号经低通滤波器去掉采样频率就可以恢复原来的模拟信号波形。图中还给出信号经采样后,在输出信号中出现的量化噪声。

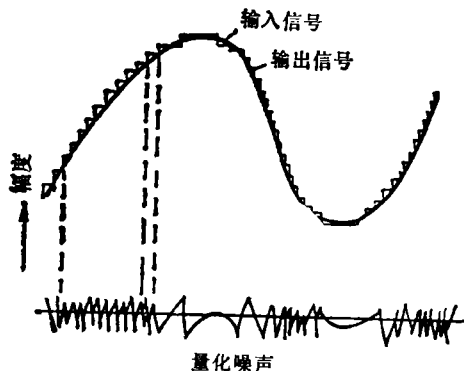


图 2 经采样和解码后输入和输出信号波形的比较

2. 数字编码音频信号的特性

早期对音频信号带宽的研究表明,在 40Hz 到 16kHz 频率范围内足以重放音乐节目而感

觉不到音质有变化。这就是目前调频广播采用的高保真放声带宽标准。最近为了发展数字音响,重新做了实验来检查人的可听声频率范围。实验结果表明,如果截止频率选择 20kHz,那么虽然受到低通滤波器相位特性的影响,主观上也感觉不出音质变坏。能够辨别出音质变化的临界频率大约在 15kHz 附近,因此设计数字音响系统时采样频率最好选择能复盖 20kHz 的带宽。在 PCM 设备中,由于采样定理的约束,采样频率必须比音频信号中最高频率分量的频率大两倍。鉴于人类听觉范围的频率上限为 20kHz,故对数字音频设备的采样频率通常选取 40—50kHz。因此对于双通路立体声系统的记录媒质要求的频率范围大约为 1.2MHz。信号采样后要量化并以二进制码表示。数字信号的动态范围为

$$D = 6n + 1.75 \quad (2)$$

式中 D 为动态范围, dB; n 为比特数。

假定数字编码音频信号的量化采用 13—16bit,则动态范围大约为 80—98dB。严格地说,量化噪声的功率是分布在很宽频率范围内,并且数字编码音频信号在解码时又必须经过一个低通滤波器,所以实际的动态范围与上述理论值稍有不同。增加量化比特数可以把这种噪声降低到容许水平。但是增加比特数是有限的,因为当输入信号幅度降低时量化噪声和信号的相关性非常明显,这种情况在 PCM 录音机中就以音质问题出现。如果用非线性量化或压缩扩展系统,则可以进一步降低量化噪声。在 PCM 设备中,由于脉冲波形可以截幅、整形并对误码纠错,因此系统的畸变很小,抗干扰能力很强。但是对频率很低的输入信号,例如 20Hz 信号,就能听出量化噪声级的变化,这种互调畸变令人感到讨厌。

通常对数字音频设备力求获得三个 90dB 的技术特性,即动态范围,信噪比和通路分离度都能达到或超过 90dB。人们的听觉强度范围大约为 0—120dB,但在实际环境中常遇到的声信号强度范围大约为 30—110dB。在音乐厅内听音乐的情况为 35—100dB,也就是说管弦乐

的动态范围大约为 65dB。人们在现场听音乐时有一种自动压缩动态范围的作用。当音量小时易受噪声掩蔽效应的影响,在音量很大时听觉系统又有保护作用,因此听到的声信号动态范围不是 35—100dB,而是 40—95dB,动态范围大约为 55dB。高质量模拟式音频设备可以达到这一指标,但是几乎没有余量。因此实际中它不能使用于最佳条件下。在听高保真度(Hi-Fi)音乐时,通常仅少数人;因此噪声较低,展宽的动态范围能显示出来。开始听数字音响系统的音质,完全有可能会不习惯;但是有较宽的运用范围还是有利的。也就是说使用数字音频设备使动态范围具有较多的储备量。

量化比特数决定了 PCM 系统的动态范围和信噪比。若用 13—16bit 编码,可以获得 80—98dB 的动态范围。如果采用线性编码,则动态范围与信噪比的数值一致。若采用非线性编码,两者就不一致。压缩和扩展技术能够改进信噪比和动态范围并且减少记录传输的信息量。例如 12bit 的非线性编码可以获得相当于 16bit 线性编码的效果。多通路传输时,串扰信号可以通过数字信号处理来消除,因此采用数字信号传输时可以把通路间串扰作用降低。目前模拟电路中的立体声分离度大约为 35dB,而数字音响系统则可以获得约 90dB 的通路分离度。这样的特性不但有助于立体声感觉,也会提高系统的音质。

3. 误码及其防止

数字音频信号的另一个特点是可以采取措施防止误码。在 PCM 信号传输和记录过程中,因为信息脱落而产生的误码会形成噪声,听

起来非常刺耳;但是可以采用误码检查和校正技术来防止。由于失落或变动引起的误码应该采用校正和误差补偿的方法来处理。解决误码常用的方法有下述四种:(1)使用纠错码将错码纠正,(2)使用检错码,检测出错码,然后用前后的正确信号插入替代,(3)同一信号记录多次,使用没有错误的信号,(4)采用交错法纠正,因为误码是成串出现的,交错法将连续的误码重新排列成随机误码以减弱对音质的影响。如果这几种方法结合起来使用,其效果将会更好。

4. 兼容性

最后简单讨论一下数字信号的兼容性问题。数字音响系统的标准化和兼容性将会影响其发展。目前使用着两种数字音频系统,一种是英国的 PCM 传输系统,另一种是日本以磁带录象机为基础的 PCM 录音机。因此数字音频设备选用采样频率的标准化将受这两方面影响。为了与彩色电视机兼容,还应考虑彩色电视制式。例如日本对 NTSC 制录象机伴音的采样频率采用了 44.0559 kHz,对 PAL 制录象机则用 44.1kHz,而欧洲卫星通信网络建议采样频率用 48kHz,激光唱片则是以激光视频唱片为基础,它的采样频率也未标准化,有用 44.1 kHz,也有采用 47kHz;编码分别采用 12, 14 和 16bit。目前编码采用 16bit 在技术上是可能的,而大多数微计算机使用 16bit 为标准码形。为了在家庭中实现由数字音响系统,彩色录象系统和个人微计算机系统组成的所谓家庭数字信息中心,统一来考虑它们之间的兼容性和标准化是非常必要的。(待续)