

信息与电子工程导论

Introduction to Information Science and Electronic Engineering

2.2 模拟和数字

章献民 主编

浙江大学出版社

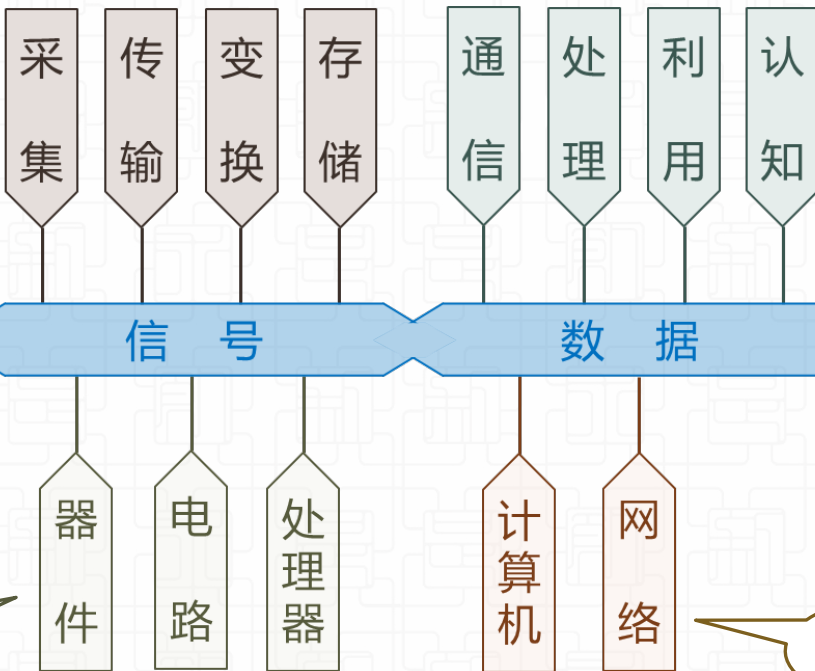
2023年9月

知识图谱

- 2.1 时域和频域
- 2.2 模拟和数字
- 2.3 编码和调制
- 2.4 电磁场与波

2 信号与数据

场与波



3 电子器件与电路

- 3.1 电路模型和基本定律
- 3.2 晶体管 and 集成电路
- 3.3 集成运算放大器

4 逻辑与数字系统

- 4.1 数字逻辑和电路
- 4.2 组合逻辑和时序逻辑
- 4.3 微处理器和计算机系统
- 4.4 嵌入式系统
- 4.5 EDA技术

1 信息与信息技术概述

- 1.1 信息
- 1.2 信息科学技术概述
- 1.3 知识图谱

5 互联与计算

- 5.1 通信与网络
- 5.2 物联与数联
- 5.3 计算与智能

内容提要

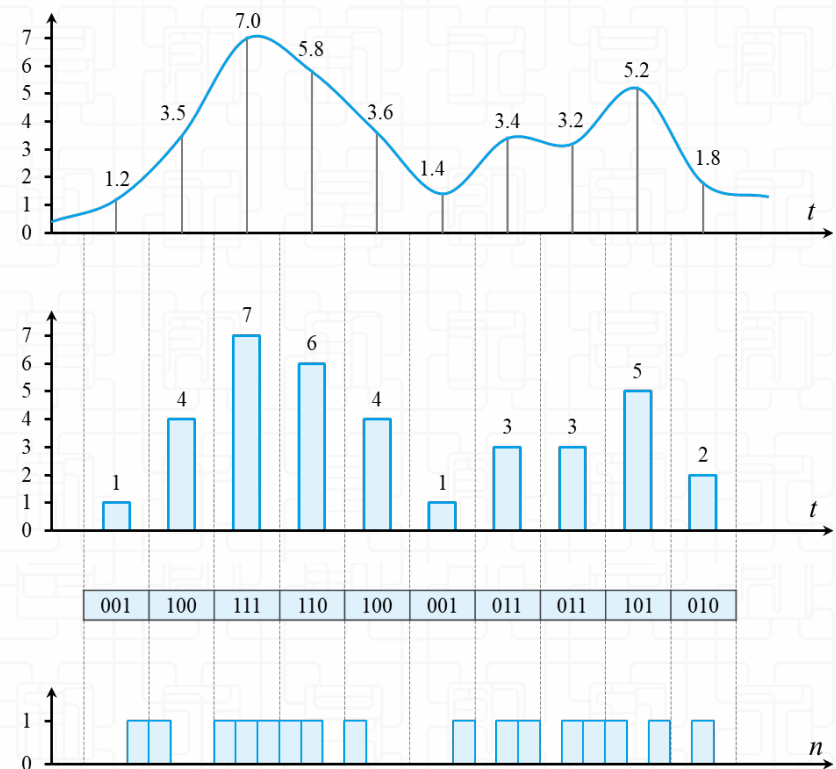
❖ 数据的数字化

❖ 信号的数字化

- 采样, 奈奎斯特采样定理
- 量化
- 编码

❖ 数字信号的优势

❖ 数字信号传输



模拟数据和数字数据

❖ 数据按表现形式可分为：

- 模拟数据和数字数据两种形式。

❖ 模拟数据 (analog data)

- 具有可连续变化的值，通常用各种测量单位表示
- 时间、身高、体重、温度、说话的声音等。

❖ 数字数据 (digital data)

- 具有离散值且可以计数的。
- 学生人数，电灯开关的状态等。

计算机中的数据表示

❖ “0” 和 “1” 的世界

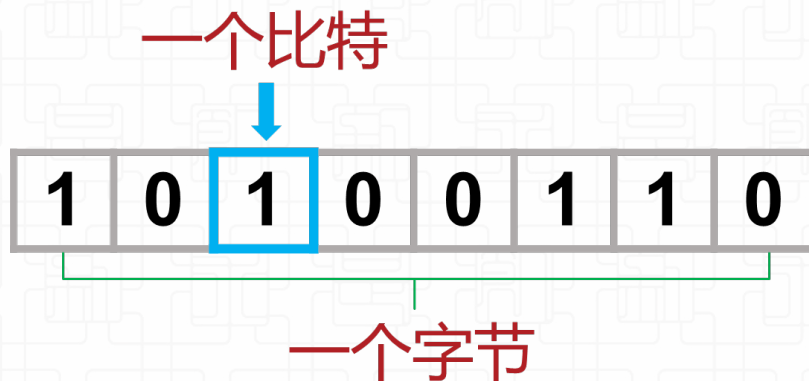
- 容易表示
- 运算简单

❖ 数在计算机内的表示，要涉及到数的长度和符号如何确定、小数点如何表示等问题。

❖ 由于二进制数的每一位数（0或1）是用电子器件的两种稳定状态来表示的，因此，**二进制位 (bit)** 是最小信息单位，一个数的长度按二进制位数（即 bit 数）来计算。

❖ **字节 (byte)** 是计算机信息技术用于计量存储容量的一种计量单位，也表示一些计算机编程语言中的数据类型和语言字符。一个字节存储8位无符号数，储存的数值范围为0-255。

❖ **字 (word)** 是计算机一次处理数据的最大单位，由若干个字节组成。字的位数叫做字长 (word size)，即计算机处理器一次处理二进制代码的位数。



ASCII码 (American Standard Code for Information Interchange)

❖ 每字符用 7 位二进制数表示，存放时占 8 位即一个字节 (byte)

高3位 低4位	000	001	010	011	100	101	110	111
0000	NUL	DEL	SP	0	@	P	.	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	DOT	DC4	\$	4	D	T	d	t
0101	ENG	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	'	7	G	W	g	w
1000	BS	CAN	(8	H	X	h	x
1001	HT	EM)	9	I	Y	l	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[k	{
1100	FF	FS	,	<	L	\	l	
1101	CR	GS	-	=	M]	m	}
1110	SO	RS	.	>	N	↑	n	~
1111	SI	US	/	?	O	↓	o	DEL

汉字编码

❖ GB2312编码

- 1980年发布的简体中文汉字编码国家标准，采用双字节编码。

❖ BIG5编码

- 台湾地区繁体中文标准字符集。

❖ GBK编码

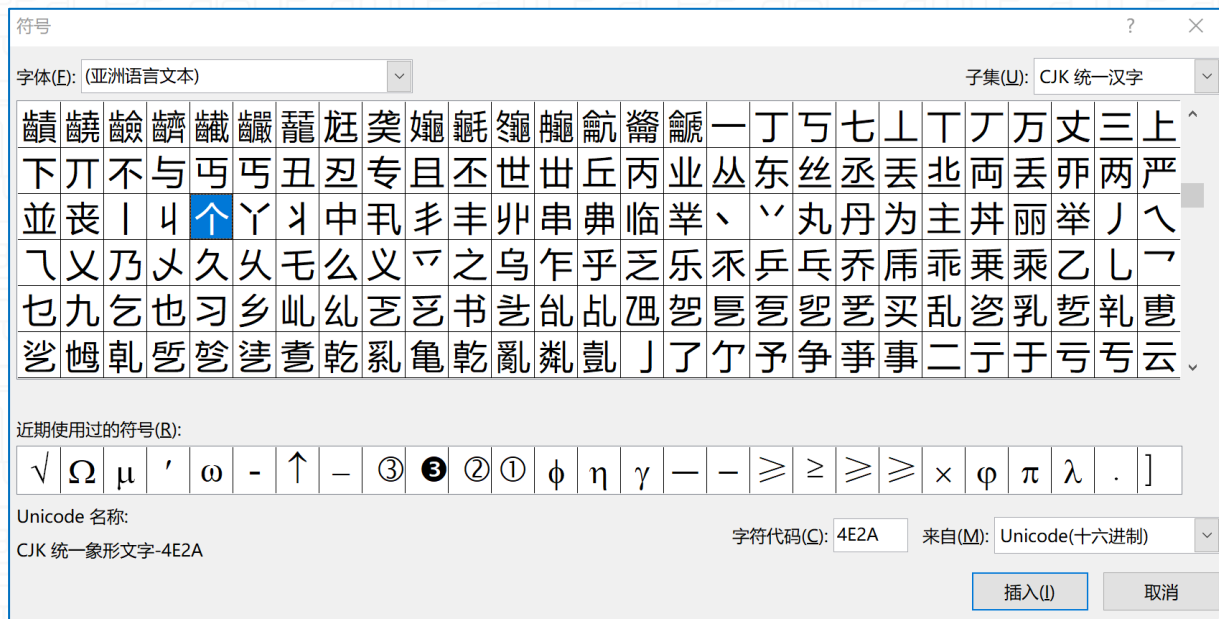
- 1995年发布，对GB2312的扩充，采用双字节编码。

❖ GB18030编码

- 2000年发布，采用单字节、双字节和四字节3种方式对字符编码。兼容GBK和GB2312。

❖ Unicode编码

- 国际标准字符集，将世界各种语言的每个字符有唯一的编码，以满足跨语言、跨平台的文本信息转换。



图像的数据表示

❖ 每个小方格称为一个像素，方格纵横乘积称为分辨率，例如：800×600

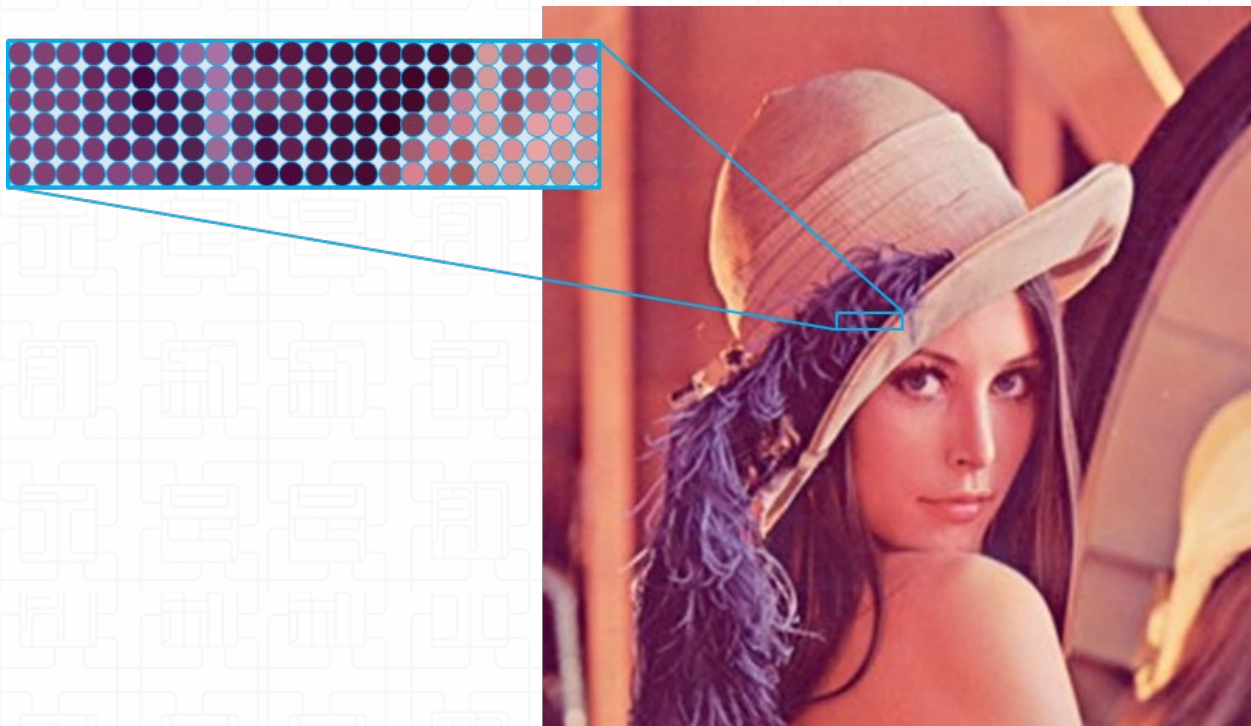
黑白图片

Black White

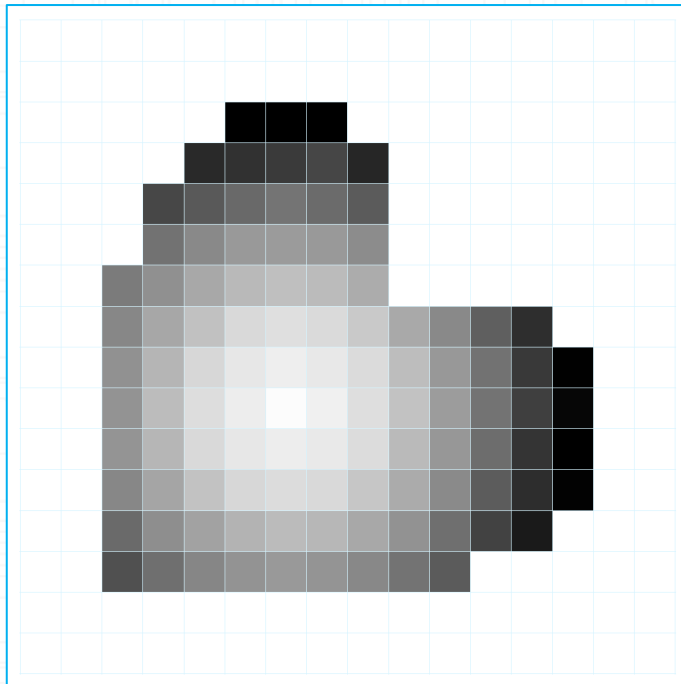
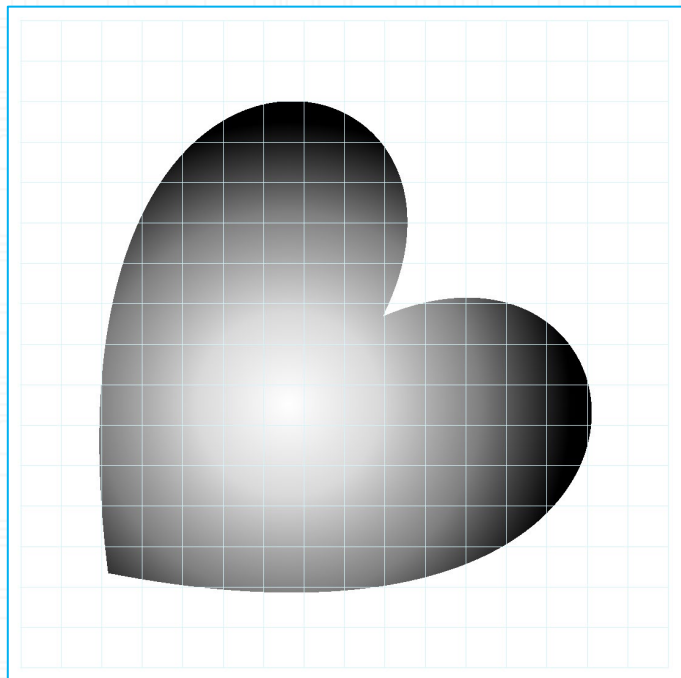
彩色图片

RGB三原色

Red Green Blue



举例：图像的数字化



0	0	0			
41	49	58	70	38	
71	89	105	116	107	91
114	137	153	155	153	140
123	144	168	185	191	187
135	167	193	217	223	218
145	181	215	231	238	232
147	188	221	237	252	240
148	182	217	231	237	233
135	165	194	215	220	217
106	142	162	179	187	183
80	111	134	147	153	148

- ❖ 主要包括取样和量化这两个过程。
- ❖ 取样过程是使图像空间坐标数字化，所形成的微小方格称为像素点。
- ❖ 量化过程是使图像函数值（灰度值）数字化。

内容提要

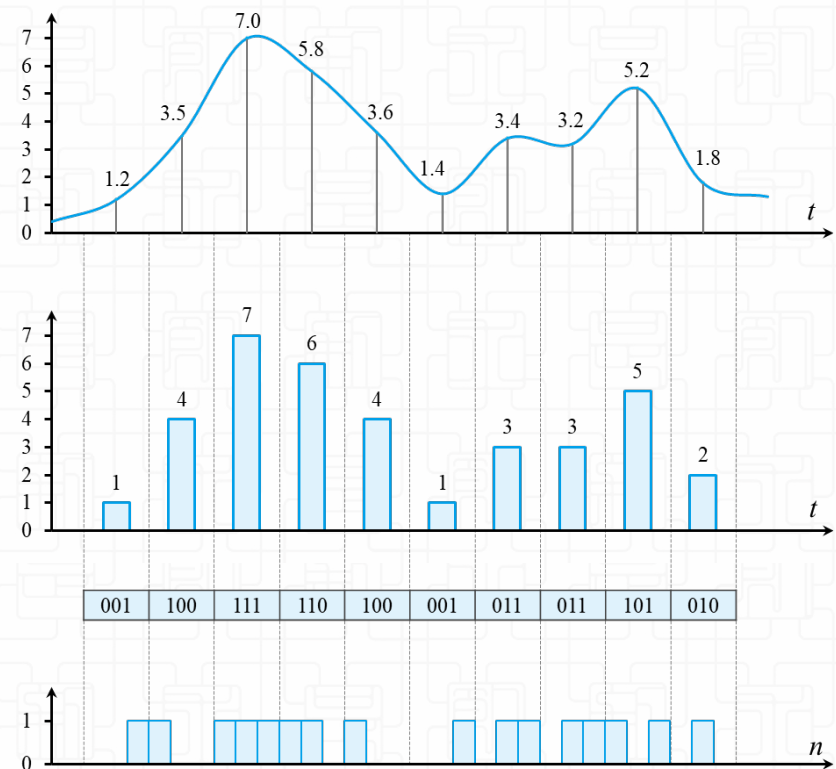
❖ 数据的数字化

❖ 信号的数字化

- 采样, 奈奎斯特采样定理
- 量化
- 编码

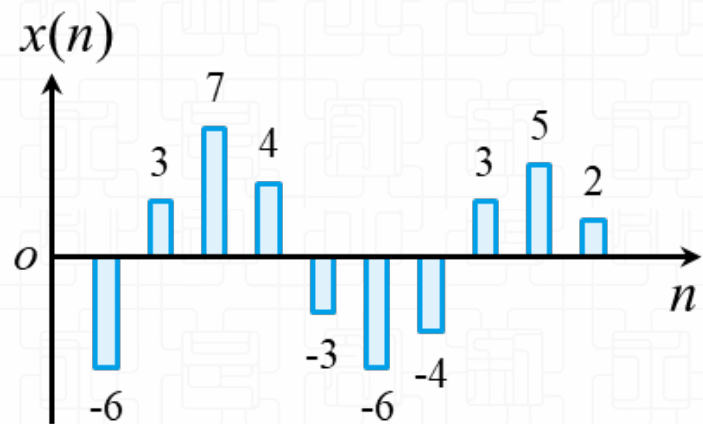
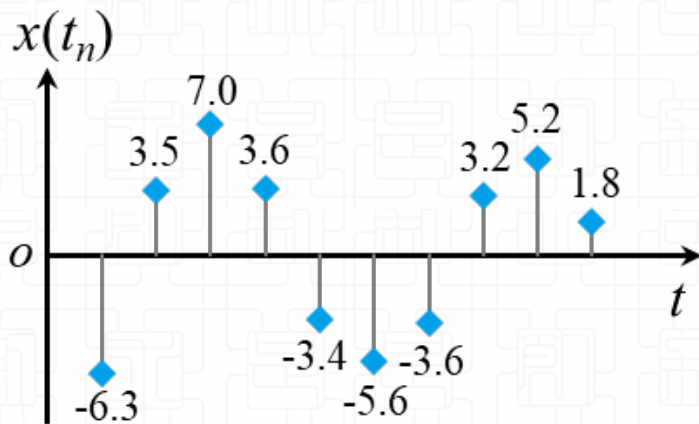
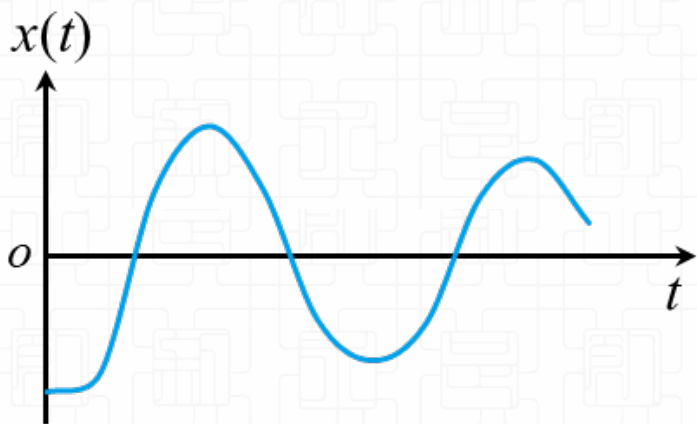
❖ 数字信号的优势

❖ 数字信号传输



模拟信号

- ❖ 模拟信号是指信息参数在给定范围内表现为**连续的信号**，或在一段连续的时间间隔内，其代表信息的特征量可以在**任意瞬间呈现为任意数值**的信号，它在一定的时间范围内可以有**无限多个不同的取值**。
- ❖ 在自然界中，我们可以感知的，在**时间和幅值**上都是**连续的物理量**称为模拟信号。
- ❖ 在电学中，用传感器将这样的物理量转变成为电信号，通常用**连续变化的电压值或电流值**表示。



模拟信号的数字化

❖ 三个关键过程：采样、量化和编码

❖ 采样

- 将模拟信号进行抽样，按照**采样定理**的要求，将时间上连续、幅度上也连续的模拟信号变换成时间上**离散**、但幅度上仍连续的已采样信号，采样完成模拟信号在时间上的**离散化**。

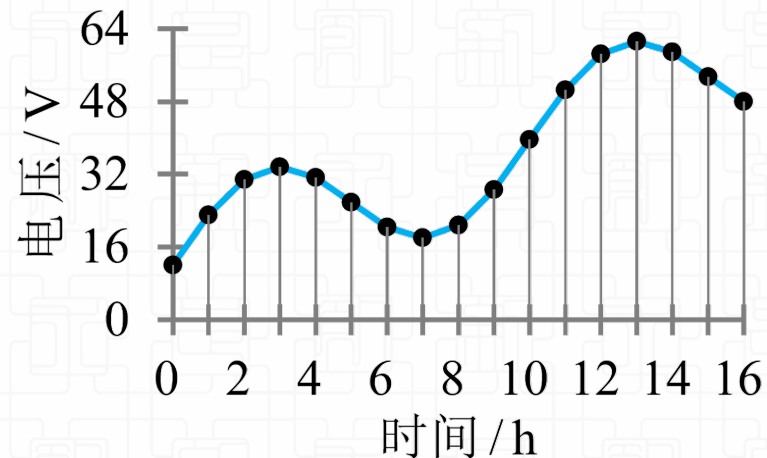
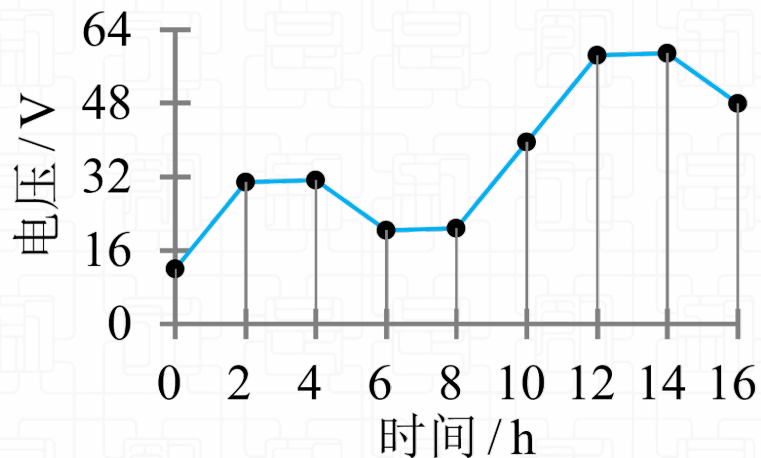
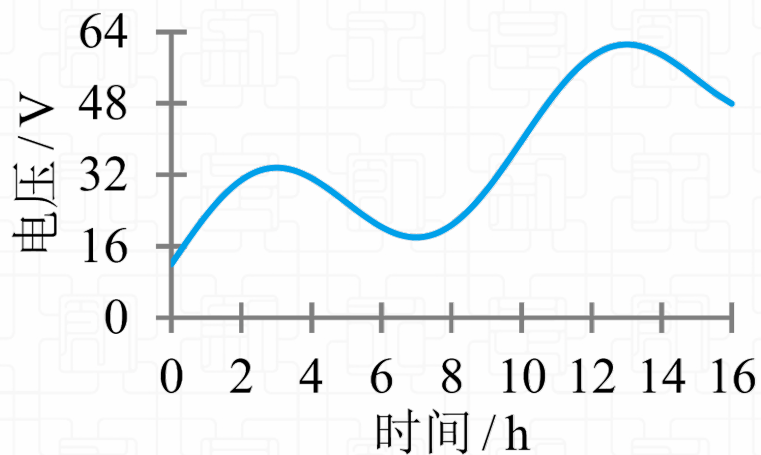
❖ 量化

- 用预先规定好了的**有限个电平值**来表示模拟抽样值，量化完成模拟信号在**幅度上的离散化**，信号经过抽样和量化后才能变成数字信号。

❖ 编码

- 通常采用**二进制编码**，即用 N 位二进制代码来表示量化值。

采样



- ❖ 在一系列离散的时间点上对物理量以一定的分辨率进行取值（采样），得到一系列离散的数字量
- ❖ 读取间隔的时间越短，越接近真实波形

采样定理（取样定理、抽样定理）

- ❖ 1928年美国贝尔实验室**哈利·奈奎斯特**推出采样定理，因此称为**奈奎斯特采样定理**。
- ❖ 1933年苏联数学家**科捷利尼科夫**首次用公式严格地表述这一定理，因此在前苏联文献中称为科捷利尼科夫采样定理。
- ❖ 1948年香农对这一定理加以明确地说明并正式作为定理引用，因此在许多文献中又称为香农采样定理。
- ❖ 采样定理有许多表述形式，但最基本的表述方式是**时域采样定理**和**频域采样定理**。



Harry Nyquist
Feb 7, 1889—Apr 4, 1976



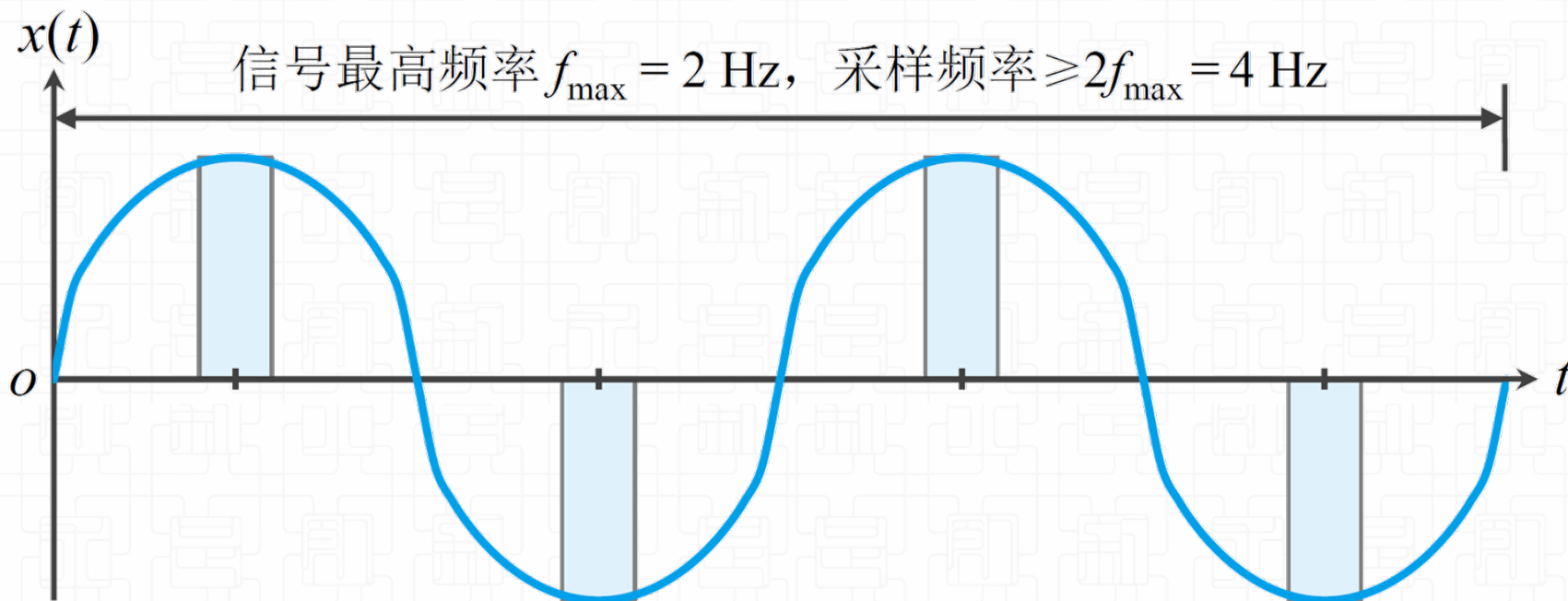
Aleksandr Petrovich Kotelnikov
Oct 20, 1865—Mar 6, 1944

时域采样定理的两种表述

- ❖ 频带为 X 的连续信号 $x(t)$ 可用一系列离散的采样值 $x(t_1)$ 、 $x(t_1 \pm \Delta t)$ 、 $x(t_1 \pm 2\Delta t)$ 、... 来表示，只要这些采样点的时间间隔 $\Delta t \leq 1/(2X)$ ，便可根据各采样值完全恢复原来的信号 $x(t)$ 。
- ❖ 当时间信号函数 $x(t)$ 的最高频率分量为 f_{\max} 时， $x(t)$ 的值可由一系列采样间隔小于或等于 $1/(2f_{\max})$ 的采样值来确定，即采样点的重复频率 $f \geq 2f_{\max}$ 。
 - 绝大多数信号都是能够进行傅里叶变换的，就意味着，不管一个信号多么复杂，总可以分解为若干个正弦信号的和，对应了信号的频率分量。
 - 因此，奈奎斯特采样定理只需找到信号最大的频率分量，再用2倍于最大频率分量的采样频率对信号进行采样，从理论上解决了，离散信号能够重建出连续信号的问题。

采样定理的结论

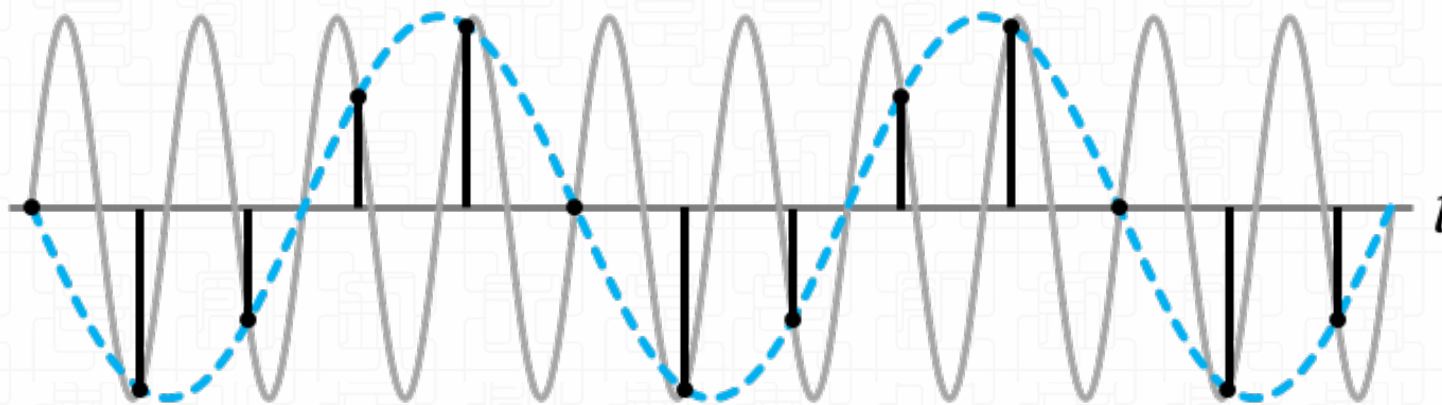
- ❖ 如果已知信号的最高频率 f_{\max} ，采样定理给出了保证完全重建信号的最低采样频率。这一最低采样频率称为**临界频率**或**奈奎斯特频率**，通常表示为 f_N 。
- ❖ 相反，如果已知采样频率，采样定理给出了保证完全重建信号所允许的最高信号频率。



混叠

❖ 如果信号的**采样率低于两倍奈奎斯特频率**，采样数据中就会出现虚假的低频成分。这种现象便称为**混叠**。

❖ 例：800 kHz的正弦波，采样频率为 1 MHz。

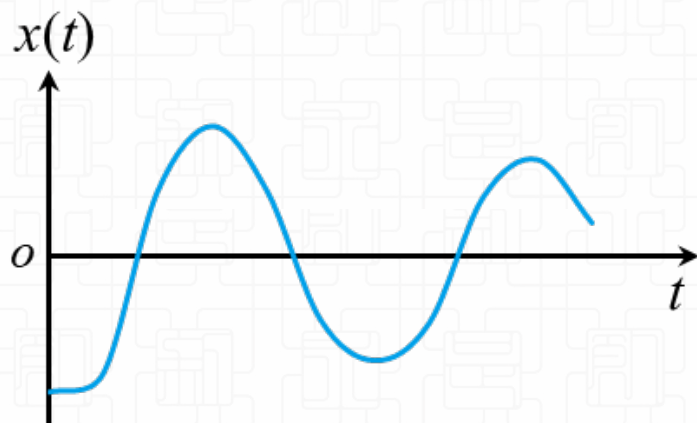


— 蓝色虚线表示该采样率时记录的混叠信号。800 kHz频率与通带混叠，错误地显示为 200 kHz正弦波。

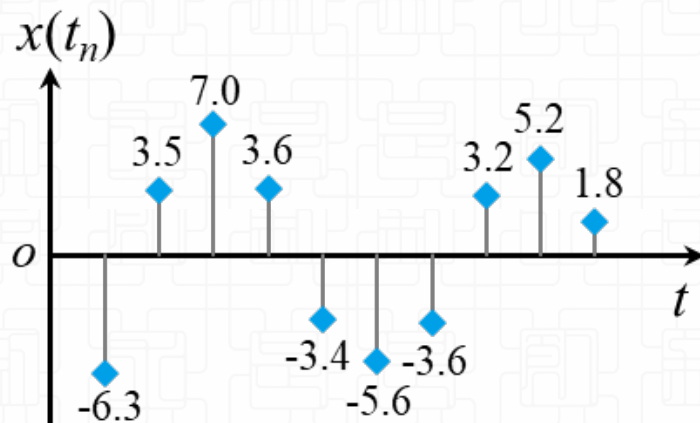
模拟信号、离散信号、数字信号

- ❖ 信号按自变量的取值是否连续可以分为**连续信号**和**离散信号**
- ❖ 按所表达的信息物理量是否连续，可分为**模拟信号**和**数字信号**
- ❖ 数字信号是指在**取值上是离散的、不连续的信号**，只有有限个特定的电压值，表现为瞬时跳变长方形。通常将这些离散数字量用 **0** 和 **1** 组成的二进制数值来表示。

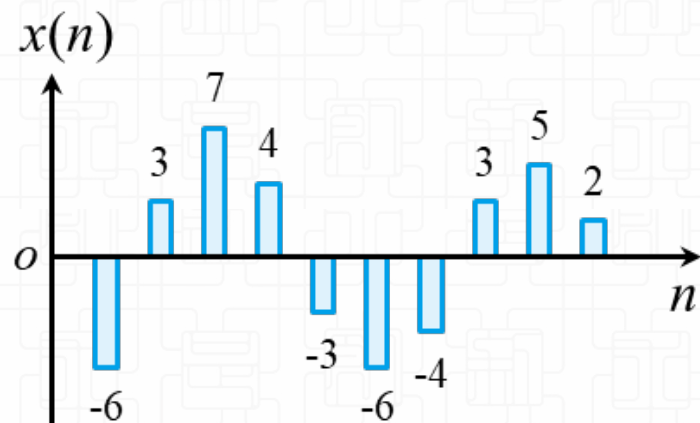
模拟信号



离散信号

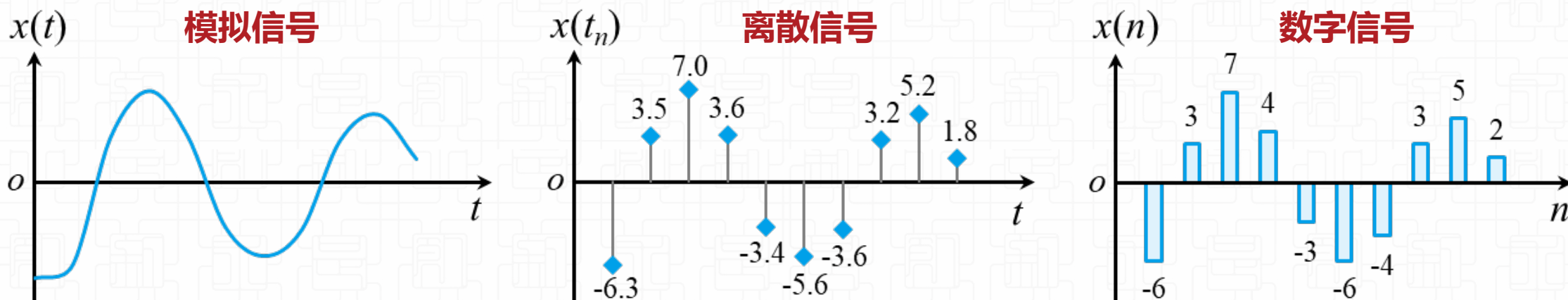


数字信号



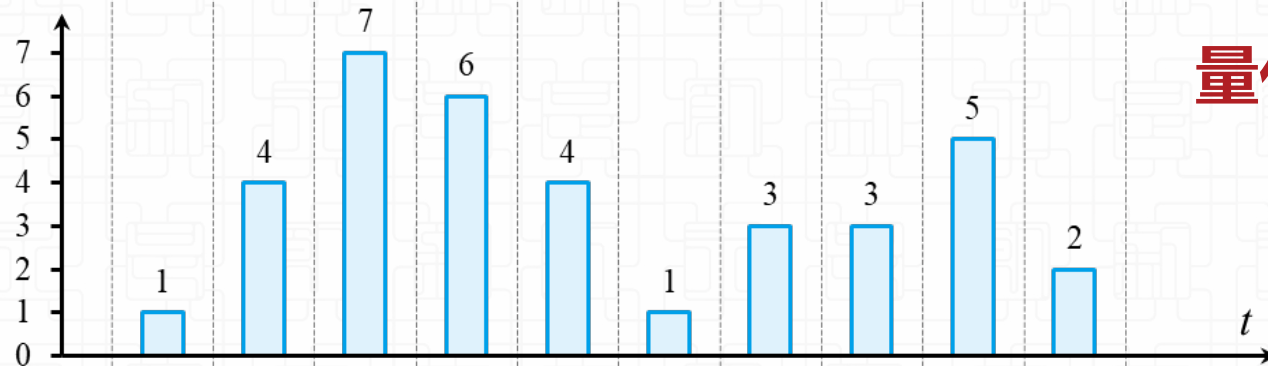
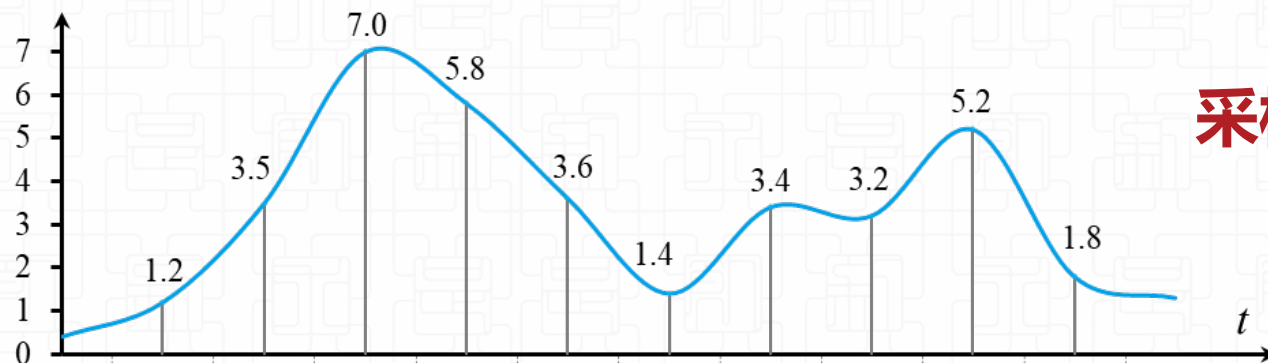
量化

- ❖ 在数字信号处理领域，量化指将信号的连续取值（或者大量可能的离散取值）近似为有限多个（或较少的）**离散值**的过程。
- ❖ 量化主要应用于**从连续信号到数字信号的转换**中。连续信号经过采样成为离散信号，**离散信号经过量化即成为数字信号**。
 - 离散信号通常情况下并不需要经过采样的过程，但可能在值域上并不离散，还是需要经过量化的过程。



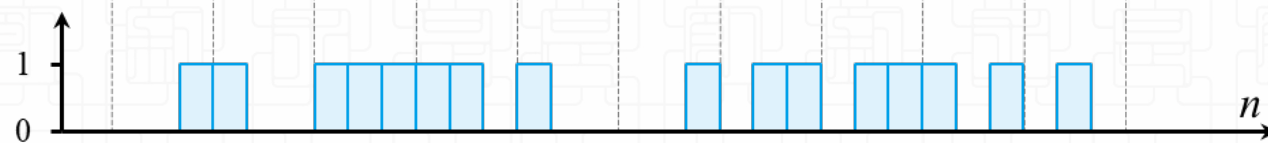
模数转换

模拟信号

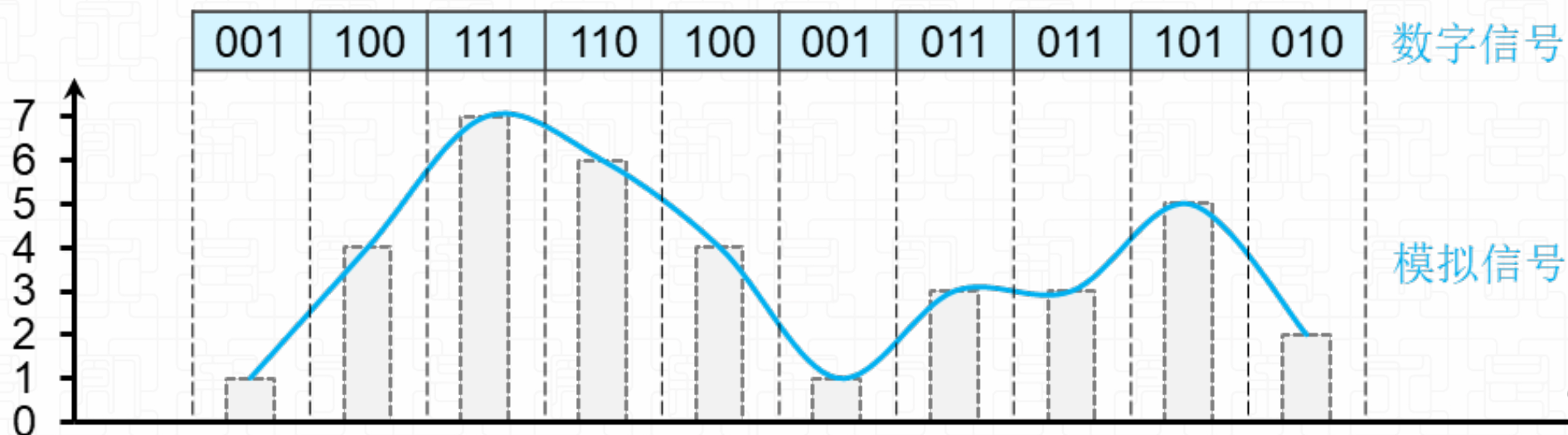


001	100	111	110	100	001	011	011	101	010
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

数字信号



数字/模拟转换

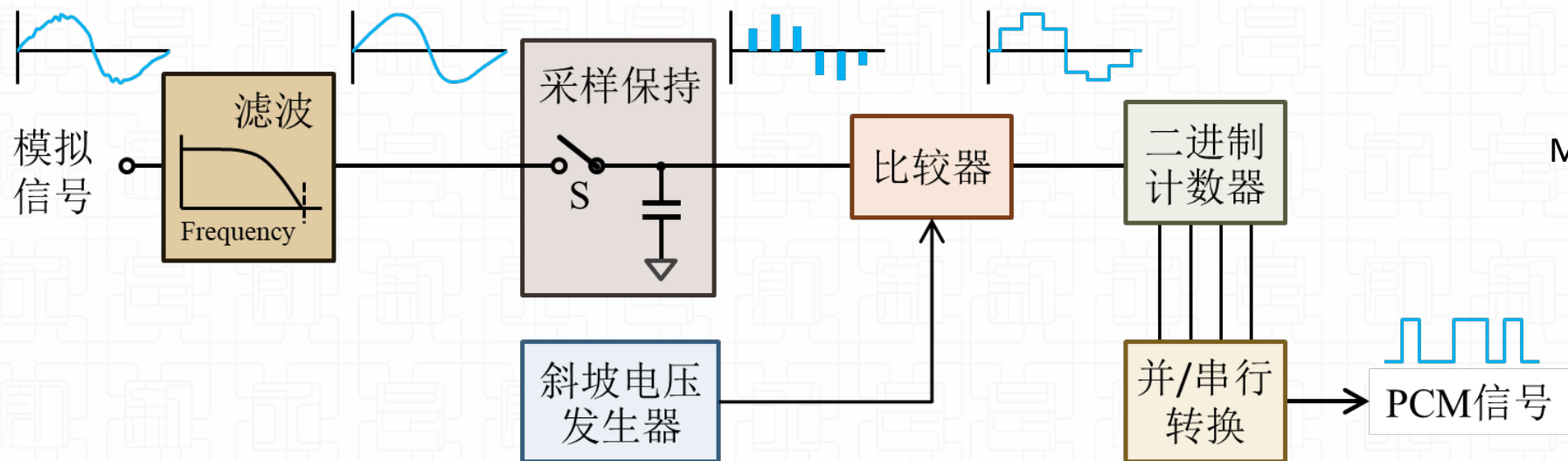


PCM (Pulse Code Modulation, 脉冲编码调制)

- ❖ 脉冲编码调制由英国科学家**亚历克·里弗斯**于1937年提出的，这一概念为数字通信奠定了基础。



Alec Harley Reeves
Mar 10, 1902—Oct 13, 1971



- ❖ 把一个**时间连续、取值连续**的**模拟信号**变换成**时间离散、取值离散**的**数字信号**后在信道中传输。
- ❖ 对模拟信号先**抽样**，再对样值幅度**量化**，**编码**的过程。

内容提要

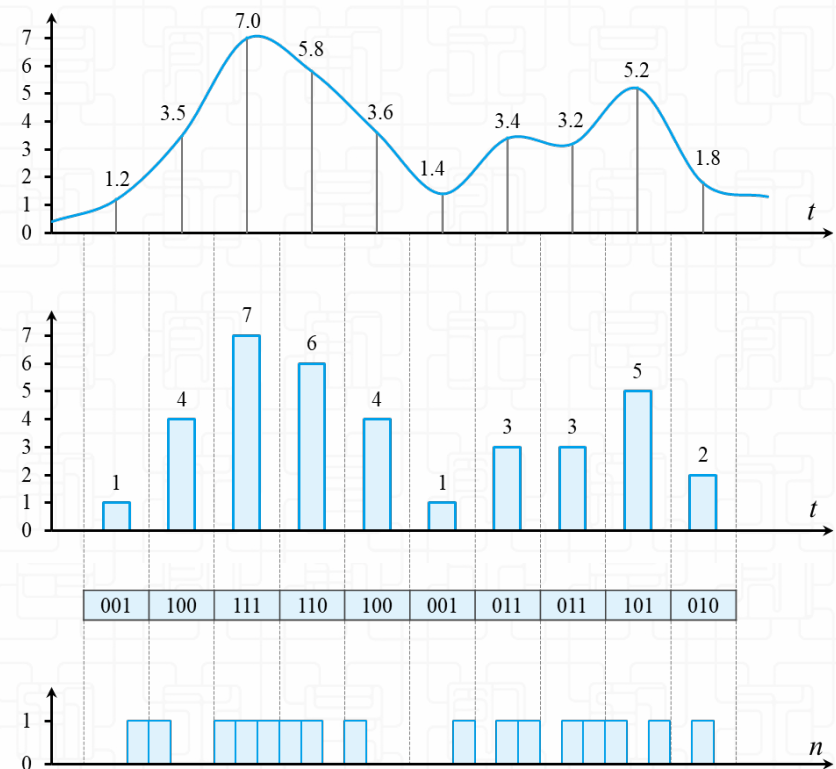
❖ 数据的数字化

❖ 信号的数字化

- 采样, 奈奎斯特采样定理
- 量化
- 编码

❖ 数字信号的优势

❖ 数字信号传输



模拟信号特点

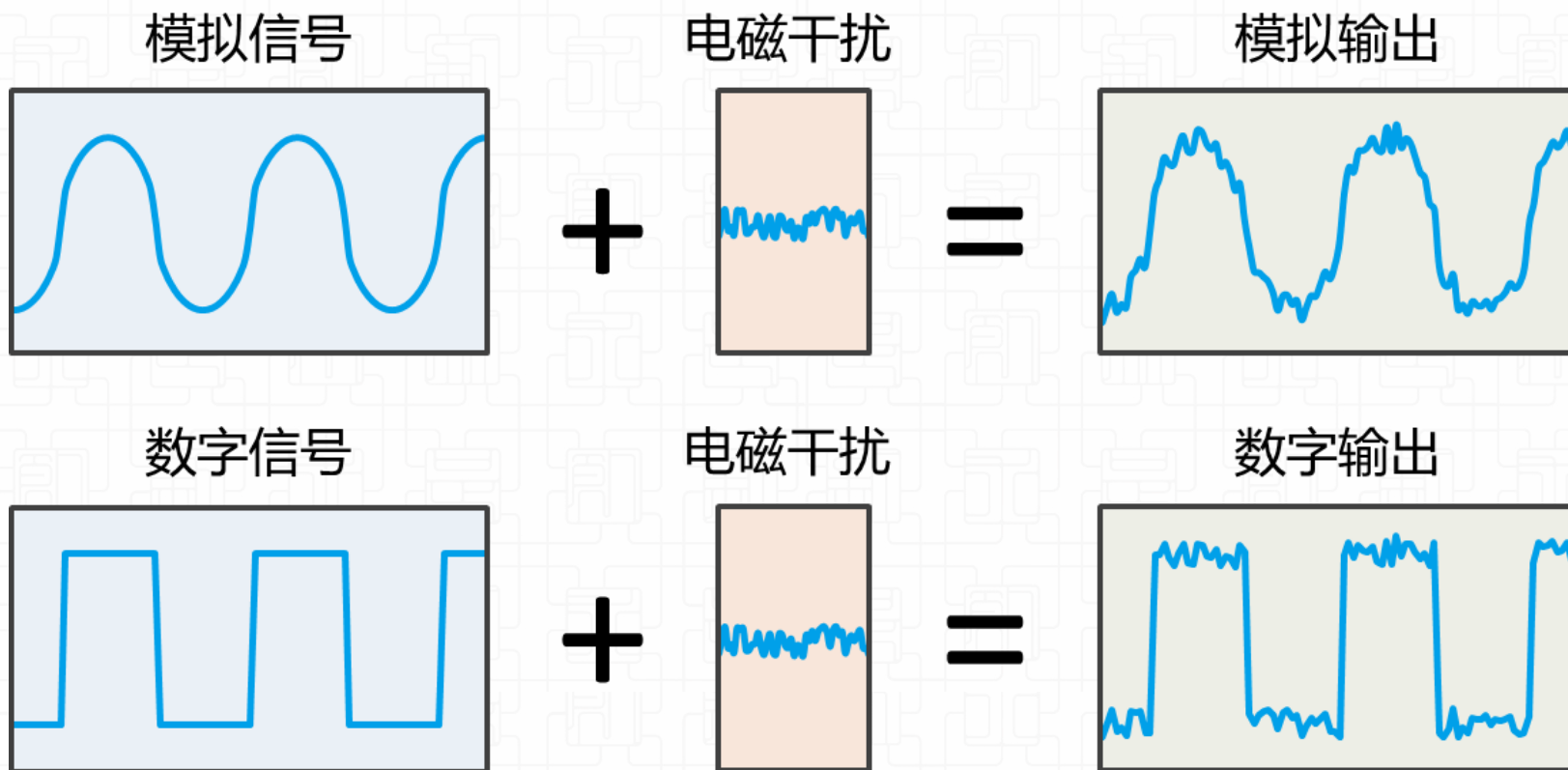
- ❖ 不易于传输。以波形的形式传输，容易受其他信号的干扰而失真变形，且在传输过程中保密性差
- ❖ 不易于存储。模拟信号用磁盘和磁带进行存储，易损坏
- ❖ 不易于运算。模拟信号电路分析难度大，容易受干扰

数字信号的特点

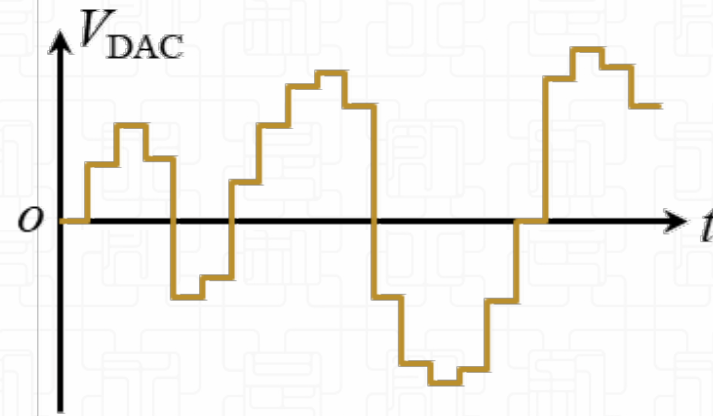
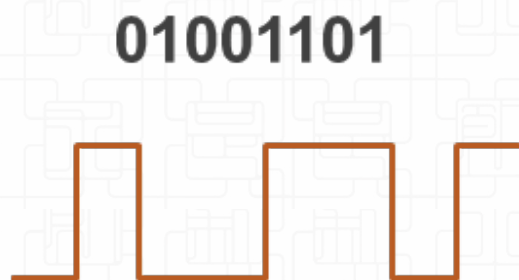
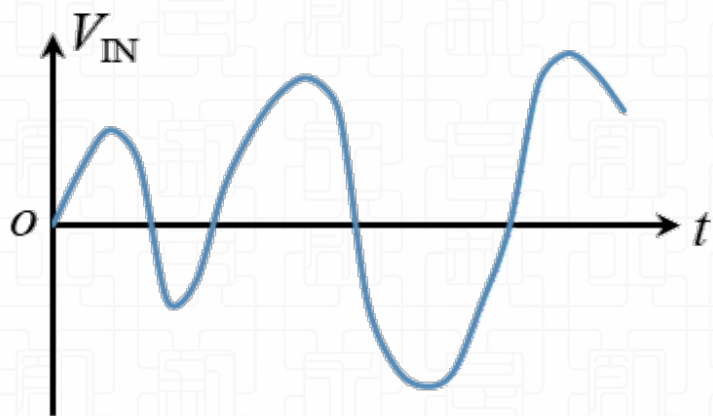
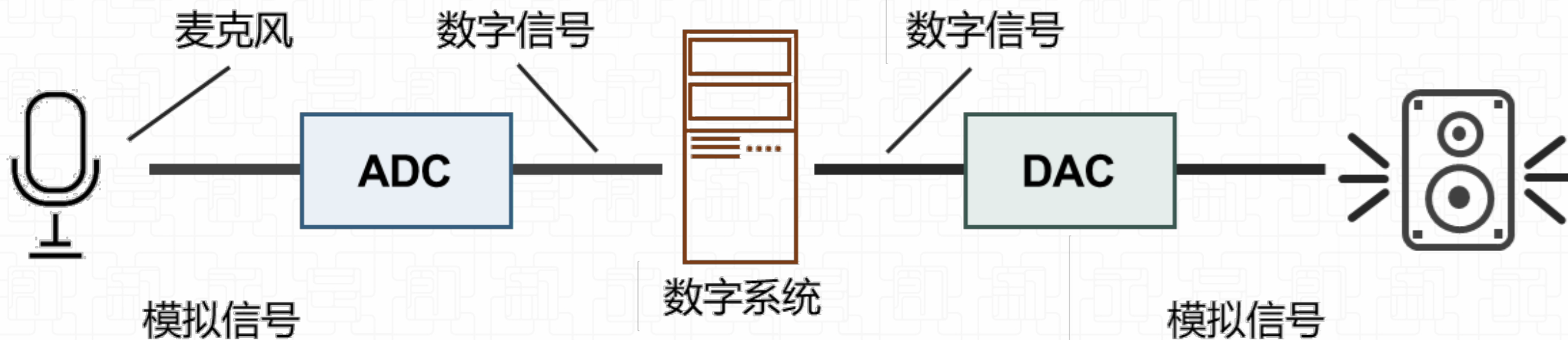
- ❖ 易于传输。以 0 和 1 的形式进行传输，不易受其它杂散信号的干扰
- ❖ 易于存储。数字信号可以存放到 Flash 和 ROM 的器件中，比如 U 盘、SD 卡等
- ❖ 易于运算。只有 0 和 1 两种代码
- ❖ 在现代电子技术中，用**模/数转换器**实现模拟信号和数字信号的转换。



模拟信号和数字信号的抗干扰能力

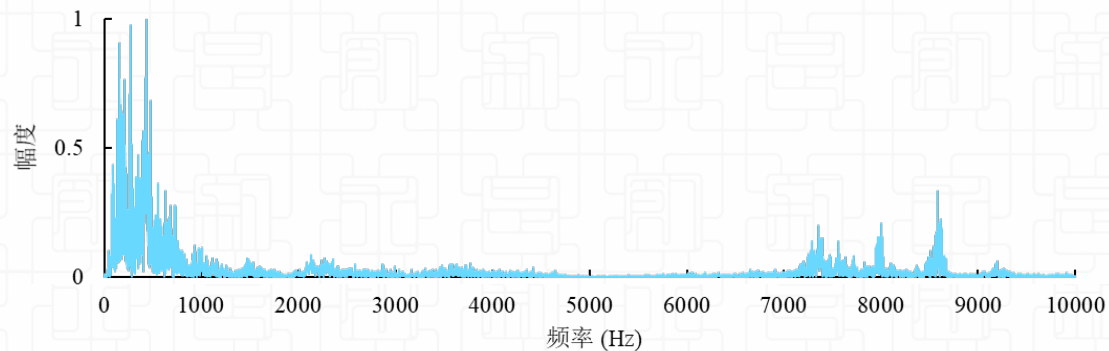
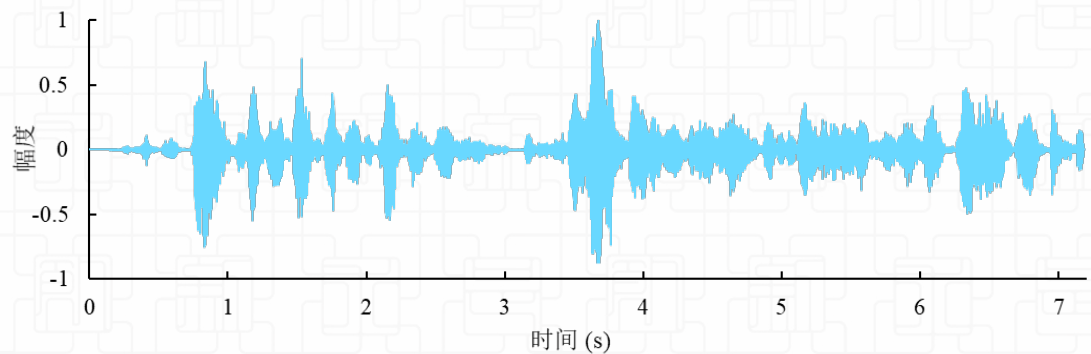


数字信号处理

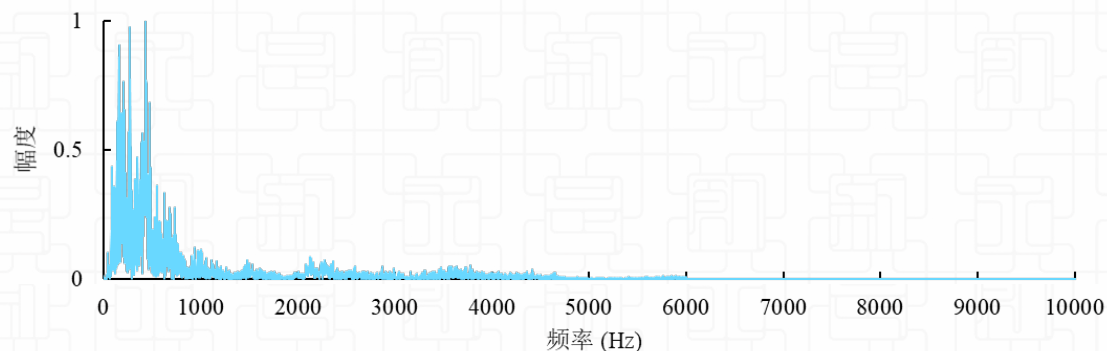
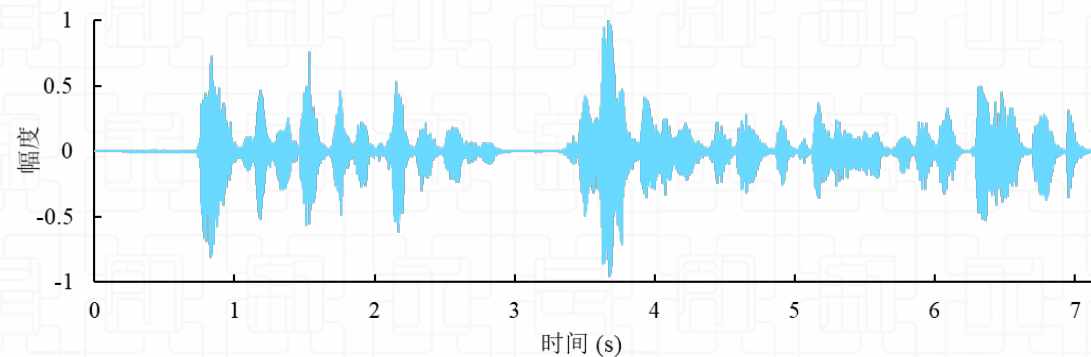


例：语音信号低通滤波去噪

语音



去噪声后的语音



❖ 将信号中特定的频率分量滤除的操作，是抑制和防止干扰的一项重要措施。

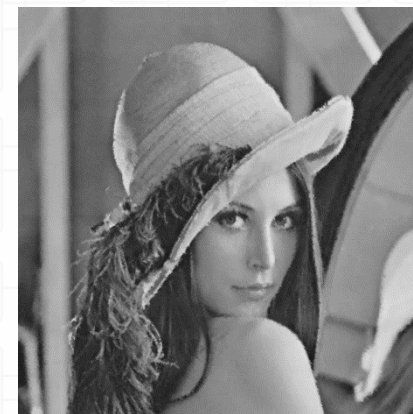
例：从被噪声污染的信号中检测出信号



中值滤波(3×3)滤波输出



中值滤波(5×5)滤波输出



适应算法滤波输出

内容提要

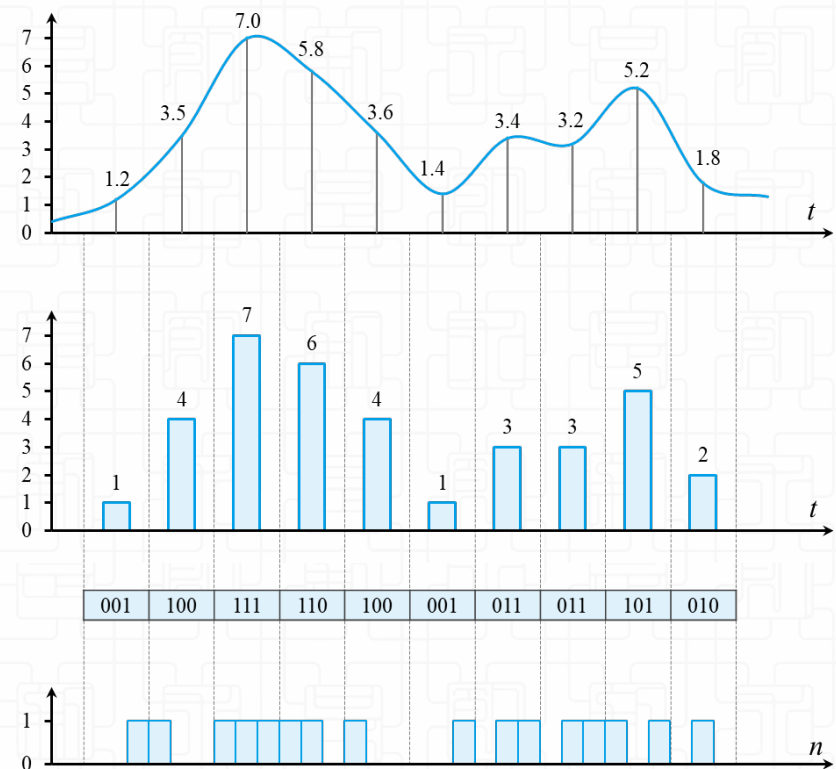
❖ 数据的数字化

❖ 信号的数字化

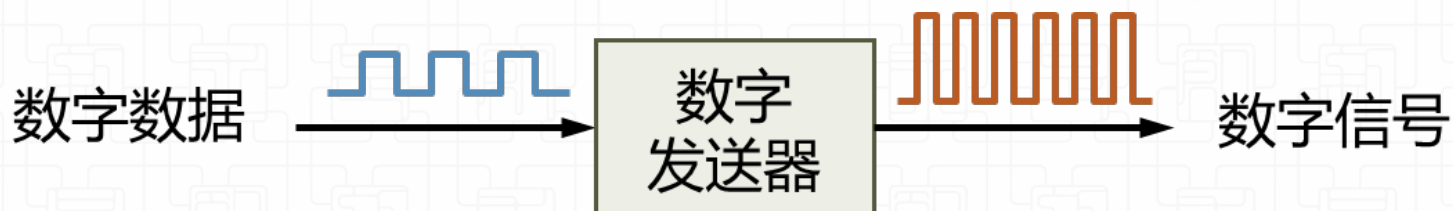
- 采样, 奈奎斯特采样定理
- 量化
- 编码

❖ 数字信号的优势

❖ 数字信号传输



基带传输



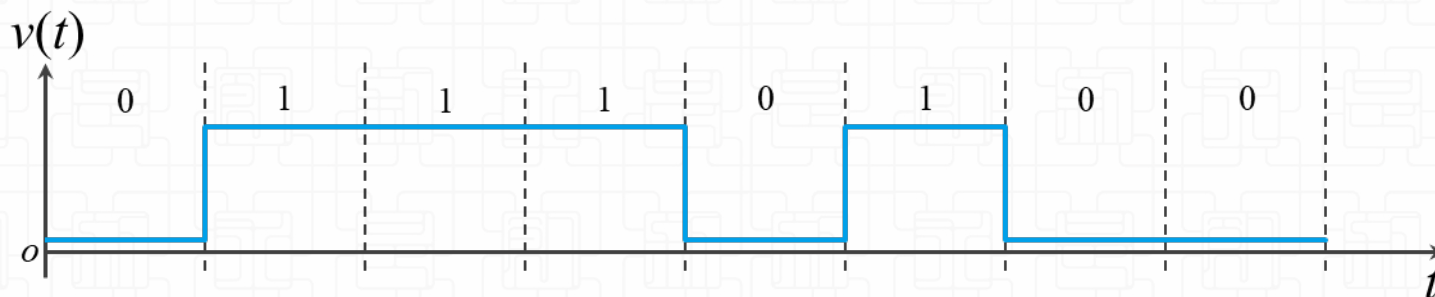
- ❖ 基带信号就是发出的直接表达了要传输的信息的信号。数字信道传输数字信息，**以0、1数字形式传输**。其特点是频率较低，信号频谱从零频附近开始，具有低通形式。



- ❖ 由于在近距离范围内基带信号的衰减不大，从而信号内容不会发生变化。因此在传输距离较近时，计算机网络都采用基带传输方式。

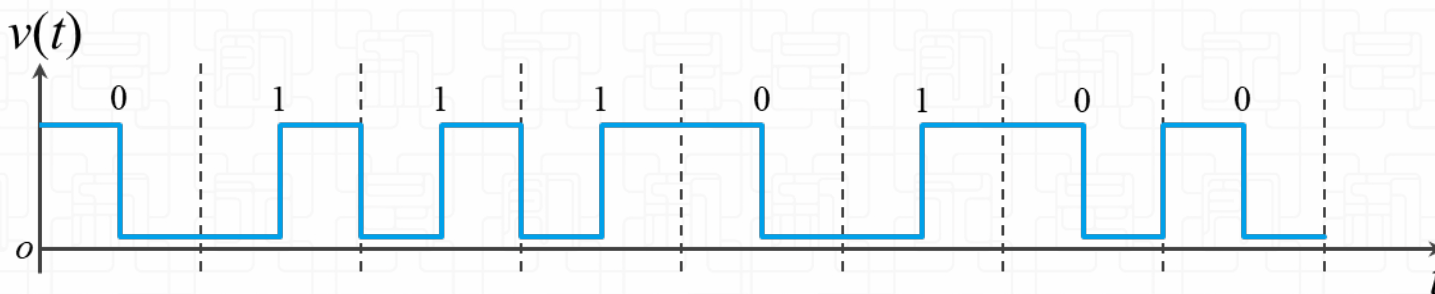
基带信号编码

非归零码



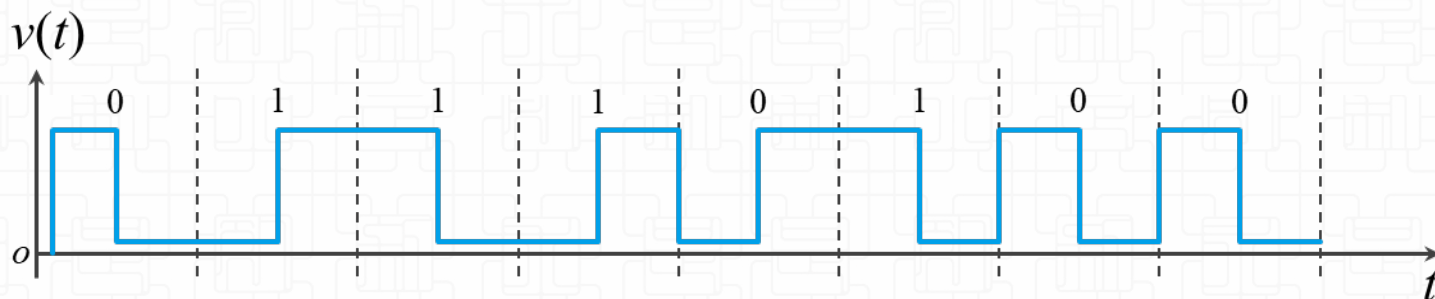
在传输中难以确定一位的结束和另一位开始，传输时必须使用外同步。

曼彻斯特编码



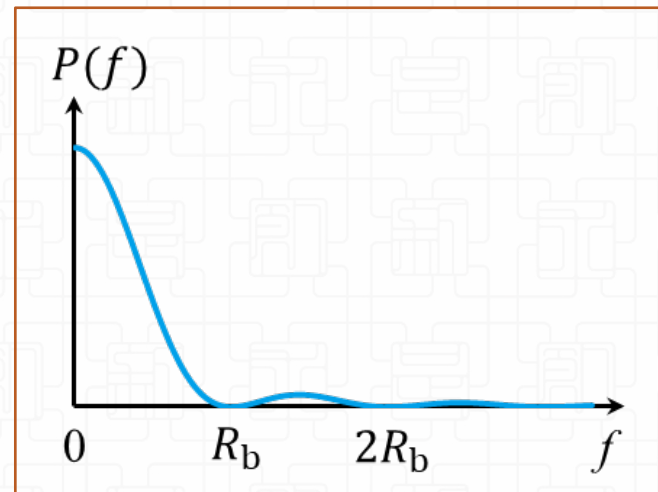
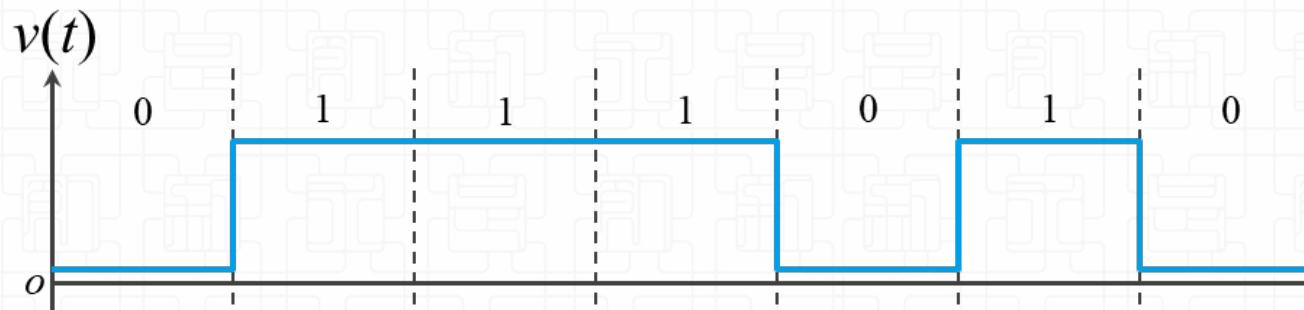
也称为**自同步码**，缺点是需要双倍的传输带宽，即信号速率是数据速率的两倍。

差分曼彻斯特编码

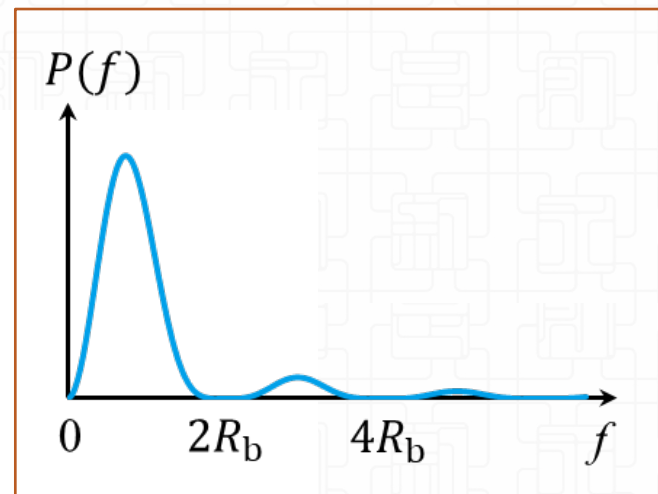
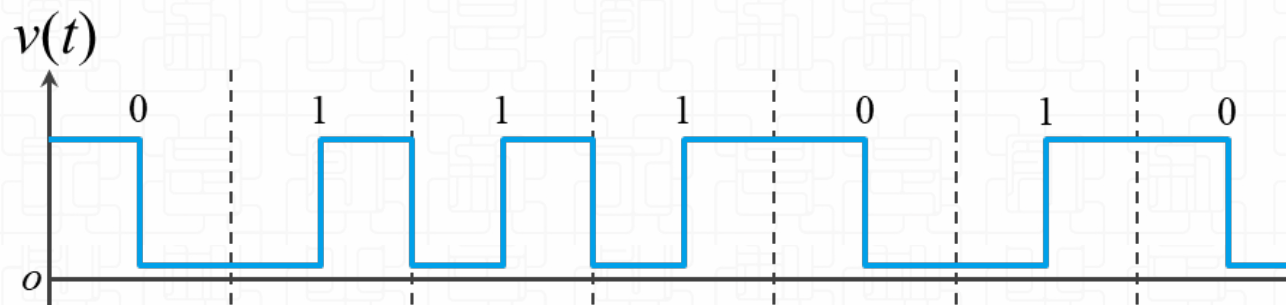


非归零码和曼彻斯特编码的功率频谱

非归零码



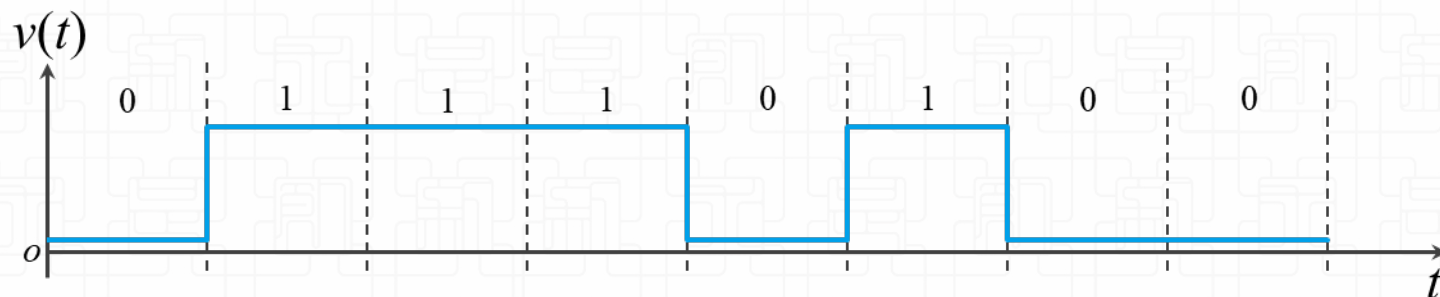
曼彻斯特编码



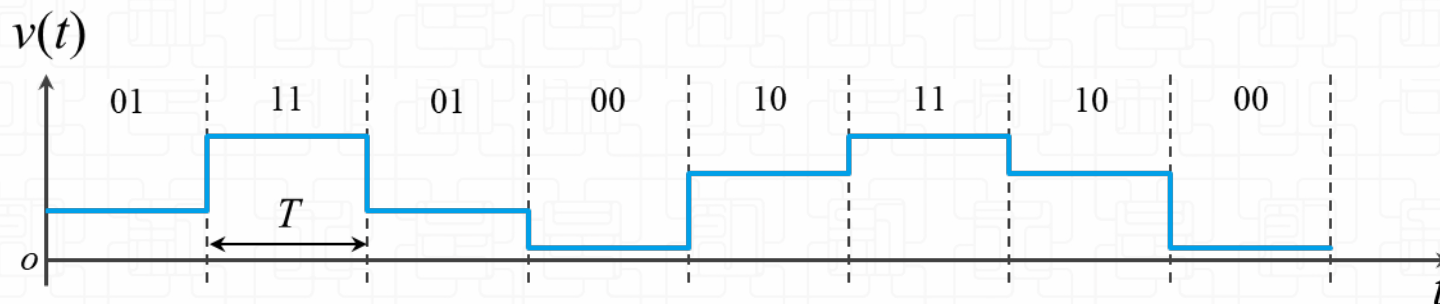
码元

- ❖ 在使用时间域的波形表示数字信号时，代表不同离散数值的基本波形就称为码元。
- ❖ 码元的不同状态（代表不同的离散数值）可以用不同的符号来表示。二进制码元只有两种状态，用两个离散数值0和1来表示。 M 进制码元则有 M 个状态，有 M 个离散数值。

二进制码元



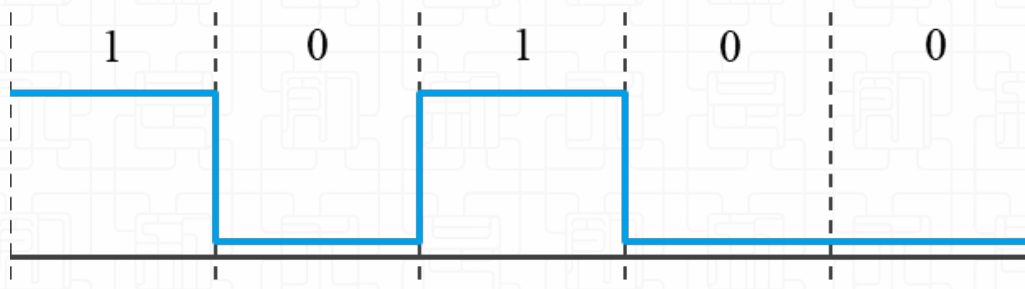
四进制码元



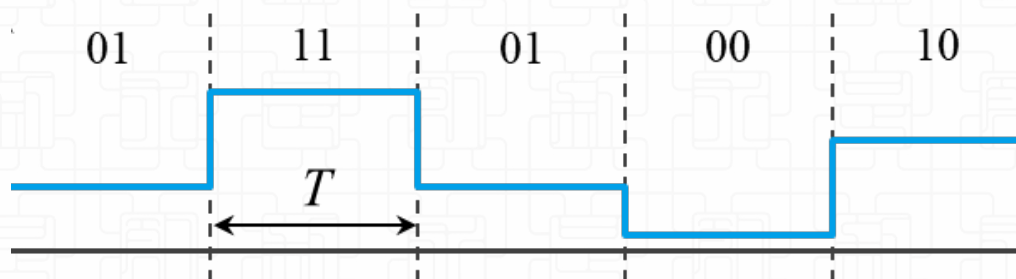
波特率 vs 比特率

❖ 波特率 (Baud Rate) 指的是信号每秒钟电平变化的次数，单位：Baud

$$B = 1/T \text{ (Baud)}$$



二进制码元



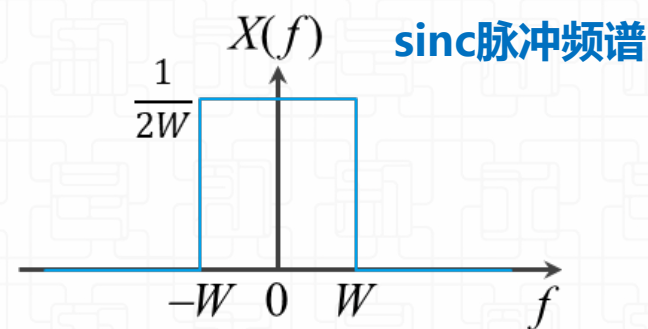
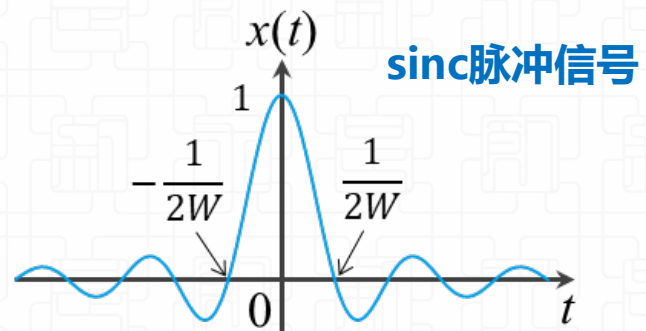
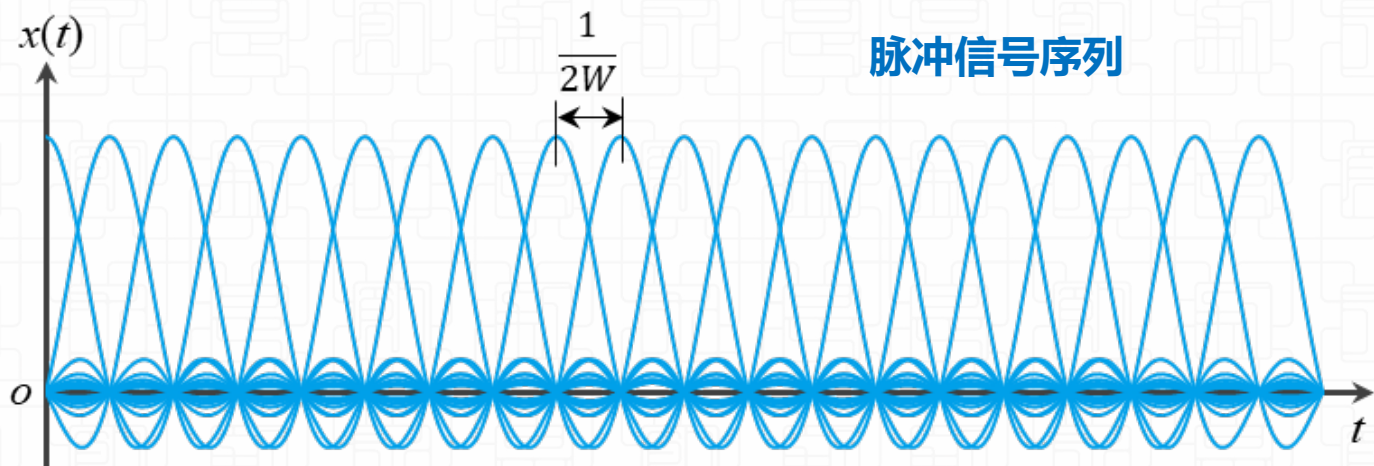
四进制码元

❖ 比特率 (Bit Rate) 用来反映数字信号的数据传输速率，即每秒传送的比特数（二进制数位），单位：bps (b/s)

$$\text{比特率} = \text{波特率} \times \log_2 M$$

奈奎斯特准则

❖ 奈奎斯特在于1924年推导出在理想低通信道下的最高码元传输速率的公式。



最高波特率: $2W$ (baud)

比特率:

$$R_{\max} = 2W \text{ (bps)} \quad (\text{二进制码元})$$

$$R_{\max} = 2W \log_2 M \text{ (bps)} \quad (M\text{进制码元})$$

❖ 如果通过带宽为 B 的无噪声通信信道传输间隔为 $1/(2B)$ 的窄脉冲信号序列, 则前后码元之间不会产生相互串扰。

有噪信道

- ❖ 因为信道总是有噪声的，因此奈奎斯特准则给出的是理论上的上限。
- ❖ 假设信道中有用信号功率是 S ，噪声功率 N ，则总功率 $P = S + N$
- ❖ 总的电压 $V_T \sim \sqrt{S + N}$
- ❖ 最低可分辨的电平电压须大于噪声电压， $V_N \sim \sqrt{N}$
- ❖ 电平数量的极限： $M_{\max} = \frac{V_T}{V_N} = \sqrt{1 + S/N}$
- ❖ 对应的比特数： $\log_2 M$
- ❖ 最大传输速率： $R_{\max} = W \log_2(1 + S/N)$

香农在其论文“通信的数学理论”中有详细的推导过程。这里给出的是一种不严格但简单的推导过程，方便理解。

香农公式

❖ 信道容量公式:

$$C = W \log_2(1 + S/N)$$

❖ C 单位为 bps, 带宽 W 单位为 Hz, 信噪比 S/N 通常以 dB 表示。

❖ 香农定理描述了有限带宽、有随机热噪声信道的最大传输速率与信道带宽、信噪比之间的关系。

❖ 若 $S/N = 30$ dB ($S/N = 1000$), 带宽 $B = 3000$ Hz, 则 $C \approx 30$ kbps。

❖ 信噪比, Singal-to-Noise Ratio (SNR 或 S/N), 是指一个电子设备或者电子系统中信号功率与噪声功率的比值。

$$S/N \text{ (dB)} = 10 \lg S/N$$

例：用香农公式计算

❖ 已知某单色显示屏大约由 3×10^5 个像素组成，每个像素有 10 个亮度等级，这些亮度等级出现的概率是相同。为了满意地显示图像，要求每秒传送 30 帧图像、信噪比为 30 dB。求传输此视频信号所需的带宽。

解：由于每个像素有10个亮度等级，且等概率，则每个像素包含的信息量为 $\log_2 10 = 3.32 \text{ bit}$

每帧有 3×10^5 个像素，故每帧的信息量为 $3.32 \times (3 \times 10^5) = 9.96 \times 10^5 \text{ bit}$

每秒传送30帧，故数据传输速率为 $R_b = 9.96 \times 10^5 \times 30 = 2.99 \times 10^7 \text{ bit/s}$

通信信道最小通信容量 C_{\min} 必须大于或等于 R_b 。由题可知，信噪比 $S/N = 30 \text{ dB} = 10^3$,

则由香农公式可求出最小带宽为
$$W_{\min} = \frac{C_{\min}}{\log_2(1 + S/N)} = \frac{2.99 \times 10^7}{\log_2 1001} = 3.02 \times 10^6 \text{ Hz}$$

所以，传送此视频图像信号所需最小带宽约为3 MHz。

“带宽”与“速率”

- ❖ 奈奎斯特准则描述了**无噪声信道**的最大数据传输速率与信道带宽的关系

$$R_{\max} = 2W$$

$$R_{\max} = 2W \log_2 M$$

- ❖ 香农定理则描述了有**随机热噪声信道**的最大传输速率与信道带宽、信噪比之间的关系

$$R_{\max} = W \log_2(1 + S/N)$$

- ❖ 由于通信信道最大传输速率与信道带宽之间存在着明确的关系，所以人们可以用“带宽”去取代“速率”。
- ❖ 例如，人们常把网络的“高数据传输速率”用网络的“高带宽”去表述。因此“带宽”与“速率”在网络技术的讨论中几乎成了同义词。

The End.



中国大学MOOC

章献民

zhangxm@zju.edu.cn