课程名称

**授课教师**

吴晓峰 高级讲师（[email@fudan.edu.cn](mailto:email@fudan.edu.cn)）

**助教**

王乐（[21210720230@m.fudan.edu.cn](mailto:21210720230@m.fudan.edu.cn)）

目录

[第一讲 信号、信息及其描述 4](#_Toc149589045)

[一、信号与图像的基本概念 4](#_Toc149589046)

[（一）信号的概念 4](#_Toc149589047)

[（二）图像与视频信号 4](#_Toc149589048)

[二、信号的来源 4](#_Toc149589049)

[（一）人体来源 4](#_Toc149589050)

[（二）电子学类来源 4](#_Toc149589051)

[（三）非电子学的自然科学与工程技术类来源 5](#_Toc149589052)

[（四）社会科学类来源 5](#_Toc149589053)

[三、信号与图像的采集——“感” 5](#_Toc149589054)

[四、信号与图像的分析处理——“知” 5](#_Toc149589055)

[（一）语音信号的分析与处理 5](#_Toc149589056)

[（二）图像的分析与处理 5](#_Toc149589057)

[（三）医学图像的分析与处理 6](#_Toc149589058)

[五、信息与消息 6](#_Toc149589059)

[（一）信息与消息的概念 6](#_Toc149589060)

[（二）信息与消息的区别 6](#_Toc149589061)

[六、香农的信息论 7](#_Toc149589062)

[（一）形式化假说 7](#_Toc149589063)

[（二）非决定论 7](#_Toc149589064)

[（三）不确定性 7](#_Toc149589065)

[七、知识与情报 7](#_Toc149589066)

[（一）知识 7](#_Toc149589067)

[（二）情报 7](#_Toc149589068)

[八、信号 7](#_Toc149589069)

[（一）信号的必要性 7](#_Toc149589070)

[（二）信源的数学模型及分类 7](#_Toc149589071)

[（三）信源的自信息量 8](#_Toc149589072)

[第二讲 信息技术中的“感”——信号采集 8](#_Toc149589073)

[一、信息技术 8](#_Toc149589074)

[（一）什么是信息技术 8](#_Toc149589075)

[（二）信息技术的发展 8](#_Toc149589076)

[二、信号图像的“感” 9](#_Toc149589077)

[（一）传感器概论 9](#_Toc149589078)

[（二）感知 10](#_Toc149589079)

[（三）传感器 11](#_Toc149589080)

[三、信息化、数字化、数智化 11](#_Toc149589081)

[（一）信号的采样 11](#_Toc149589082)

[（二）信号的量化 12](#_Toc149589083)

[（三）典型的数字信号处理系统 12](#_Toc149589084)

[（四）采样定理 13](#_Toc149589085)

[（五）信号采样中如何选择元器件 13](#_Toc149589086)

[第三讲 医学信号的识别和监测 13](#_Toc149589087)

[一、医学信号的基本概念 13](#_Toc149589088)

[（一）医学信号分类 13](#_Toc149589089)

[（二）医学信号研究的正问题和逆问题 13](#_Toc149589090)

[（三）医学信号特点：复杂信号 14](#_Toc149589091)

[二、常见的医学信号 14](#_Toc149589092)

[三、心电信号的检测与分析 14](#_Toc149589093)

[四、超声多普勒血流信号的检测与分析 14](#_Toc149589094)

[第四讲 医学图像及其分析处理 14](#_Toc149589095)

[一、医学图像的基本概念 14](#_Toc149589096)

[（一）医学图像的感与知 14](#_Toc149589097)

[（二）目前常用的医学图像 15](#_Toc149589098)

[第五讲 物联网技术及应用——智能化网络化的信息采集 15](#_Toc149589099)

[一、物联网概述 15](#_Toc149589100)

[（一）物联网（IoT）提出的背景 15](#_Toc149589101)

[（二）物联网的概念 15](#_Toc149589102)

[（三）物联网的特征 15](#_Toc149589103)

[二、物联网的网络架构 16](#_Toc149589104)

[三、EPC物联网：系统结构 16](#_Toc149589105)

[第六讲 图像与视频 16](#_Toc149589106)

[一、人眼视觉 16](#_Toc149589107)

[（一）彩色图像的获取模型 16](#_Toc149589108)

[（二）人眼视觉对视觉信息处理的启示 16](#_Toc149589109)

[（三）人眼视觉的特性与错觉 17](#_Toc149589110)

[二、计算机视觉 17](#_Toc149589111)

[（一）计算机视觉发展简史 17](#_Toc149589112)

[（二）计算机视觉与相关学科的关系 17](#_Toc149589113)

[（三）计算机视觉技术的应用 18](#_Toc149589114)

[三、图像文件格式及其转换 18](#_Toc149589115)

[（一）常用图形、图像文件的格式 18](#_Toc149589116)

[（二）常用动态图像的格式 19](#_Toc149589117)

[第七讲 视频获取与视频处理 20](#_Toc149589118)

[一、视觉媒体：图像与视频 20](#_Toc149589119)

[（一）位图 20](#_Toc149589120)

[（二）分辨率 20](#_Toc149589121)

[（三）视频 20](#_Toc149589122)

第一讲 信号、信息及其描述

2023.9.4 / 2023.9.11

一、信号与图像的基本概念

（一）信号的概念

**信号**是随时间变化的某种物理量（声、光、电[磁]等）。

**电信号 x(t) ：**随时间变化的电压或电流。

***例如：***心电信号——电压随时间变化的函数。

***例如：***语音信号——空气压力随时间变化的函数。

（二）图像与视频信号

**灰度图像 I(x, y)** 是随空间位置变化的亮度信号。

**彩色图像**是随空间位置变化的三基色（RGB）信号。

4.**视频**是随时间变化的灰度图像或彩色图像。

二、信号的来源

（一）人体来源

**人类获取信息的方式：**视觉＞听觉＞触觉＞嗅觉＞味觉

**视觉的形成：**眼球成像→信号转换→神经处理传输→大脑形成视觉

（二）电子学类来源

交流电电流信号、脑电信号、Chip信号、FM信号等。

**医学图像：**投影X射线、X射线计算机断层（X-CT）、超声、磁共振（MRI）、放射性核素、红外、分子、光声图像等。

（三）非电子学的自然科学与工程技术类来源

微粒质量浓度信号、大坝水位（海平面）信号等。

（四）社会科学类来源

人口数量信号、股票信号等。

三、信号与图像的采集——“感”

**“感”的本质：**信号、图像的采集过程

**形式：**

1. 物理参量和逻辑值：物理参量——人口、股票等数据；逻辑值——症状有无（可用0/1表示）；

2. 一维波形：如心电图、脑电图、语音信号、地震波等；

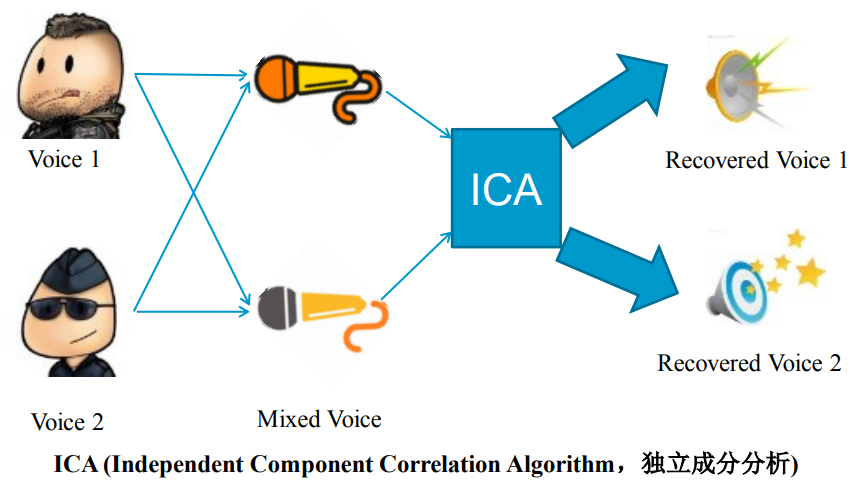
3. 二维图像：如文字、指纹、地图、照片等。

**采集方式：**传感器

四、信号与图像的分析处理——“知”

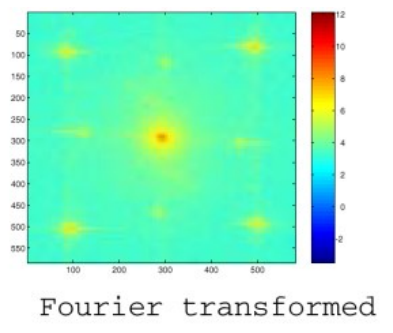
**“知”的本质：**信号、图像的分析与处理

（一）语音信号的分析与处理



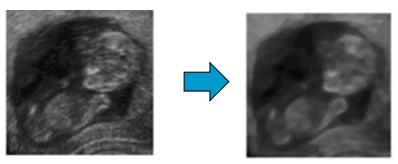
（二）图像的分析与处理

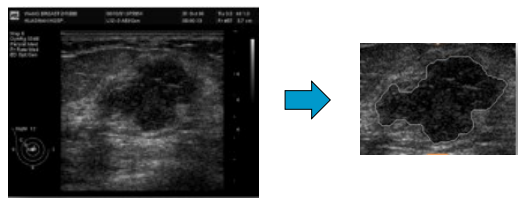
**傅里叶变换：**时间域→频率域



（三）医学图像的分析与处理

**图像降噪：**提高成像质量，抑制噪声，增强对比度



**图像分析：**感兴趣病灶的分割与特征提取

**手术导航：**医学影像与微创、无创手术结合

**疗效评估：**术前、术后药物治疗或手术治疗评估

五、信息与消息

（一）信息与消息的概念

**信息**是人和自然界中需传送、交换、存贮和提取的**抽象内容**。为了传送和交换，通过**语言、文字、图像和数据**表示出来。

**消息**是表示**信息**的**语言、文字、图像和数据**等。有时仍不便传送和交换，需要借助**信号**（电、光、声等物理量）来运载。（**消息**是用语言、文字、数据、符号、音符、图像等能被人们的感觉器官所感知的形式，对客观物质运动和主观思维活动的一种**表述**。）

（二）信息与消息的区别

人们通过通信接收到消息后，得到的是关于描述事物状态的具体内容。例如听气象广播，天气预报为“晴间多云”，这就告诉了我们某地的气象状态，而“晴间多云”这广播语言则是对气象状态的具体表述。

消息中包含信息，是信息的载体。得到消息，从而获得信息。

六、香农的信息论

香农认为，信息是用来消除不确定性的东西。香农对信息具有三个观点：

（一）形式化假说

通信的基本问题是在消息的接收端精确或近似地复制发送端发出的消息。通信中消息的意义和效果与通信技术本身没有关系。

（二）非决定论

一个实际的消息总是从可能发生的消息集合中选择出来的；因此，系统必须对每种选择都能工作。

（三）不确定性

人们只有在两种情况下采用通信的需要：一是自己有某种形式的消息要告知对方，而估计对方“不知道”这个消息；二是自己有某种“疑问”要询问对方，而估计对方能做出一定的解释。

七、知识与情报

（一）知识

**知识**是人们根据某种目的，从自然界收集得来的数据中，整理、概括、提取得到有价值的、人们所需的信息。知识是一种具有普遍和概括性质的高层次的**信息**。

**知识**是信息，但不等于信息的全体。**知识**是以实践为基础，通过抽象思维，对客观事物规律性的概括。知识信息只是人类社会中客观存在的部分信息。

（二）情报

根据情报学中的定义，**情报**是人们对于某个特定对象所见、所闻、所理解而产生的**知识**。

八、信号

（一）信号的必要性

在各种实际通信系统中，为克服时间或空间的限制而进行通信，往往必须对消息进行加工处理。把消息变换成适合信道传输的物理量，这种物理量称为信号（如电信号、光信号、声信号、生物信号等)。**信号携带着消息，是消息的运载工具。**

（二）信源的数学模型及分类

1. 离散信源

离散信源指信源输出的都是单个符号的消息，符号集的取值是有限的或可数的。数学上表示为：

且满足

其中，*X*为样本空间；*P*(*x*)为每个符号出现的概率，称为先验概率。

2. 连续信源

连续信源指信源输出的都是单个符号的消息，符号集的取值是连续的或取值是实数集*R*= (-∞, +∞)。

并满足

（三）信源的自信息量

1. 自信息量的定义

不确定性*I*是概率的函数*I*[*P*(*x*)]。

一次测量获得的信息量为*I*[*P*1(*x*)]-*I*[*P*2(*x*)]。

如果事件*ai*发生的概率为*P*(*ai*)，事件*ai*发生所含有的信息量，就称为**自信息量**，表示为*I*[*P*(*ai*)]。计算公式如：

2. 自信息量的特点

（1）如果*P*1(*a*1)>*P*2(*a*2)，则*I*[*P*1(*a*1)]<*I*[*P*2(*a*2)]

（2）当*P*(*ai*)=1，则*I*[*P*(*ai*)]=0

（3）当*P*(*ai*)=0，则*I*[*P*(*ai*)]=∞

（4）两个独立事件联合信息量等于他们分别的信息量之和

第二讲 信息技术中的“感”——信号采集

2023.9.18 / 2023.9.25

一、信息技术

（一）什么是信息技术

广义上，信息技术是对信息的获取、存储、加工、传递和利用的技术。

狭义上，信息技术是指利用电子计算机和现代通信手段获取、存储、加工、传递和利用的技术（IT）。

（二）信息技术的发展

1. 第一次信息技术革命：语言的使用

人类最初通过手势、表情等肢体动作来表达传递信息。经长期演化，逐渐产生了语言的沟通和表达，语言逐渐成为思想交流和信息传播不可缺少的工具。

2. 第二次信息技术革命：文字的出现与使用

语言的局限在于转瞬即逝，难以保存。文字的出现和使用，让人类对信息的保存和传播超越了时间和地域的局限。

3. 第三次信息技术革命：印刷术的发明与使用

我国是印刷术的发源地，印刷术和造纸术、指南针、火药并称为我国古代科技的四大发明。在公元7世纪初的隋唐之际，我国发明了雕版印刷术；北宋时期的平民毕昇，在我国和世界上最早发明了活字印刷术。

4. 第四次信息技术革命：电报、电话、电视等的发明与普及使用

19世纪30年代，以塞缪尔·莫尔斯为代表的一批发明家发明了电报；1876年，亚历山大·贝尔发明了电话。电报、电话、电视等技术实现了信息的远距离传输。

5. 第五次信息技术革命：计算机与互联网的使用

计算机处理速度快，存储容量大，计算精度高,通用性强，大大扩展和延伸了人们的信息处理和存储能力。互联网使信息的交流和传播在时间和空间上大大缩短和缩小，使整个世界成为地球村。

二、信号图像的“感”

（一）传感器概论

1. 传感器的相关用语与定义

**传感器**是将物理、化学、生物等自然科学和机械、土木、化工等工程技术中的非电信号转换成电信号的换能器。相应的英文单词为**Sensor**或**Transducer**。注意, 若在英文文献中**Sensor**和**Transducer**，甚至还有**Actuator**同时出现时，则**Transducer**应译为“换能器”，是指将自然科学和工程技术中的非电能量转换成电能的设备。

2. 传感器的作用

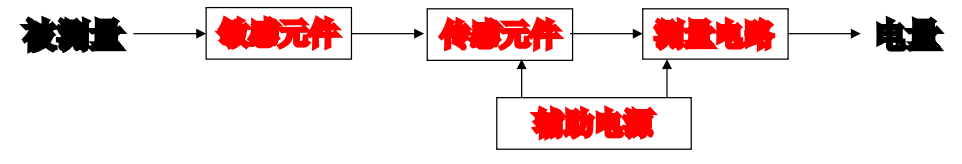
传感器的作用是将非电学量（如角度、位移、速度、加速度、压力、温度、湿度、声强、光照强度等）转化为电学量（如电压、电流、电阻、电容等）（即将各类物理量变为电量）。

3. 传感器的组成

传感器一般由**敏感元件**与**转换元件**组成，这二者是传感器的基本部分。

**敏感元件（sensing element）**是指传感器中能直接感受或响应被测量，并输出与被测量成确定关系的其他量（一般为非电量）部分。

**转换元件（transduction element）**是指传感器中能将敏感元件感受或响应的被测量转换成适于传输或测量的可用输出信号（一般为电信号）部分。



4. 传感器的特性

传感器是信息系统的源头。在客观对象的测量、测试、检测、监测、分析、定位、跟踪、导航、制导、控制及健康管理等系统中,传感器是不可缺少的且在一定程度上是决定系统性能的重要部件。

传感器是科学和工程结合产物，既依赖于科学的新现象和新规律，又依赖于新技术和工艺。

传感器分类方法很多，而且互相交叉，一般以被测量参数来分类和以测量原理两种分类为主：

被测量参数分类可分为温度、压力、流量、位移、速度、加速度、粘度、湿度等传感器，又除去模拟量以外，还有离散量（开关等）传感器等。

按测量原理分类可分为根据电阻定律的电位计式、应变式传感器，根据变磁阻原理的电感式、差动变压器式、电涡流式传感器，根据半导体理论的半导体力敏、热敏、光敏、气敏等固态传感器。

5. 传感器的发展趋势

**（1）新型传感器的开发**

采用新原理、开发新材料、采用新工艺；从结构型为主变为物性型为主；采用仿生原理（机器人研究中仿人视觉、触觉等感觉的传感器）。

**（2）传感器的集成（系统）化、多功能化、微型化**

点→线→面→体；单一类型→多类型；单一功能→多功能；MEMS技术。

**（3）传感器的智能化**

“电五官”+“电脑”=智能化传感器；信号检测转换功能+记忆、存储、解析、处理及自诊断、自适应。

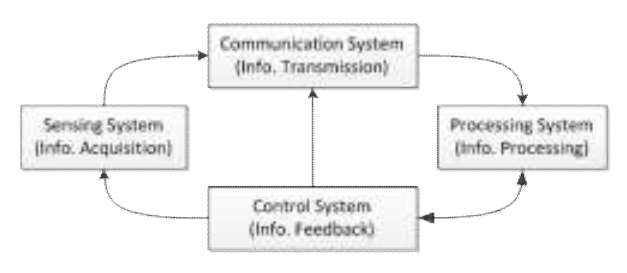
**（4）网络化**

空间、时间上的大跨度，信号图像的传输→物联网。

（二）感知

1. 智能感知技术

人从自然界和环境中获得信息的过程，可以抽象和表示为信息在系统中流动的过程，这个过程包括信息的**“获取”**、**“传输”**、**“处理”**和**“反馈”**，承担这些任务的（典型）信息系统分别是**“感知系统”**、**“通信系统”**、**“处理系统”**和**“控制系统”**，其关系如下图所示：



**智能感知（Intelligent Sensing）**区别于传统的感知技术，具有以下三点基本特征：

* 在感知的过程中形成对感知对象的知识积累和一定的推理规则；
* 具有自适应能力，感知能力能够随环境变化而相应调整；
* 在主动传感系统中，具有信息反馈结构和能力，信息处理的结果能够影响系统发射。

2. 物联网之感知社会论

**物联网**是一个以感知物理世界为目的的物物互联的综合信息系统。简单来说就是八个字：物物互联，感知世界。物联网服务的对象是人类社会，所关注的重点是物理世界，**感知**是物联网的**核心**。

**物联网的感知社会理论体系**是建立在已有的智能化与网络化基础之上的全新理论体系。物联网将在人类社会与物理世界之间建立起一套社会化的感知体系，使人类文明实现新的升华。

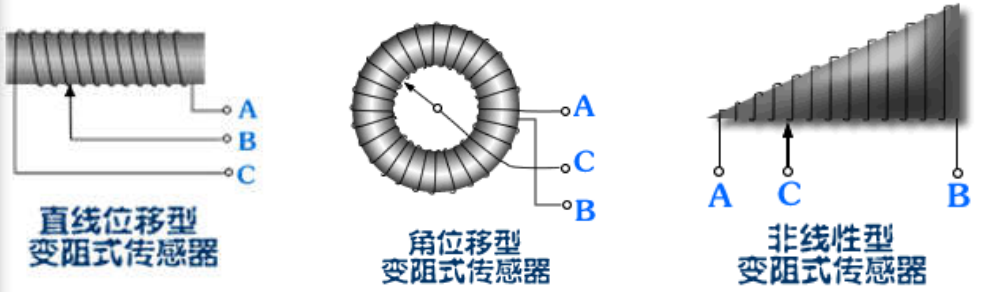
3. 常见的被测物理量

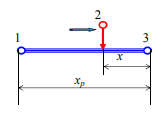
* **机械量：**长度、厚度、位移、速度、加速度、旋转角、转数、质量、重量、力、压力、真空度、力矩、风速、流速、流量;
* **声：**声压、噪声
* **磁：**磁通、磁场
* **温度：**温度、热量、比热
* **光：**亮度、色彩

（三）传感器

1. 电阻式传感器

**电阻式传感器**是把被测量转换为电阻变化的一种传感器,按工作的原理可分为:变阻器式、电阻应变式、热敏式、光敏式、电敏式。



 如图为直线位移型电阻式传感器，当被测位移量*x*变化时，电刷触点2沿变阻器移动，则任意位置2点与3点之间的电阻值为：

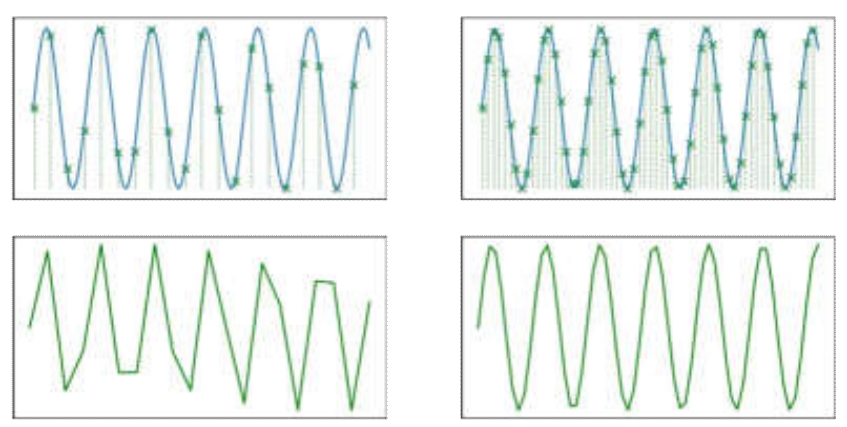
传感器灵敏度公式为：

其中*k*1是单位长度的电阻值。

三、信息化、数字化、数智化

数字化就是把自然界的物理量转化为虚拟的消息体。

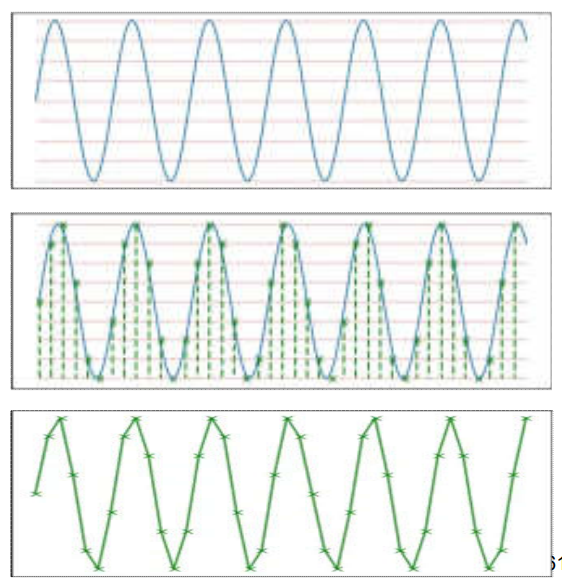
（一）信号的采样



如图，采样的频率越高，就越逼近原本的情况，获得的声音品质就越好。

（二）信号的量化

自然界中的量的范围可以很大，但测量仪器上量的范围相对较小，因此，要把自然界的量转化为可测量的量。通过放大器，就可以实现幅度的缩放。这一过程就是量化，而量化中的基本概念就是逼近，逼近的一个基本方法是对分逼近。



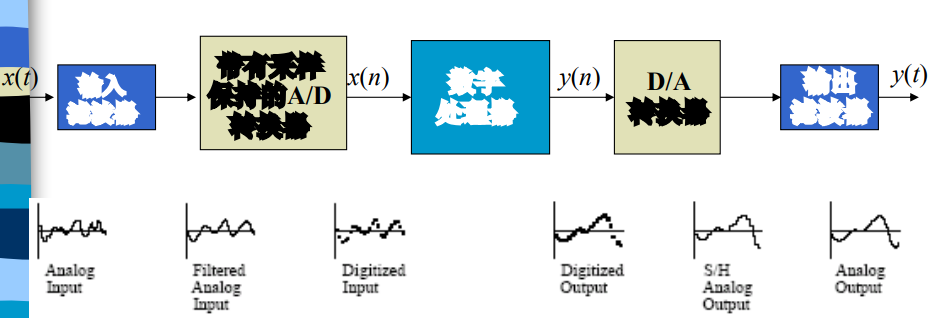
如图，在采集了自然量后，还需要将其量化。有一段正弦声波，假设量化位数为3，即存储的数据只有000、001、010、011、100、101、110、111这8种可能。现在还是等距离采样，不过采样的点只能落在最近的红线上。此时，每个点纵坐标的取值只有23，即只有8种可能。

量化位数越大，声音效果越好。

（三）典型的数字信号处理系统

**模拟信号*x*(*t*)：**时间和幅度上都是连续的信号。自然界的大多数物理量都是模拟量，具有无限精度。

**数字信号*y*(*n*)：**时间和幅度上都是离散的信号。计算机可以处理的量是数字量，可以用整数来表示。



（四）采样定理

如何用最少的抽样获取最精确的、可以还原的量？采样定理（抽样定理）回答了这个问题。

设*x*(*t*)是带限实信号，则抽样后信号频谱不混叠的充分条件为：

抽样频率*fs*满足

或抽样间隔*T*满足

是频谱不混叠最小抽样频率；是频谱不混叠最大抽样间隔。

（五）信号采样中如何选择元器件

信号特性包括幅度范围、频率范围、允许误差；元器件参数包括可测范围、转换速度、转换精度。

第三讲 医学信号的识别和监测

2023.10.9

一、医学信号的基本概念

（一）医学信号分类

人作为活着的生物体，存在本能的反应，会对外界发出主动信号，包括电（如脑电、心电、肌电等）、磁、光、压力等。同时，也存在外源激励人体、人体响应而被动发生的信号，如X射线、超声、激光、红外线等。

（二）医学信号研究的正问题和逆问题

在医学信号研究中，通过人体考察信号的问题是正问题，通过信号得知人体信息的问题是逆问题。例如，心电正问题是“心脏活动在体表面的心电分布”，心电逆问题是“从体表心电来推断心脏的活动（即医生应用心电图进行诊断）”。

逆问题，即解读信号的问题，通常要比正问题更加复杂。

（三）医学信号特点：复杂信号

1. 随机性

医学信号不可能用确定的数学函数来描述，其规律从大量统计结果中呈现。此外，生物系统在外在因素作用下具有适应能力，处于不断的调整之中，因此医学信号具有非平稳性。

2. 背景噪声强——强噪声弱信号

噪声与信号的结合方式包括：

* **加法性：***x*(*t*)=*s*(*t*)+*n*(*t*)
* **乘法性：***x*(*t*)=*s*(*t*)·*n*(*t*)
* **卷积性：***x*(*t*)=*s*(*t*)\**n*(*t*)

二、常见的医学信号

常见的医学信号有：

* **生理电信号：**心电、脑电、肌电、胃电、眼电等。
* **声音信号：**心音信号、打鼾声信号、呼吸音信号、肠鸣音信号等。
* **超声多普勒血流信号**（测胎心音）。
* **压力/压强信号：**血压信号、脉搏波信号、胎动信号、结肠压力信号、足底压力信号灯。
* **光信号**（测血氧饱和度）。
* **磁信号：**心磁信号、脑磁信号等。

三、心电信号的检测与分析

四、超声多普勒血流信号的检测与分析

第四讲 医学图像及其分析处理

2023.10.9

一、医学图像的基本概念

（一）医学图像的感与知

**感：**医学成像系统——图像形成过程。借助于某种介质（如X射线、电磁场、超声波、放射 性核素等）与人体相互作用，将人体内组织器官的结构、功能信息传递出来，通过传感器接收及其相应的算法，最终以图像的方式表现出来，提供给诊断医生

**知：**医学图像分析与处理——获得图像之后。

（二）目前常用的医学图像

* 放射科
  + X射线透射图像
  + CT（计算机断层扫描）图像
  + MR（磁共振）图像
* 超声科
  + 超声图像
* 核医学科
  + PET（正电子发射断层扫描）图像
  + SPECT（单光子发射断层扫描）图像

第五讲 物联网技术及应用——智能化网络化的信息采集

2023.10.16

一、物联网概述

（一）物联网（IoT）提出的背景

**经济危机下的推手：**经济长波理论；过去10年间，互联网技术取得巨大成功；

**传感技术的成熟：**随着微电子技术的发展，涉及人类生活、生产、管理等方方面面的各种传感器已经比较成熟；

**网络接入与信息处理能力大幅提高：**目前，随着网络接入多样化、IP宽带化和计算机软件技术的飞跃发展，基于海量信息收集和分类处理的能力大大提高。

（二）物联网的概念

物联网（Internet of Things）——Internet代表计算机通过标准协议连接形成的全球性网络；Things代表客观世界的物理实体；Internet of Things就是由可唯一标识的物理实体通过标准协议形成的全球性网络。

物联网是由具有自我标识、感知和智能的物理实体基于通信技术相互连接形成的网络，这些物理设备可以在无需人工干预的条件下实现协同和互动，为人们提供智慧和集约的服务。传感器是物联网的基础，是物联网的组成部分。

（三）物联网的特征

**感知：**全面感知，即利用RFID、传感器、二维码等能够随时随地采集物体的动态信息；

**传输：**可靠传输，即通过网络将感知的各种信息进行实时传送；

**智能：**智能处理，即利用计算机技术，及时对海量的数据进行信息控制，形成人与物的沟通。

感知包括对物体静态数据及属性的感知（如RFID、红外感应器、激光扫描、二维码等），对物理固定属性的动态感知（如传感器网、GPS等），对环境模糊信息的感知（如视频探头等）。

二、物联网的网络架构

* 上层：互联网络
  + 物联网的信息存储
  + 物联网的计算决策
* 中层：泛在接入
  + 无线网络（蜂窝、WIFI等）
  + 有线网络
* 下层：物物网络
  + 智能嵌入式设备
  + 感知、标识和通信
  + 协同、互动

三、EPC物联网：系统结构

物联网概念一经提出，立即受到各国政府、企业和学术界的重视，在需求和研发的相互推动下，迅速热遍全球。目前国际上对物联网的研究逐渐明朗起来，最典型的解决方案有欧美的EPC（Electronic Product Code，电子产品码）系统和日本的UID（Universal identification，泛在识别）系统。这里着重介绍EPC物联网。

EPC系统是一个先进的、综合性的和复杂的系统。它由**EPC编码体系、RFID系统及信息网继系统三个部分**组成，主要包括**六个方面**：EPC编码、EPC标签、读写器、EPC中间件、对象名称解析服务（ONS）和EPC信息服务（EPCIS）。

第六讲 图像与视频

2023.10.23 / 2023.10.30

在人类从大自然获取的信息中，图像（视觉）信息总是最多的，占全部信息的过半，其次是语音（听觉）信息，再次是触觉信息，最后是嗅觉、味觉等其他信息。

一、人眼视觉

（一）彩色图像的获取模型

光是一种电磁波，而只有处于一定波长范围内的光才是人眼可见光。通过光的色散，我们得以看到彩色。

人眼接收视觉信息的基本过程是：存在一个光源，对不发光的物体进行光照，不发光的物体对光进行反射（通常是漫反射），反射入人眼中，投射到视网膜上，被各类视神经细胞（锥状细胞、杆状细胞）感受并传输到大脑中，视觉在大脑形成。

以上过程可被抽象为彩色图像的获取模型，其中有三个要素：入射光（光源）、反射特性（被观测物体）和感光特性（光学感知物体）。不同光照条件、不同感光特性，会影响对图像的获取。

（二）人眼视觉对视觉信息处理的启示

人眼视觉的复杂性，启示了视觉信息处理。视觉信息处理的开始是仿生学，仿生的已知条件有：

* 眼睛是视觉信息的接收装置
* 大脑是视觉信息的处理装置
* 仿造眼睛，可以用摄像机获取图像
* 仿造大脑，可以在电脑上通过一定算法实现视觉信息处理

（三）人眼视觉的特性与错觉

人眼视觉具有局部自适应性、选择性、马赫带效应、同时对比效应、相对性、整体性、封闭性等特性，也因此会产生视觉错觉。

二、计算机视觉

（一）计算机视觉发展简史

**1950年代：**二维图像分析和识别，如光学字符识别，工件表面、显微图片和航空图片的分析和解释等。是模式识别的重要内容。

**1960年代：**MIT的Roberts通过计算机程序从数字图像中提取出诸如立方体、楔形体、棱柱体等多面体的三维结构，并对物体形状及物体的空间关系进行描述。这项研究开创了以理解三维场景为目的的三维计算机视觉的研究。Roberts对积木世界的创造性研究给人们以极大的启发，许多人相信，一旦由白色积木玩具组成的三维世界可以被理解，则可以推广到理解更复杂的三维场景。

（二）计算机视觉与相关学科的关系

1. 计算机视觉涉及的学科方面

**图像处理（Image Processing）：**通常是把一幅图像变换成另外一幅图像，即图像处理系统的输入是图像，输出仍然是图像，信息恢复任务则留给人来完成。

**计算机图形学（Computer Graphics）：**通过几何基元，如线、圆和自由曲面，来生成图像，属于图像合成，在可视化（visualization）和虚拟现实（virtual reality）中起着很重要的作用。计算机视觉正好是解决相反的问题，即从图像中估计几何基元和其它特征，属于图像分析。

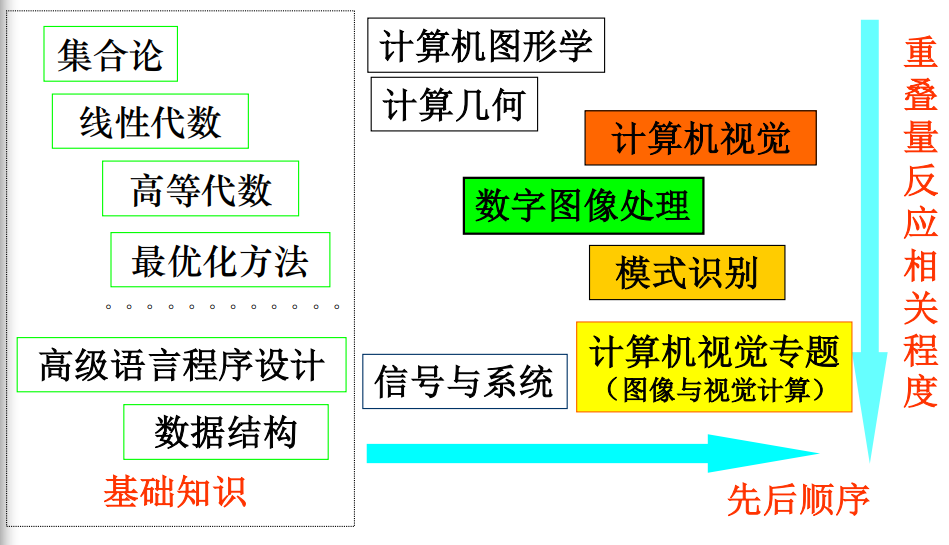
**模式识别（Pattern Recognition）：**研究分类问题，确定符号、图画、物体等输入对象的类别。强调一类事物区别于其它事物所具有的共同特征。一般不关心三维世界的恢复问题。

**人工智能（Artificial Intelligence）：**涉及到智能系统的设计和智能计算的研究。在经过图像处理和图像特征提取过程后，接下来要用人工智能方法对场景特征进行表示，并分析和理解场景。

**媒体计算（Multimedia Computing）：**文字\图形\图像\动画\视频\音频等各类感觉媒体的共性基础计算理论、计算方法，以及媒体系统实现技术。以实现下一代计算机能听、能看、会说、会学习为目标。

**认知科学与神经科学（Cognitive science and Neuroscience）：**将人类视觉作为主要的研究对象。计算机视觉中已有的许多方法与人类视觉极为相似。许多计算机视觉研究者对研究人类视觉计算模型比研究计算机视觉系统更感兴趣，希望计算机视觉更加自然化，更加接近生物视觉。

2. 计算机视觉相关学科与研究基础



（三）计算机视觉技术的应用

* 工业领域（生产装配、质量检验）
* 机器人（星球探测机器人）
* 遥感图像分析（植被分析）
* 医学图像分析（骨骼定位）
* 安全鉴别、监视与跟踪（门禁系统、视频监控）
* 国防系统（目标自动识别与目标跟踪）
* 图像与视频检索（基于内容的检索）
* 文物保护（数字博物馆）
* 其他（游戏、动画、体育、人机交互等）

三、图像文件格式及其转换

多媒体计算机通过彩色扫描仪能把各种印刷图像及彩色照片数字化后送到计算机存储器中；通过视频信号数字化器能把摄像机、录像机、激光视盘等彩色全电视信号数字化存到计算机存储器中；还有计算机本身可以通过计算机图形学的方法编程，生成二维、三维彩色几何图形及三维动画，存在计算机存储器中。

一般分两类图像文件格式：一类是静态图像文件格式；另一类是动态视频图像文件格式。对于静态图像文件格式，将介绍6种当前比较流行的图像格式：GIF、TIFF、TGA、BMP、PCX及MMP；对于动态视频图像文件格式，将简单介绍一下MPG、AVI等文件格式。

（一）常用图形、图像文件的格式

1. 图像文件的一般结构

**文件头：**软件ID、软件版本号、图像分辨率、图像尺寸、图像深度、彩色类型、编码方式；

**文件体：**压缩算法、图像数据；

**文件尾：**彩色变换表、用户名、注释、开发日期、工作时间。

2. BMP文件格式

图文件（Bitmap-File，BMP）格式是Windows采用的图像文件存储格式，在Windows环境下运行的所有图像处理软件都支持这种格式。

BMP图像文件由以下三部分组成：位图文件头（BITMAPHEADER）数据结构、位图信息（BITMAPINFO）数据结构和位图阵列。

3. TIF文件格式

TIF是TIFF（Tagged Image Format File）文件的扩展名，该格式由美国Aldus Developer’s Desk和Microsoft Windows Marketing Group制订。

TIFF支持任意大小的图像，从单色的二值图像到24位的真彩色图像；支持灰度图像，也支持EGA/VGA上最常见的调色板式图像。

TIF格式的优点主要是适合于广泛的应用程序，它与计算机体系结构、操作系统和图形处理的硬件无关。

4. GIF文件格式

GIF（Graphics Interchange Format）是CompuServe公司开发的图像文件存储格式，称为图形交换格式。

1987年开发的GIF文件格式版本号是GIF87a，1989年进行了扩充，扩充后的版本号定义为GIF89a。

一个GIF文件由表示图形/图像的数据块、数据子块以及显示图形/图像的控制信息块组成。

GIF文件格式采用了LZW（Lempel-Ziv Walch）压缩算法来存储图像数据。

GIF文件格式可在一个文件中存放多幅彩色图形/图像。

5. PNG文件格式

PNG是20世纪90年代中期开始开发的图像文件存储格式，其目的是企图替代GIF和TIFF文件格式，同时增加一些GIF文件格式所不具备的特性。

PNG使用从LZ77派生的无损数据压缩算法。

6. 图像的种类

图像的种类包括：标准单色图、256色标准彩色图、256色灰度图、标准灰度图、24位真彩色图、24位真彩色图转换成的灰度图等。

（二）常用动态图像的格式

1. 动态图像的分类

动态图像一般可分为：视频和动画。

**视频：**每一帧图像是实时获取的自然景物的真实图像。和音频一样，在多媒体计算机中使用的是数字视频。数字视频的应用于VCD/DVD、数字电视、远程教学及视频会议等。

**动画：**每一帧图像是由计算机或人工制作的具有真实感的图像。若画面仅为二维透视效果时，则为二维动画；若画面具有空间效果时，则为三维动画；若加上真实的光照效果和质感，则为三维真实感动画。

2. AVS和AVI文件格式

AVS和AVI是Intel和IBM公司共同研制的数字视频交互（Digital Video Interactive，DVI）系统动态图像文件格式，AVS文件格式只能在DVI系统硬件支撑下才能读写，这样系统的硬件设备投资比较大，为了降低系统造价，Intel公司最近又推出了Indeo系统，它可以在Microsoft公司的Video for Windows支持下，用软件播放AVI 文件。

AVS文件格式比起图像文件格式能够提供较多的灵活性，它能够支持多个数据流同时操作。

AVS文件还提供了三种附加数据流的类型：底层数据、数据和图像。

AVS文件的一帧可能没有固定的尺寸，特别是经过64压缩编码的视频数据，某些帧可能比另外的一些帧多些或少些数据。

3. RIFF文件格式

资源交换文件格式（Resource Interchange File Format，RIFF），是IBM和Microsoft公司推荐了为多媒体资源文件而开发的一种带标记的文件结构。实际上它不是一种文件格式，而是一种定义标准的范式。目前，很多多媒体资源文件都是按RIFF范式定义的，如多媒体影片文件（MMM）、波形音频文件（WAV）和位图文件（RDIB）。Microsoft公司在Windows 3.1中提供了对多媒体文件I/O（MMIO）的支持，为RIFF文件的应用开辟了新的途径。

第七讲 视频获取与视频处理

2023.10.30

一、视觉媒体：图像与视频

利用视觉传递信息的媒体称为视觉媒体，如文字和符号、位图、矢量图形、视频、动画等。

（一）位图

位图是由n×m个（像素）点排列成的图片，也称点阵图像或光栅图像。

一幅位图就像是一个矩阵，矩阵中的每一个元素对应图片中的一个像素点，相应的值表示该点的灰度或颜色等级。

像素是赋予颜色和亮度的最小单位。

位图适用于逼真照片或要求精细细节的图片。

（二）分辨率

1. 空间分辨率——分辨率

**显示分辨率（设备分辨率）：**屏幕在横纵两个方向上能够显示的点数，如800×600（dpi）。

**图像分辨率：**图像在横纵两个方向上的像素数，是像素密度的度量方法，单位一般为ppi。

对同样尺寸的一幅图，如果图像的分辨率越高，则说明组成该图的像素数目越多，看起来就越逼真。

2. 颜色分辨率——像素深度

位图矩阵中，每个元素的值就是对应像素点的颜色/灰度值，其所用的存储位数就称为像素深度。

像素深度决定了灰度图像中每个像素可能有的灰度级数，或者彩色图像中每个像素可能有的颜色数。

（三）视频

视频由一幅幅单独的画面序列组成，这些画面以一定的速率依次显示，使观察者具有图像连续运动的感觉。从数学角度看，视频是随时间变化的图像，或称为时变图像、运动图像、活动图像。视频是携带的信息量最为丰富、直观、生动、具体的一种媒体。

视频的两个特点是：内容随时间而变化；伴有与画面同步的声音（伴音）。

视频分为模拟视频和数字视频。

1. 模拟视频

模拟视频是一种用于传输图像和声音的并且随时间连续变化的电信号。早期视频的记录、存储和传输都采用模拟方式。

例如：电视信号就是以一种模拟电信号的形式来记录的，录制时使用盒式磁带进行存储，并依靠模拟的调制手段在空间传播。

2. 模拟电视信号

**①扫描方式**

电视信号是一种重要的视频。电视机播放视频，需要将一维的电信号转换成二维的图像，这一过程是通过光栅扫描实现的。

扫描的方式主要有两种：逐行扫描、隔行扫描。

**逐行扫描：**每次显示整幅图像，画面清晰稳定。

**隔行扫描：**利用了人眼视觉暂留的心理作用，每次只需一半带宽资源就可获得较稳定的图像，但显示会有闪烁。

**②电视信号的类型**

按信号源，可以对电视信号进行分类：分量视频信号、复合视频信号、射频信号（RF）、S-Video等。

**③电视信号的制式**