

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



第2课 图像数据处理

Image Data Processing

主讲人：张宗健

悉尼科技大学博士

主要研究方向： 计算机视觉、视觉场景理解、图像&语言、深度学习
图像检索CbIR、Human ReID等

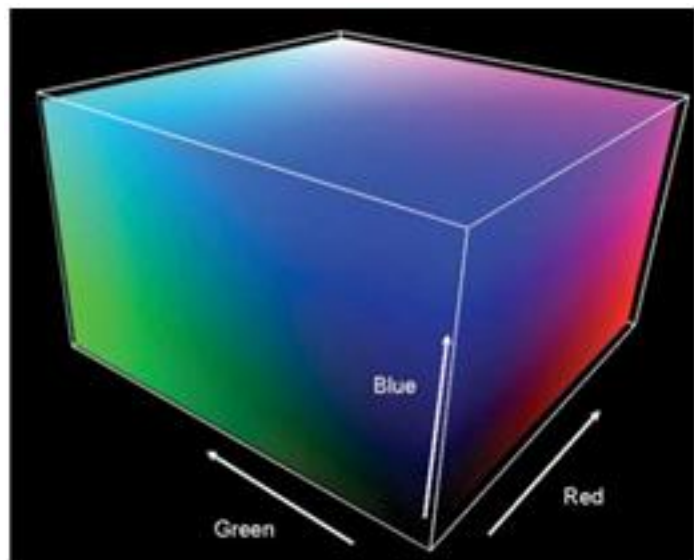
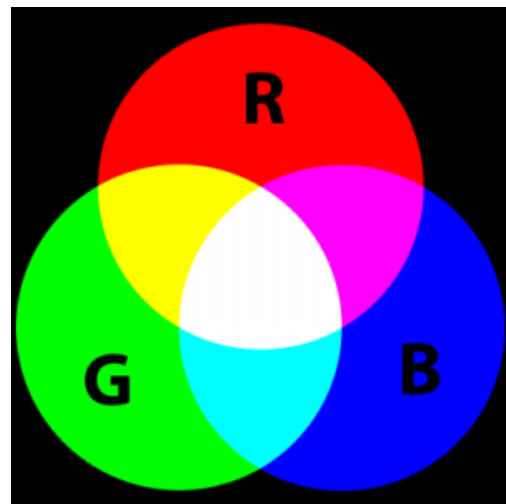
本章结构

- 图片存储原理
- 空域分析及变换
- 频域分析及变换
- 金字塔
- 模板匹配
- 代码实践

图片存储原理

RGB颜色空间

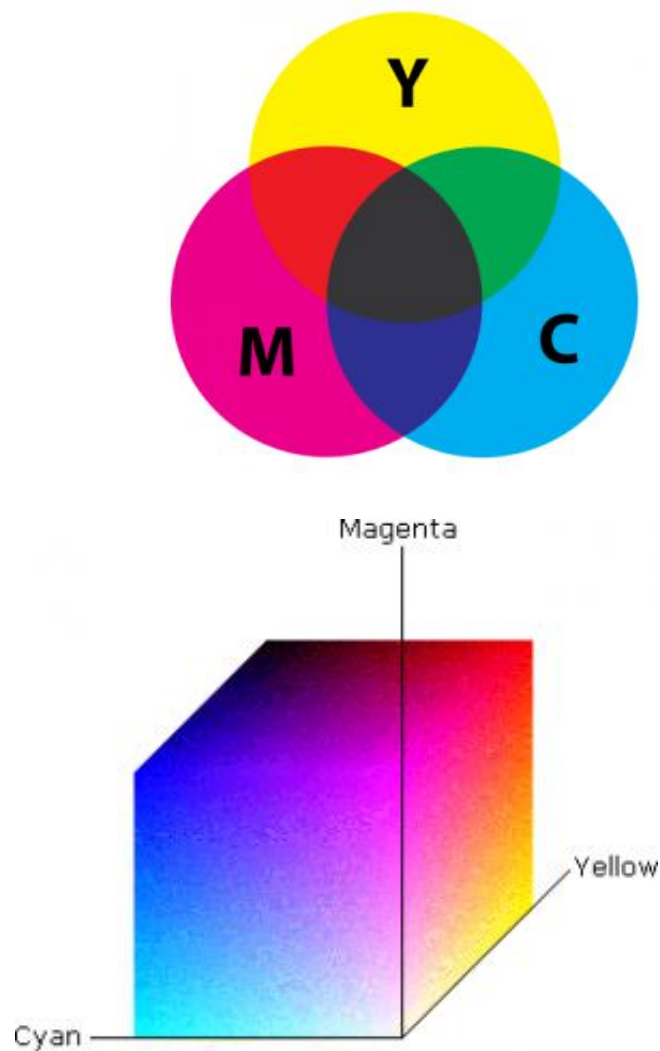
- 加法混色，彩色显示器
- 3通道
 - Red通道
 - Green通道
 - Blue通道
- 一个像素颜色值
 - (b, g, r)
- 取值范围
 - $[0, 255]$
 - $[0.0, 1.0]$



图片存储原理

CMY(K)颜色空间

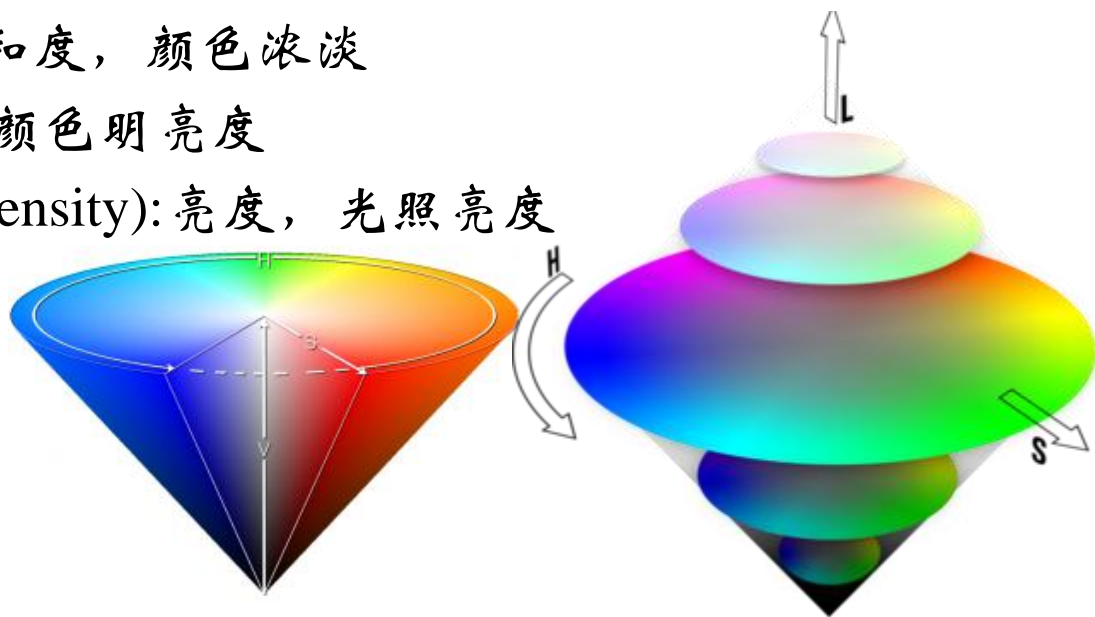
- 减法混色，印刷
- 4通道
 - Cyan通道
 - Magenta通道
 - Yellow通道
 - Key通道
- 一个像素颜色值
 - (c, m, y, k)
- 取值范围
 - [0, 255]
 - [0.0, 1.0]



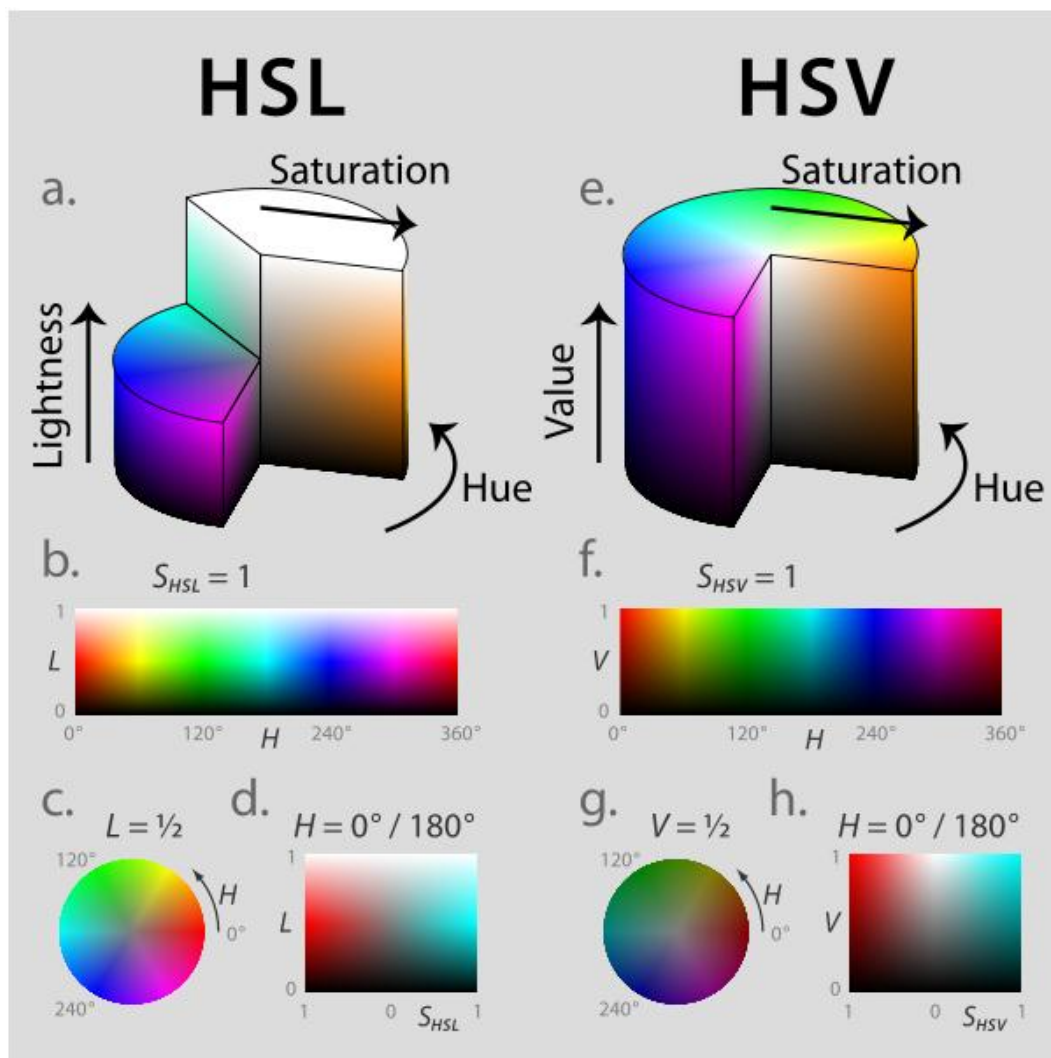
图片存储原理

HSV/HSL(I)颜色空间

- 人类视觉概念，画家配色
- 3通道
 - H/Hue通道：色调，颜色种类
 - S/Saturation：饱和度，颜色浓淡
 - V/Value:明度，颜色明亮度
 - L/Lightness(I/Intensity):亮度，光照亮度
- 一个像素颜色值
 - (h, s, v/l/i)
- 取值范围
 - [0, 255]
 - [0.0, 1.0]



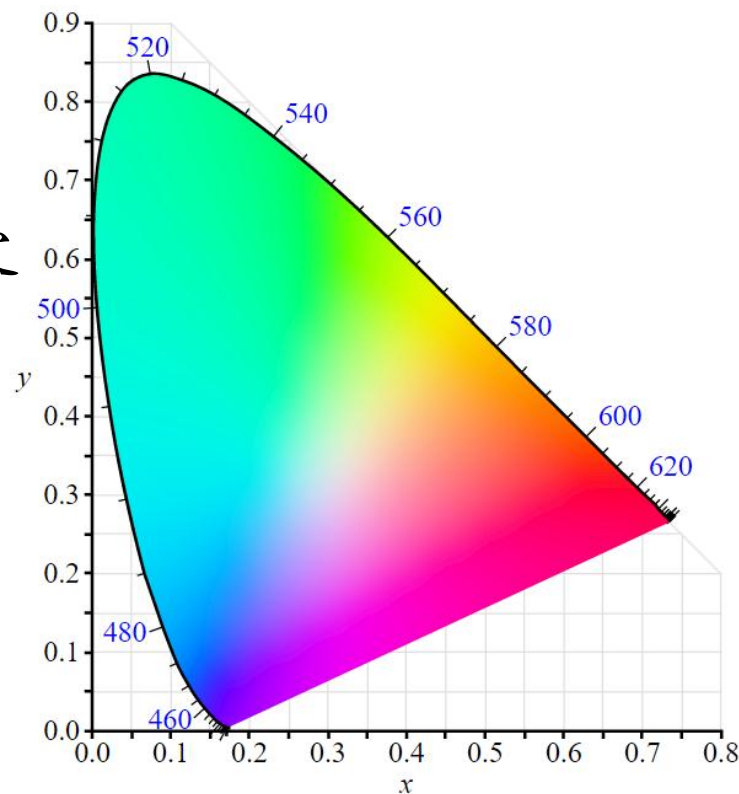
图片存储原理



图片存储原理

CIE-XYZ颜色空间

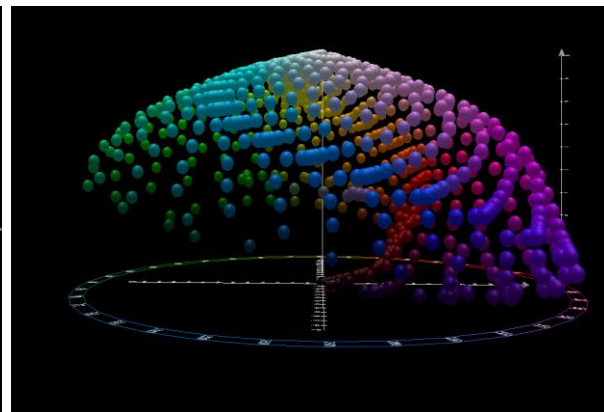
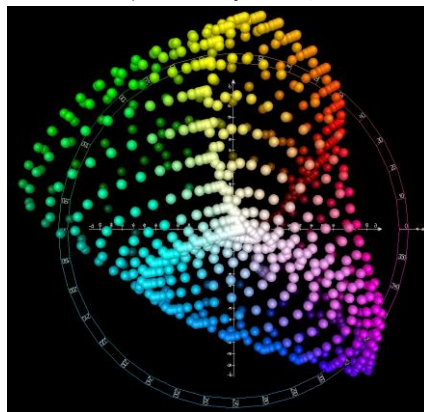
- 国际照明协会，1931
- 基于人类颜色视觉的直接测定
- 其他颜色空间基础
- 人类视觉系统-视锥细胞
 - 短波 (S, 420-440nm)
 - 中波 (M, 530-540nm)
 - 长波 (L, 560-580nm)
- 3色刺激值通道
 - X, Y, Z约略对应于红色、绿色、蓝色
 - 一种波的刺激等于几种波的混合刺激



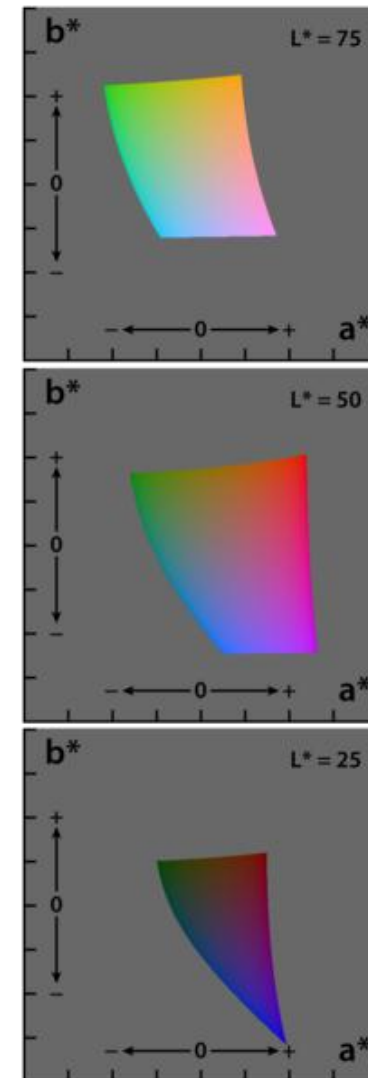
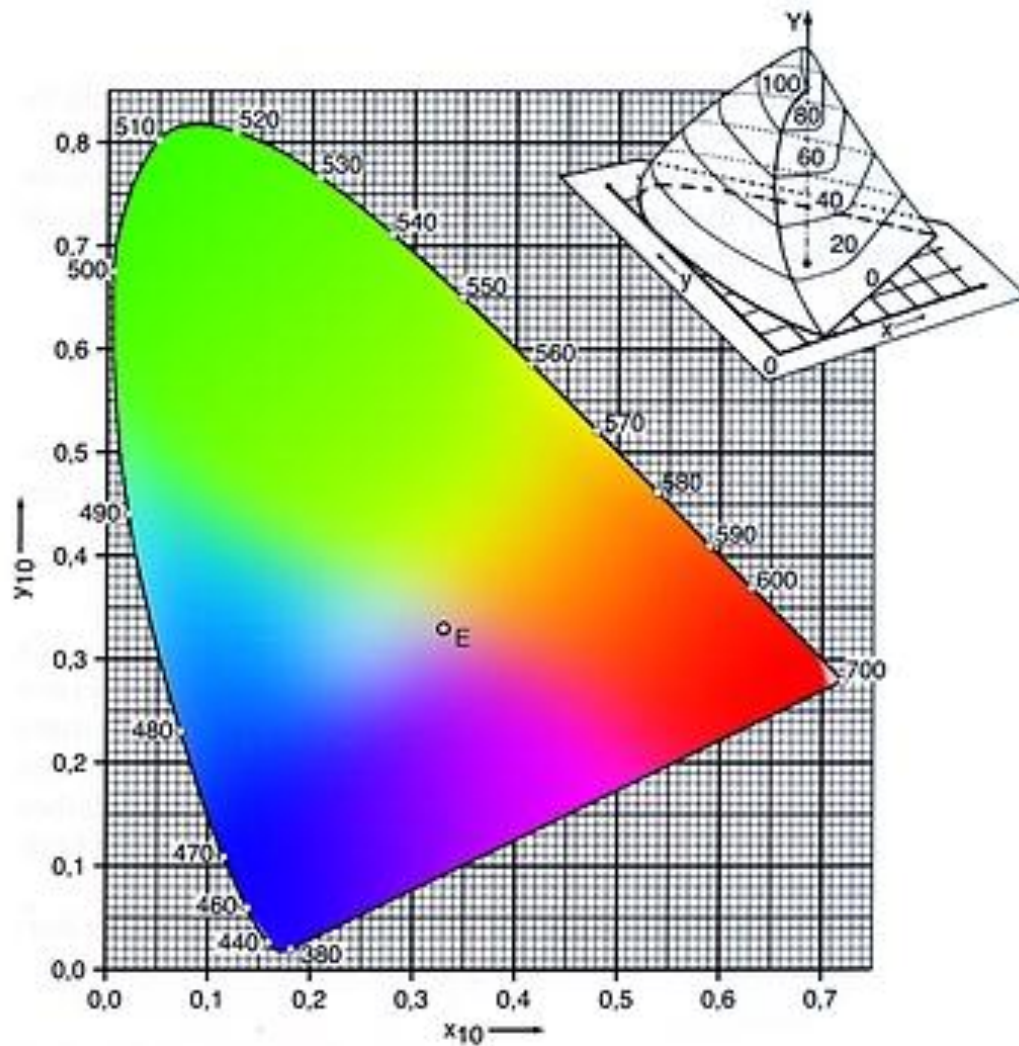
图片存储原理

CIE-Lab对色空间

- 国际照明协会，1976
- 接近人类视觉、致力于感知均匀性
- 颜色空间本身不均匀
 - 非线性关系→人类眼睛的非线性响应
- 3通道
 - L: 亮度
 - a: 红 / 绿
 - b: 黄 / 蓝
- 色差距离CIE2000
 - 基于人类色差实验数据，均匀性好



图片存储原理

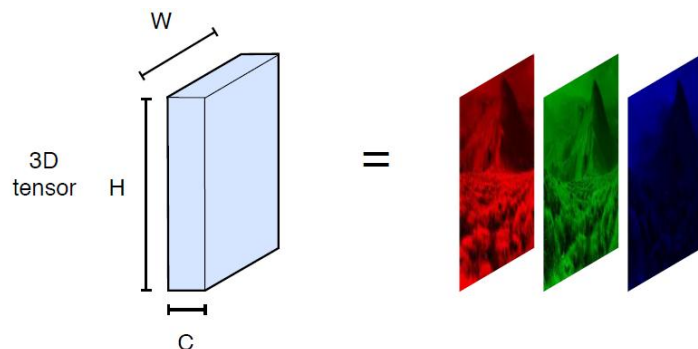


图片存储原理

主流颜色空间

- RGB 三通道彩色图

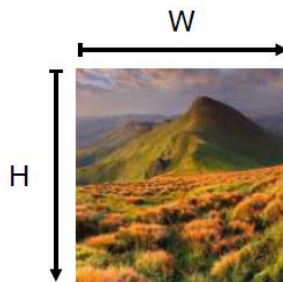
- 图片 \rightarrow 3维矩阵 $([0,255])$



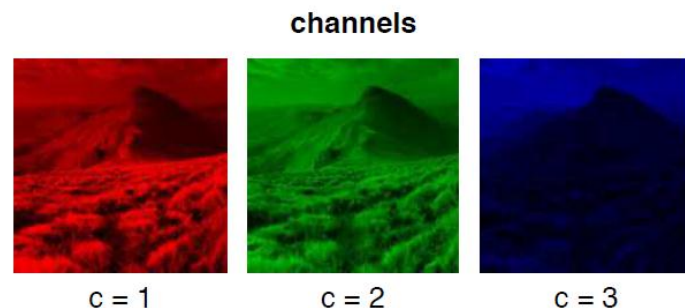
- 单通道灰度图

- 亮度信息 $([0,255])$

- $\text{Gray} = R * 0.3 + G * 0.59 + B * 0.11$



=



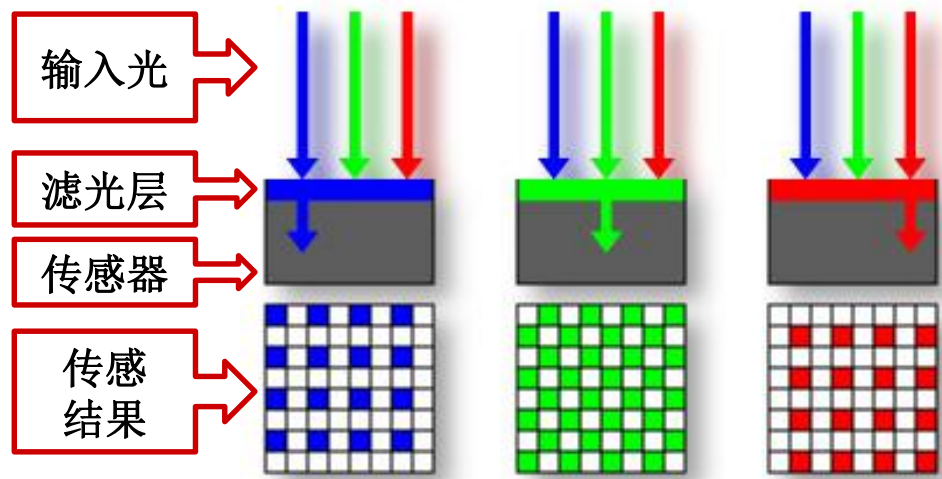
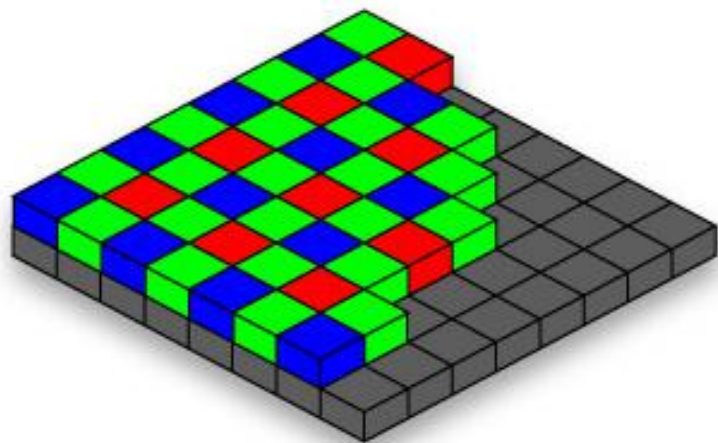
灰度化



图片存储原理

颜色传感器：Bayer网格

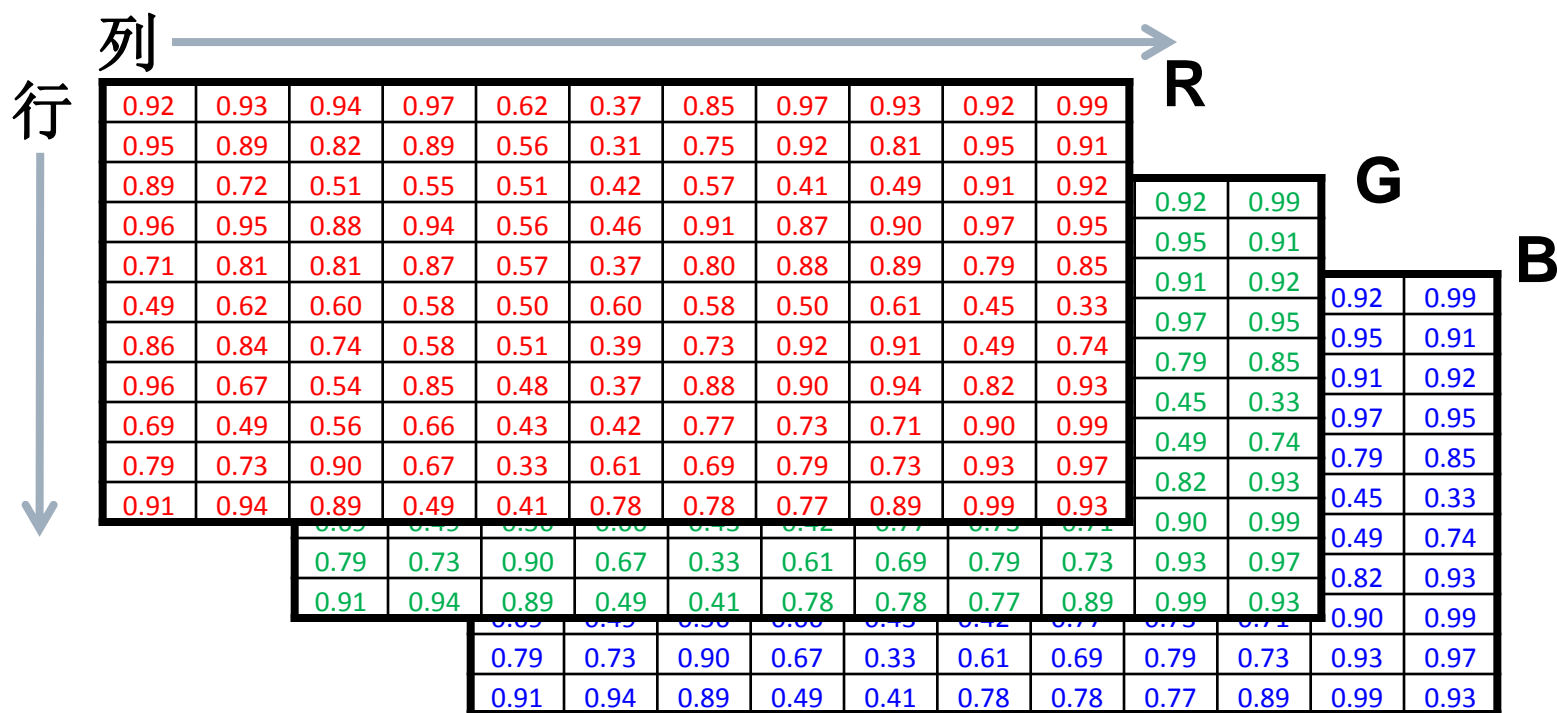
- 在G格估计RGB
- 基于周边9个色值



图片存储原理

RGB图片数据格式

- 3维矩阵



空域分析及变换

滤波/卷积

- 在每个图片位置 (x, y) 上进行基于邻域的函数计算

- 滤波函数 \rightarrow 权重相加

- 卷积核、卷积模板

滤波结果

- 滤波器、滤波模板

- 扫描窗

滤波函数

像素邻域值

$$h[x, y] = \sum_{k, l} f[k, l] I[x + k, y + l]$$

- 不同功能需要定义不同的函数

- 图像增强

- 平滑/去噪

- 梯度/锐化

- 信息提取、检测

- 边缘、显著点、纹理

- 模式

空域分析及变换

滤波/卷积

滤波结果

滤波函数

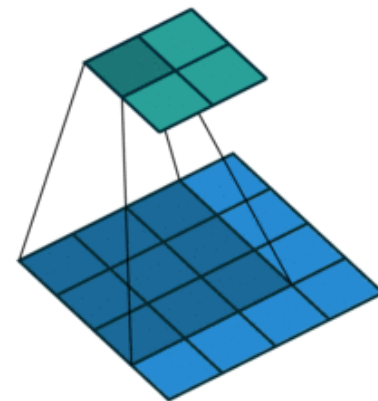
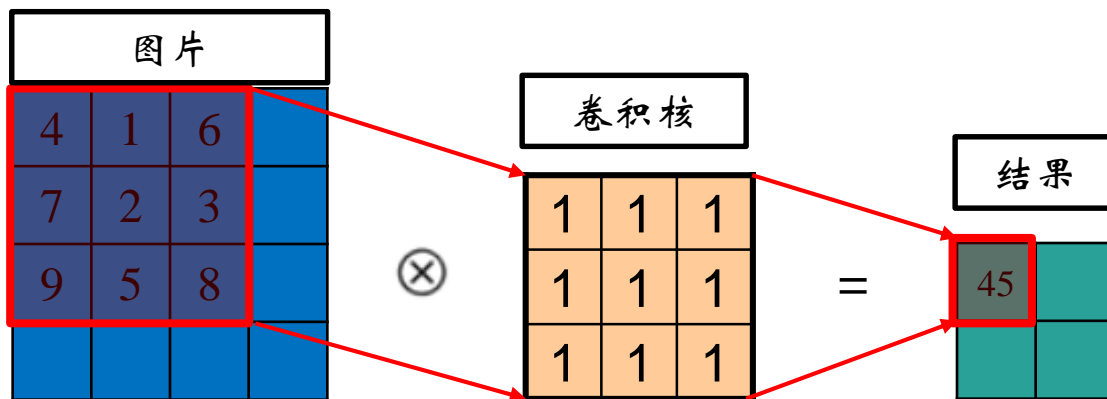
像素邻域值

$$h[x, y] = \sum_{k, l} f[k, l] I[x + k, y + l]$$

参数解释

- x, y 是像素在图片中的位置/坐标
- k, l 是卷积核中的位置/坐标
 - 中心点的坐标是 $(0, 0)$
- $f[k, l]$ 是卷积核中在 (k, l) 上的权重参数
- $I[x+k, y+l]$ 是与 $f[k, l]$ 相对应的图片像素值
- $h[x, y]$ 是图片中 (x, y) 像素的滤波/卷积结果

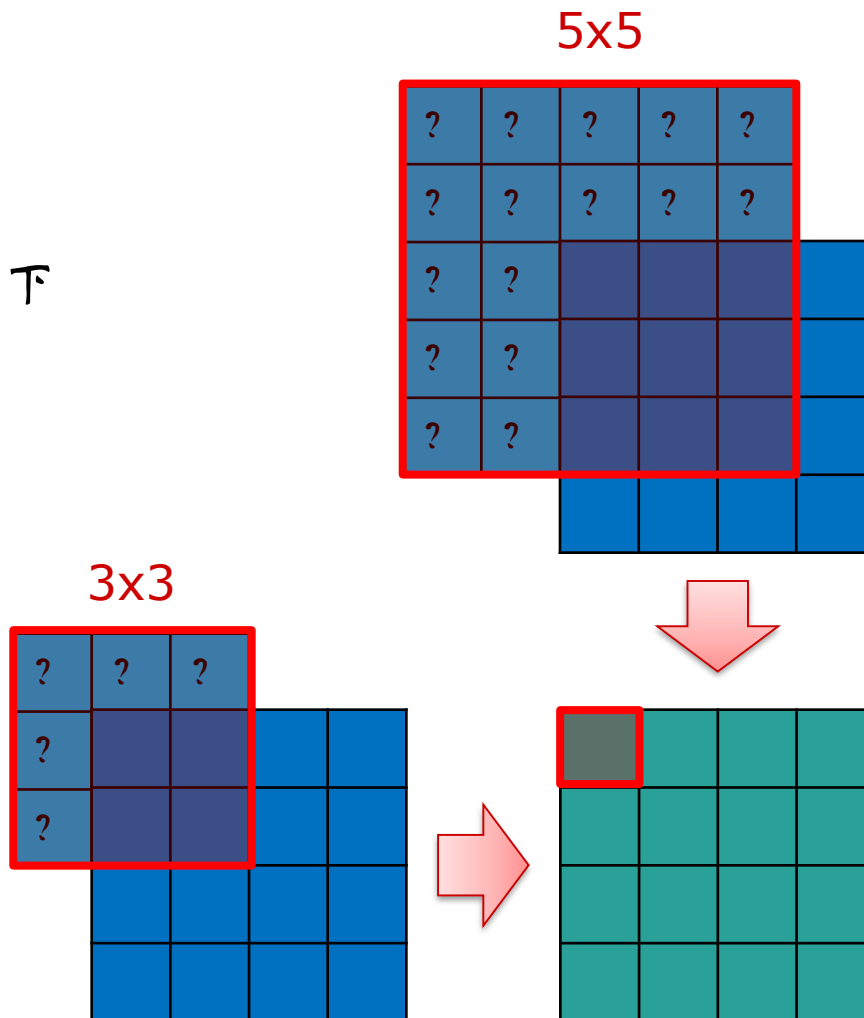
$(-1, -1)$	$(-1, 0)$	$(-1, 1)$
$(0, -1)$	$(0, 0)$	$(0, 1)$
$(1, -1)$	$(1, 0)$	$(1, 1)$



空域分析及变换

滤波/卷积

- 边界补充 (Padding)
 - 获得同尺寸输出的情况下
 - 卷积核越大, 补充越多
- 补充类型
 - 补零 (zero-padding)
 - 边界复制 (replication)
 - 镜像 (reflection)
 - 块复制 (wraparound)



空域分析及变换

滤波/卷积

- 补零 (zero-padding)
- 边界复制 (replication)
- 镜像 (reflection)
- 块复制 (wraparound)

0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	2	3	0	0	0
0	0	0	4	5	6	0	0	0
0	0	0	7	8	9	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

空域分析及变换

滤波/卷积

- 7x7卷积: $3 \times 3 \rightarrow 9 \times 9$

补零 (zero-padding)

0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	1	2	3	0	0	0
0	0	0	4	5	6	0	0	0
0	0	0	7	8	9	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0

边界复制 (replication)

1	1	1	1	2	3	3	3	3
1	1	1	1	2	3	3	3	3
1	1	1	1	2	3	3	3	3
1	1	1	1	2	3	3	3	3
4	4	4	4	5	6	6	6	6
7	7	7	7	8	9	9	9	9
7	7	7	7	8	9	9	9	9
7	7	7	7	8	9	9	9	9
7	7	7	7	8	9	9	9	9

空域分析及变换

滤波/卷积

- 7x7卷积: $3 \times 3 \rightarrow 9 \times 9$

镜像 (reflection)

9	8	7	7	8	9	9	8	7
6	5	4	4	5	6	6	5	4
3	2	1	1	2	3	3	2	1
3	2	1	1	2	3	3	2	1
6	5	4	4	5	6	6	5	4
9	8	7	7	8	9	9	8	7
9	8	7	7	8	9	9	8	7
6	5	4	4	5	6	6	5	4
3	2	1	1	2	3	3	2	1

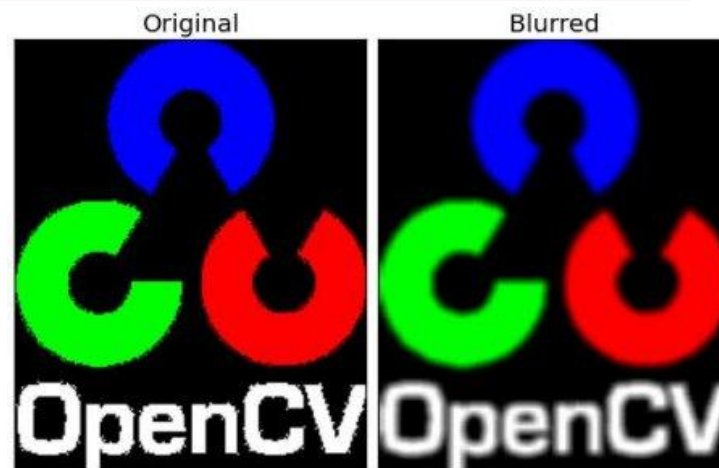
块复制 (wraparound)

1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9
1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9
1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9

空域分析及变换

平滑均值滤波/卷积

- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 参数和为: 1



$$\otimes \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} =$$



空域分析及变换

平滑均值滤波/卷积

- 3x3
- 扫描步长: 1
- 边框补零

输入图片
6x6

0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

4					

空域分析及变换

平滑均值滤波/卷积

- 3x3
- 扫描步长: 1
- 边框补零

输入图片
6x6

0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

4	6				

空域分析及变换

平滑均值滤波/卷积

- 3x3
- 扫描步长: 1
- 边框补零

输入图片
6x6

0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

4	6	8	10	12	8

空域分析及变换

平滑均值滤波/卷积

- 3x3
- 扫描步长: 1
- 边框补零

输入图片
6x6

0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

4	6	8	10	12	8
6					

空域分析及变换

平滑均值滤波/卷积

- 3x3
- 扫描步长: 1
- 边框补零

输入图片
6x6

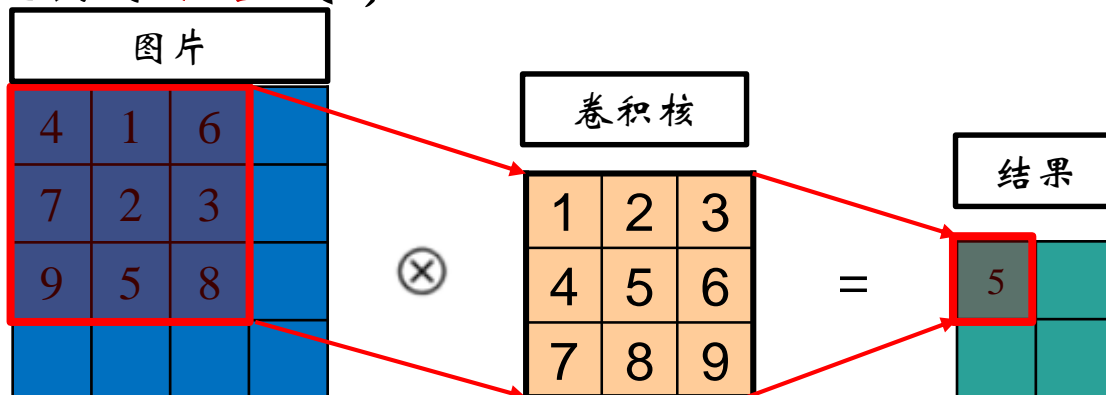
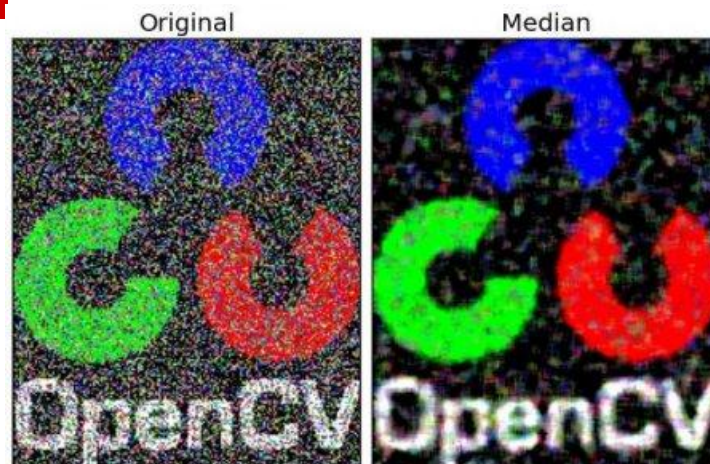
0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

4	6	8	10	12	8
6	9	12	15	18	12
8	12	13	14	15	10
10	15	14	13	12	8
12	18	15	12	9	6
8	12	10	8	6	4

空域分析及变换

平滑 中值滤波/卷积

- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 操作原理
 - 卷积域内的像素值从小到大排序
 - 取中间值作为卷积输出
- 有效去除椒盐噪声



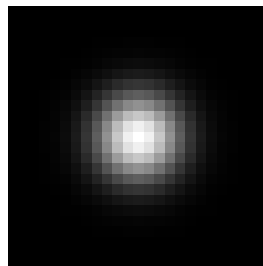
空域分析及变换

平滑 中值滤波/卷积



空域分析及变换

平滑高斯滤波/卷积

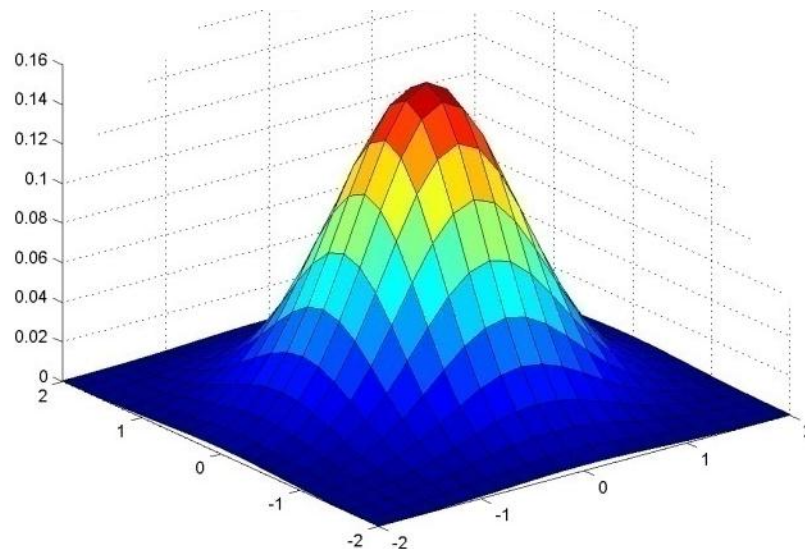


- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 模拟人眼, 关注中心区域
- 有效去除高斯噪声
- 参数
 - x, y 是卷积参数坐标
 - 标准差 σ

$$G_{\sigma} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

0.003	0.013	0.022	0.013	0.003
0.013	0.059	0.097	0.059	0.013
0.022	0.097	0.159	0.097	0.022
0.013	0.059	0.097	0.059	0.013
0.003	0.013	0.022	0.013	0.003

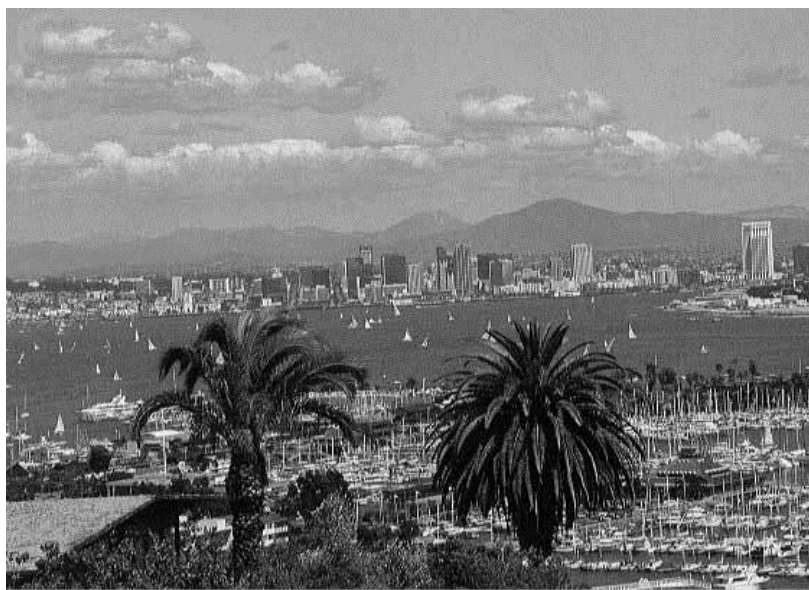
5 x 5, $\sigma = 1$



空域分析及变换

平滑高斯滤波/卷积

- 人眼特性：离关注中心越远，感受精度越模糊



摄像头



人类视觉

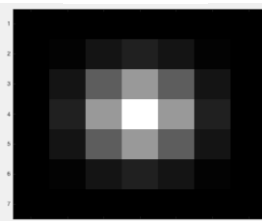
空域分析及变换

平滑高斯滤波/卷积

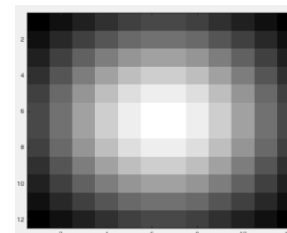
- σ 越小
- 关注区域越集中

原图

$\sigma=1$



$\sigma=4$



空域分析及变换

平滑高斯滤波/卷积

- 分解特性（级联高斯）
 - 2D卷积拆分成两个相同的1D卷积
 - 列卷积
 - 行卷积
- 降计算
 - 2D卷积：KxK次计算
 - 2x1D卷积：2K次计算

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

$$\begin{aligned} G_{\sigma}(x, y) &= \frac{1}{2\pi\sigma^2} \exp^{-\frac{x^2 + y^2}{2\sigma^2}} \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{x^2}{2\sigma^2}} \right) \left(\frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{y^2}{2\sigma^2}} \right) \end{aligned}$$

空域分析及变换

梯度Prewitt滤波/卷积

- 水平梯度/垂直边缘

水平梯度

均值平滑

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

水平梯度

均值平滑

- 垂直梯度/水平边缘

均值平滑

垂直梯度

$$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

均值平滑

垂直梯度

空域分析及变换

梯度Sobel滤波/卷积

- 水平梯度/垂直边缘

水平梯度

高斯平滑

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

水平梯度

高斯平滑

- 垂直梯度/水平边缘

高斯平滑

垂直梯度

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

高斯平滑

垂直梯度

空域分析及变换

梯度Laplacian滤波/卷积

- 二阶微分算子
- 一阶导数极值

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- 作用
 - 团块检测：周边高于（低于）中心点
 - 边缘检测：像素值快速变化的区域

1	-2	1
---	----	---

1D

0	1	0
1	-4	1
0	1	0

2D 4邻域

1	1	1
1	-8	1
1	1	1

2D 8邻域

0	0	1	0	0
0	1	2	1	0
1	2	-16	2	1
0	1	2	1	0
0	0	1	0	0

5x5

空域分析及变换

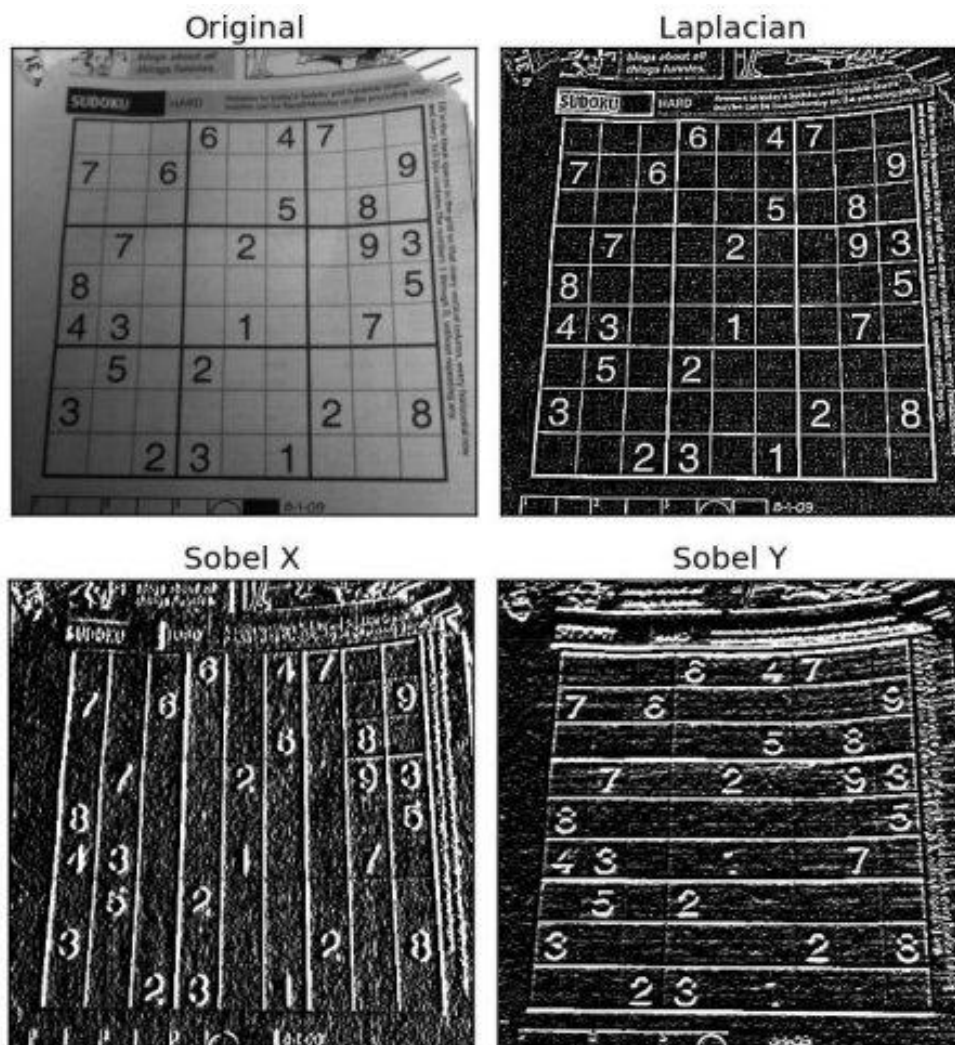
Laplacian滤波锐化



空域分析及变换

梯度滤波/卷积


- Laplacian 滤波
- Sobel-X 滤波
- Sobel-Y 滤波




空域分析及变换

其他滤波/卷积


- (左)移位

 \otimes

0	0	0
0	0	1
0	0	0

 $=$ 

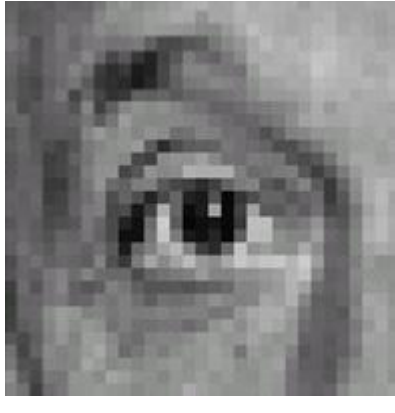
- 锐化

 \otimes

0	0	0
0	2	0
0	0	0

 $- \frac{1}{9}$

1	1	1
1	1	1
1	1	1

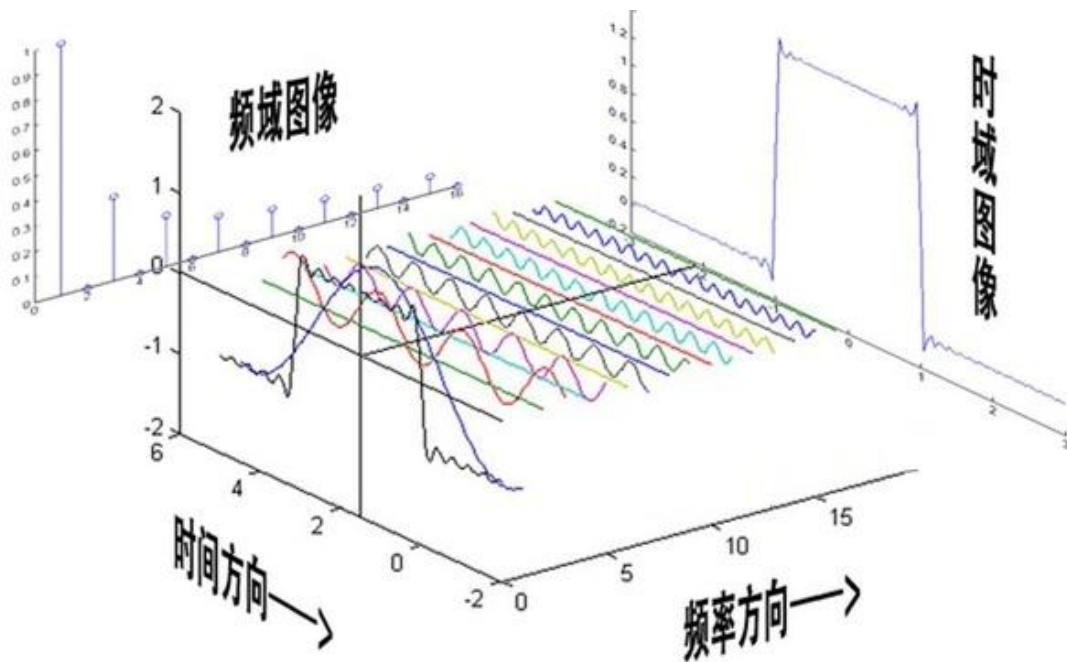
 $=$ 

频域分析及变换

傅里叶变换

- 一个信号可以由足够多个不同频率和幅值的正余弦波组成

$$A \sin(\omega x) + B \cos(\omega x)$$

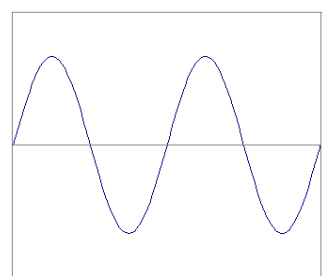
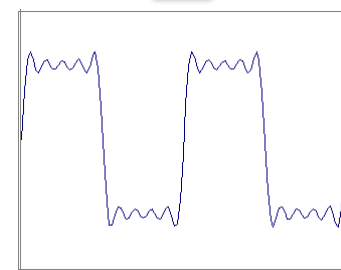
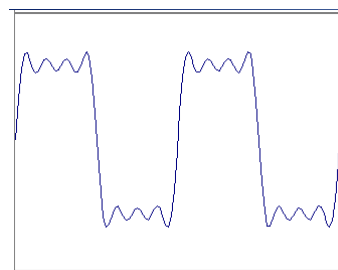
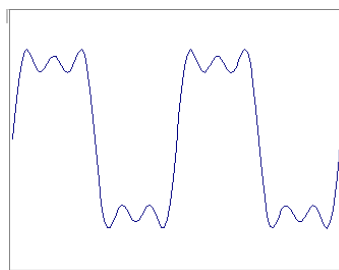
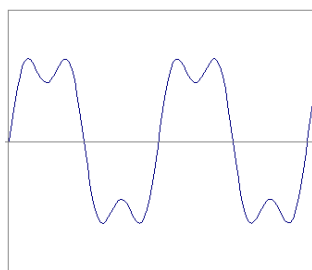
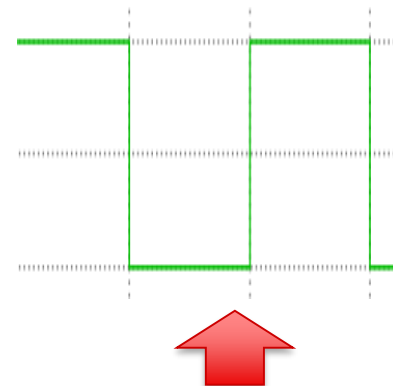


频域分析及变换

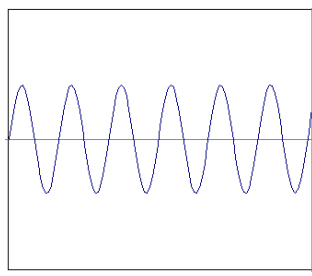
傅里叶变换

- 矩形波

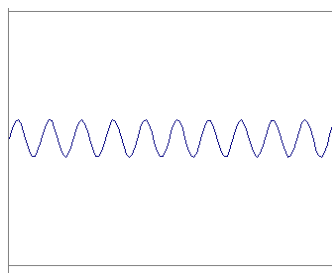
$$A \sum_{k=1}^{\infty} \frac{1}{k} \sin(2\pi kt) =$$



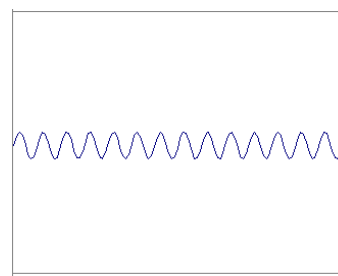
+



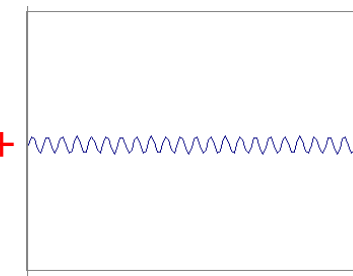
+



+



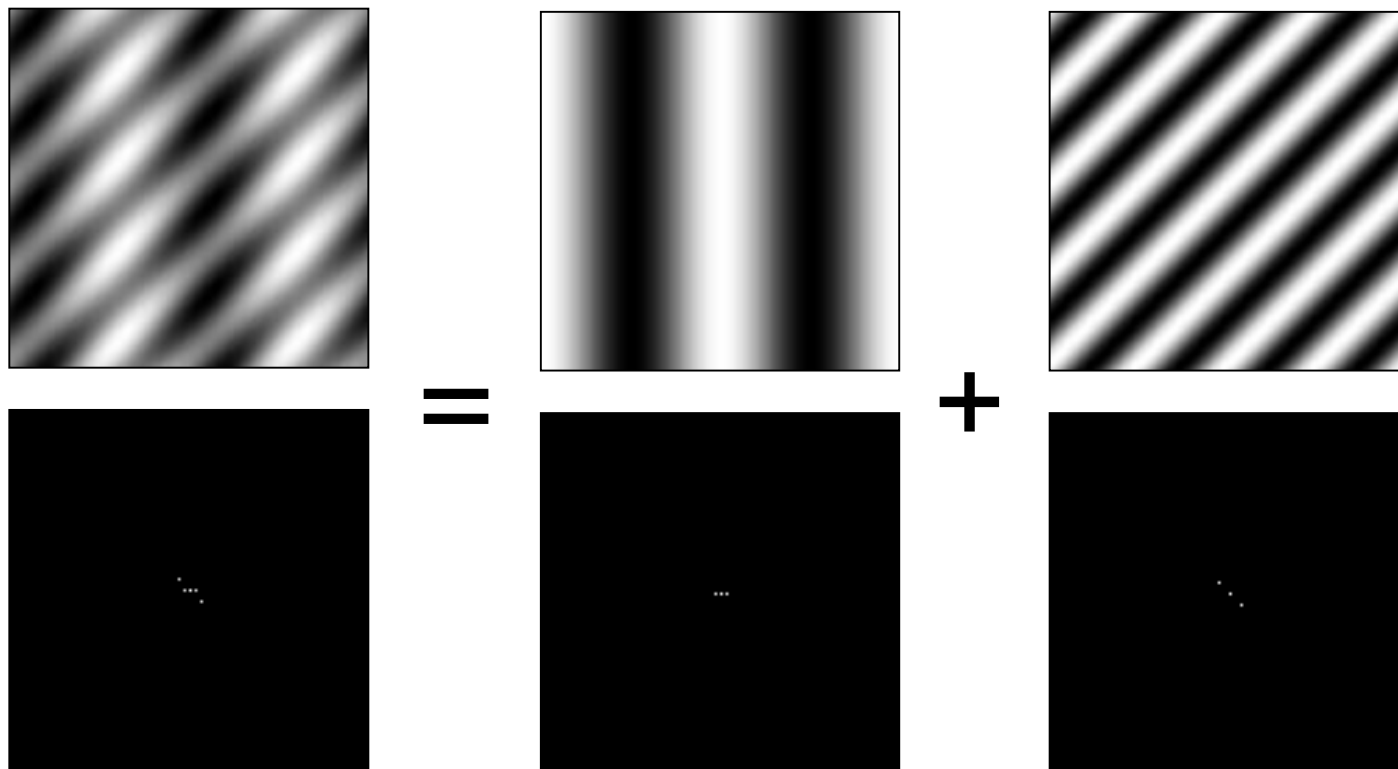
+



频域分析及变换

傅里叶变换

- 信号分解



频域分析及变换

傅里叶变换

- 连续变换 $H(\omega) = \int_{-\infty}^{\infty} h(x)e^{-j\omega x} dx$

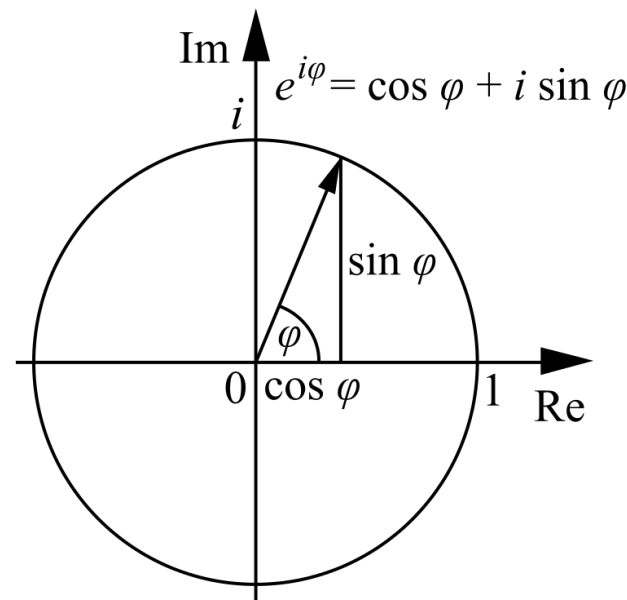
- 离散变换 $H(k) = \frac{1}{N} \sum_{x=0}^{N-1} h(x)e^{-j\frac{2\pi kx}{N}}$

$$k = -N/2 \dots N/2$$

- 存储每个频率段上的：

- 幅值 $A = \pm \sqrt{\text{Re}(\omega)^2 + \text{Im}(\omega)^2}$

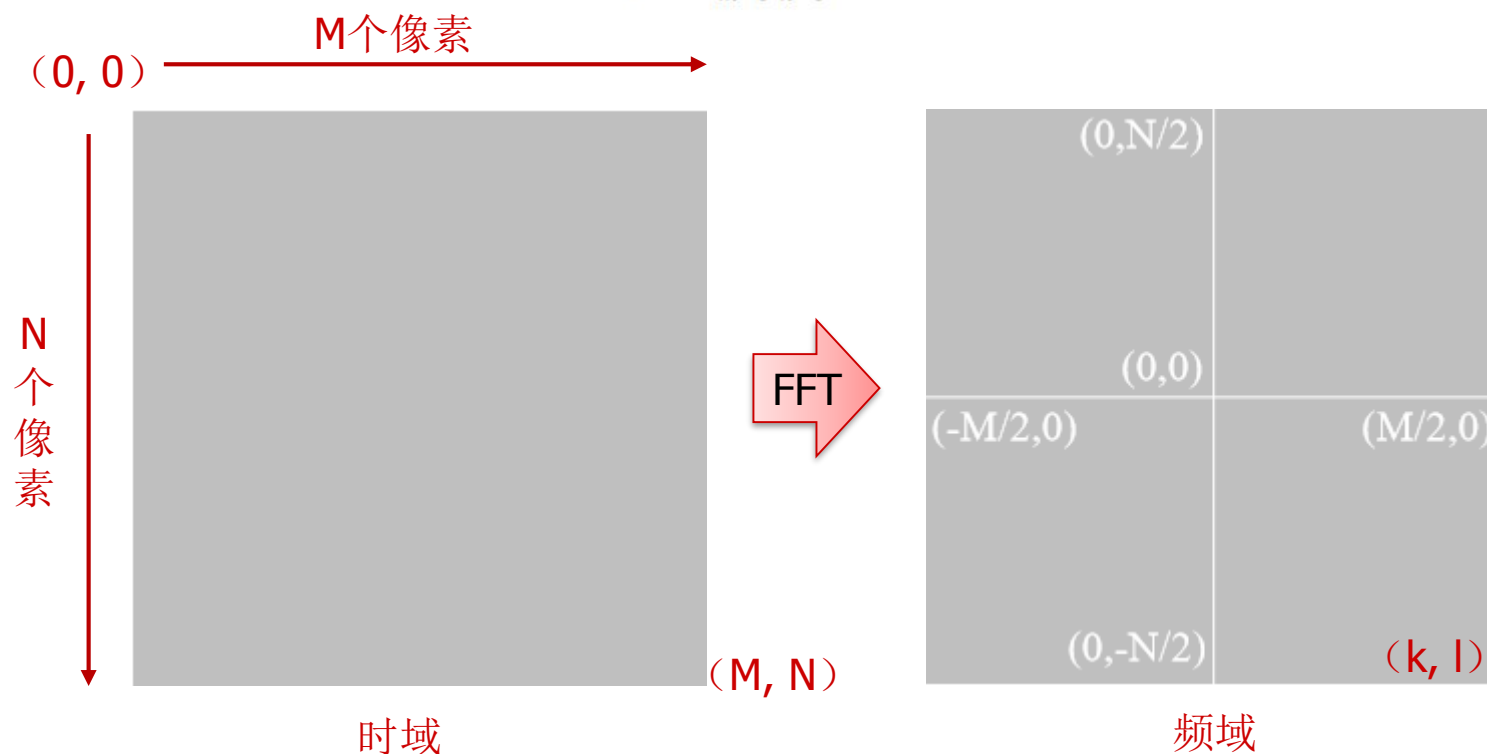
- 相位 $\phi = \tan^{-1} \frac{\text{Im}(\omega)}{\text{Re}(\omega)}$



频域分析及变换

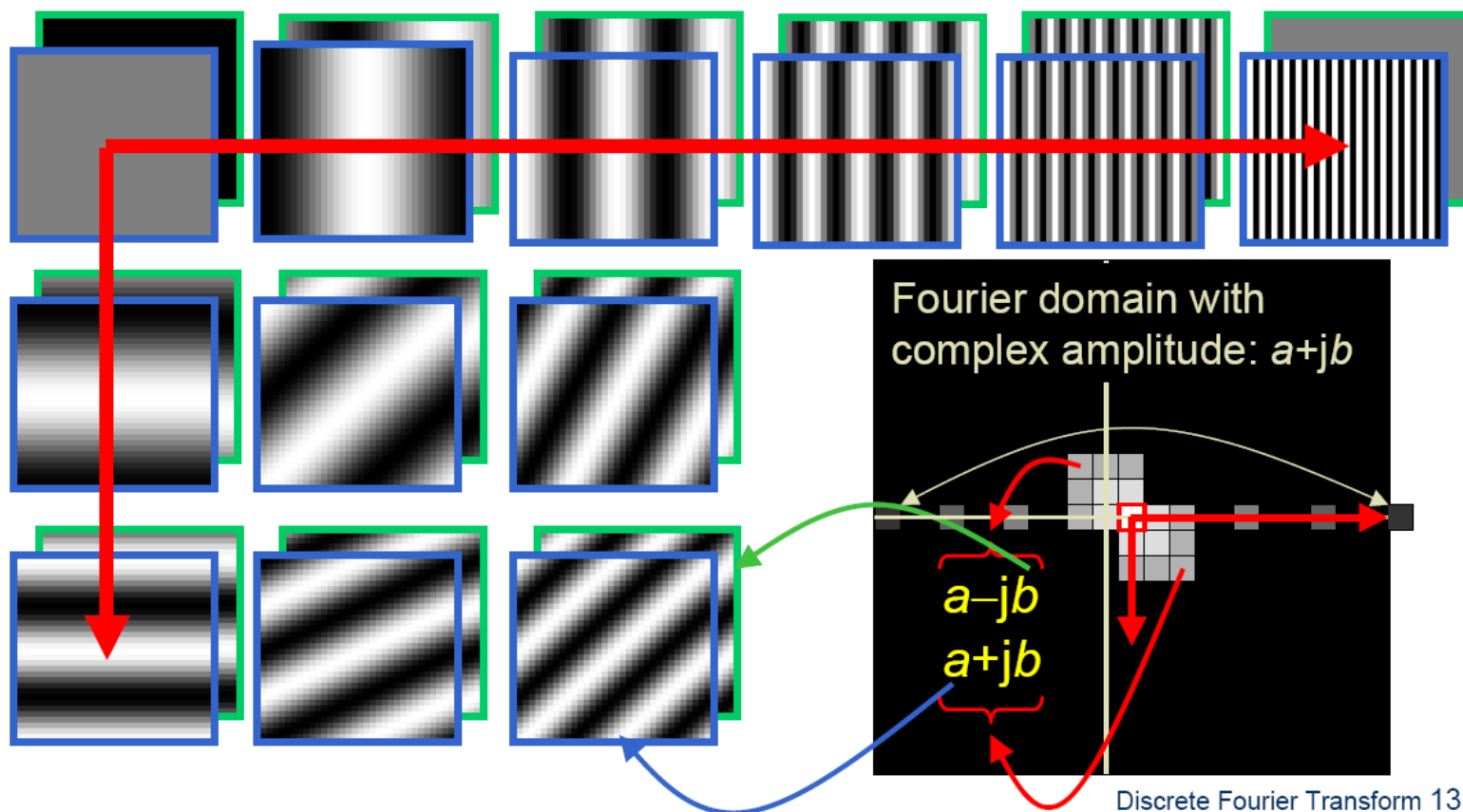
2D离散傅里叶变换

$$F[k, l] = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m, n] e^{-j2\pi \left(\frac{k}{M}m + \frac{l}{N}n \right)}$$



频域分析及变换

2D傅里叶基图片



频域分析及变换

滤波效果

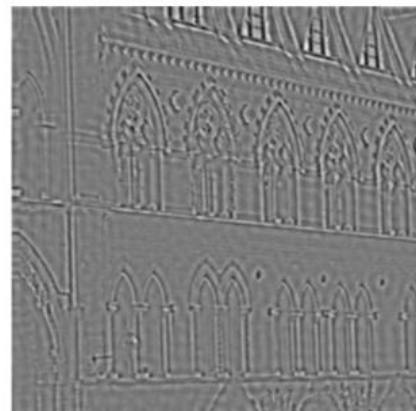
原图



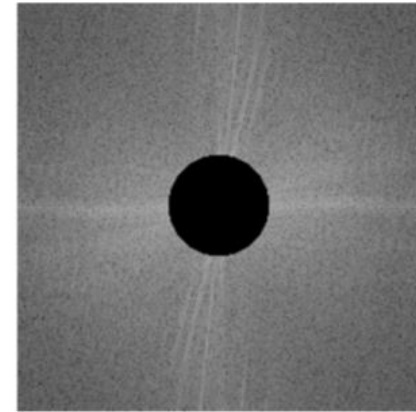
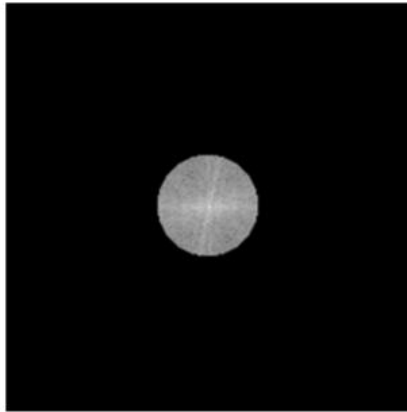
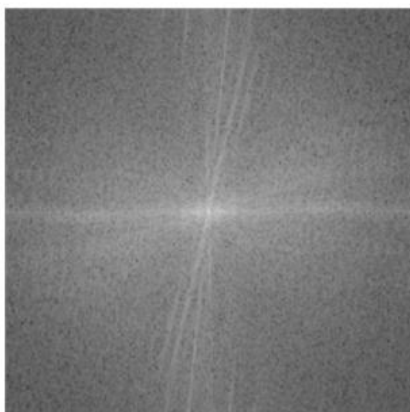
低通



高通

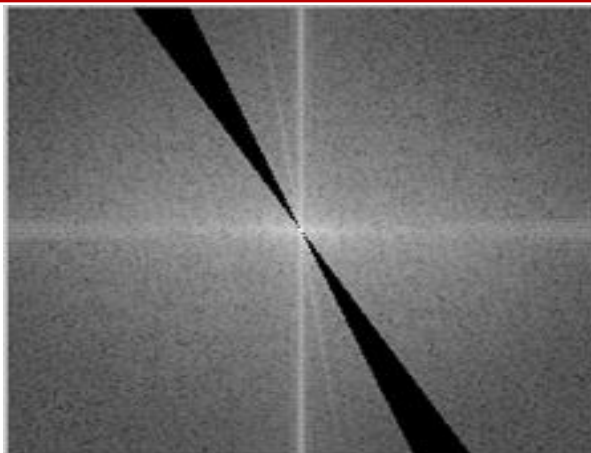
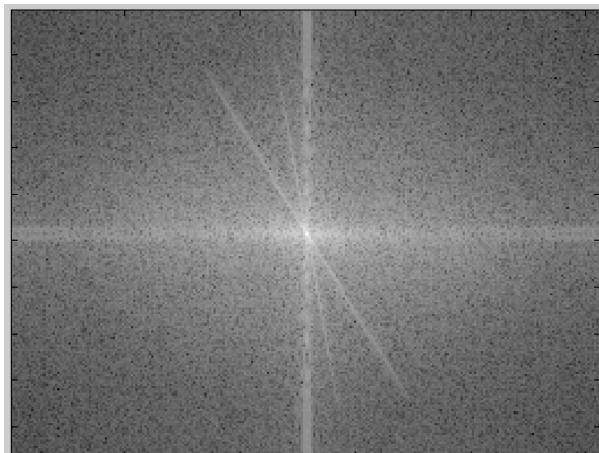
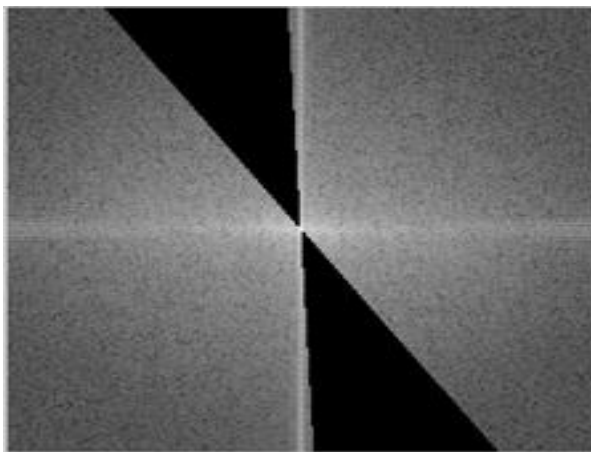


FFT



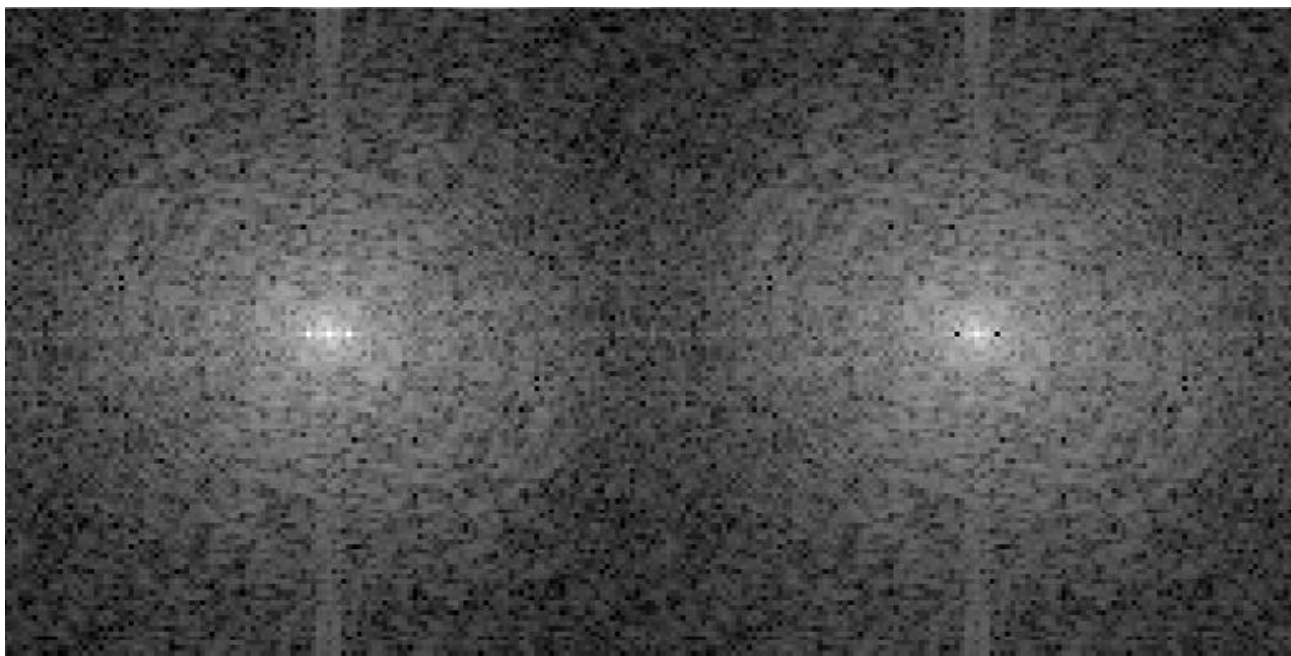
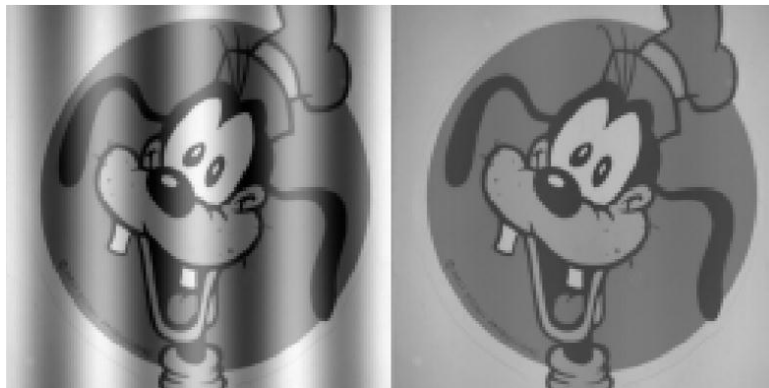
频域分析及变换

相位滤波



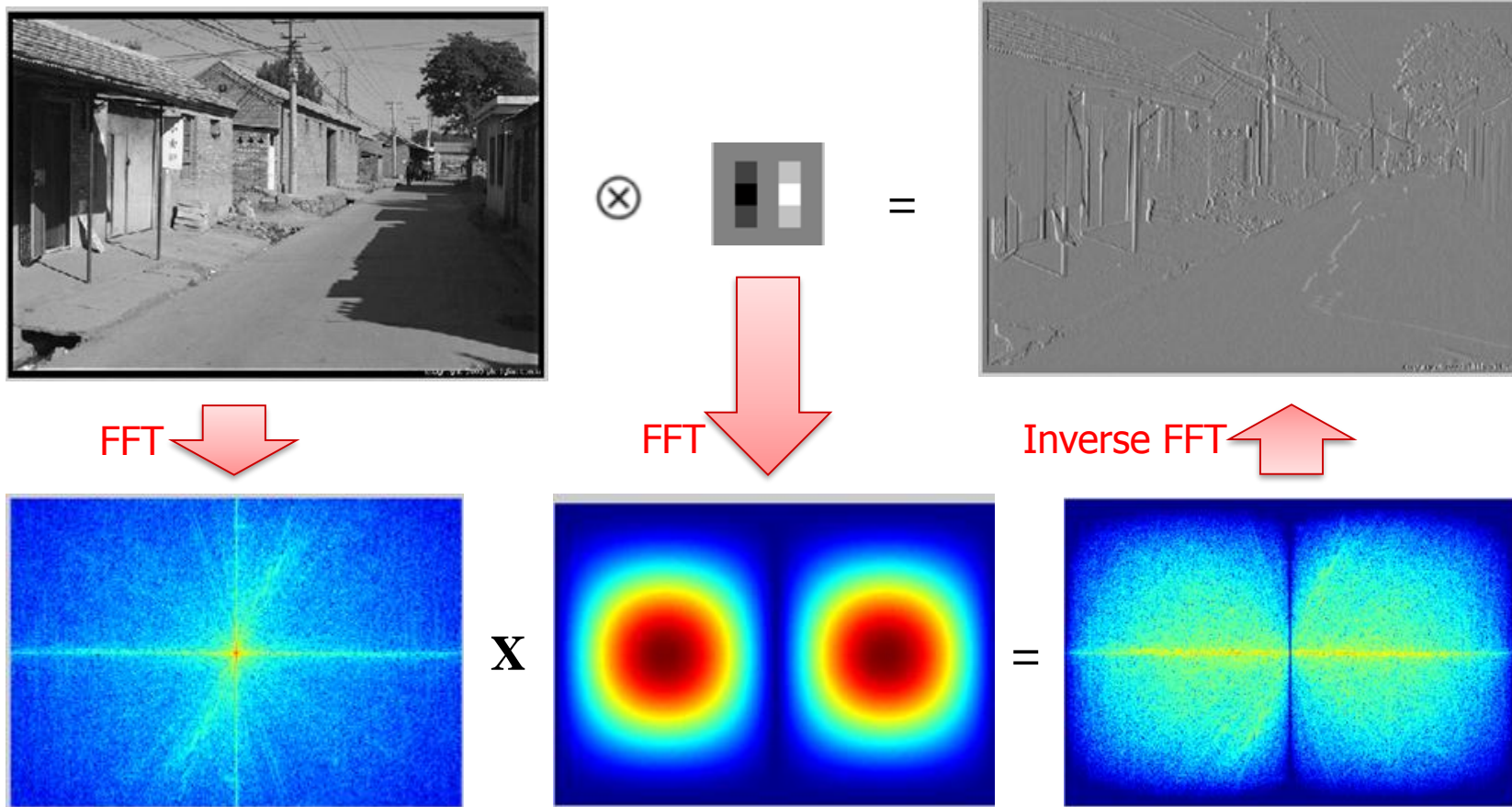
频域分析及变换

频段滤除



频域分析及变换

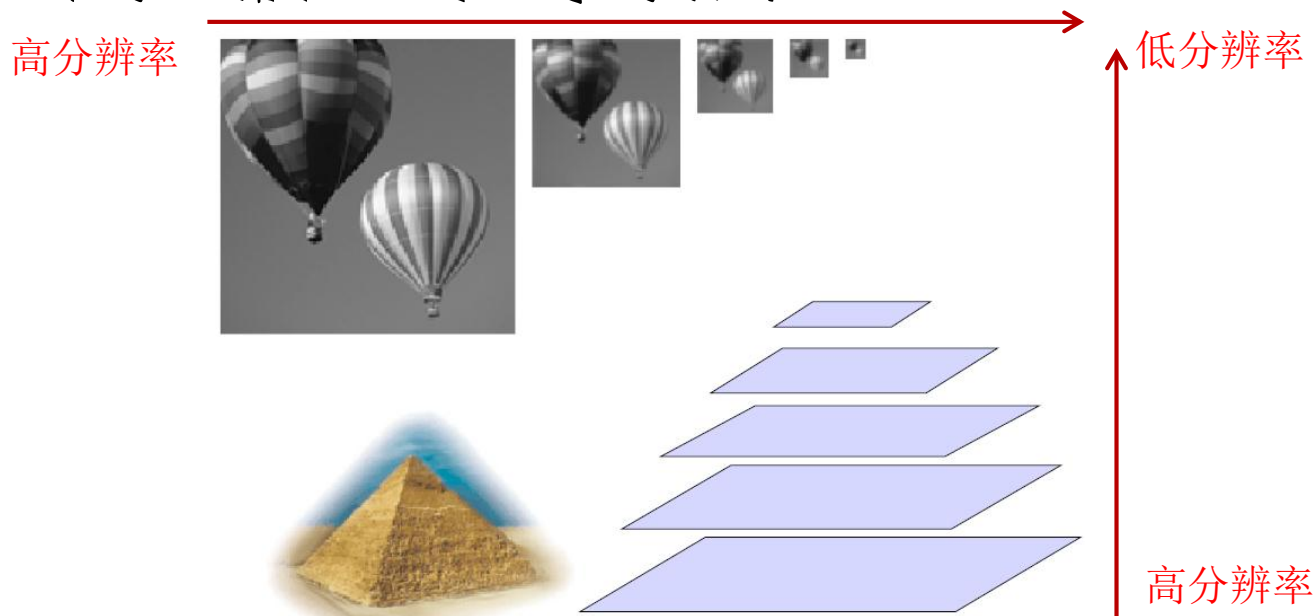
空域卷积=频域乘积



金字塔

高斯金字塔

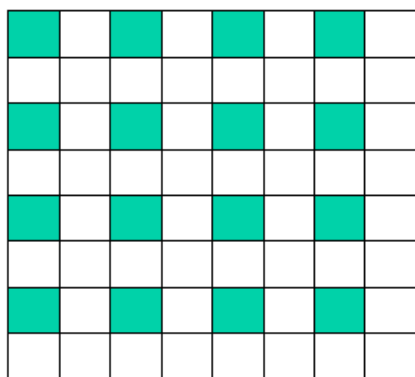
- 原因：多次高斯卷积之后，一些像素是多余的
- 操作： n 次（高斯卷积 $\rightarrow 2\times$ 降采样） $\rightarrow n$ 层金字塔
- 目的：捕捉不同尺寸的物体



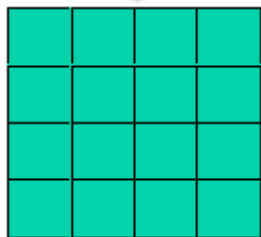
金字塔

高斯金字塔

- 高斯滤波的必要性
- 直接降采样损失信息



降采样



原图



直接降采样

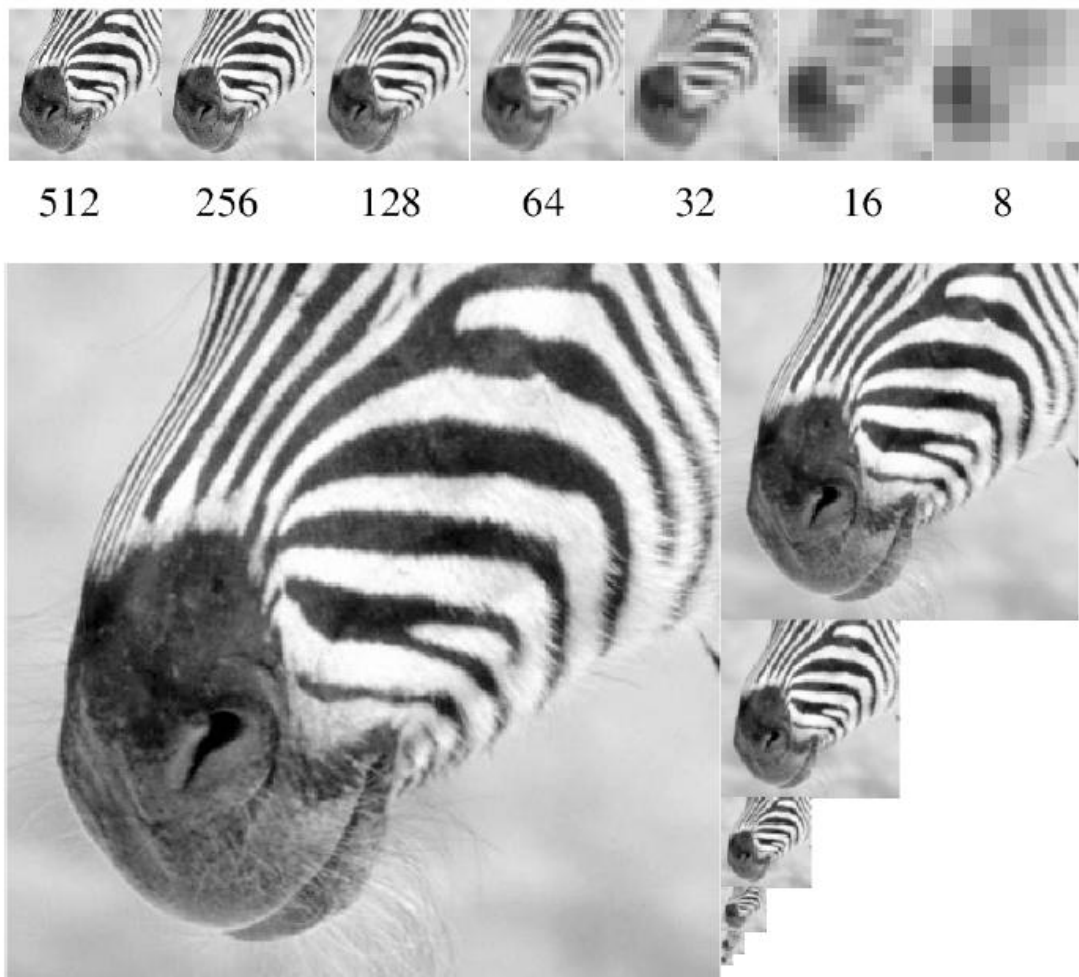
高斯+降采样

金字塔

高斯金字塔

- 尺度空间

- 不同尺度适合不同尺寸的物体
- 合适的尺度永远未知



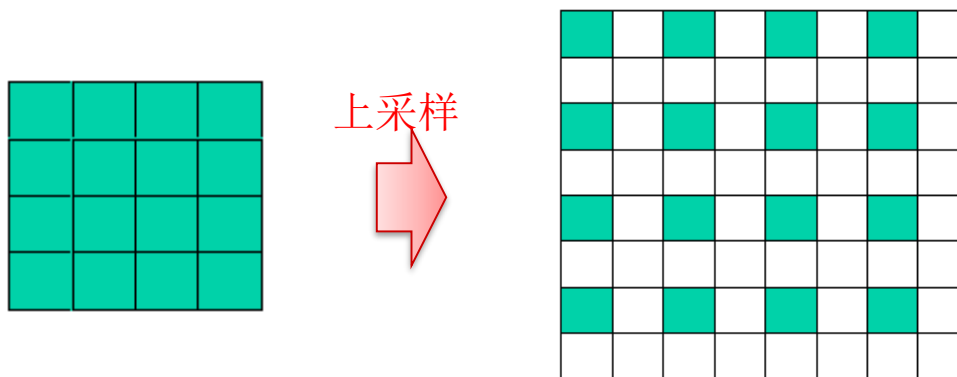
金字塔

拉普拉斯金字塔 (Laplacian)

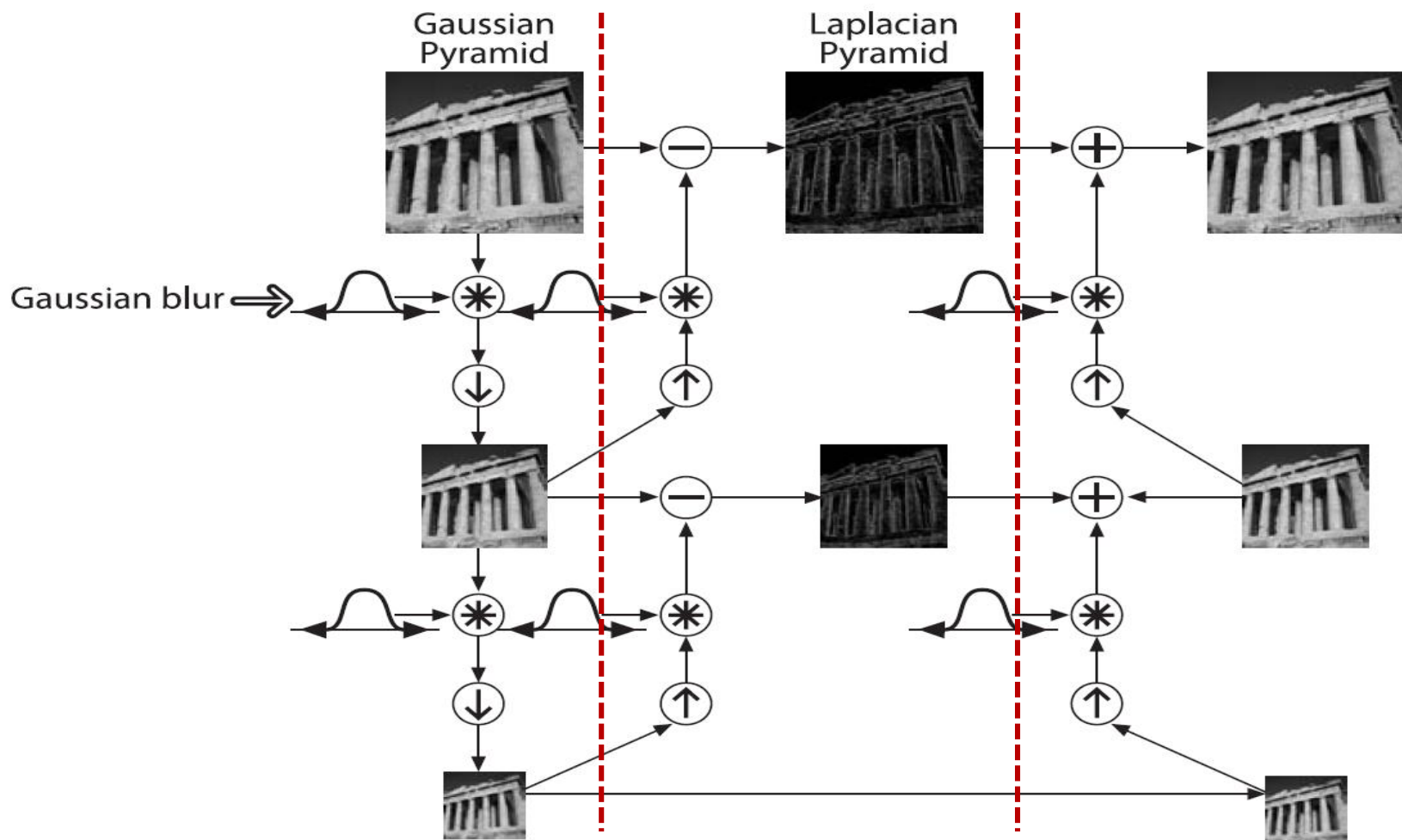
- 高频细节信息在卷积和下采样中丢失
- 保留所有层所丢失的高频信息，用于图像恢复

$$L_i = G_i - \text{UP}(G_{i+1}) \otimes \mathcal{G}_{5 \times 5}$$

http://blog.csdn.net/poem_qianmo



金字塔



http://blog.csdn.net/poem_qianme

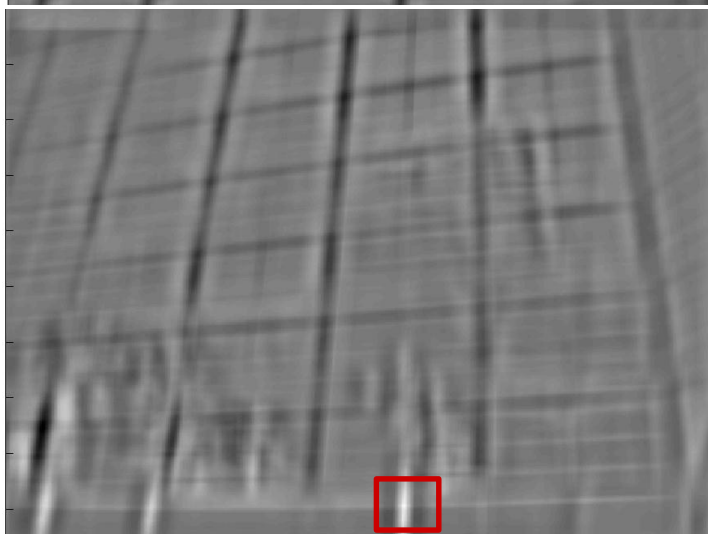
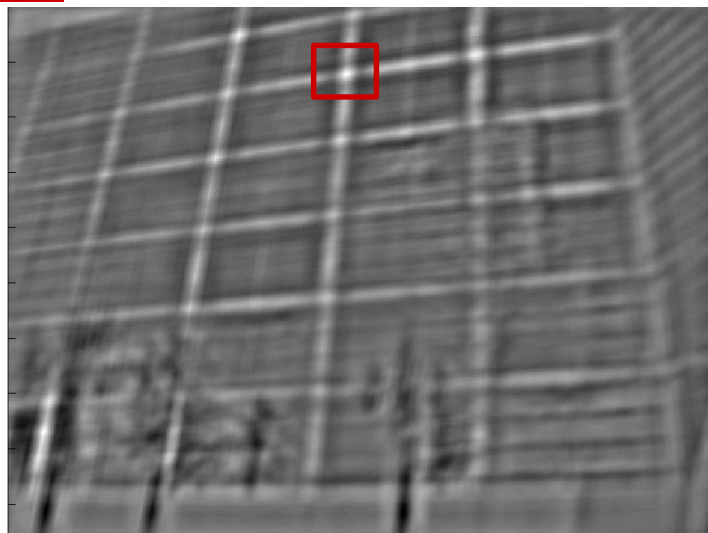
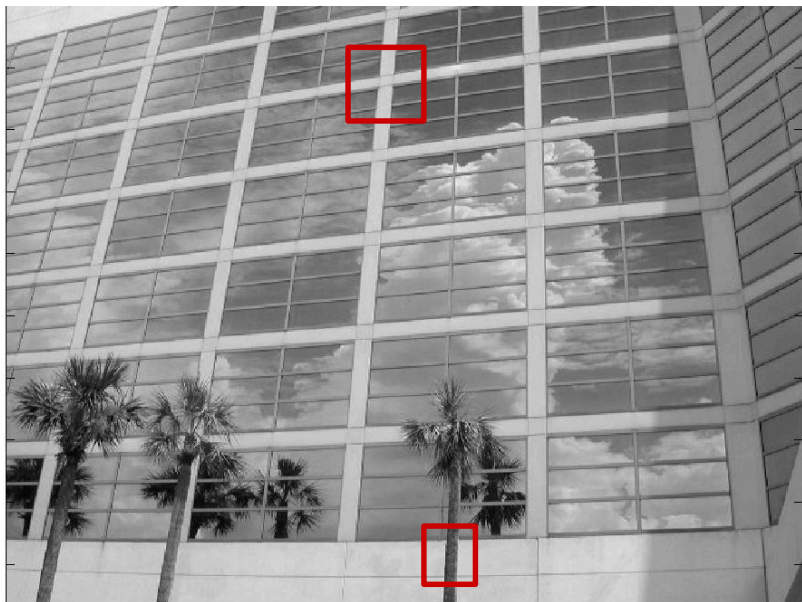
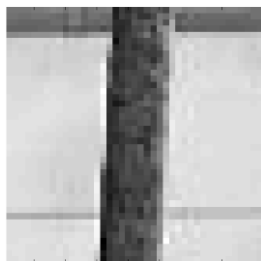
模板匹配

模板图片匹配 VS 卷积

- 作用：同尺度目标检测
- 模板：真实图片 VS 卷积核
- 操作：使用模板图片扫描整个图片 VS 卷积扫描
- 匹配结果：相似度量 VS 权重相加
 - 返回相似度图
 - 6中相似距离计算
 - (标准化) 欧式距离
 - (标准化) 相关
 - (标准化) 去均值相关

模板匹配

模板图片匹配



演示环节

- Github

- <https://github.com/349zzjau>

- 链接推荐

- 平滑滤波

- http://docs.opencv.org/master/d4/d13/tutorial_py_filtering.html

- 梯度滤波

- http://docs.opencv.org/master/d5/d0f/tutorial_py_gradients.html

- 傅里叶变换

- http://docs.opencv.org/master/de/dbc/tutorial_py_fourier_transform.html

- 模板匹配

- http://docs.opencv.org/master/d4/dc6/tutorial_py_template_matching.html

- 金字塔

- http://docs.opencv.org/master/dc/df/tutorial_py_pyramids.html

疑问

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回复问题

Q & A

小象账号：349zzjau

课程名：基于深度学习的计算机视觉

课后调查问卷：<http://cn.mikecrm.com/0D9JujS>

联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

