

# 机器学习

计算机视觉基本理论

DAY01

# 计算机视觉基础

## 计算机视觉基础

### 计算机视觉概述

什么是计算机视觉

计算机视觉的应用

计算机视觉相关学科

计算机视觉与人工智能

### 数字图像处理基础

人眼成像原理

计算机成像原理

图像采样与分辨率

灰度级与灰度图像

彩色图像与色彩空间

颜色空间变化

图像灰度化

二值化与反二值化

# 计算机视觉概览

---

# 什么是计算机视觉

- 计算机视觉在广义上是和图像相关的技术总称。包括图像的采集获取，图像的压缩编码，图像的存储和传输，图像的合成，三维图像重建，图像增强，图像修复，图像的分类和识别，目标的检测、跟踪、表达和描述，特征提取，图像的显示和输出等等。
- 随着计算机视觉在各种场景的应用和发展，已有的图像技术也在不断的更新和扩展。



# 计算机视觉的应用

- 计算机视觉技术已经在许多领域得到了广泛的应用，以下是一些典型的例子：

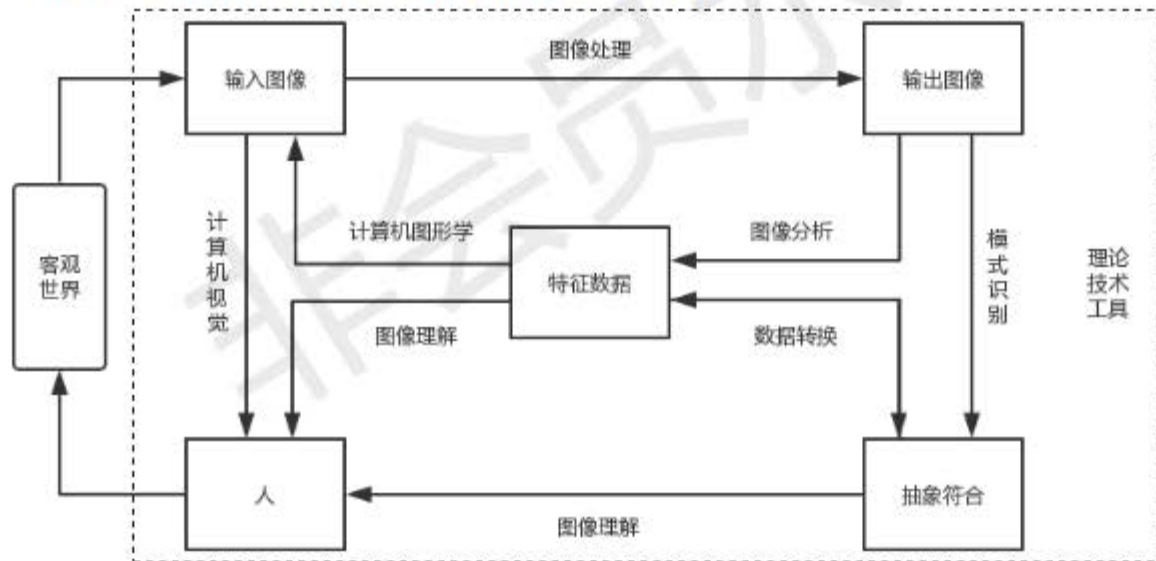
- ✓ 公安安防：人脸识别，指纹识别，场景监控，环境建模。
- ✓ 生物医学：染色体分析，X光、CT图像分析，显微医学操作。
- ✓ 文字处理：文字识别，文档修复，办公自动化，垃圾邮件分类。
- ✓ 国防军事：资源探测，军事侦察，导弹路径规划。
- ✓ 智能交通：公路交通管理，电子警察执法抓拍系统，自动驾驶车辆。
- ✓ 休闲娱乐：电影特效，视频编辑，人像美颜，体感游戏，VR。





# 计算机视觉相关学科

- 计算机视觉是一门研究图像理论、技术和应用的交叉学科。计算机视觉不仅和传统的数学、物理学、生理学、心理学、计算机科学、电子工程等学科相关。并且还涉及到计算机图形学、图像模式识别、图像工程等专业技术，这几个技术名词相互关联，经常混合使用，在许多情况下他们只是专业背景不同的人习惯使用的不同称呼术语。





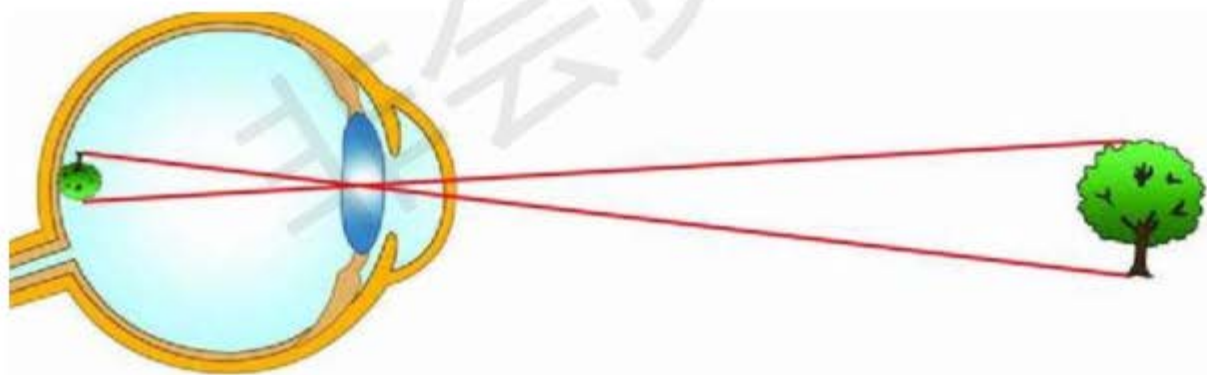
# 数字图像处理基础

---



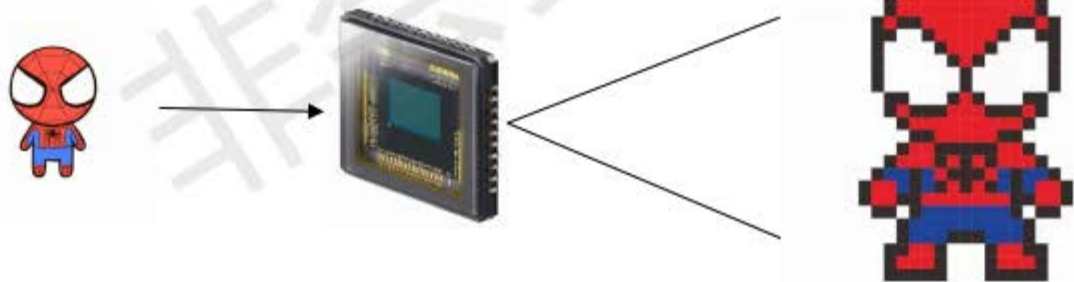
# 人眼成像原理

- 人的眼睛近似为一个球体。物体的光线通过角膜和晶状体的折射，在视网膜上成倒立缩小的实像。
- 视网膜上分布光线接收的神经细胞，分为锥状体和杆状体。每只眼睛有600万-700万个锥状体，其对颜色灵敏度很高，负责亮光视觉。有7500万-15000万杆状体，杆状体没有颜色感觉，负责暗视觉。



# 计算机成像原理

- 数字图像的采集过程类似人眼，使用大量的光敏传感器构成的阵列获取图像。成像的质量由传感器的单元数，尺寸和传感性能决定。
- 多数传感器的输出是连续的电压波形，图像数字化就是将一副画面的数据转换为计算机能够处理的数字形式。
- 图像数字化包括两种处理过程：采样和量化。



# 图像采样与分辨率

- 将空间上连续的图像变换成离散点的操作称为**采样**。
- 采样是按照某种时间间隔或空间间隔，采集模拟信号的过程，即空间离散化。
- 图像数字化的采样过程是将空间上连续的图像变化为离散的点。
- 采样的效果由传感器的采样间隔和采样孔径决定，采样间隔和采样孔径的大小是两个很重要的参数。



# 图像采样与分辨率（续1）

- 采样后得到离散图像的尺寸称为图像分辨率。分辨率是数字图像可辨别的最小细节。
- 分辨率由宽（width）和高（height）两个参数构成。宽表示水平方向的细节数，高表示垂直方向的细节数。
- 例如：
  - ✓ 一副640\*480分辨率的图像，表示这幅图像是由  
 $640 \times 480 = 307200$ 个点组成。
  - ✓ 一副1920\*1080分辨率的图像，表示这幅图像是由  
 $1920 \times 1080 = 2073600$ 个点组成。





# 图像采样与分辨率（续2）

- 采样间隔越小，所得图像像素数越多，空间分辨率高，图像质量好，但数据量大。

下图展示了lena图的分辨率从512x512依次降低到8x8的图像效果。



# 灰度级与灰度图像

- 灰度级 (depth) 表征了每个采样点的传感器输出中可分辨的最小变化。
- 灰度级通常是2的整数次幂。我们用m级或者n位来表示灰度级。图像数据的灰度级越多视觉效果就越好。计算机中最常用的是8位图像。
- 例如：
  - ✓ 一副8位的图像，表示每个采样点有 $2^8=256$ 级。从最暗到最亮，可以分辨256个级别。
  - ✓ 一副32级的灰度图像，每个采样点从最暗到最亮，可以分辨32个级别。





# 灰度级与灰度图像（续1）

- ✓ 量化等级越多，所得图像层次越丰富，灰度分辨率高，图像质量好，但数据量大。下图展示了lena图的灰度级从256级依次降低到4级的图像效果。





# 彩色图像与色彩空间

- 为了表征彩色图像，我们需要使用多通道数字图像。最普遍的方式是使用RGB颜色空间。RGB颜色空间中每个像素点有三个维度，分别记录在红（Red）、绿（Green）、蓝（Blue）三原色的分量上的亮度。

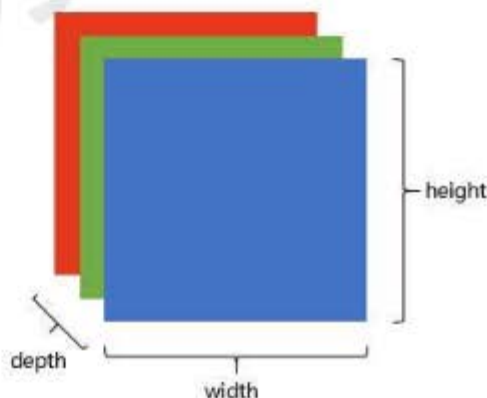
例如：

✓ 按照(r,g,b)的方式：

✓ (255,0,0) 纯红

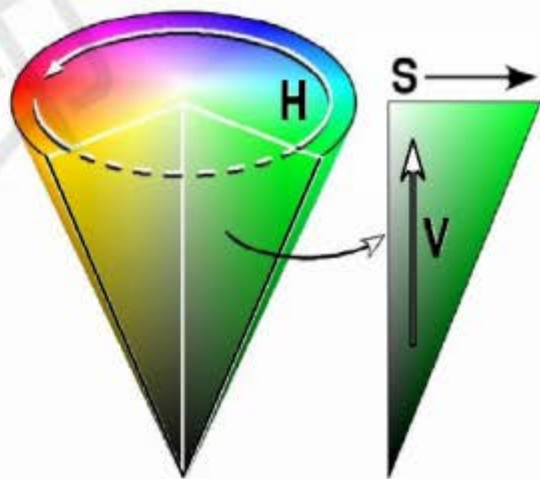
✓ (124,252,0) 草坪绿

✓ (135,206,235) 天蓝色



# 彩色图像与色彩空间（续1）

- 另一种常用的颜色空间是HSV，该颜色空间可以用一个圆锥来表示。
- HSV表示色相(hue)、饱和度(saturation)和亮度(value)。
- H表示颜色的相位角(hue)，取值范围是0---360；S表示颜色的饱和度(saturation)，范围从0到1，它表示成所选颜色的纯度和该颜色最大的纯度之间的比率；
- V表示色彩的明亮程度(value)，范围从0到1。





# 彩色图像与色彩空间（续2）

- YUV：亮度信号Y和两个色差信号R - Y、B - Y，最后发送端将亮度和色差三个信号分别进行编码。采用YUV色彩空间的重要性是它的亮度信号Y和色度信号U、V是分离的。如果只有Y信号分量而没有U、V分量，那么这样表示的图就是黑白灰度图。
- CMYK：CMYK颜色空间应用于印刷工业,印刷业通过青(C)、品(M)、黄(Y)、黑(BK)四色油墨的不同网点面积率的叠印来表现丰富多彩的颜色和阶调。
- Lab：Lab的色彩空间要比RGB模式和CMYK模式的色彩空间大，自然界中任何一点色都可以在Lab空间中表达出来。



# 颜色空间变化

- 在计算机视觉中，尤其是颜色识别相关的算法设计中，各种颜色空间混合使用是常见的方法。RGB，HSV，YUV等常见颜色空间可以通过计算公式实现相互转化，这个过程叫做颜色空间变化。颜色变换的计算公式比较复杂，通常图像处理库会提供颜色空间变化的API给用户调用。





# 图像灰度化

- 在RGB模型中，如果 $R=G=B$ 时，则彩色表示一种灰度颜色，其中 $R=G=B$ 的值叫灰度值，因此，灰度图像每个像素只需一个字节存放灰度值（又称强度值、亮度值），灰度范围为0-255。将RGB图像转换为灰度图像的过程称为图像灰度化处理。



# 图像灰度化（续）

- 灰度化处理方法包括：
  - ✓ 分量法。将彩色图像中的三分量的亮度作为三个灰度图像的灰度值，可根据应用需要选取一种灰度图像。
  - ✓ 最大值法。将彩色图像中的三分量亮度的最大值作为灰度图的灰度值。
  - ✓ 将彩色图像中的三分量亮度求平均得到一个灰度值。
  - ✓ 根据重要性及其它指标，将三个分量以不同的权值进行加权平均。例如，由于人眼对绿色的敏感最高，对蓝色敏感最低，因此，按下式对RGB三分量进行加权平均能得到较合理的灰度图像。如：

$$f(i, j) = 0.30 R(i, j) + 0.59 G(i, j) + 0.11 B(i, j)$$



# 二值化与反二值化

- 二值化阈值处理是将原始图像处理为仅有两个值的二值图像，对于灰度值大于阈值 $t$ 的像素点，将其灰度值设定为最大值。对于灰度值小于或等于阈值的像素点，将其灰度值设定为0。



## 二值化与反二值化（续）

- 反二值化阈值处理的结果也是仅有两个值的二值图像，对于灰度值大于阈值的像素点，将其值设定为0；对于灰度值小于或等于阈值的像素点，将其值设定为255。



# 图像计算

## 图像计算

图像计算基本概念

数字图像处理基础

坐标变换

什么是图像计算

图像计算类别

图像加法

图像减法

平移

镜像

旋转

缩放



# 图像计算基本概念

---



# 什么是图像计算

- 数字图像使用矩阵来表示，所以一切矩阵的计算都适用于图像，但按照实际意义，图像的计算一般分为算术运算（点操作）和坐标变换（几何操作）。



# 图像计算类别

- 算术运算：通过一些运算规则，包括加减乘除和逻辑运算，改变了像素点的灰度值，实现一些图像变化的应用。
- 坐标变换：通过像素点的坐标几何变换来改变像素的位置，实现图像的一些改变。



# 算数运算

---

# 图像加法

- 图像加法可以用于多幅图像平均去除噪声：

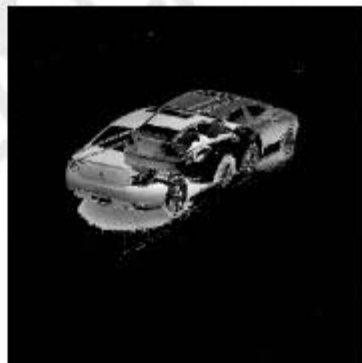


- 图像加法实现水印的叠加：



# 图像减法

- 图像减法是找出两幅图像的差异，可以在连续图像中可以实现消除背景 and 运动检测：



# 坐标变换

---



# 平移



# 镜像



原图



水平镜像



垂直镜像



# 旋转



# 缩放

- 图像缩放 (image scaling) 是指对数字图像的大小进行调整的过程。将分辨率为(w,h)的图像, 缩放为(w', h')的图像, 水平方向系数为 $S_x = w' / w$ , 垂直方向缩放系数为 $S_y = h' / h$ 。缩放变换矩阵为:

$$A = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- 示例:



原始图像



放大图像

# 缩放（续1）

- 图像缩小可以通过删除矩阵中的元素来实现，例如：下面的例子进行隔行、隔列删除后，高度、宽度均减小为原来的一半

1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8



1	1	1	1
3	3	3	3
5	5	5	5
7	7	7	7



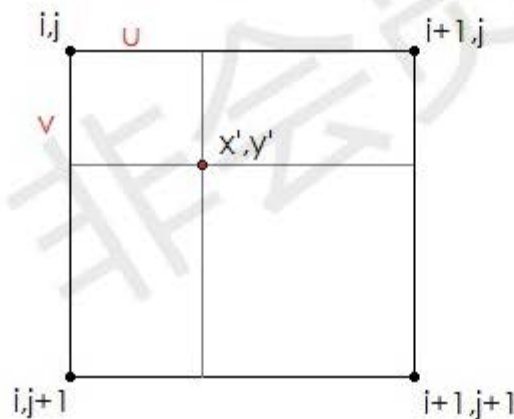


# 缩放（续2）

• 图像放大需要进行像素插入，常用的插值法有最邻近插值法和双线性插值法

– **最邻近插值法**：直接使用新的像素点  $(x', y')$  最近的整数坐标灰度值作为该点的值，该方法计算量小，但精确度不高，并且可能破坏图像中的线性关系

– **双线性插值法**：使用新的像素点  $(x', y')$  最邻近的四个像素值进行插值计算，假设为  $(i, j)$ ,  $(i+1, j)$ ,  $(i, j+1)$ ,  $(i+1, j+1)$ , 则  $u = x' - i$ ,  $v = y' - j$ .



# 缩放（续3）

- 下图是最邻近插值法和双线性插值法的效果对比



最邻近插值



双线性插值

# 图像预处理

## 图像预处理

图像预处理概述

什么是图像预处理

常用预处理技术

直方图

直方图均衡化

模板运算

均值滤波

高斯滤波

边缘检测

锐化

色彩处理

亮度调整

饱和度调整

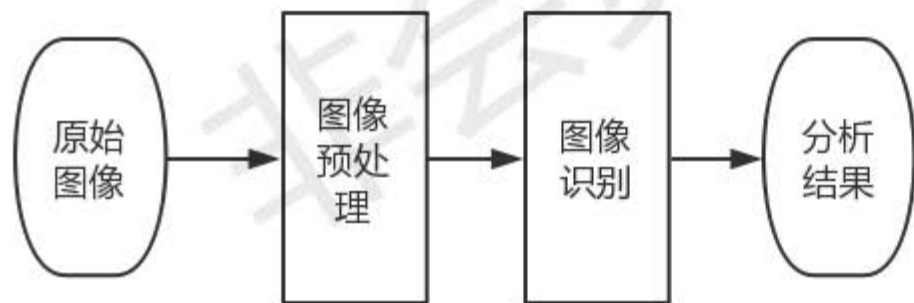
色调调整

# 图像预处理概述

---

# 什么是图像预处理

- 图像预处理的主要目的是消除图像中无关的信息，恢复有用的真实信息，增强有关信息的可检测性、最大限度地简化数据，从而改进特征提取、图像分割、匹配和识别的可靠性。
- 通常，图像预处理的流程包括：灰度变化，几何矫正，图像增强，图像滤波等操作。





# 常用预处理技术

---

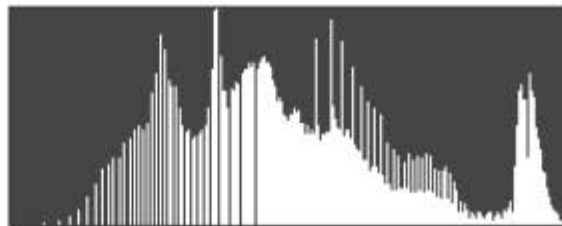
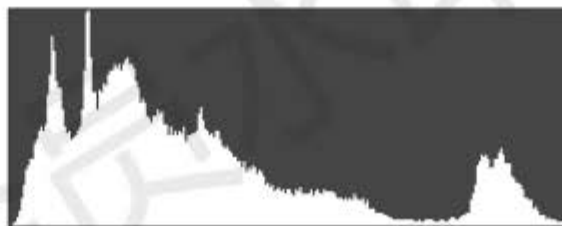
# 直方图

- 灰度直方图反映的是一幅图像中各灰度级像素出现的频率。以灰度级为横坐标，纵坐标为灰度级的频率，绘制频率同灰度级的关系图就是灰度直方图。它是图像的一个重要特征，反映了图像灰度分布的情况。
- 使用直方图进行图像变换是一种基于概率论的处理方法，通过改变图像的直方图，修改图像中各像素的灰度值，达到增强图像视觉效果的目的。
- 相对于灰度变化只针对单独的像素点操作，直方图变化综合考虑了全图的灰度值分布。



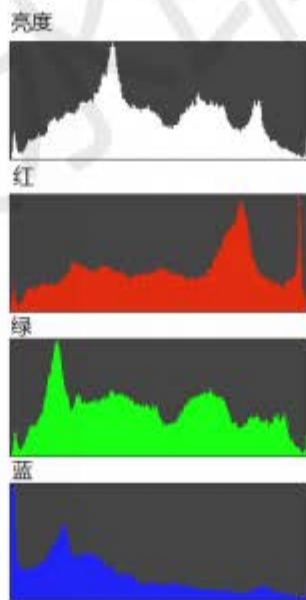
# 直方图（续1）

- 下面是两幅灰度图像的直方图，直方图的形状能反映图像的视觉效果。



# 直方图（续2）

- 对于彩色图像，可以对不同的通道分别统计直方图



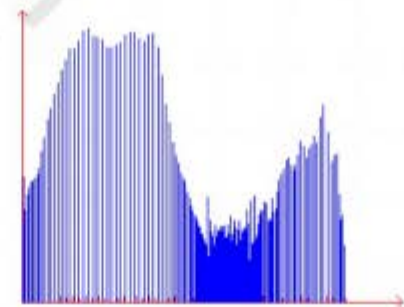
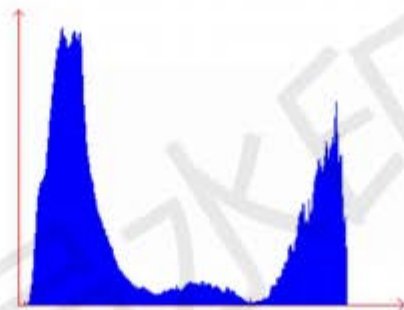
# 直方图均衡化

- 直方图均衡化将原始图像的直方图，即灰度概率分布图，进行调整，使之变化为均衡分布的样式，达到灰度级均衡的效果，可以有效增强图像的整体对比度。
- 直方图均衡化能够自动的计算变化函数，通过该方法自适应得产生有均衡直方图的输出图像。能够对图像过暗、过亮和细节不清晰的图像得到有效的增强。
- 在常用的图像处理库中，直方图操作都有API直接调用实现。





# 直方图均衡化（续）



# 模板运算

- 模板（滤波器）是一个尺寸为 $n \times n$ 的小图像 $W$ （ $n$ 一般取奇数，称为模板尺寸），每个位置上的值 $w$ 被称为权重。在进行计算时，将模板的中心和像素 $P$ 对齐，选取原始图像中和模板相同范围的邻域 $N$ 的像素值作为输入。
- 模板卷积的计算是将对齐后的对应位置像素相乘，再进行累加作为像素 $P$ 位置的输出值。记原始图像的像素灰度值为 $s$ ，计算后的值为 $d$ ，则 $P$ 点的输出值 
$$d = \frac{\sum w_i s_i}{\sum w_i}$$
- 模板排序的计算时将邻域 $N$ 的像素值进行排序，选择特定次序的灰度值，作为像素 $P$ 位置的输出值，如最大值、最小值、中位数等。



# 均值滤波

- 均值滤波指模板权重都为1的滤波器。它将像素的邻域平均值作为输出结果，均值滤波可以起到图像平滑的效果，可以去除噪声，但随着模板尺寸的增加图像会变得更加模糊。经常被作为模糊化使用。

1	1	1
1	1	1
1	1	1



原图



3x3均值滤波



5x5均值滤波



# 高斯滤波

- 为了减少模板尺寸增加对图像的模糊化，可以使用高斯滤波器，高斯滤波的模板根据高斯分布来确定模板系数，接近中心的权重比边缘的大。5的高斯滤波器如下所示：

1	4	7	4	1
4	16	26	16	4
7	26	41	26	7
4	16	26	16	4
1	4	7	4	1



原图



5x5均值滤波



5x5高斯滤波



# 中值滤波

- 中值滤波属于模板排序运算的滤波器。中值滤波器将邻域内像素排序后的中位数值输出代替原像素值。它在实现降噪操作的同时，保留了原始图像的锐度，不会修改原始图像的灰度值。
- 中值滤波的使用非常普遍，它对椒盐噪声的抑制效果很好，在抑制随机噪声的同时能有效保护边缘少受模糊。但中值滤波是一种非线性变化，它可能会破坏图像中线性关系，对于点、线等细节较多的图像和高精度的图像处理任务中并不太合适。



原图



3x3均值滤波



3x3中值滤波



# 边沿检测

- 通过梯度计算可以获取图像中细节的边缘。为在锐化边缘的同时减少噪声的影响，通过改进梯度法发展出了不同的边缘检测算子：
  - ✓ 一阶梯度：Prewitt梯度算子、Sobel梯度算子
  - ✓ 二阶梯度：Laplacian梯度算子。

-1	0	1
-1	0	1
-1	0	1

Prewitt算子

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-2	0	2
-1	0	1

Sobel算子

-1	-2	-1
0	0	0
1	2	1

0	1	0
1	-4	1
0	1	0

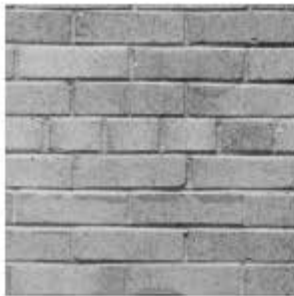
Laplacian算子

1	1	1
1	-8	1
1	1	1

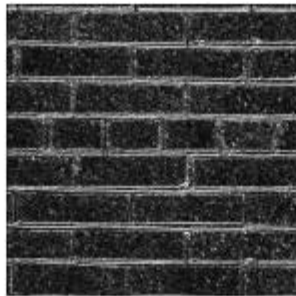


# 边沿检测（续）

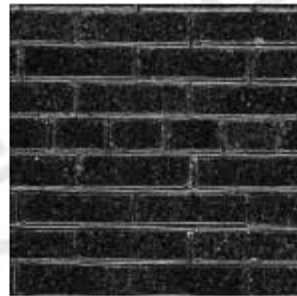
- 边沿检测效果



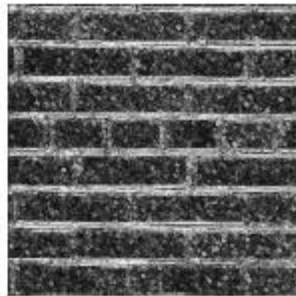
原图



Prewitt算子



Sobel算子



Laplacian算子

# 锐化

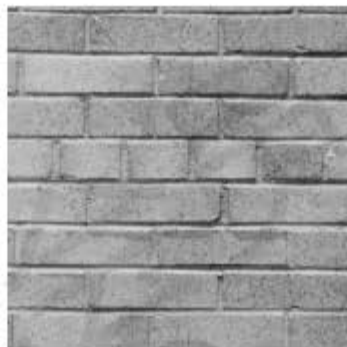
- 图像锐化与图像平滑是相反的操作，锐化是通过增强高频分量来减少图像中的模糊，增强图像细节边缘和轮廓，增强灰度反差，便于后期对目标的识别和处理。锐化处理在增强图像边缘的同时也增加了图像的噪声。



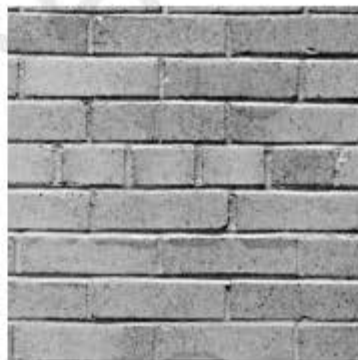
# 锐化（续）

- 将求取的边缘按照一定系数比例叠加到原始图像上，即可实现对图像的锐化操作。  
例如使用Laplacian梯度算子进行锐化操作的模板，其中A是大于等于1的系数：

0	-1	0
-1	A+4	-1
0	-1	0



原图



Laplacian锐化后的效果



# 色彩处理

---



# 亮度调整

- 对HSV空间的V分量进行处理可以实现对图像亮度的增强。
- 直接将彩色图像灰度化，也可以得到代表图像亮度的灰度图进行图像处理，计算量比HSV颜色空间变化低。但在HSV空间中进行处理可以得到增强后的彩色图像。



# 饱和度调整

- 对HSV空间的S分量进行处理可以实现对图像饱和度的增强。
- 饱和度的调整通常是在S原始值上乘以一个修正系数。
- 修正系数大于1，会增加饱和度，使图像的色彩更鲜明；
- 修正系数小于1，会减小饱和度，使图像看起来比较平淡。



# 色调调整

- 对HSV空间的H分量进行处理可以实现对图像色调的增强。
- 色相H的值对应的是一个角度，并且在色相环上循环。所以色相的修正可能会造成颜色的失真。
- 色相的调整通常在H原始值上加上一个小的偏移量，使其在色相环上有小角度的调整。调整后，图像的色调会变为冷色或者暖色。



## 今日总结

- 图像处理基本原理
- 图像处理技术