



# 人工智能基础与进阶

## 计算机与环境感知

上海交通大学

# 目录 content



第一节 声音是一维信号

第二节 图像是二维信号

第三节 一线激光雷达简介

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

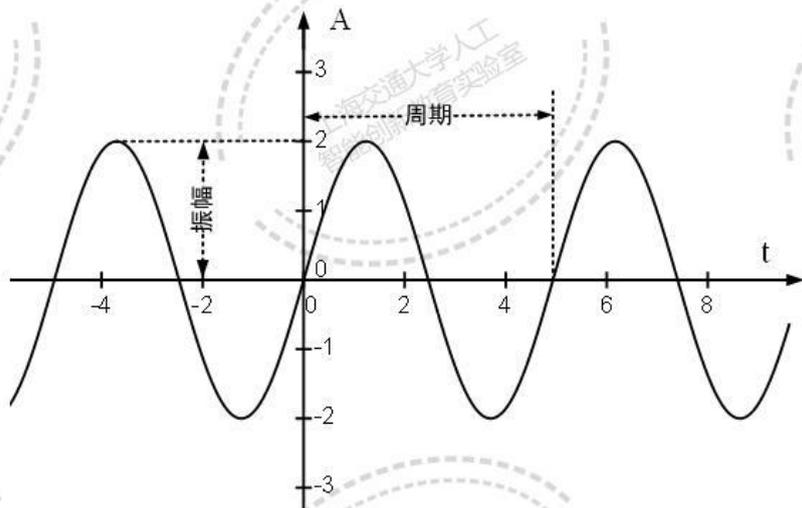
## 第一节 声音是一维信号

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

## 声音的基本属性介绍



- 声源在一秒钟之内振动的次数叫做**频率**，单位是赫兹（Hz），常记作 $f$ 。
- **周期**是声源振动一次所经历的时间，记作 $T$ ，单位为秒（s），频率和周期的关系为 $T=1/f$ 。
- **波长**是沿声波传播的方向，震动一个周期所传播的距离。
- **振幅**是声源振动时，声源离开原位置最大的位移距离，单位为米（m）。

# 声音的主要性质

## 响度



- 人耳感受到的声音的大小叫做**响度**。
- **响度**是人耳主观上的听觉效果，**音强**是声音的客观强度。
- **响度**（音强）反映的是物体振动时幅度的大小。
- 振幅大，声音响；振幅小，声音轻。

# 声音的主要性质

## 音调



- 音高又称为音调。
- 振动频率的高低决定了音高：频率越高，音高越高；频率越低，音高越低。
- 不同音高的搭配方式决定声音的听觉效果，当高音和低音搭配适当时，听起来就会感到声音清晰而柔和，有自然的感觉。

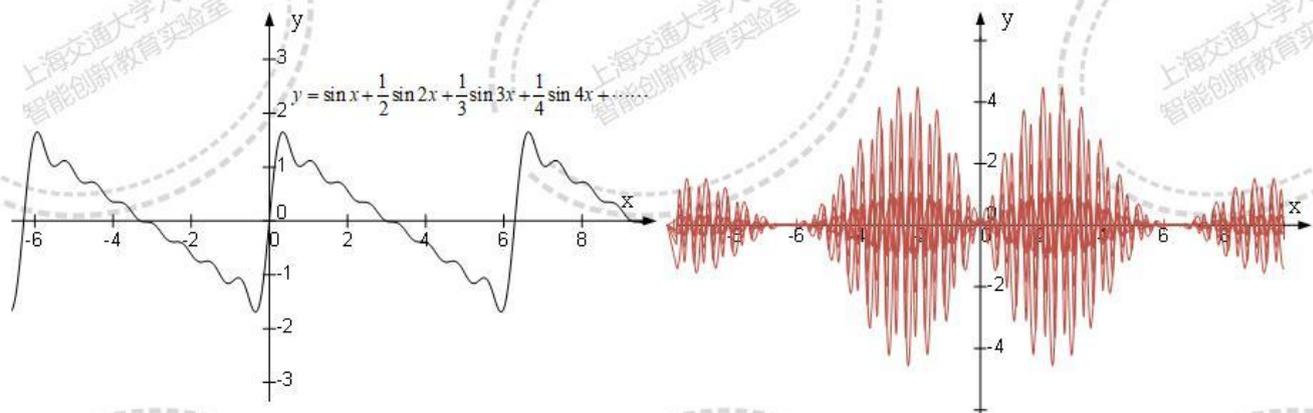
# 声音的主要性质

## 音色



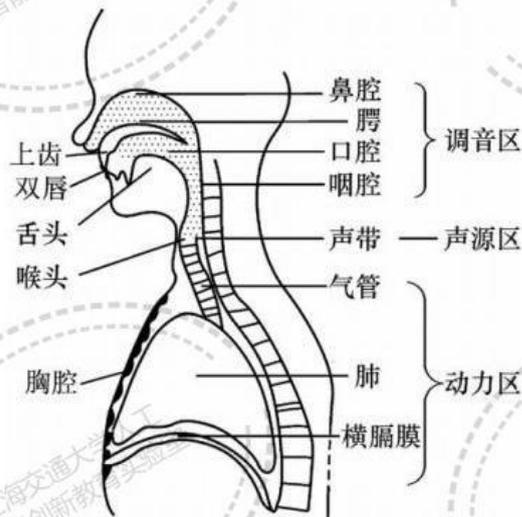
- 人耳在主观感觉上，区别相同响度和音高的两类不同声音的主观听觉特性称为音色。
- 每个人讲话的声音及钢琴、小提琴、笛子等各种乐器所发出的不同的声音，都是由于音色不同造成的。
- 使用各种不同的乐器演奏同样一个乐音，虽然响度、音高可以调节至相同或者相近，但是人在听起来的时候还是会觉得有显著的区别，这就是音色不同所导致的。

# 复合音



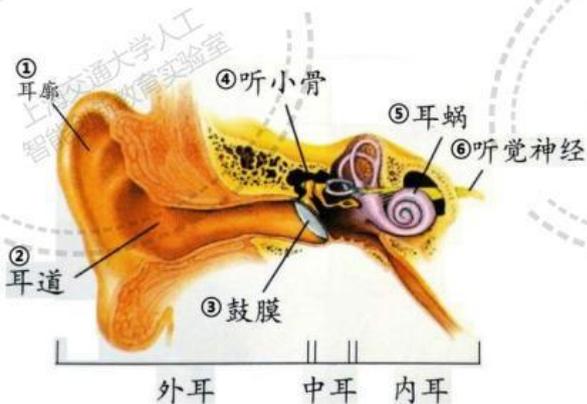
- 现实世界中的各种声音并不是单一频率的，绝大多数声音都是由多个频率的声音组成的复合音。
- 复合声波的组成包括基波和基波上的多次谐波。
- 谐波的频率一般为基波频率的若干倍。

# 人的发声机制



- 大脑下达发声指令
- 呼吸系统呼出的气流带动声带振动，发出声音
- 声音再通过舌、腭、咽、齿、唇、下颌等发音器官进行调节最终发出语音。

# 人耳结构



- 人的听觉系统由外耳、中耳、内耳以及神经构成。
- 外耳起集音作用，并将声音从外耳道传入中耳
- 中耳起传音作用，可将传入的声音振动放大后传声至耳蜗
- 内耳的毛细胞受到刺激，引起了细胞的生物电变化，释放化学递质，由此将波动转化为电生理信号
- 刺激传至各级听觉中枢，经过多层次的信息处理，最后在大脑皮层引起听觉

# 声音传感器



- 传感器是一种能够把物理量或化学量转变成便于利用的电信号的器件。
- 声音传感器就是一种可以将声音信号转换成电信号的转换器，使用的是与人类耳朵具有相似频率反映的电麦克风，因此声音传感器就是计算机的耳朵。
- 上图所示是常见的四种声音传感器，也就是俗称的麦克风。

# 常见声音传感器的应用



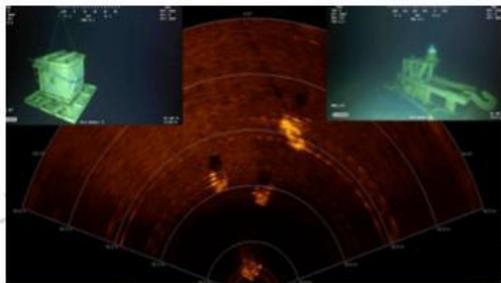
智能音箱

激光雷达

声控灯

手机语音交互

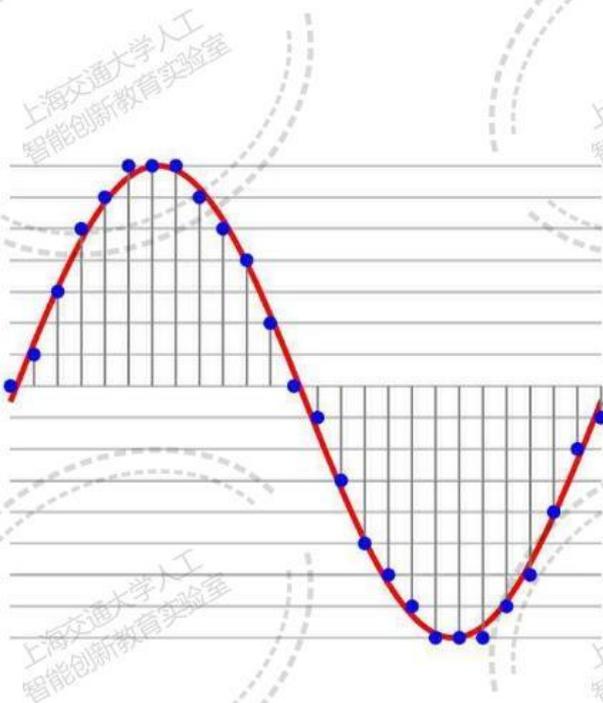
交互机器人



Siri



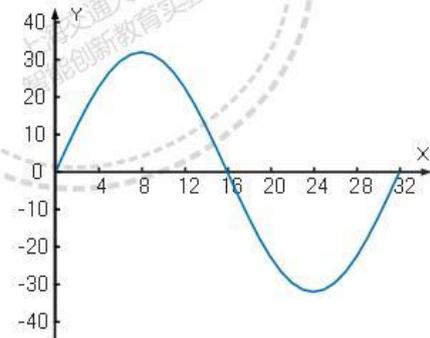
# 声音的数字化过程



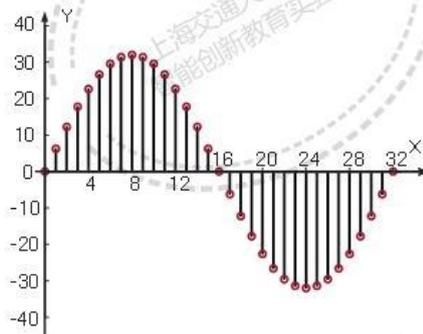
- 所谓**采样**，就是从时间轴上等间隔的取N个时间点，然后再取对应于该点的N个信号的幅值，这是对模拟信号从时间上的离散化。
- 采样频率过低时，丢失的连续信号中包含的信息就越多；采样频率过高，对计算机的内存，计算量等性能要求较高，也容易造成资源浪费。
- **奈奎斯特定理**：当采样的样本足够密集的时候，就可以无失真的还原信号。
- 依据奈奎斯特定理，当语音信号采样频率为8kHz时就基本可以满足无失真还原信号的要求。

# 声音的数字化过程

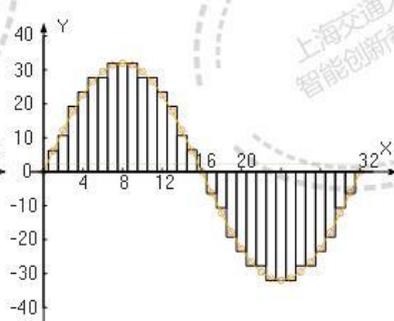
## 连续的模拟声音信号



## 声音信号的采样



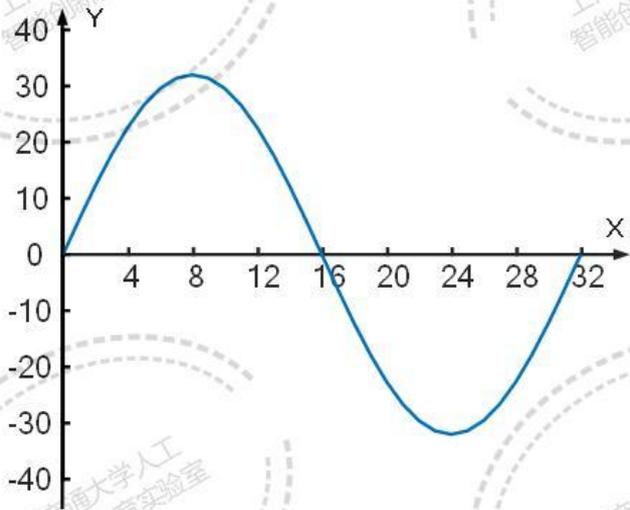
## 离散的声音信号



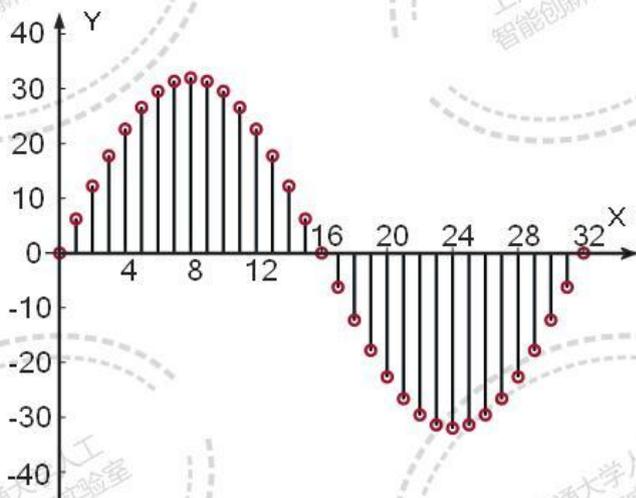
- 数字信号是时间和幅值上的离散信号，因此采样后的信号并不是真正的数字信号。
- 当前采样点到下一个采样点之前，信号幅值保持不变，采样后得到的是幅值是连续，但时间上离散的音频信号，并不是真正的数字信号
- 左图中下图所示是采样信号经保持后所得的离散的音频信号。

# 声音的数字化过程

## 连续的模拟声音信号

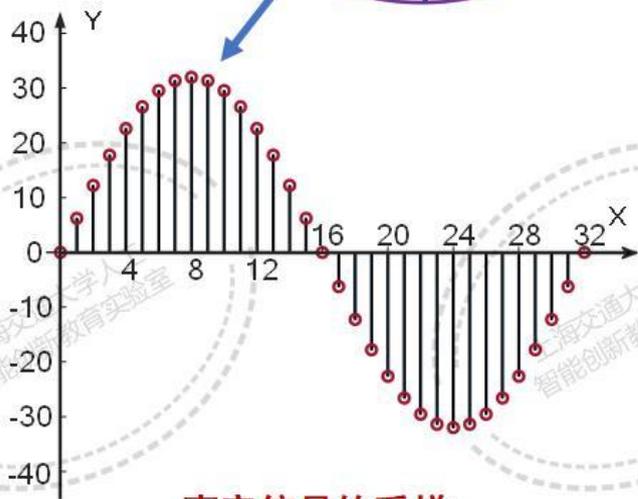


## 声音信号的采样

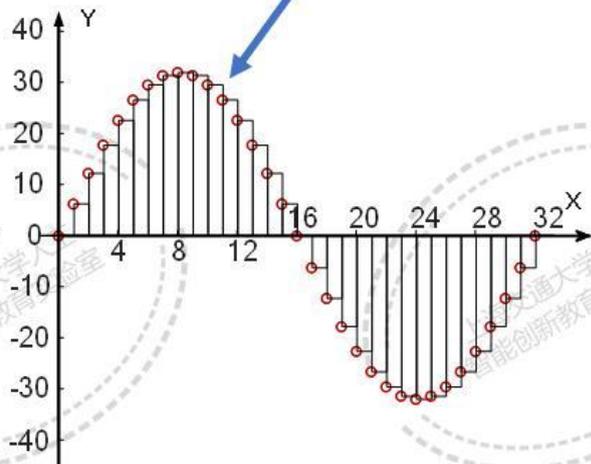


- 所谓**采样**，就是从时间轴上等间隔的取N个时间点，然后再取对应于该点的N个信号的幅值，这是对模拟信号从时间上的离散化。

# 声音的数字化过程



声音信号的采样



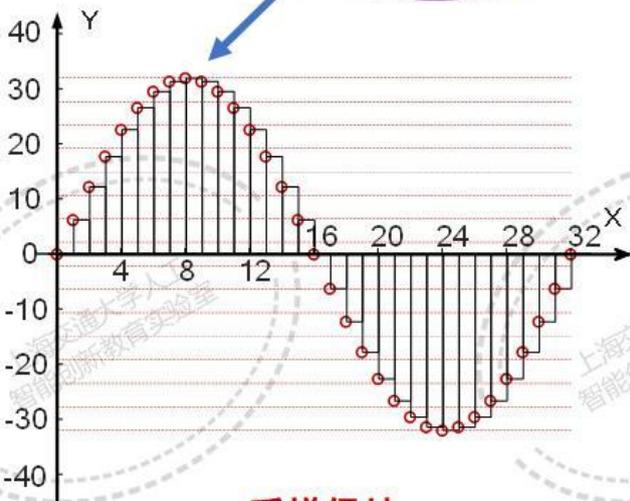
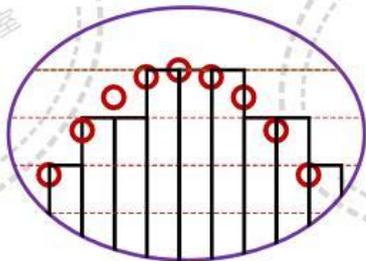
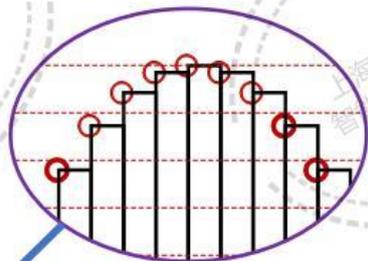
采样保持

# 声音的数字化过程

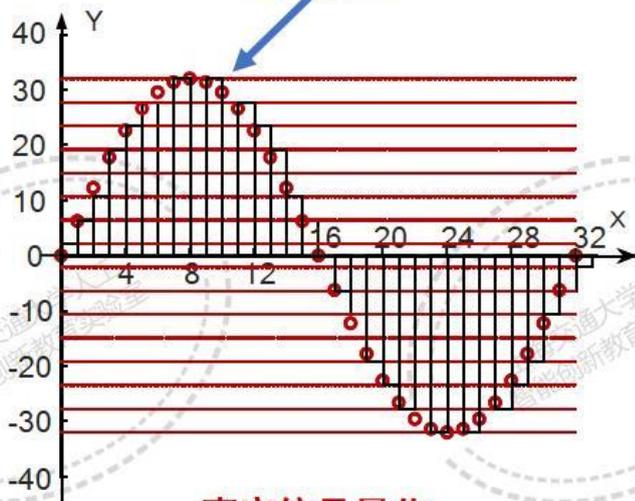
上海交通大学人工智能创新教育实验室

上海交通大学人工智能创新教育实验室

上海交通大学人工智能创新教育实验室



采样保持



声音信号量化

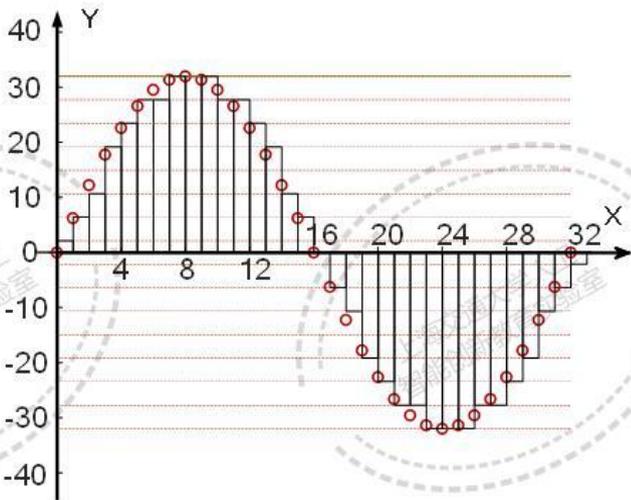
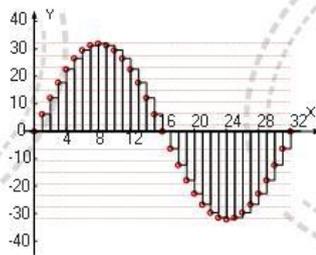
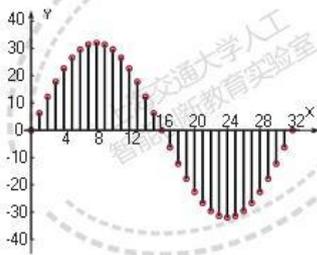
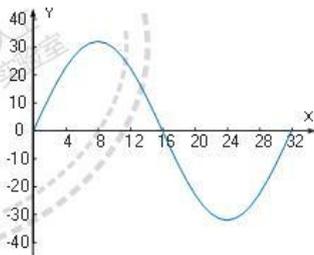
上海交通大学人工智能创新教育实验室

上海交通大学人工智能创新教育实验室

上海交通大学人工智能创新教育实验室

# 声音的数字化过程

0
6.24289
12.24587
17.77825
22.62742
26.60703
29.56415
31.38513
-32
31.38513
29.56415
26.60703
22.62742
17.77825
12.24587
6.24289
0
-6.24289
-12.2459
-17.7782
-22.6274
-26.607
-29.5641
-31.3851
-32
-31.3851
-29.5641
-26.607
-22.6274
-17.7782
-12.2459
-6.24289
0



32	0111
27.734	0110
23.468	0101
19.202	0100
14.936	0011
10.67	0010
6.404	0001
2.138	0000
-2.128	1000
-6.394	1001
-10.66	1010
-14.926	1011
-19.192	1100
-23.458	1101
-27.724	1110
-31.99	1111

2.138	0000
6.404	0001
14.936	0011
19.202	0100
23.468	0101
27.734	0110
27.734	0110
32	0111
32	0111
32	0111
27.734	0110
27.734	0110
23.468	0101
19.202	0100
14.936	0011
6.404	0001
2.138	0000
-6.394	1001
-10.66	1010
-19.192	1100
-23.458	1101
-27.724	1110
-27.724	1110
-31.99	1111
-31.99	1111
-31.99	1111
-27.724	1110
-27.724	1110
-23.458	1101
-19.192	1100
-10.66	1010
-6.394	1001
-2.128	1000

## 声音的数字化过程

- 经采样和保持后的信号还需要进行**量化**，才能得到幅值上也离散的信号
- 现在用一个N位二进制数来表示某一次采样的电平值大小，N位二进制只能表征  $M = 2^N$  个电平值，不能与无穷多个电平值相对应。（前例 N=4, M=16）
- 采样值必须被划分为M个离散电平值，这样的电平才被称作为**量化电平**。

● 量化单位: 
$$q = \frac{Rank}{2^N - 1} (V) \quad Rank = U_{\max} - U_{\min}$$

前例:

$$q = \frac{32 - (-32)}{2^4 - 1} = 4.266$$

- 判断某一采样值u值中有多少个量化单位q。

# 声音的数字化过程

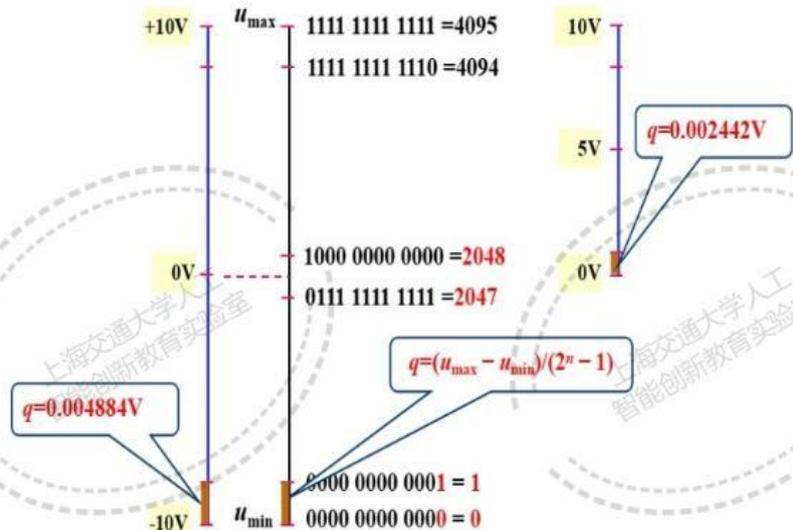
二进制数	十进制数	双极性 -10V~+10V量程	单极性 0~+20V量程
1111 1111 1111	4095	+10V	20V
1111 1111 1110	4094		
...	...	...	...
1000 0000 0000	2048	+0.002442V	10.002442V
0111 1111 1111	2047	-0.002442V	9.997558V
...	...	...	...
0000 0000 0001	1		
0000 0000 0000	0	-10V	0V

从广义的角度来说，每一种语言，都可以认为对应为一种编码的方式。

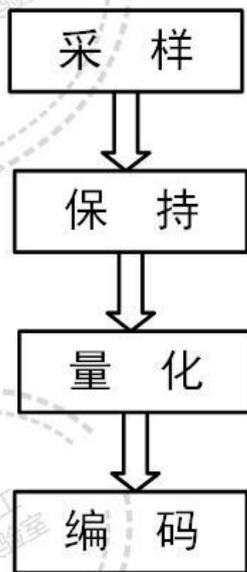
计算机的世界中常见的编码方式有二进制编码、十六进制编码等。

右图上部分是某个12位的A/D转换器所获得的模拟信号大小与二进制的对应关系的表格

右图下部分中q是该A/D的量化单位，下图所示是二进制的编码过程。



# 声音的数字化过程



上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

## 第二节 图像是二维信号

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

# 图像处理技术应用前景



高铁刷脸进站



指纹识别



无人驾驶



引自<https://blog.csdn.net/GarfieldEr007/article/details/86813549>

# 图像处理技术发展历史

1957年，第一张数码照片出现了，图像处理系统可以将照片直接传输至计算机，产生数码照片。

七八十年代，随着现代电子计算机的出现，计算机视觉技术也初步萌芽。

进入21世纪，得益于互联网兴起和数码相机的普及，带来了海量数据，加之机器学习方法的广泛应用，计算机视觉发展迅速。

1970年

1990年

2010年

1950年

1960年

1980年

2000年

60年代末，数字图象处理形成一个比较完整的理论与技术体系，从而构成了一门独立的技术。

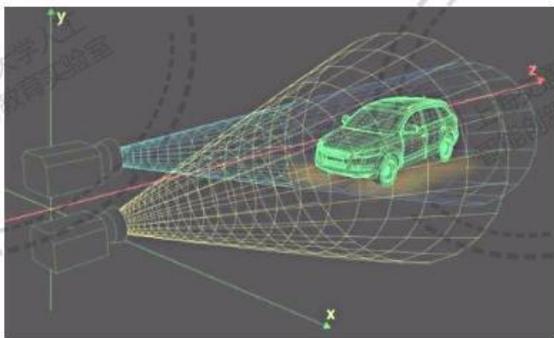
九十年代计算机视觉技术取得了更大的发展，开始广泛应用于工业领域。硬件技术和软件算法的进步促进了计算机视觉的发展。

2010年以后，借助于深度学习的力量，计算机视觉技术得到了爆发增长和产业化。

# 图像处理技术发展历史



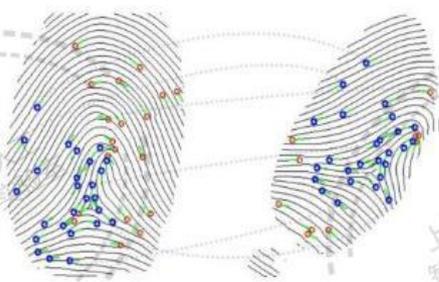
1957年诞生的第一张数码照片



90年代后的三维重建

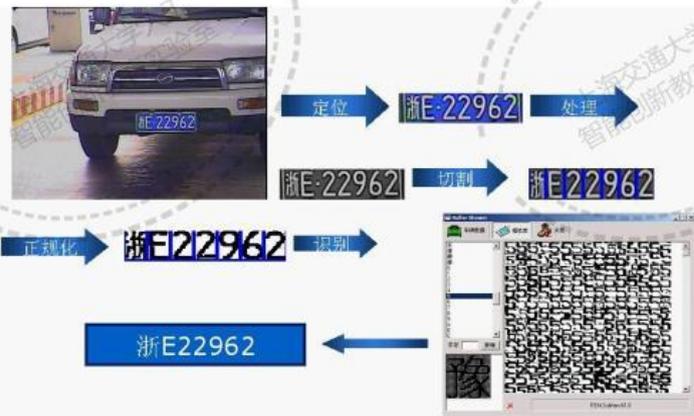


1964年获取的第一张月球表面的照片

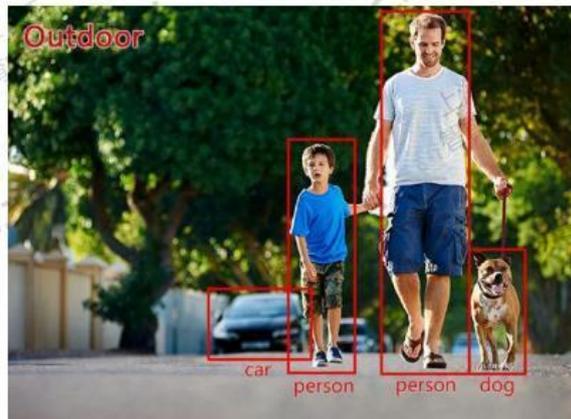


90年代后的局部特征与匹配

# 图像处理技术发展历史



21世纪后机器学习应用于计算机视觉，车牌识别和人脸检测识别问题得以进一步解决

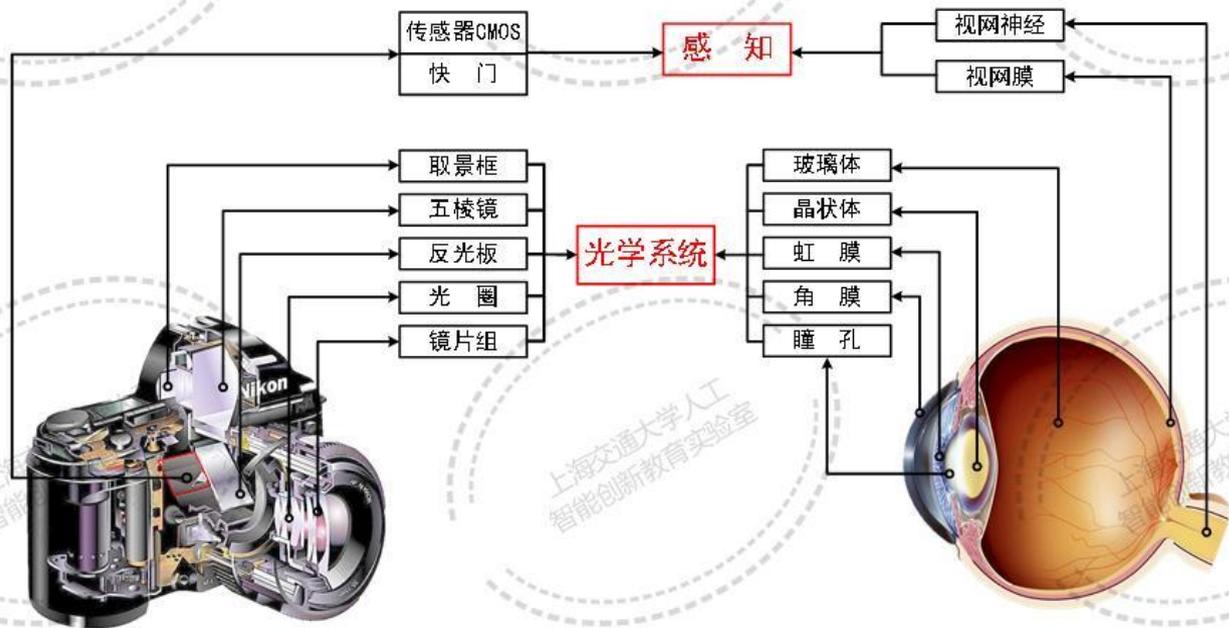


2010年以后，借助于深度学习，物体识别准确率达到前所未有的高度

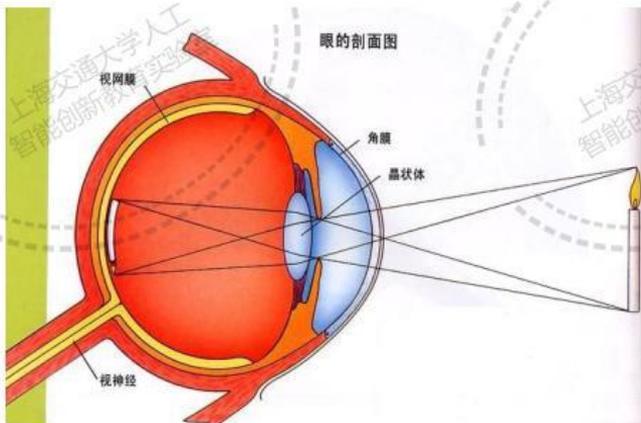


# 人眼结构和人眼成像

## 人眼结构与相机成像结构对比关联



# 人眼结构和人眼成像

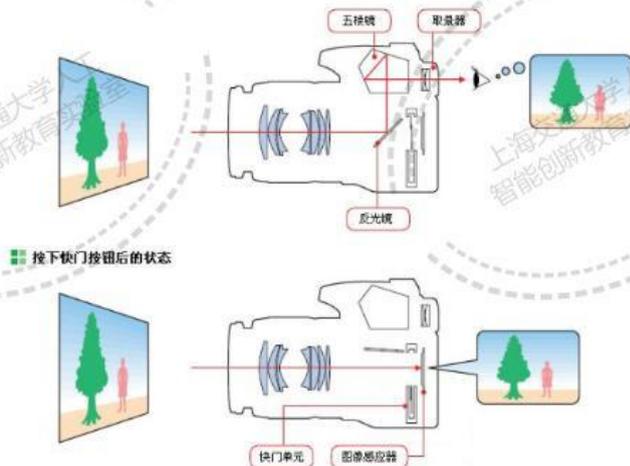


## 人眼成像原理

人眼结构中的晶状体相当于一个凸透镜  
人眼结构中的视网膜则相当于可以用来承接凸透镜实像的光屏。

视觉成像就是物体的反射光通过晶状体折射成像于视网膜上，再由视觉神经感知传给大脑，这样人就看到了物体。

人眼通过调节晶状体来控制成像的大小和清晰程度



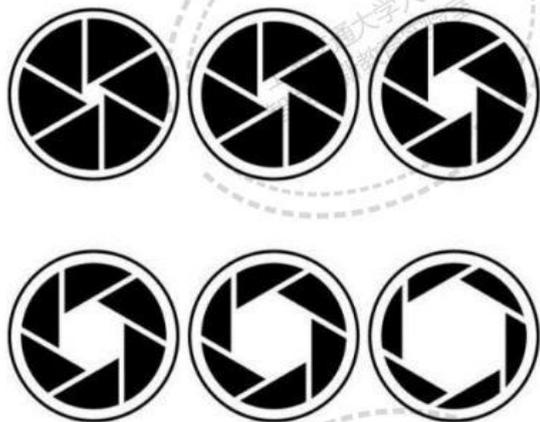
## 单反相机成像过程

在快门按钮按下之前，通过镜头的光线由反光镜反射至取景器内部；

当按下快门按钮时，反射镜迅速抬起，光线通过透镜后直接射向快门帘幕，与此同时，快门帘幕打开让光线通过进入图像传感器形成图像；

之后快门关闭，反射镜落到初始位置，完成拍摄过程。

## 相机组件介绍



**光圈：**作用类似瞳孔，通过调节光圈大小控制进入相机光线数量，光圈大小用 $f$ 表示， $f/1.4$ 的镜头比 $f/2$ 的镜头更亮， $f$ 比 $f/2.8$ 的镜头更亮，以此类推。

## 相机组件介绍



不同快门速度对照片质量与风格的影响

快门速度越慢，曝光时间越长，照片会越亮；

快门速度还会影响在运动中的拍摄效果，如果快门速度不够快，会因运动导致物体上的同一点在多个像素上曝光而引起运动模糊。

# 相机组件介绍



不同感光度下照片

## 相机组件介绍

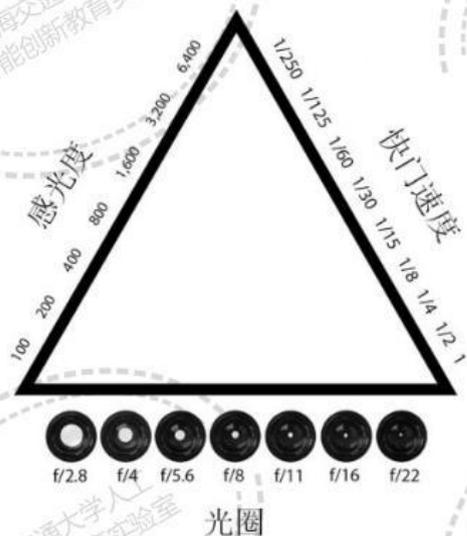


感光度是一个用来衡量数码相机中感光元件对光线的灵敏度的值

感光度常用ISO值表示，ISO值越大，感光度越强

一般来说，在低感光度下拍摄出的照片画质更为细腻，色彩也更为真实；而高感光度下的照片更容易出现噪声的问题，降低照片画质。如右图所示。

# 相机组件介绍

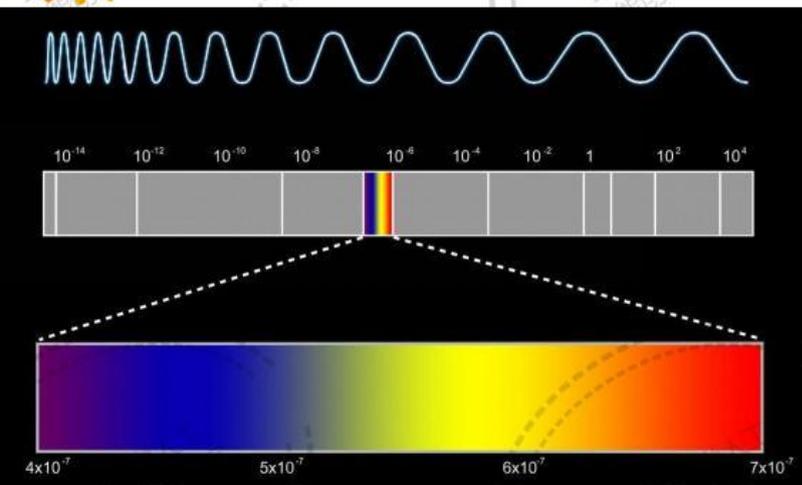


- 影响曝光的三个要素：感光度、光圈和快门速度
- 三者之间相互补偿，例如，为了更好的成像效果需要调整其中某个因素，就会增加或减少曝光度，这时可通过其他因素进行补偿。

# 图像采集系统



能量源



可见光波长与频率

颜色	波长	频率
红色	625-740纳米	480-405兆赫
橙色	590-625纳米	510-480兆赫
黄色	565-570纳米	600-530兆赫
绿色	500-565纳米	620-600兆赫
紫色	485-500纳米	680-620兆赫
蓝色	380-440纳米	790-680兆赫

- 图像获取系统的组成通常包括三个部分：**能量源**，**光学系统**和**图像传感器**
- 上图所示是常见是能量源
- 光学系统是指由透镜、反射镜等多种光学元件按一定次序组合成的系统，通常用来成像或者做一些简单的光学信息处理。

# 图像采集系统 空间采样



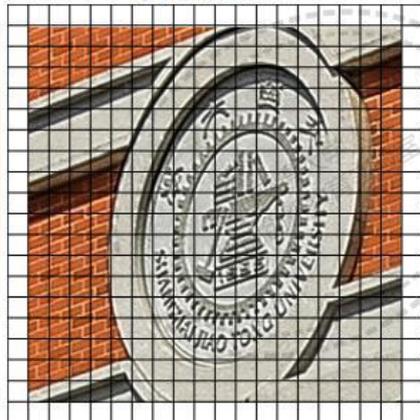
5472\*3648



20\*20



300\*300

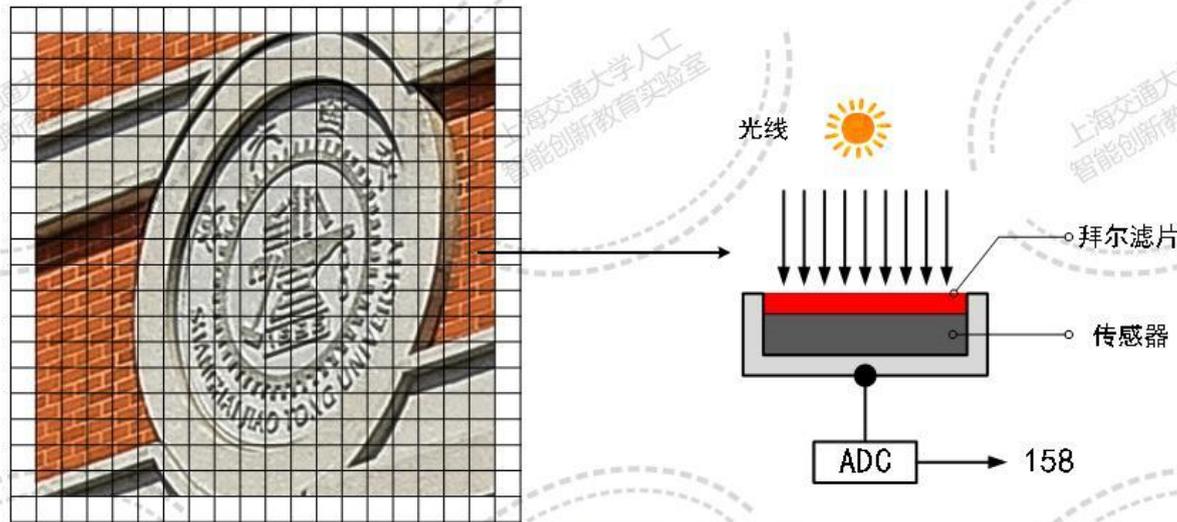


空间采样:

图像传感器的像元的多少决定了空间的分辨率;

像元越多, 空间分辨率越高, 图像越细腻越清晰

# 图像获取系统 传感器的采样与量化



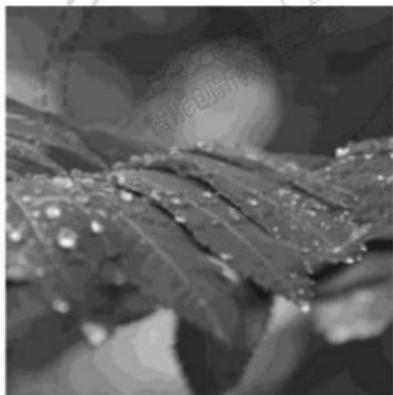
图像传感器采样量化原理

- 图像获取系统的组成通常包括三个部分：能量源，光学系统和图像传感器
- 图像传感器由一个二维阵列组成，里面的每个元素称为像素，每个像素都能测量投射到它上面的光线数量并把它转换成电压信号，然后再通过后面的AD转换器转换成相应的数字信号。
- 像素的总数目称为空间分辨率。

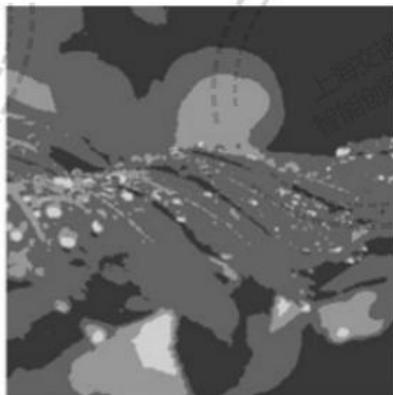
## 图像获取系统 传感器的采样与量化



256 gray-levels



16 gray-levels

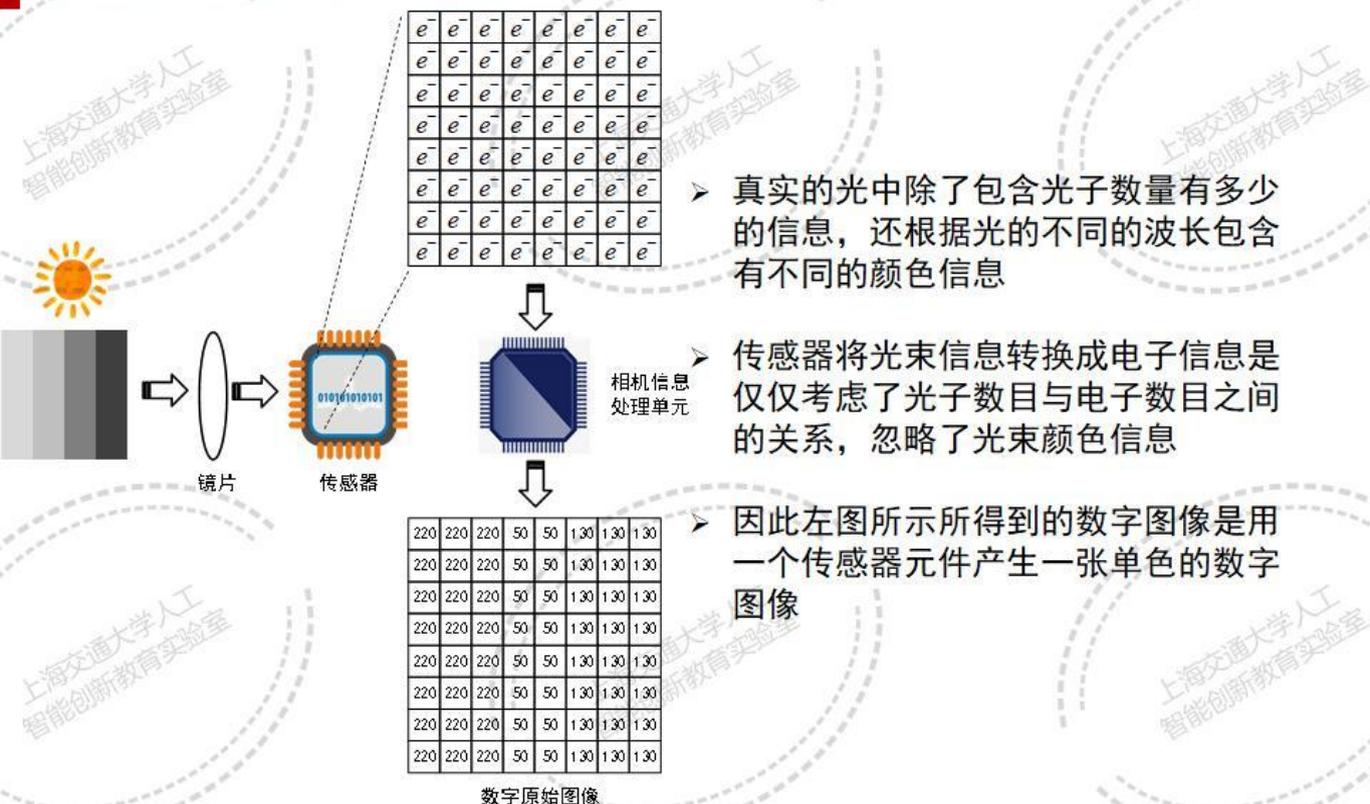


4 gray-levels

- 像素的总数目称为空间分辨率
- 不同灰度级分辨率对图像影响不同
- 最常用的是8位256级的灰度级分辨率
- 上图所示是不同灰度级表示的图像



# 相机传感器的工作原理



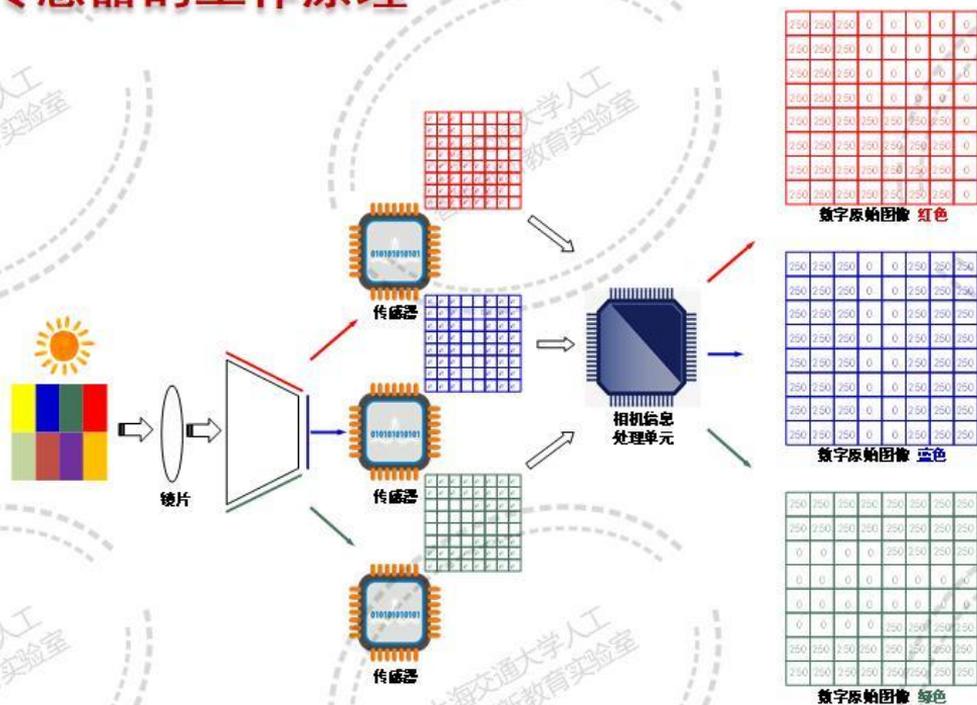
➤ 真实的光中除了包含光子数量有多少的信息，还根据光的不同的波长包含有不同的颜色信息

➤ 传感器将光束信息转换成电子信息是仅仅考虑了光子数目与电子数目之间的关系，忽略了光束颜色信息

➤ 因此左图所示所得到的数字图像是用一个传感器元件产生一张单色的数字图像

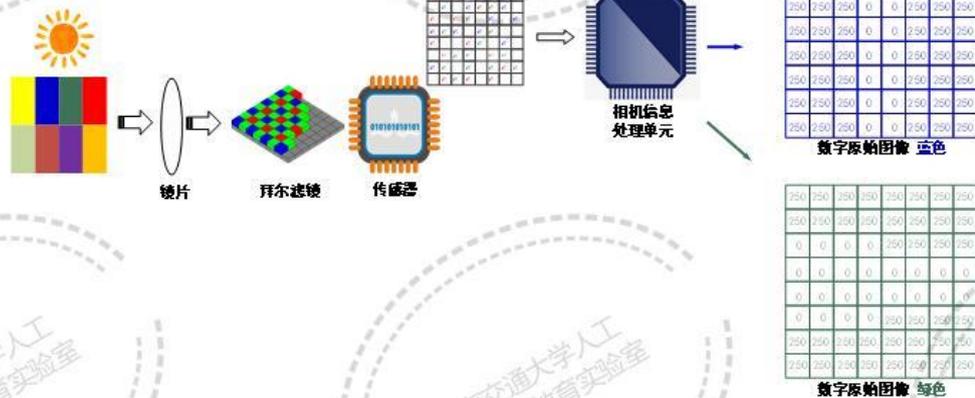
用一个元件产生一张单色的数字图像的过程

# 相机传感器的工作原理



- 可以用一个分光棱镜将光线中的红、绿、蓝三个基本色分开,对每个颜色都单独投射到一个传感器元件上。
- 最后将三个传感器元件收集到的三种单色的信息合成可以产生效果非常好的彩色图像
- 但是这样获得的彩色数字图像成本非常高,无法得到广泛的推广

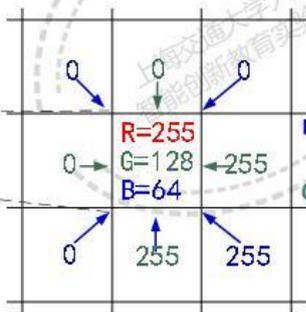
# 相机传感器的工作原理



- 改进后用1个传感器元件就可以产生彩色数字图像
- 如左图所示，将彩色滤光片像马赛克一样分布在传感器所有的像素上
- 使每个像素只能产生红、绿或蓝色三种颜色中的一种颜色的值
- 再经过相机处理单元处理以后依然可以输出一幅彩色的数字图像

# 相机传感器的工作原理

0	0	0	0	255	255	0	255
255	0	255	255	0	0	0	0
0	0	0	0	255	255	0	255
255	0	255	255	0	0	0	0
255	0	255	255	0	255	255	255
255	255	255	0	255	255	0	255
255	0	255	255	0	255	255	255
255	255	255	0	255	255	0	255

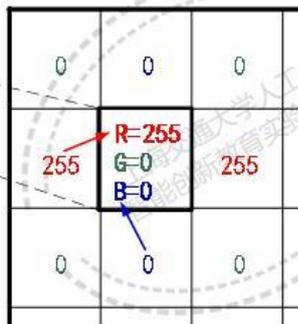


临近像素均值法  
(双线性插值)

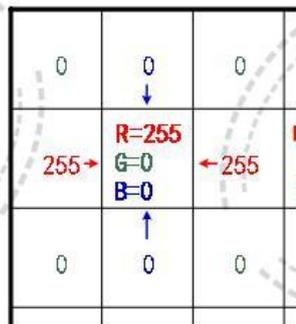
$$R = \frac{0 + 0 + 0 + 255}{4}$$

$$G = \frac{0 + 0 + 255 + 0}{4} \cdot 255$$

0	0	0	0	255	255	0	255
255	0	255	255	0	0	0	0
0	0	0	0	255	255	0	255
255	0	255	255	0	0	0	0
255	0	255	255	0	255	255	255
255	255	255	0	255	255	0	255
255	0	255	255	0	255	255	255
255	255	255	0	255	255	0	255



临近像素复制法

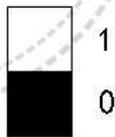


$$R = \frac{255 + 255}{2} = 255$$

$$B = \frac{0 + 0}{2} = 0$$

# 计算机看到的图像

**黑白图像：**图像的每个像素只能是黑或者白，没有中间的过渡，故又称为二值图像。



**灰度图像：**灰度图像指的是每个像素的信号值是由一个量化后的灰度级来描述的图像，没有彩色信息，0级表示黑色，255级表示白色

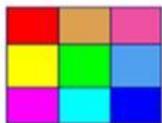


216	184	160	215
159	195	164	244
198	177	175	227
197	148	177	209
112	209	176	233
161	158	162	233
187	188	185	245
182	186	174	216
186	242	223	251

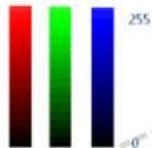


# 计算机看到的图像

彩色图像：在RGB空间中，每幅彩色图像可以看成是由红(R)、绿(G)、蓝(B)三个不同颜色通道的灰度图像组合而成的。



$$R = \begin{bmatrix} 255 & 240 & 240 \\ 255 & 0 & 80 \\ 255 & 0 & 0 \end{bmatrix} \quad G = \begin{bmatrix} 0 & 160 & 80 \\ 255 & 255 & 160 \\ 0 & 255 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 80 & 160 \\ 0 & 0 & 240 \\ 255 & 255 & 255 \end{bmatrix}$$



## 计算机看到的图像



彩色图像分离出的R、G、B通道分量

- 一张RGB空间的彩色图片可以分离出R、G、B三个通道分量，每个通道分离出的图片都是灰度图片。
- 三个通道叠加后才产生彩色图像

## 计算机看到的图像

0	1	2	2	0
0	1	1	0	0
2	2	0	2	1
0	1	1	0	2
1	2	2	1	1

红色分量

1	0	2	1	1
1	2	2	1	0
1	1	0	2	1
0	0	0	1	1
2	1	1	2	0

绿色分量

0	2	2	0	2
1	1	1	1	1
0	0	1	2	2
0	0	1	2	2
0	2	2	2	0

蓝色分量

- 但是在计算机眼里，每个通道实际上就是一个二维矩阵，矩阵中每个元素值就是一个像素值
- 计算机眼里看到的实际上是一系列的二维数字

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

## 第三节 一线雷达信号简介

上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

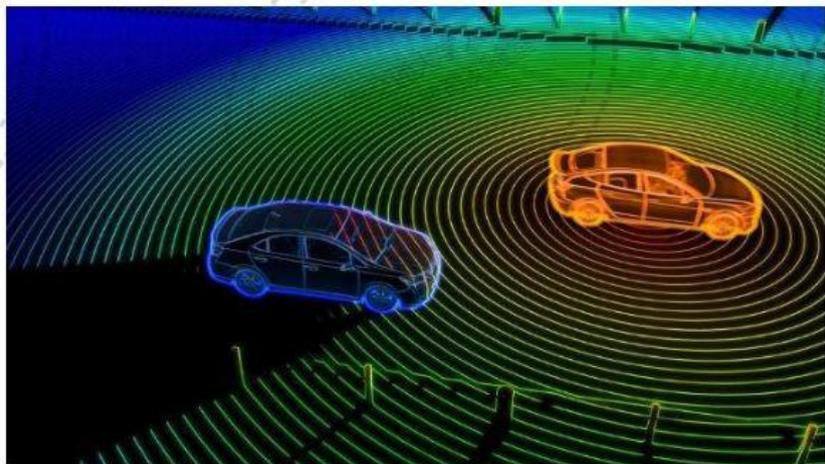
上海交通大学人工  
智能创新教育实验室

# 激光雷达的介绍



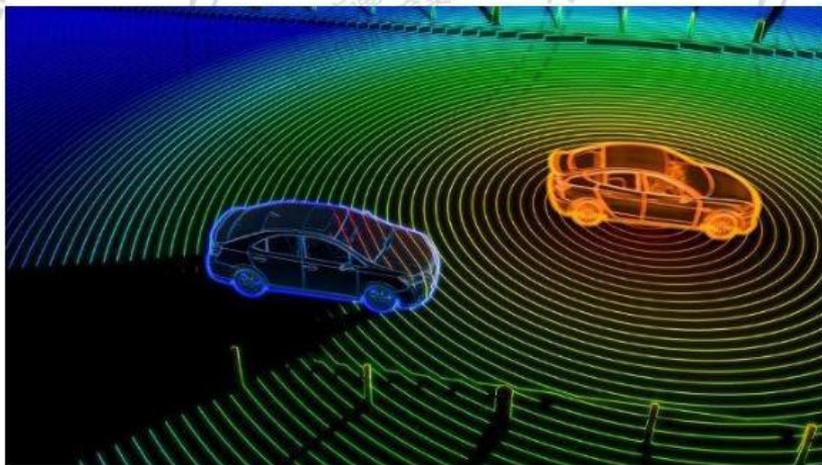
- 激光雷达用于扫描获取周围环境信息
- 本课程所介绍微缩车采用的是1线激光雷达
- 除1线激光雷达以外，还有多线激光雷达，如32线激光雷达、64线激光雷达等

# 激光雷达的介绍



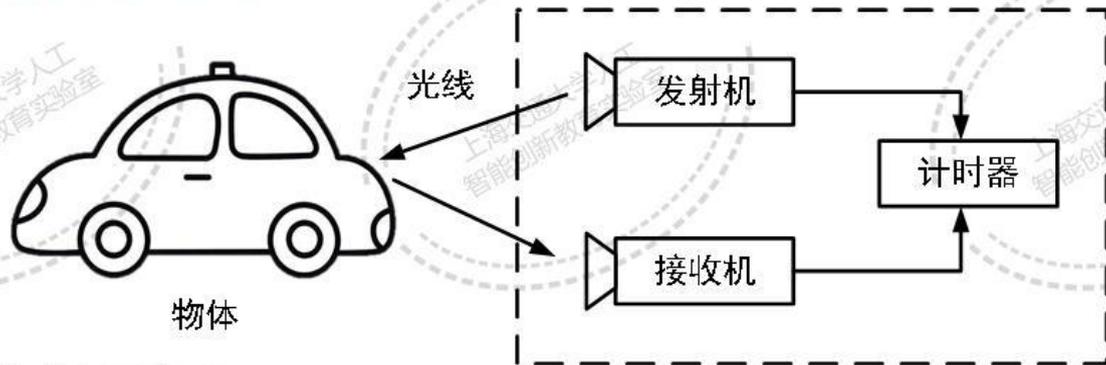
- 激光雷达的系统主要分成三部分：
- 激光发射器，用于发射出波长为600nm到1000nm之间的激光射线
- 扫描与光学部件，主要用于收集反射点距离与该点发生反射的时间和水平角度
- 感光部件，主要检测返回光的强度。

## 激光雷达的介绍



- 激光雷达检测到的每一点都包括了空间坐标信息  $(x, y, z)$  以及光强度的信息  $(i)$
- 不同物体的光反射一般都不一样，根据检测到的光强度就可以对检测到的物体有初步判断

# 激光雷达的介绍



1线激光雷达测距原理

一线激光雷达

一线激光雷达发出一个脉冲波式的激光，光速用 $c$ 表示

激光从发射时开始计时

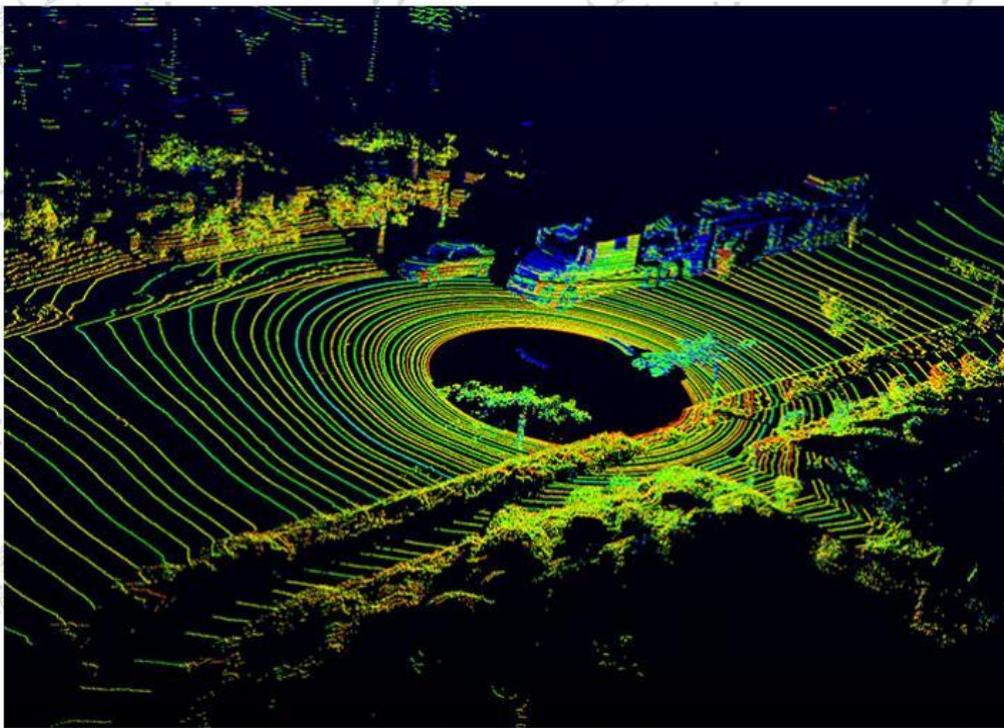
$$D = \frac{c \times t}{2}$$

激光到达物体后反射回来再次被激光雷达接收

往返一次所需的时间为 $t$

该物体距离一线激光雷达的距离为

# 激光雷达的介绍



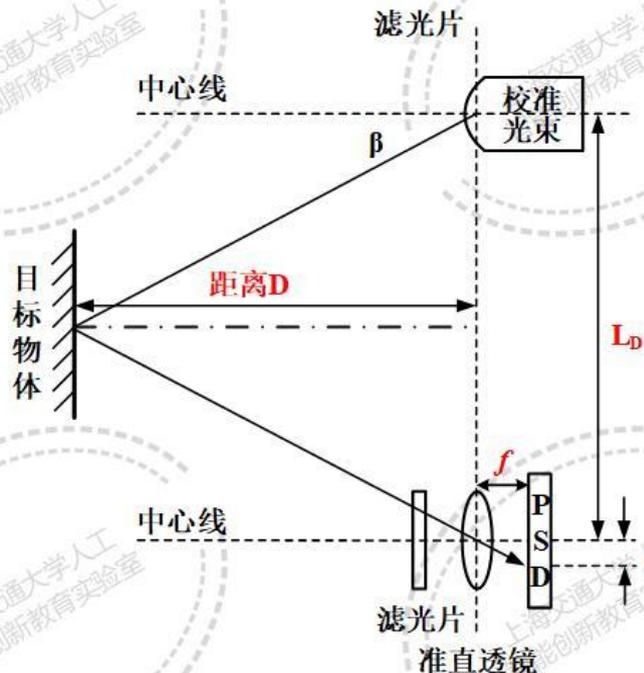
64维激光雷达用于场景还原

## 红外测距原理

红外发射器按照一定的角度发射红外光束，当遇到物体以后，光束会反射回来，如图所示。

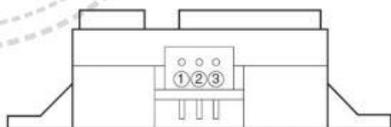
反射回来的红外光线被CCD检测器（PSD为位置敏感检测装置，12位的A/D转换器）检测获得一个偏移值  $x$

利用三角关系，发射角度，偏移距  $x$ ，中心矩，以及透镜的焦距以后，传感器到物体的距离就可以通过几何关系计算出来了。



# 红外测距原理

输出电压与距离的关系是非线性的（图中曲线可知），为便于程序中使用距离信息，必须将模拟电压信号转换为相应的距离值

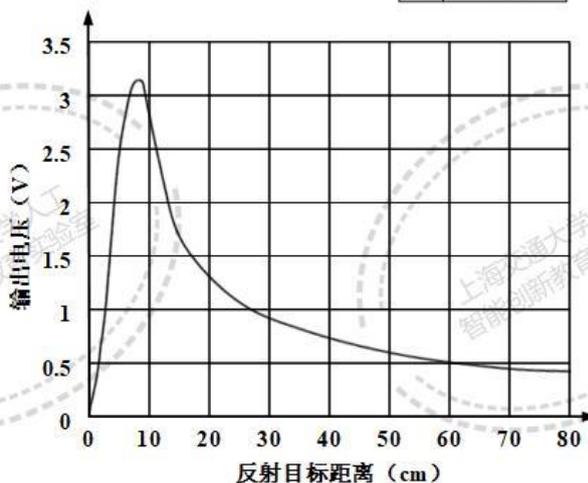


名称	
①	V <sub>0</sub>
②	GND
③	V <sub>CC</sub>

$$(float) R = \frac{1}{A2D * m + b} - k$$

相应A/D转换后的数值A2D，通过学到的直线拟合方法（线性回归），得到直线的斜率 $m$ 和截距 $b$ ；

$k$ 是个常数， $k=4$ 。



## 本章总结

- ▶ 计算机通过麦克风获取到包含声音信息的数字信号，获得“耳朵”
- ▶ 计算机通过图像传感器获取到数字图像信息，获得“眼睛”
- ▶ 计算机还可以通过一线激光雷达和红外测距获取到外界环境中其他信息，获得其他的“感官”
- ▶ 计算机逐渐成为“眼明心亮”，能够感知周围环境的计算机
- ▶ 计算机的世界里只有“0”和“1”，它所有“看到的”、“听到的”、或者其他方式获得的信息都只能用“0”和“1”来表示
- ▶ 最终目的是实现计算机的智能化



上海交通大学人工  
智能创新教育实验室

上海交通大学人工  
智能创新教育实验室

稍后继续

THANKS FOR YOUR ATTENTION