

SHUZIYINPIN
JISHUJICHU

数字音频技术基础

○ 刘嘉彧◎编著

天津科学技术出版社

SHUZIYINPIN
JISHUJICHU

数字音频技术基础

刘嘉彧◎编著

<DEBFHACKL



天津科学技术出版社

图书在版编目 (CIP) 数据

数字音频技术基础/刘嘉彧编著. —天津: 天津科学技术出版社,
2008. 10

ISBN 978-7-5308-3984-3

I. 数… II. 刘… III. 数字技术-应用-音频设备 IV. TN912. 2

中国版本图书馆 CIP 数据核字 (2008) 第 158985 号

责任编辑: 刘 颖

责任印制: 王 莹

天津科学技术出版社出版

出版人: 胡振泰

天津市西康路 35 号 邮编 300051

电话: (022) 23332400 (编辑室) 23332393 (发行部)

网址: www.tjkjcs.com.cn

新华书店经销

廊坊市博远印务有限公司印刷

开本 889×1194 1/32 印张 4.5 字数 122 000

2008 年 10 月第 1 版第 1 次印刷

定价: 25.00 元

目 录

第一章 音频信号的数字化 /1

- 第一节 音频信号数字化的历史 /1
- 第二节 模拟信号的数字化 /2
- 第三节 调制方式 /11
- 第四节 检错与纠错 /16

第二章 数字录音的传声器布置 /24

第三章 数字调音台 /34

第四章 数字录音机 /46

- 第一节 数字磁带录音机的种类 /46
- 第二节 数字磁带录音机的构成 /47
- 第三节 固定磁头录音机与旋转磁头录音机的比较 /48
- 第四节 固定磁头方式录音机 /50
- 第五节 旋转磁头方式录音机 /62
- 第六节 小型盒式数字磁带录音机(DAT) /63
- 第七节 数字磁带录音机使用注意事项 /82
- 第八节 磁带以外的数字录音机概要 /83

第五章 光记录技术 /94

第一节 光盘存储器 /94

第二节 CD唱机的组成 /99

第三节 光盘 /101

第四节 光盘的利用形式 /124

第六章 数字音频工作站 /126

第一节 数字音频工作站的组成 /126

第二节 数字音频工作站的特点 /129

第三节 数字音频工作站的功能 /132

第四节 数字音频工作站的系统构成 /133

参考文献 /140

第一章 音频信号的数字化

第一节 音频信号数字化的历史

音频技术从 1877 年美国爱迪生发明圆筒留声机算起,已经经历了一百多年。在这其中的前一百年间一直是以模拟音频技术(Analogue Audio Technique)为主的。1953 年日本等国开始研制数字录音机,直到 1976 年日本市场上开始出售与 β -max 型家用录像机配合的 PCM 音频适配器(Adapter),才实现了数字录音。可以认为,音频技术在经历整整一百个年头后,开始进入数字音频技术(Digital Audio Technique)的新阶段。目前数字延时器、混响器等音频处理设备、激光唱片(CD)、专业用 PCM 录音机、小型盒式磁带录音机(DAT)等以诱人的优良性能,正在不断取代模拟设备,加快了音频技术数字化的过程。

目前,在音频领域,CD 和它的软件、硬件都已经得到了广泛的普及;继专业用的数字磁带录音机之后,家用小型盒式磁带录音机(DAT)也开始出售;PCM 技术被引用到广播中继、卫星直播以及 8 mm VTR 的声音系统;用于声场控制的效果器也采用了数字信号处理器(DSP)装置。数字音频从此进入了真正的实用阶段。

数字信号处理的概念始于 17 世纪发达起来的数值解析。但是使数字信号达到今天这样广泛应用的原动力是:

第一, 20 世纪 40、50 年代信息论的发展和电子计算机的实用化;

第二, 以 LSI(大规模集成电路)为中心的半导体技术的飞跃发展,特别是数字集成电路性能的提高和价格的下降令人瞩目。

无论你是否喜欢,数字化的浪潮正在冲击过来。

数字信号处理技术在音频领域的应用大致可以划分如下。

- ① 应用电子计算机的信号处理。
- ② 模拟处理、模拟信号的数字化。
- ③ 新型数字处理系统的研制。

以上划分并不一定与历史流程相对应,但一开始是利用电子计算机进行信号处理的;然后为了提高设备性能和操作方便或降低价格而逐步开始单个设备的数字化;而后在 LSI 的飞跃发展等基础上相继产生了新型系统。

在音频领域,最先试用数字化技术的硬件是厅堂及演播室的一些效果设备;数字延时器、混响器于 20 世纪 60 年代后期在美国和欧洲相继诞生。信号传输系统的数字化始于 1970 年前后,它被用于广播中继、PCM 录音、测量数据的收集等方面。

第二节 模拟信号的数字化

将模拟信号变成数字信号,需要经过取样、量化和编码三个基本步骤。

一、取样与取样频率

模拟量是指在时间上和幅度上都是连续的量。例如声波就是声压幅度随时间连续变化的模拟量,它由传声器转换成音频信号后,也是时间和电压幅度都连续变化的模拟信号。如图 1-1(a)所示。

如果幅度只是一些特定值的阶梯波,则是时间连续而幅度不连续(离散)的信号。如图 1-1(b)所示。

幅度连续而时间不连续(离散)的信号是脉冲幅度调制(PAM)信号。如图 1-1(c)所示。

时间和幅度都不连续(都离散)的信号是数字信号。如图 1-1(d)所示。

要将模拟信号数字化,需要取样和量化过程,通常用取样保持电路(S/H)来进行取样,然后用 A/D 变换器量化。

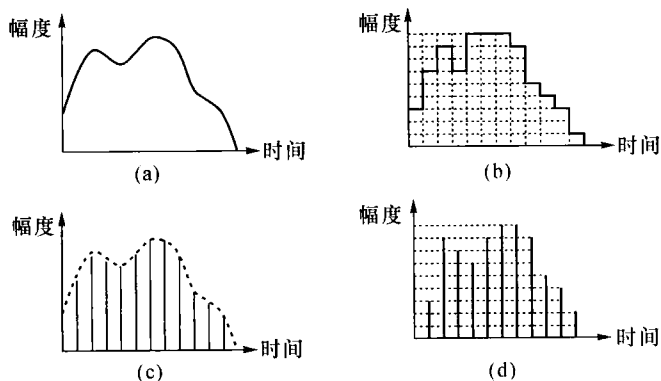


图 1-1 信号的分类

所谓取样,顾名思义,就是将连续信号在某时刻的数值读取出来的过程。取样后的一系列样值称为取样序列;取样的时间间隔称为取样周期(Period of sampling);每秒内抽取的次数称为取样频率(Frequency of sampling)。根据奈奎斯特(Nyquist)的取样定理,当取样频率 f_s 为被取样信号最高频率 f_b 的两倍时,则被取样信号可以被恢复,即 $f_s \geq 2f_b$ 。由于音频信号的最高频率为 20 kHz,所以优质的 A/D 变换,取样频率应大于 40 kHz。对于最高频率限制到 15 kHz 的一般质量信号的 A/D 变换,取样频率也应大于 30 kHz。目前有四种取样频率:①48 kHz;②44.1 kHz;③44.056 kHz;④32 kHz。其中,①、②、④为音频工程协会(AES)推荐的取样频率。

二、量化与量化噪声

量化就是将具有任意数值的取样值,经“四舍五入”的方法转换成有限的离散值的过程。

在数字音频技术中,我们采用二进制表示一个数,即用“1”和“0”来表示一个数,逢 2 进 1。因为“1”可以相当于电路接通,“0”可

以相当于电路断开。这就使利用电路来进行二进制数的运算有了可能。

采用二进制时的有效位数称为比特数(bit number)或位数。由取样定理知道,如果取样频率能满足这个定理,取样就能完全恢复原波形;但要真正完全恢复原波形,则需要无穷多位数。在通常的数字系统中,每个取样点都会产生取样误差,并且存在与这种舍入误差相应的失真和噪声,这种噪声或失真被称为量化噪声(Quantization noise)或量化失真(Quantization distortion)。

为了使量化噪声转变为与输入不相关的白噪声,就要先在输入信号上加上一种称为高频脉动信号的噪声后再加以量化,如图 1-2 所示。

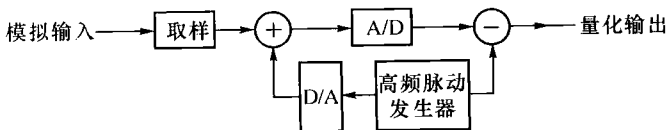
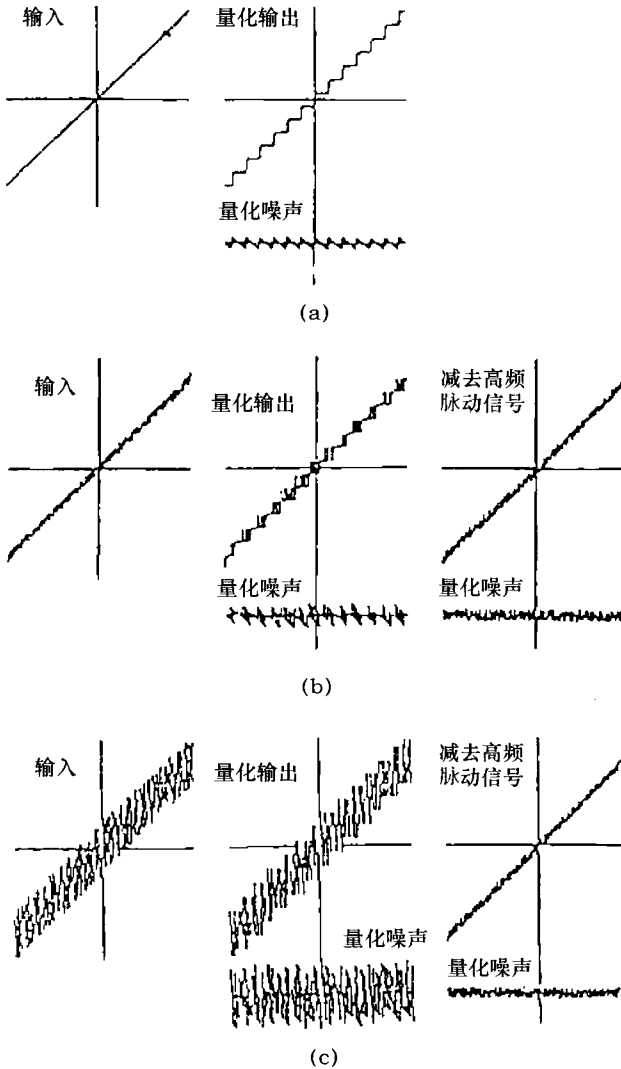


图 1-2 高频脉动信号的叠加和去除

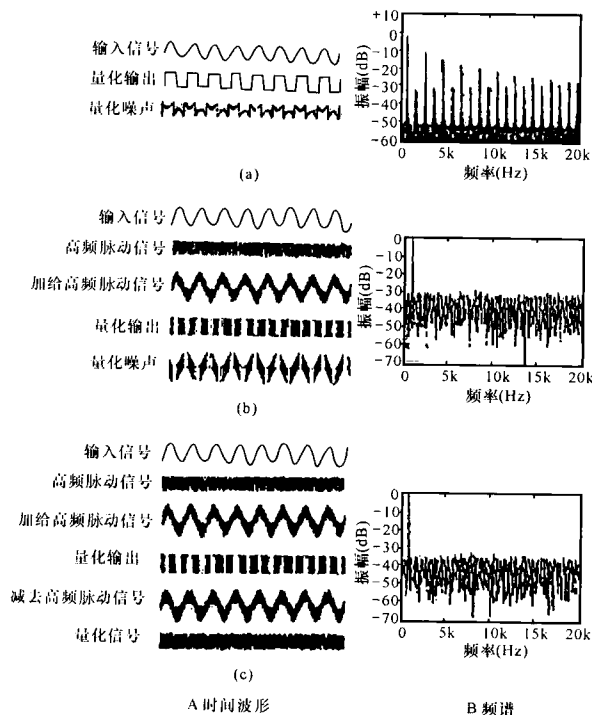
在理论上,量化前在信号上叠加一种在量化步长上均匀分布的噪声(高频脉动信号),再从量化后的信号中减去相同的高频脉动信号,就可使量化噪声成为在幅度上均匀分布的白噪声。如图 1-3、图 1-4 所示。

取样量化彼此关联。通常可以认为,取样频率决定信号频带,量化特性决定动态范围,它们是分别各自决定的。但实际上两者有着密切的关系。的确,取样频率决定了能处理的频带,量化步级数越多,在数字上的动态范围越大。但是如上节所述,采用适当的高频脉动信号进行处理,则量化噪声将在频带内均匀分布。因此,可以这样认为,当信号频带相同时,取样频率越高,量化噪声分布频带越广,于是在一定的频带内的噪声将变小,信号频带内分布的量化噪声功率也将变小。



(a) 没有高频脉动时 (b) 加有 $\pm\Delta/2$ 高频脉动时 (c) 加有 $\pm 2\Delta$ 高频脉动时

图 1-3 高频脉动信号分布以及量化噪声与量化输出间的关系



(a)未加高频脉动信号的量化 (b)叠加高频脉动信号的量化 (c)量化后去除高频脉动信号

图 1-4 正弦波的量化

三、编码与高效率编码

(一) 编码

二进制使用“0”和“1”两个数字，逢 2 进 1，用二进制数表示某一数值时，这个二进制数成为字(word)；这种表示过程称为二进制编码。字内各个位(bit)最左端称为最高位(MSB: the most significant bit)，以下依次称为第二位(2SB)、第三位(3SB)……最右端的位称为最低位(LSB: the least significant bit)。

常用的各种二进制码如表 1-1 所示。其中自然二进制码(natural binary code)只有正数，偏移二进制码(offset binary code)将-8 作为

☆第一章 音频信号的数字化☆

0,而编码规律与自然二进制码相同;补码(2's complement code)是将偏移二进制码的最高位做了相反的改变得出的码。在各种数字设备之间进行信号交换时,如果有一台设备出现故障,那么这台设备内所有的位都会变为1或0,但补码中所有位为1或0时,量化电平相当于0附近,所以不会将这种故障的损害给予其他设备。因此,音频工程学会(AES)规定,在专业数字设备中,应使用补码。

表 1-1 几种二进制码

十进制码	二进制码		
	自然二进制码	偏移二进制码	补码
+7	111	1111	0111
+6	110	1110	0110
+5	101	1101	0101
+4	100	1100	0100
+3	011	1011	0011
+2	010	1010	0010
+1	001	1001	0001
0	000	1000	0000
-1		0111	1111
-2		0110	1110
-3		0101	1101
-4		0100	1100
-5		0011	1011
-6		0010	1010
-7		0001	1001
-8		0000	1000

(二)香农的杰出定理

如果信号的平均信息量小于所用传输通路的传输容量,则存在着能使信号正确传输的编码方法。这就是信息论先驱者香农得出的定理。

这个定理指出,要有效地利用传输通路来高效率地传输信号,就必须知道信号的信息量和传输通路的容量,以及找出最佳的编码方法。

(三) 高效率编码

关于传输通路的节约,一种方法是除去信号所具有的冗余度而不使信号变坏;另一种方法是利用听觉特性,允许信号有一定程度的变坏。也就是说前者是按照信号实际的信息量,减少传输容量,使信号信息量与传输容量相匹配的技术;而后者则是将数据本身截去一部分,即数据压缩技术。

1. 利用冗余度节约传输通路

(1) 传输通路的传输容量。

传输系统的传输容量由系统的带宽和动态范围决定。当所能传输的信号功率为 P , 带宽为 W , 传输通路白色热噪声功率为 N 时, 即动态范围为 $D = (P + N) / N$ 时, 该传输通路的传输容量 $C = W \log 2D$ (b/s)。

(2) 预测编码。

预测编码如图 1-5 所示:

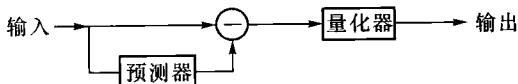


图 1-5 预测编码的构成

它是从过去的几个取样值预测出现在的取样值,并将真正的取样值与预测值之差(称为预测误差)进行编码后再进行传输的方法。

(3) 熵编码。

在数字化的信号中,通常小电平比大电平出现的机会要多,因此,可以像莫尔斯电报码那样,对出现频率高的电平用短码,出现频率低的电平用长码,从而进一步节约比特率。这种编码称为熵编码。

2. 利用数据压缩技术节约传输容量

前面介绍了利用信号冗余度,无失真地充分利用传输通路的方法。下面研究一下利用人耳的听觉特性,允许某种程度失真来节约传输通路的问题。

(1) 瞬时压扩与准瞬时压扩。

众所周知,音乐及人声等音频信号的振幅是随时间大幅度变

☆ 第一章 音频信号的数字化 ☆

化的。利用这一性质,在输入时将较宽的动态范围进行压缩,而输出时再予以扩张,这种按信号的振幅改变输入输出特性的过程称为压缩扩张,简称“压扩”。压扩操作可在模拟阶段进行,也可在A/D变换或D/A变换时进行,还可以在数字阶段通过数字运算来进行。杜比及dBX降噪是在模拟阶段的压扩。

压扩可分为相应于每个采样的幅度来进行压缩扩张的瞬时压扩和根据对某一范围内信号的检测结果来进行的准瞬时压扩。

图 1-6 所示为两种压扩方式编码的构成。

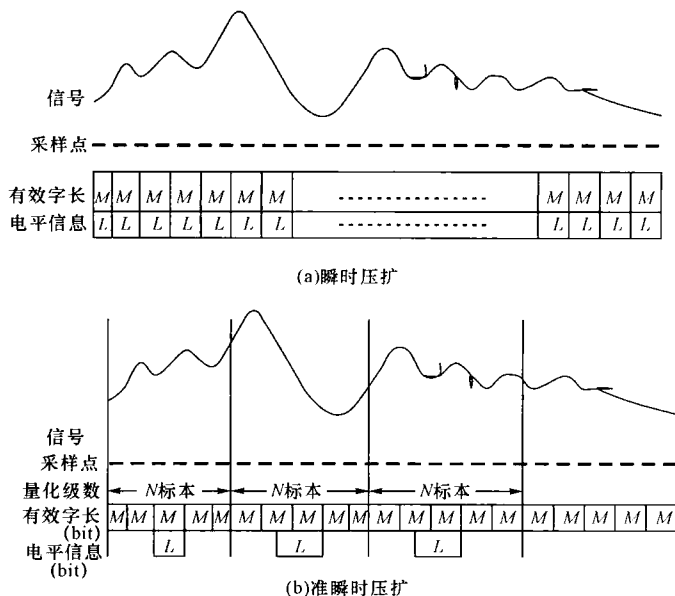


图 1-6 瞬时压扩与准瞬时压扩

数字阶段的瞬时压扩方式中,需要对每个取样传输 M bit 的有效数字和 L bit 的电平信息,即一个取样要用 $M+L$ bit;而准瞬时压扩方式中,以每 N 个取样形成的一个块来决定量化步长,用此步长对块内所有取样进行量化,因此每个取样需要 M bit 有效数字,每 N 个取样需要 L bit 的电平信息,即传输一个取样所需的比特数为 $(L/N) + M$ (bit/取样)。 N 取值较大,就可以大幅度节约比特数。

(2) 动态预加重。

准瞬时压扩方式是按块进行电平控制的,如果再按每块的频率成分改变预加重成分则称为动态预加重。动态预加重方式需要在电平信息之外传输预加重选择信息,即在预先规定的多个预加重特性中选择一种特性,以便在解调时加给去加重来恢复原频率特性。

(3) 自适应编码。

① 自适应 PCM(APCM)。

如前所述,音频信号的振幅和频率分布随时间有比较平缓的变化,但变化的幅度却很大。按照附近信号的性质来改变量化步长的编码方法就是自适应 PCM,即它是一种量化级大小跟随相邻信号幅度变化的编码方法。准瞬时压扩及动态预加重可以看成是一种 APCM。

② 自适应差分 PCM(ADPCM)。

自适应差分 PCM 是将自适应量化步长引入差分 PCM 中的编码,即不直接将信号量化,而是对信号与预测值之差加以自适应量化,比 APCM 效率更高,作为中等程度音质的高效率编码是很有效的。这种自适应编码广泛用于多功能电话机中信息录音等时间短又不使用磁带的固体录音机。

(四) 高效率编码的比较

表 1-2 所示为上述各种高效率编码应用于钢琴独奏时的例子。这些高效率编码都是为了节约传输通路而允许有一定程度失真和噪声的。表中列出了对 4 bit、5 bit 编码的信噪比和主观评价的比较结果。

准瞬时压扩比瞬时压扩改善 3~4 dB;APCM 比准瞬时压扩改善 3~4 dB;而 ADPCM 与 APCM 相比,信噪比有 15~16 dB 的大幅度改善,主观评价也有大幅度提高。由此可知,ADPCM 是适合于听觉特性的有效节约传输通路的方法。

表 1-2 采用各种编码时的音质比较

编码方法	4 bit 取样		5 bit 取样	
	SN 比(dB)	主观音质	SN 比(dB)	主观音质
瞬时压扩	9.9	1	14.3	2
准瞬时压扩	12.6	1	18.4	2
APCM	16.1	2	21.1	3
ADPCM	31.1	4	36.3	5

第三节 调制方式

调制是一种对信号进行变换的手段,调制可将信号变换为适合于传输和记录的形式。数字信号的传输、记录方式有:将数字化的信号波形直接或按某种规则进行变换后再传输、记录的方法,对载波进行调制后再进行传输、记录的方法。前者包括 NRZ、NRZI、PE、MFM、GCR、EFM、8-10 变换等,后者包括与模拟信号的调幅(AM)、调频(FM)、调相(PM)相对应的数字信号的振幅键控(ASK)、移频键控(FSK)、移相键控(PSK)以及将调幅调相组合起来的 QAZ 等。

一、希望的条件

在信号的传输中,如何将信息尽可能多地正确快速传输是很重要的问题。如果在单位时间内能够传输大量的信息,就可以节约传输经费。这就是调制的目的之一。

另一方面,在传输通路内多少会混入噪声,使接收端接收的信号与发送端发出的信号形状不同,导致信号质量下降。因此调制的目的之一是克服噪声,改善通信质量。

确定调制方式时,必须考虑传输通路和记录系统的特性。下面研究对记录用调制方式所希望的条件。

1. 容易提取比特同步信息

为了能从接收或重放信号中提取读出信息所需用的时钟信号,希望使用不会在调制波形中连续产生长时间空白的调制方式。

2. 不易受直流截止特性的影响

在磁记录这种微分信号检出型的重放系统中传输直流很困难,而且在传输通路中有时需要叠加直流,因此希望调制后的信号不含直流分量或希望直流变化要小。

3. 所需的传输频带宽度要小

因为记录、重放,传输通路的传输特性通常是频率越高,衰减量越大,所以希望调制所需要的频带宽度尽可能小。

4. 抗噪声,耐抖动

考虑到传输、记录系统的噪声及偏移抖动,希望使用能使解调有一定裕度的调制方式。

二、各种调制方式

图 1-7 所示为常用的几种调制方式波形图。下面对各种调制方式的调制规则做简单说明。

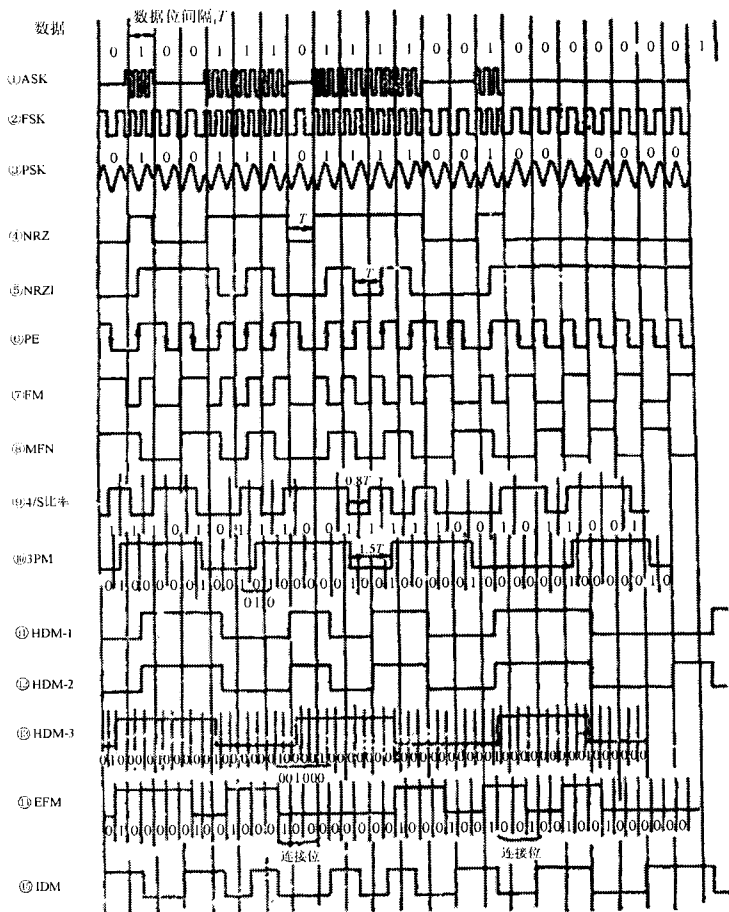


图 1-7 几种调制方式的调制波形