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。



关注 小象学院

第2讲　图像预处理

计算机视觉

主讲人：叶梓

上海交通大学博士
主要研究方向：机器学习、深度学习、人工智能

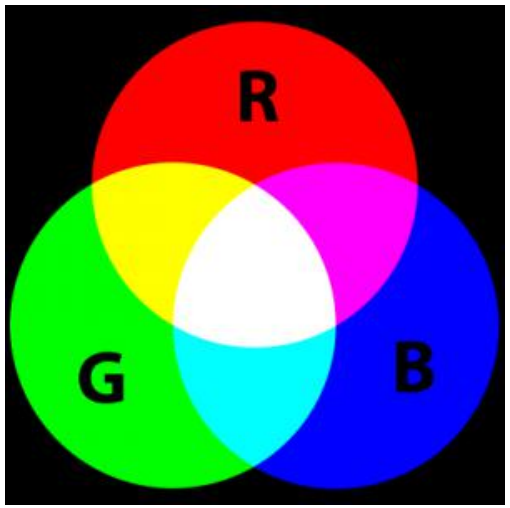
本章内容

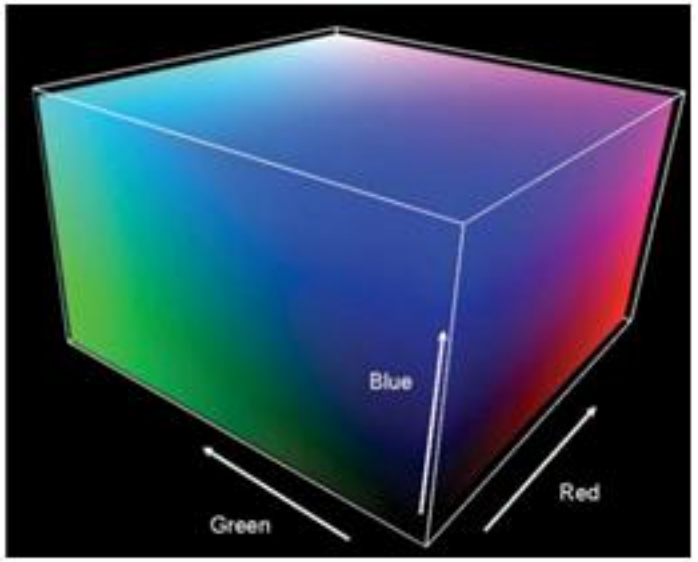
- 图像显示与存储原理
- 图像增强的目标
- 点运算：基于直方图的对比度增强
- 形态学处理
- 空间域处理：卷积
- 卷积的应用（平滑、边缘检测、锐化等）
- 频率域处理：傅里叶变换、小波变换
- 应用案例：
  - 平滑、边缘检测、CLAHE等

颜色空间

RGB颜色空间

- 加法混色，彩色显示器
- 3个通道
  - 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]

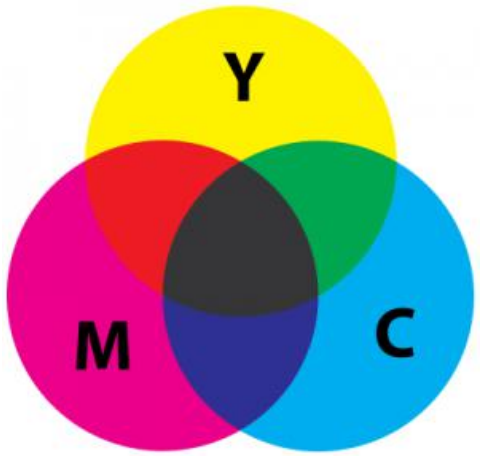


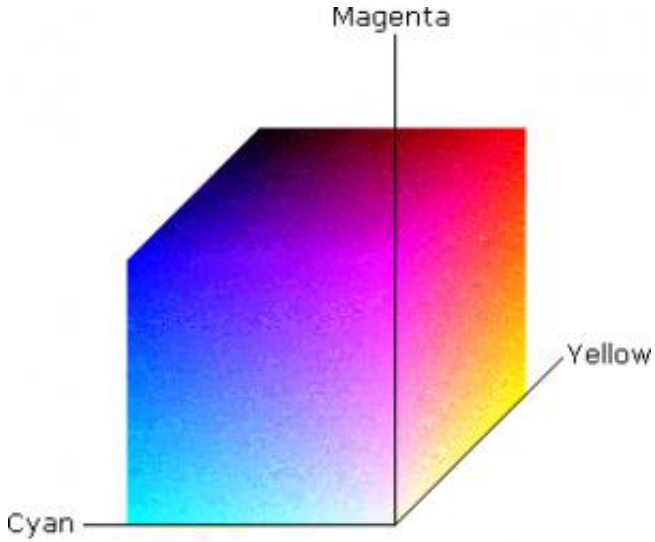


颜色空间

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]

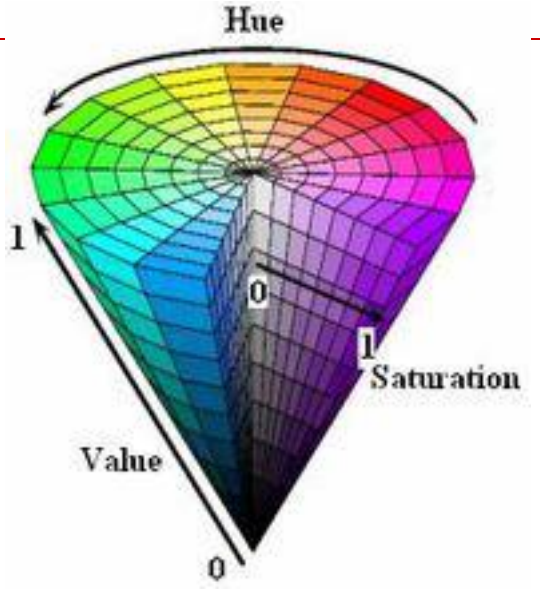




颜色空间

HSV颜色空间

- 人类视觉概念，画家配色
- 3个要素
  - H/Hue：色调，颜色种类
  - S/Saturation：饱和度，颜色的纯度
  - V/Value：明度，颜色明亮度
- 一个像素颜色值
  - （h, s, v）
- 取值范围
  - [0, 255]
  - [0.0, 1.0]



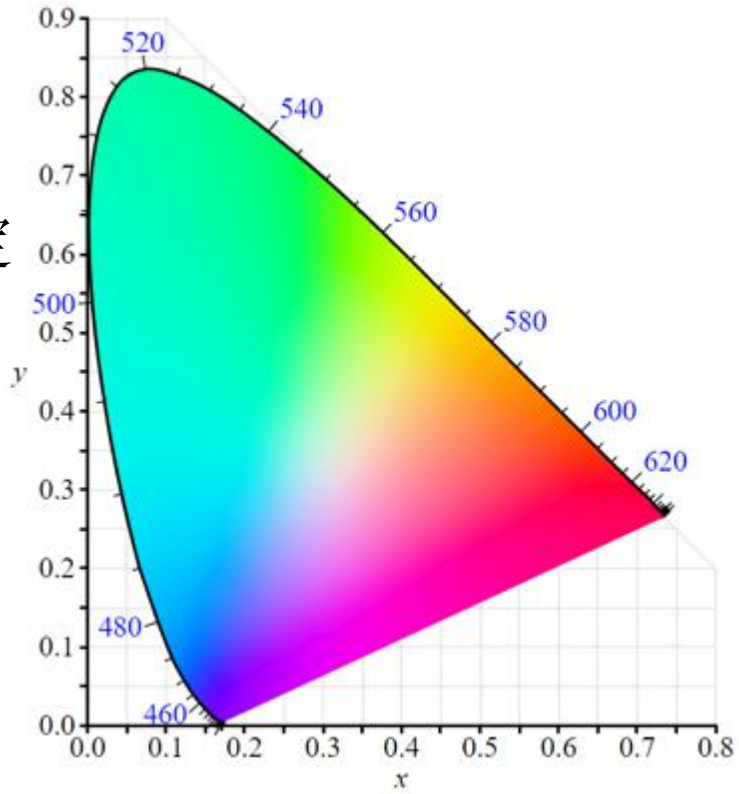




颜色空间

CIE-XYZ颜色空间

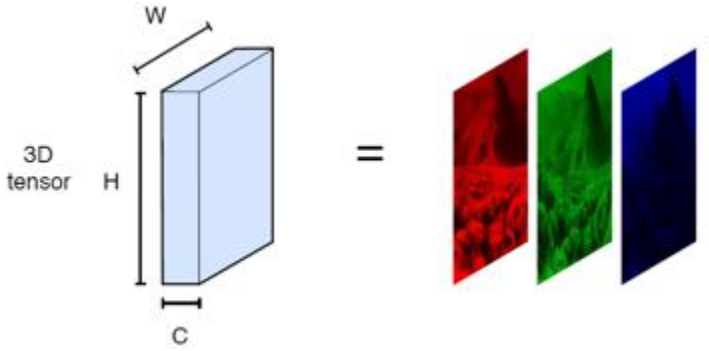
- 国际照明协会，1931
- 基于人类颜色视觉的直接测定
- 其他颜色空间基础
- 人类视觉系统-视锥细胞
  - 短波（S, 420-440nm）
  - 中波（M, 530-540nm）
  - 长波（L, 560-580nm）
- 3色刺激值通道
  - X，Y，Z约略对应于红色、绿色、蓝色
  - 一种波的刺激等于几种波的混合刺激

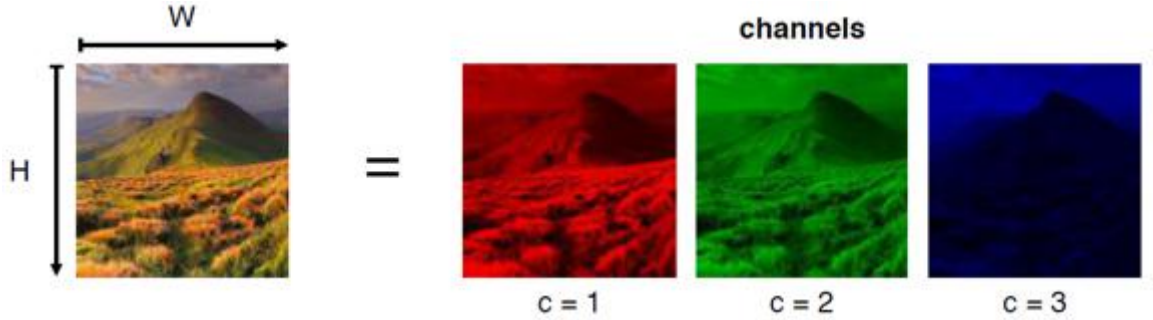


图片存储原理

主流颜色空间

- RGB三通道彩色图
  - 图片→3维矩阵（[0,255]）





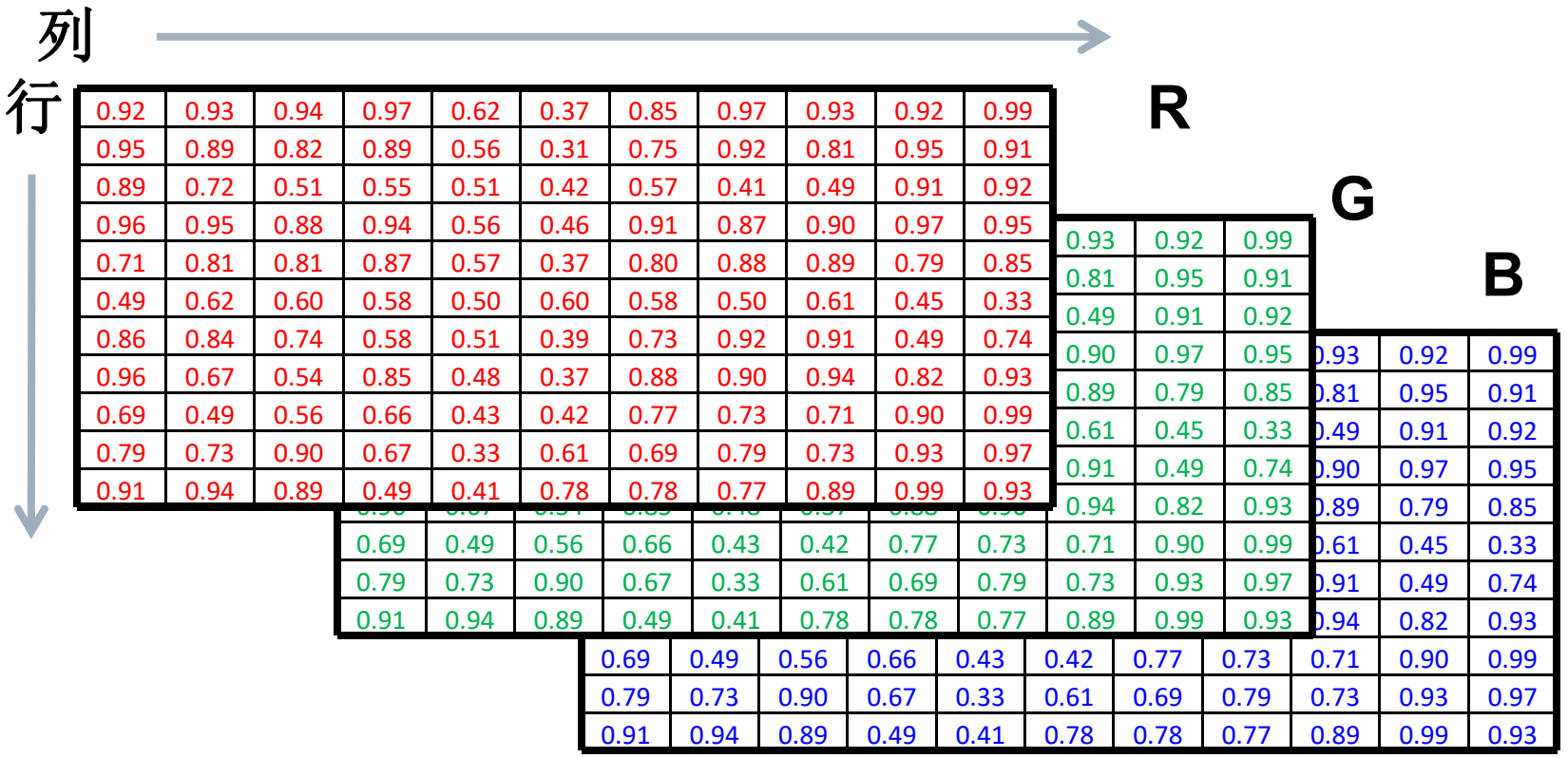
- 单通道灰度图
  - 亮度信息（[0,255]）
  - Gray=R*0.3+G*0.59+B*0.11



图片存储原理

RGB图片数据格式

- 3维矩阵



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

0.93	0.92	0.99								
0.81	0.95	0.91								
0.49	0.91	0.92								
0.90	0.97	0.95								
0.89	0.79	0.85								
0.61	0.45	0.33								
0.91	0.49	0.74								
0.94	0.82	0.93								
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

0.93	0.92	0.99								
0.81	0.95	0.91								
0.49	0.91	0.92								
0.90	0.97	0.95								
0.89	0.79	0.85								
0.61	0.45	0.33								
0.91	0.49	0.74								
0.94	0.82	0.93								
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

图片存储原理

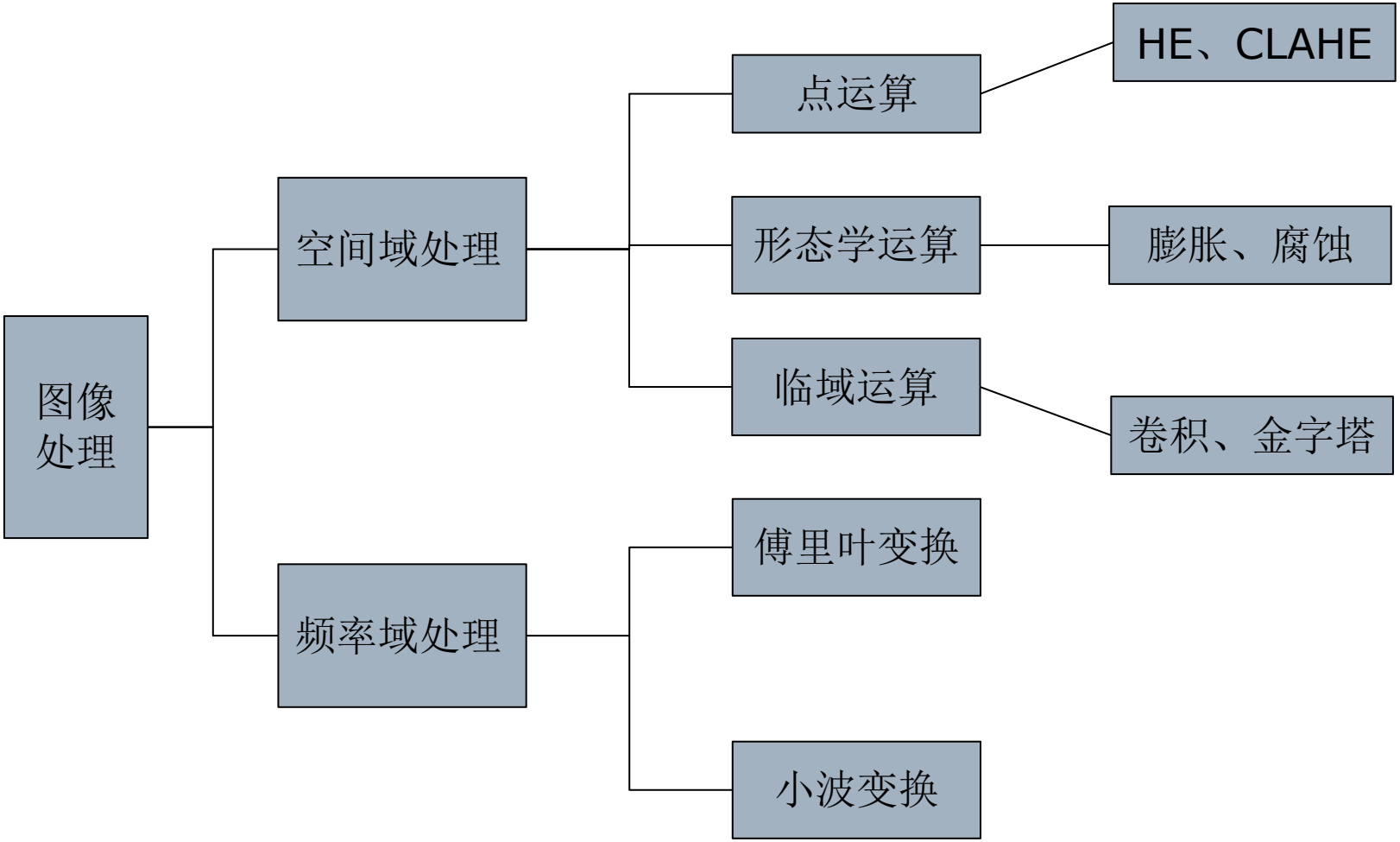
- 常见的存储的格式有：bmp, jpg, png, tiff, gif, pcx, tga, exif, fpx, svg, psd, cdr, pcd, dxf, ufo, eps, ai, raw, WMF, webp等
- BMP：采用位映射存储格式，不采用其他任何压缩，所占用的空间很大。
- JPG：最常见的有损压缩格式，能够将图像压缩到很小的空间，压缩比可达10:1到40:1之间。
- GIF：基于LZW算法的连续色调的无损压缩格式，其压缩率一般在50%左右。
- PNG：是比较新的图像文件格式，能够提供长度比GIF小30%的无损压缩图像文件。

图像增强的目标

- 改善图像的视觉效果；
- 转换为更适合于人或机器分析处理的形式；
- 突出对人或机器分析有意义的信息；
- 抑制无用信息，提高图像的使用价值；

- 具体的说：
- 包括图像锐化，平滑、去噪，灰度调整（对比度增强）

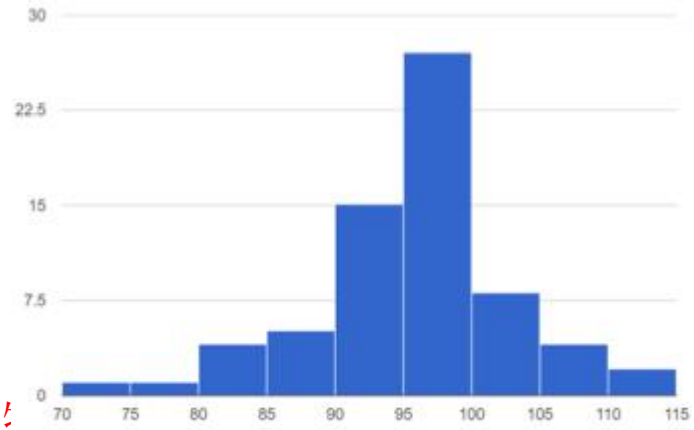
图像处理方法


图像处理
空间域处理
频率域处理
点运算
形态学运算
临域运算
傅里叶变换
小波变换
HE、CLAHE
膨胀、腐蚀
卷积、金字塔

特征提取方法

直方图（Histogram）

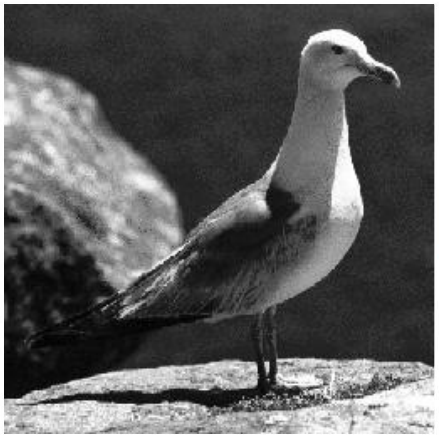
- 对图片数据/特征分布的一种统计
  - 灰度、颜色
  - 梯度/边缘、形状、纹理
  - 局部特征点、视觉词汇
- 区间（bin）
  - 具有一定的统计或物理意义
  - 一种数据或特征的代表
  - 需要预定义或基于数据进行学习
  - 数值是一种统计量：概率、频数、

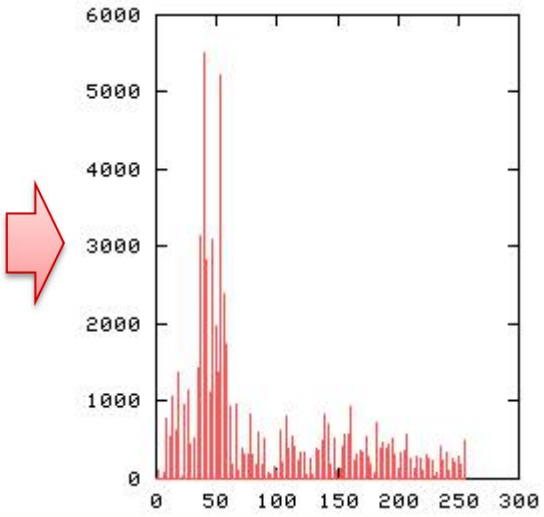


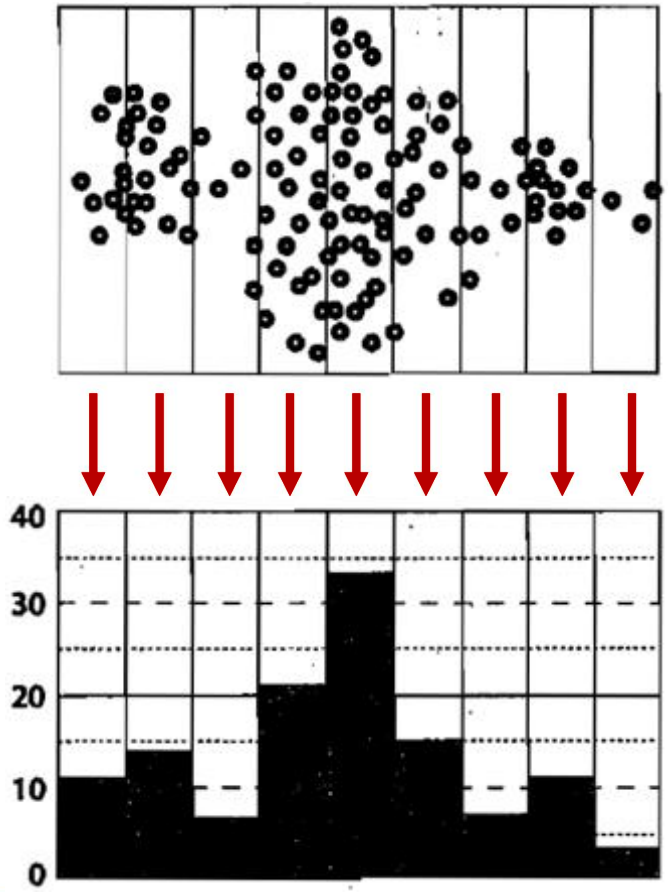
特征提取方法

直方图

- 对数据空间（bin）进行量化







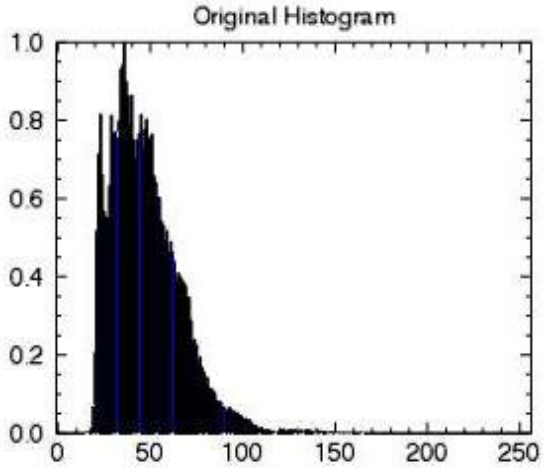
直方图均衡化

- 直方图均衡化是指：利用图像直方图对对比度进行调整的方法。
- 直方图均衡化通常用来增加许多图像的局部对比度，尤其是当图像的有用数据的对比度相当接近的时候。
- 直方图均衡化以后，亮度可以更好地在直方图上分布。这样就可以用于增强局部的对比度而不影响整体的对比度，直方图均衡化通过有效地扩展常用的亮度来实现这种功能。

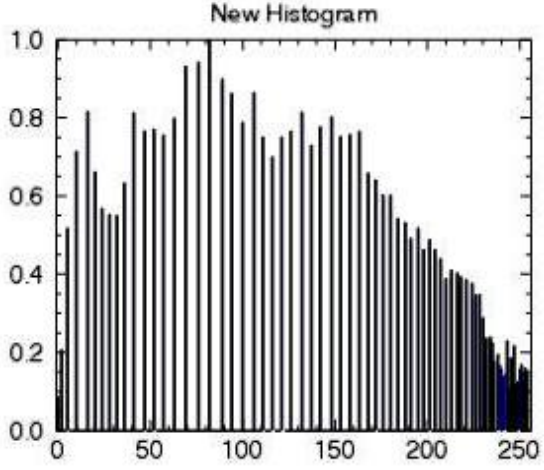
直方图均衡化

- 直方图均衡化：实质上是对图像进行非线性拉伸
- 重新分配各个灰度单位中的像素点数量，使一定灰度范围像素点数量的值大致相等。







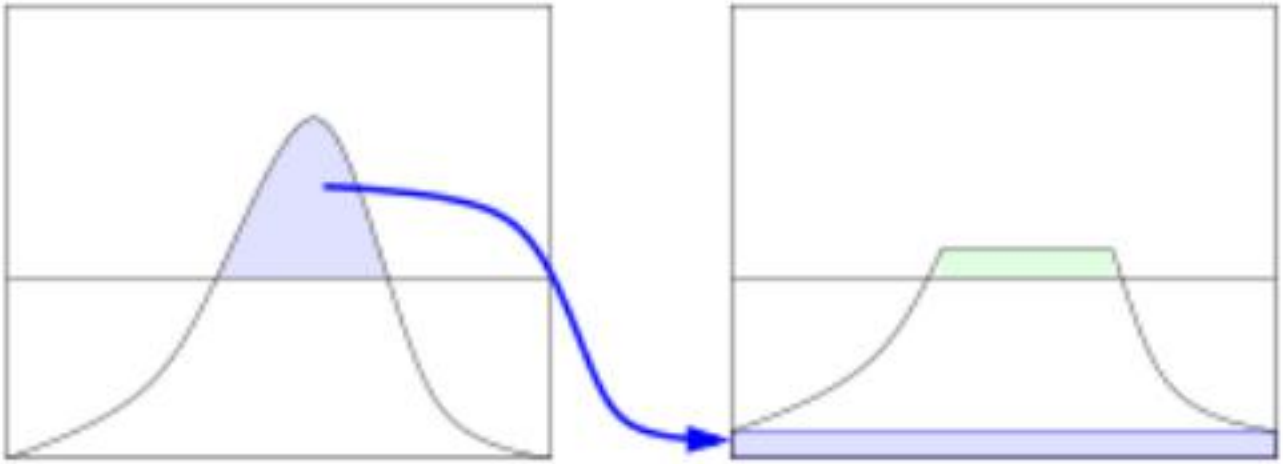


自适应直方图均衡

- 直方图均衡的经典算法对整幅图像的像素使用相同的变换，如果图像中包括明显亮的或者暗的区域，则经典算法作用有限。
- 自适应直方图均衡（AHE）算法通过对局部区域进行直方图均衡，来解决上述问题。
  - 移动模板在原始图片上按特定步长滑动；
  - 每次移动后，模板区域内做直方图均衡，映射后的结果赋值给模板区域内所有点，
  - 每个点会有多次赋值，最终的取值为这些赋值的均值。

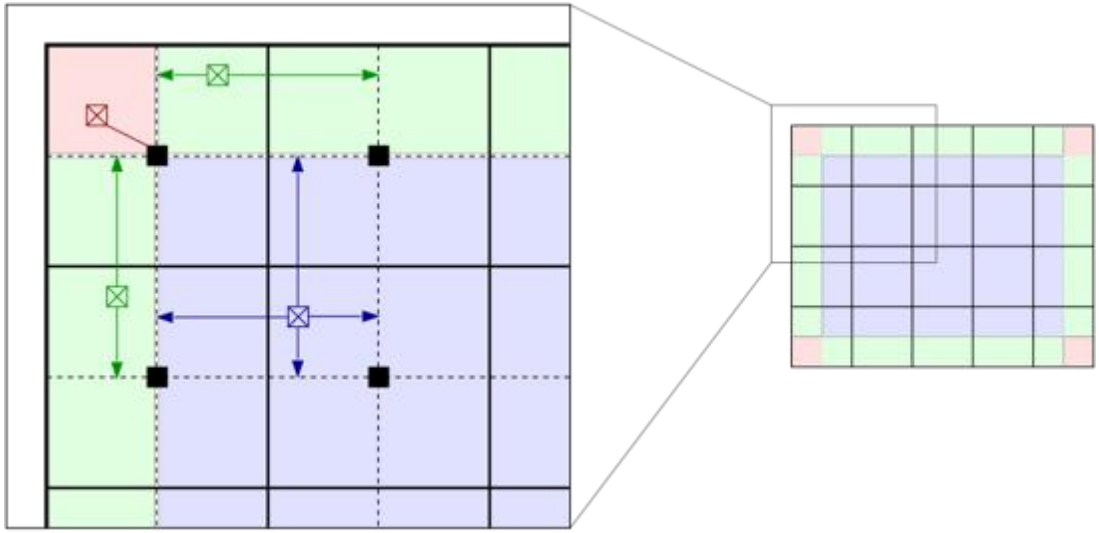
CLAHE

- AHE会过度放大图像中相对均匀区域的噪音，可采用限制对比度自适应直方图均衡(CLAHE) 。
- 与普通的自适应直方图均衡相比，CLAHE的不同地方在于直方图修剪过程，用修剪后的直方图均衡图像时，图像对比度会更自然。



CLAHE

- 小黑点的灰度直接由映射函数计算得到；
- 粉色区域内点的灰度由映射函数计算而得；
- 绿色区域内点的灰度由由相邻2块灰度映射值线性插值而得；
- 其他区域所有点的灰度由相邻4块的灰度映射值双线性插值而得。



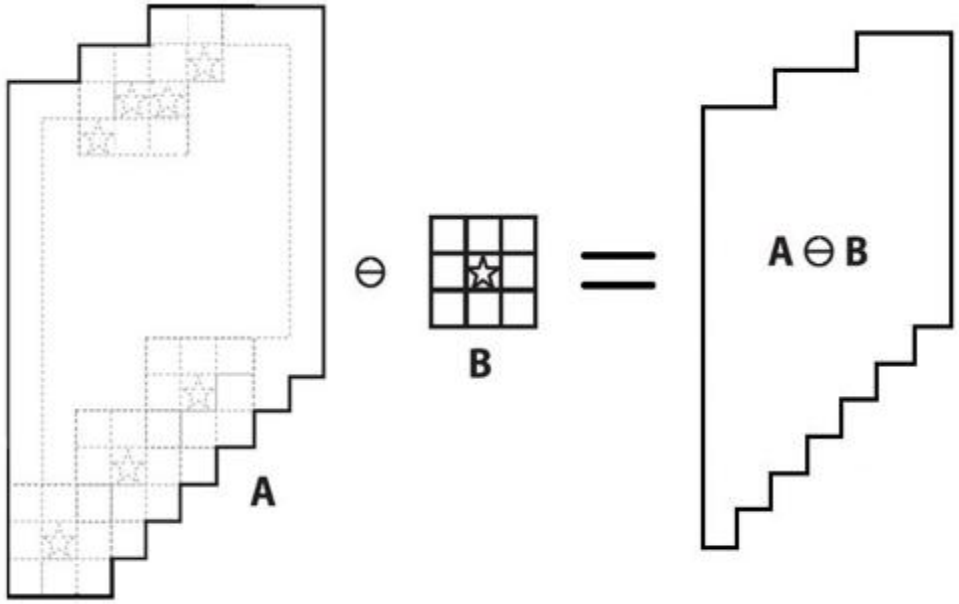
CLAHE算法步骤

- 1.图像分块，以块为单位；
- 2.先计算直方图，然后修剪直方图，最后均衡；
- 3.遍历操作各个图像块，进行块间双线性插值；
- 4.与原图做图层滤色混合操作。（可选）



形态学运算

- 膨胀是图像中的高亮部分进行膨胀，类似于领域扩张。
- 腐蚀是原图的高亮部分被腐蚀，类似于领域被蚕食。



形态学运算

- **开运算：**先腐蚀再膨胀，可以去掉目标外的孤立点。
- **闭运算：**先膨胀再腐蚀，可以去掉目标内的孔。
- 通常，当有噪声的图像用阈值二值化后，所得到的边界是很不平滑的，物体区域具有一些错判的孔洞，背景区域散布着一些小的噪声物体，连续的开和闭运算可以显著的改善这种情况。

空间域处理及其变换

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→权重相加
    - 卷积核、卷积模板
    - 滤波器、滤波模板
    - 扫描窗

滤波结果 滤波函数 像素邻域值

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

- 不同功能需要定义不同的函数
  - 平滑/去燥
  - 梯度/锐化
  - 边缘、显著点、纹理
  - 模式检测

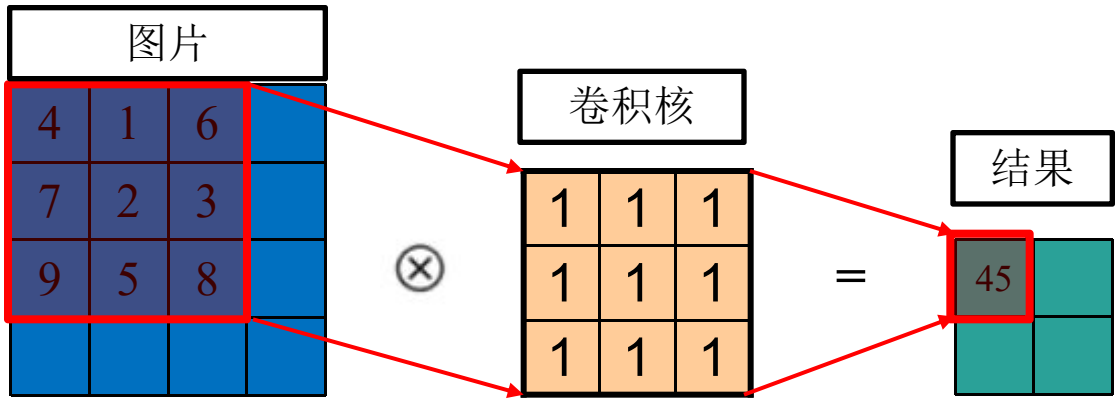
空域分析及变换

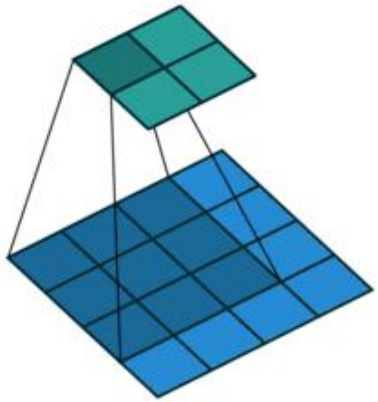
滤波/卷积

滤波结果 $h[x,y]=\sum_{k,l} f[k,l]\, I[x+k,y+l]$ 滤波函数 像素邻域值

- 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)

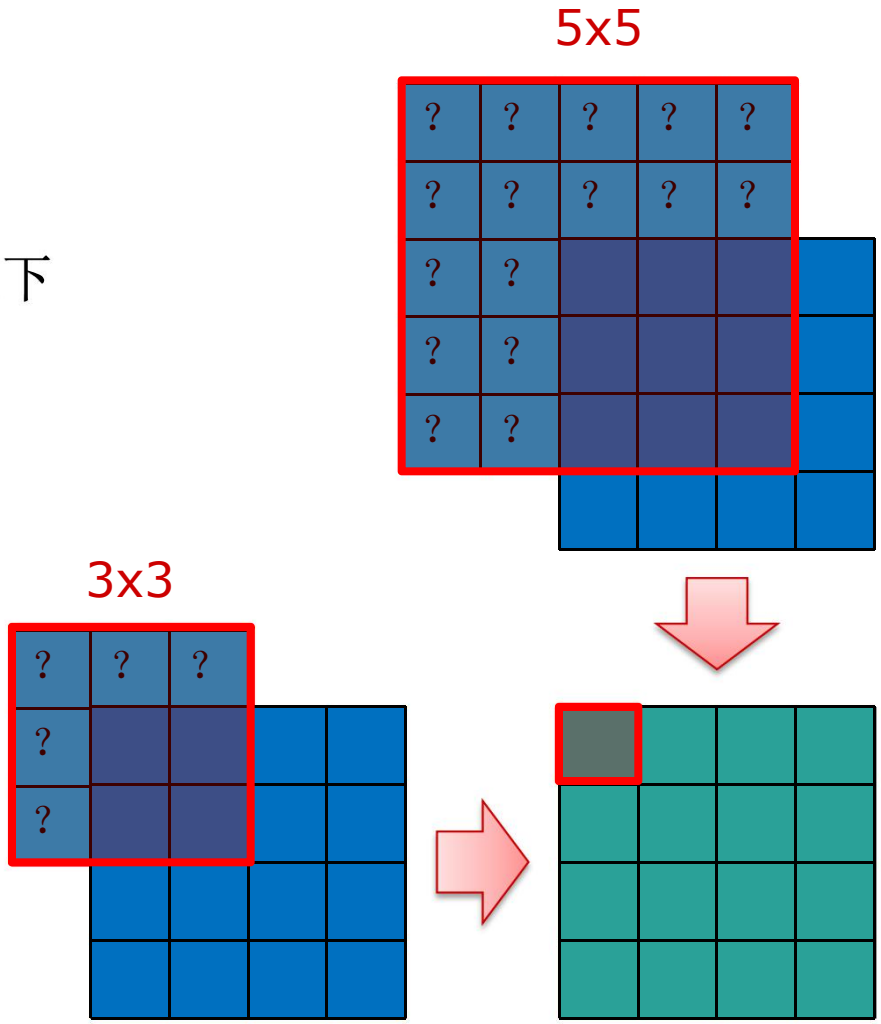




空域分析及变换

滤波/卷积

- 边界填充（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）



边界填充策略

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

边界填充策略

滤波/卷积

- 7x7卷积：3x3 → 9x9

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

边界填充策略

滤波/卷积

- 7x7卷积：3x3 → 9x9

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

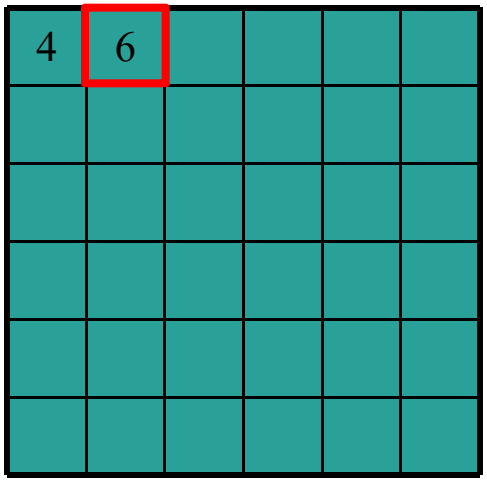
均值滤波示意

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



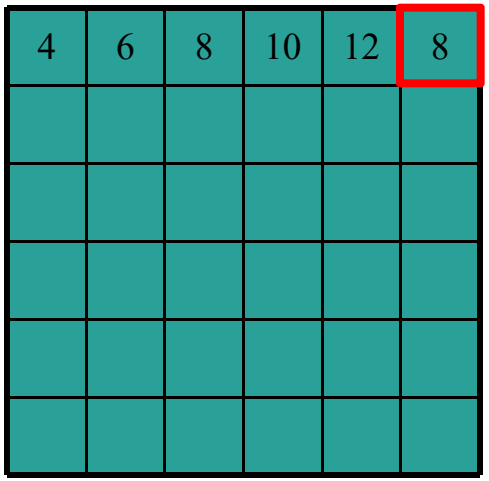
均值滤波示意

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



均值滤波示意

平滑均值滤波/卷积

- 3x3
- 扫描步长：1
- 边框补零

均值滤波本身存在缺陷，既没有很好地去除噪声点，也破坏了图像的细节反而使图像变得模糊。

输入图片
6x6

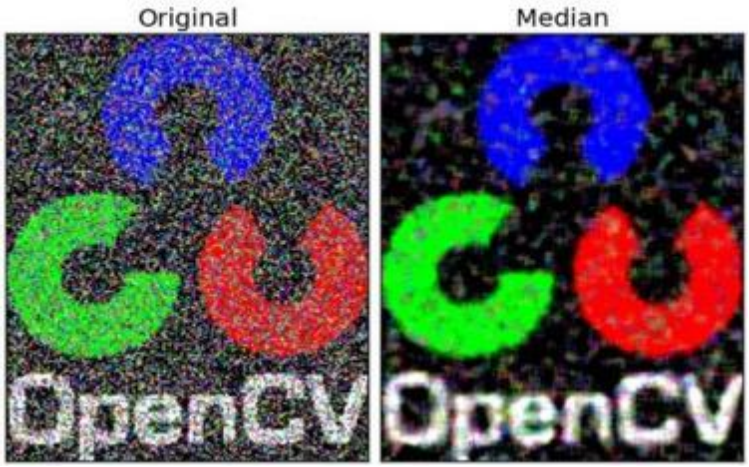
0	0	0	0	0	0	0	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	9	9	9	18	18	18	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	18	18	18	9	9	9	0
0	0	0	0	0	0	0	0

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

均值滤波

平滑均值滤波/卷积

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





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

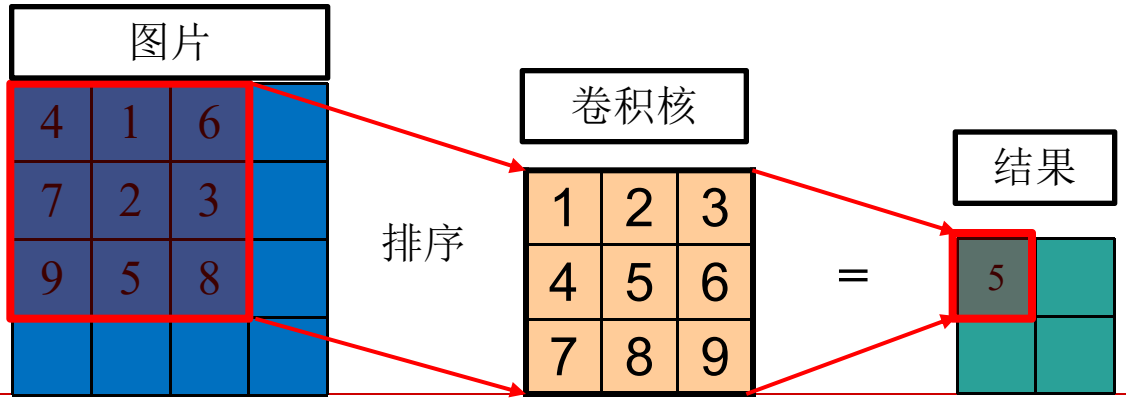


空域分析及变换

平滑中值滤波/卷积

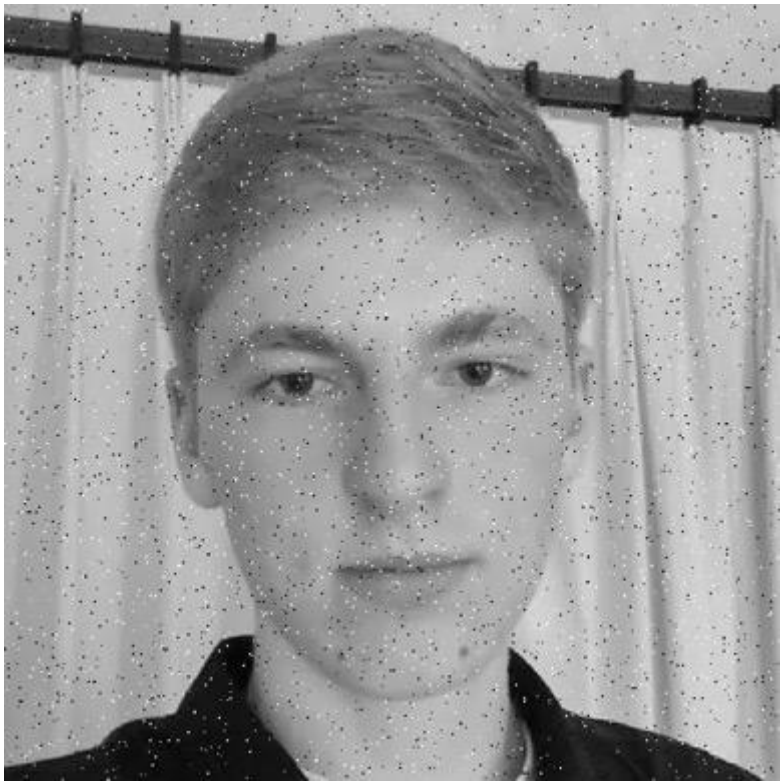
- 奇数尺寸
  - 3x3，5x5，7x7，2n-1 x 2n-1
- 操作原理
  - 卷积域内的像素值从小到大排序
  - 取中间值作为卷积输出
- 有效去除椒盐噪声

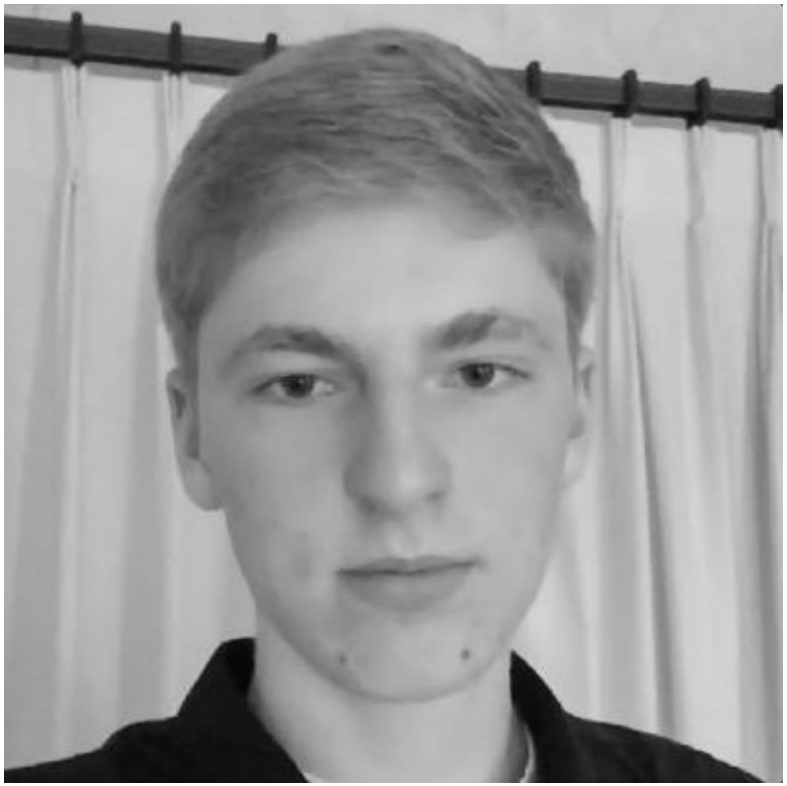
将领域矩阵中的N个像素进行排序，并将这个矩阵的中心点赋值为这N个像素的中值。



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

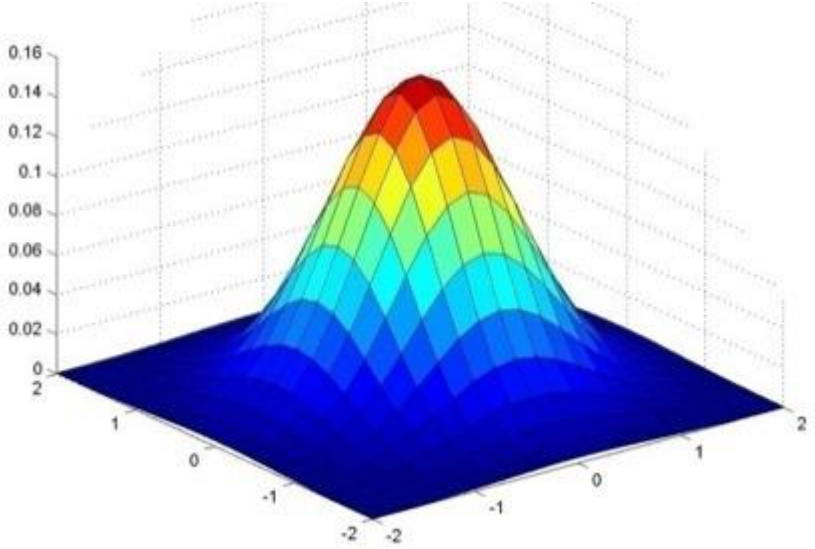


- 奇数尺寸
  - 3x3，5x5，7x7，2n-1 x 2n-1
- 模拟人眼，关注中心区域
- 有效去除高斯噪声
- 参数
  - x, y是卷积参数坐标
  - 标准差 σ

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

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, σ = 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



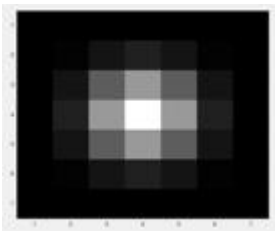
人类视觉

空域分析及变换

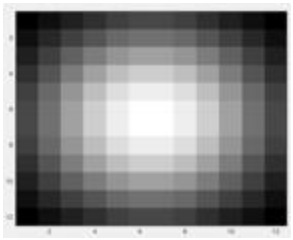
平滑高斯滤波/卷积

- σ越小
- 关注区域越集中

σ=1



σ=4



原图







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

1	2	1
2	4	2
1	2	1

=

1
2
1

X

1	2	1

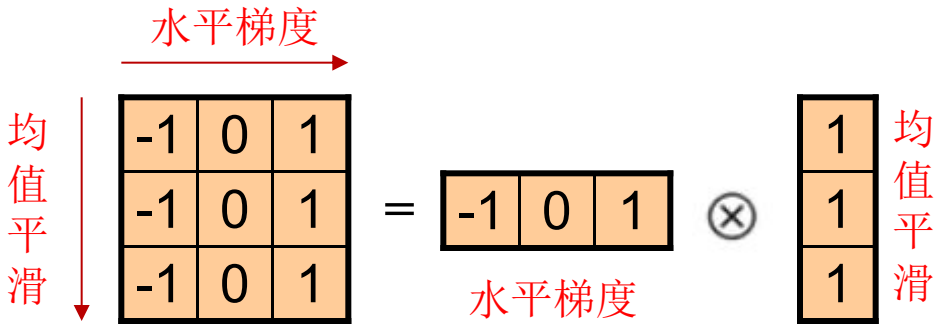
$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp^{-\frac{x^2+y^2}{2\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi}\sigma}\exp^{-\frac{x^2}{2\sigma^2}}\right)\left(\frac{1}{\sqrt{2\pi}\sigma}\exp^{-\frac{y^2}{2\sigma^2}}\right)$$

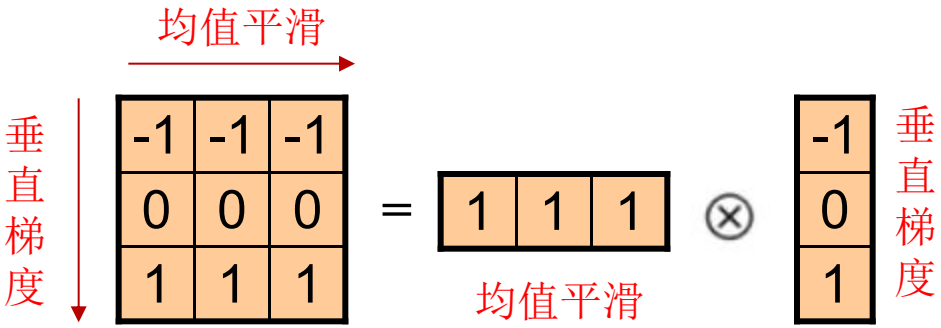
空域分析及变换

梯度Prewitt滤波/卷积

- 水平梯度/垂直边缘



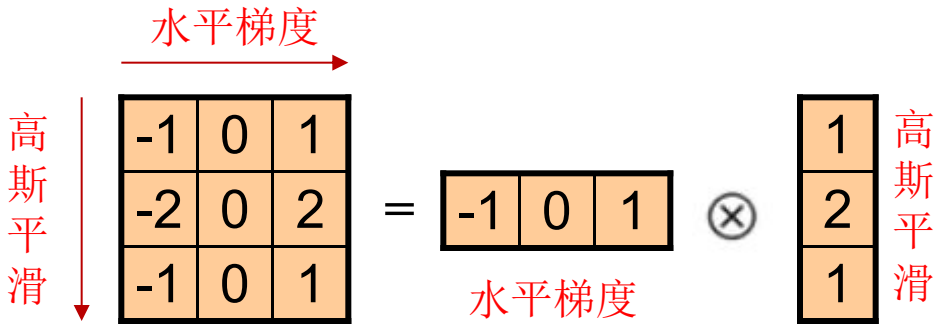
- 垂直梯度/水平边缘



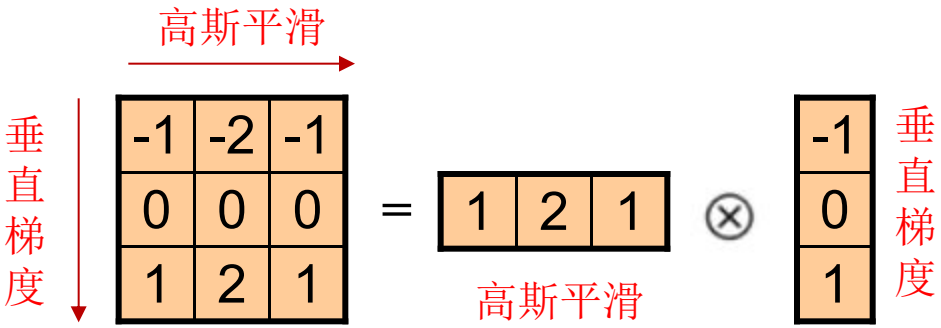
空域分析及变换

梯度Sobel滤波/卷积

- 水平梯度/垂直边缘



- 垂直梯度/水平边缘



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

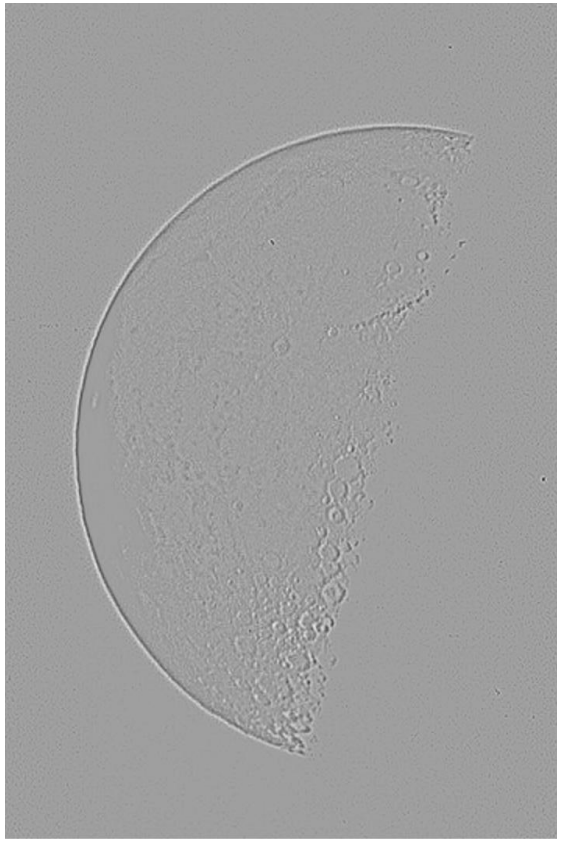
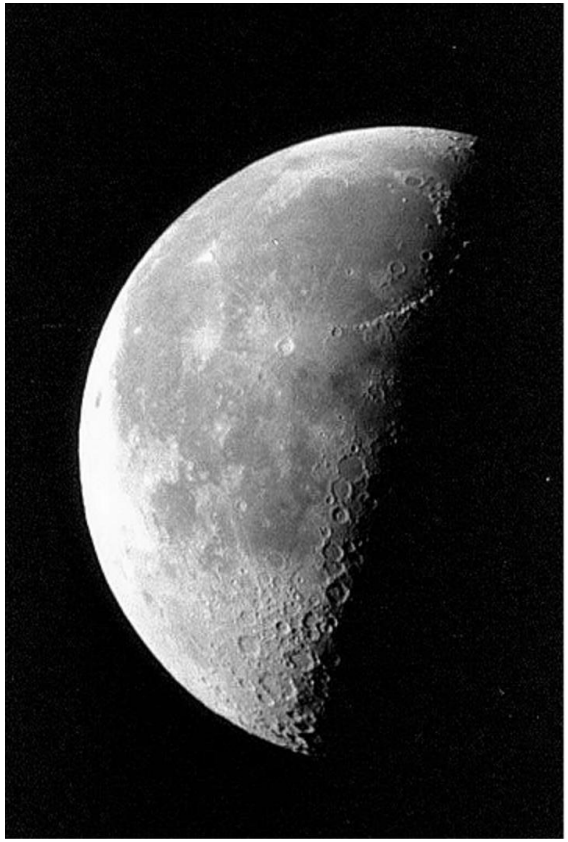
2D 8邻域

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

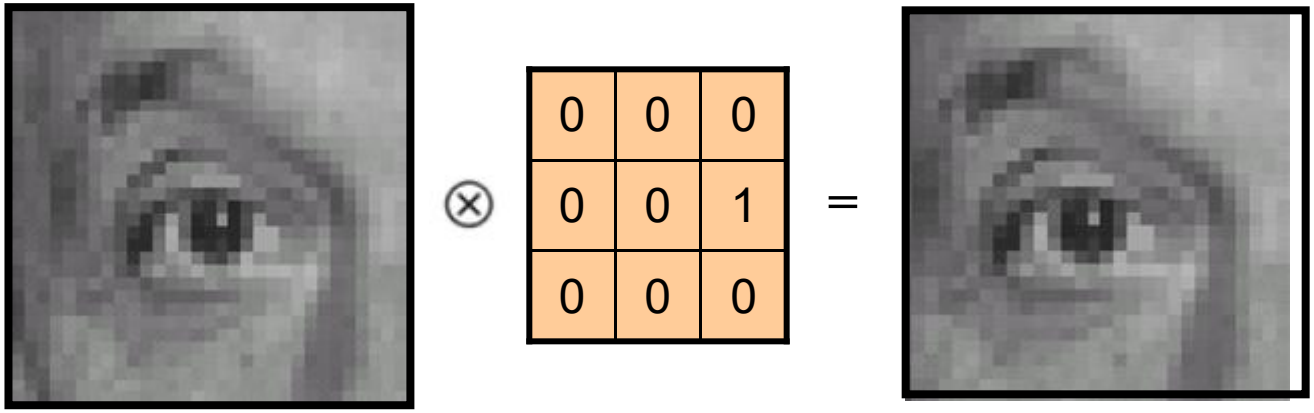
Laplacian滤波锐化

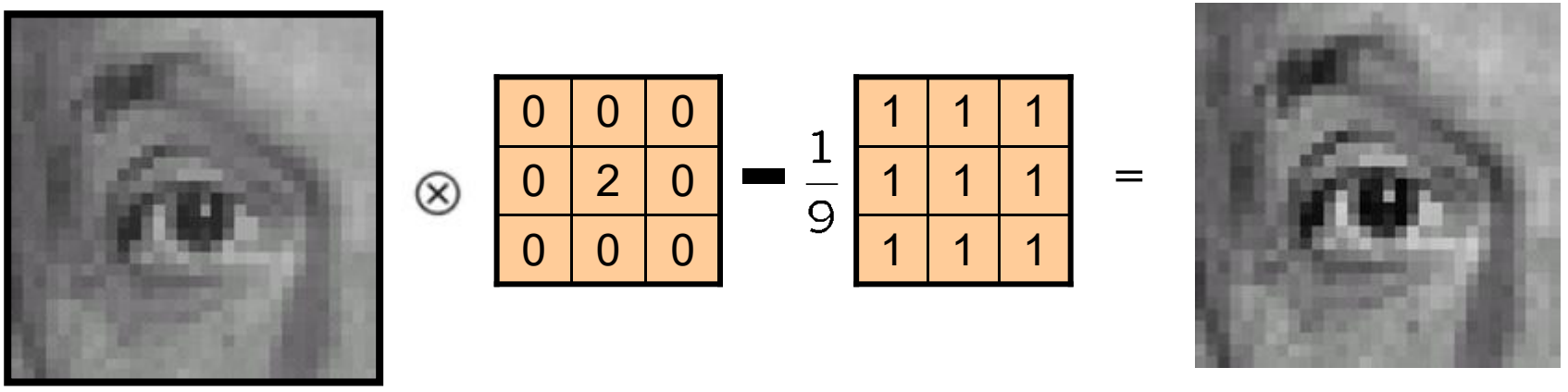
空域分析及变换

其他滤波/卷积

- (左)移位

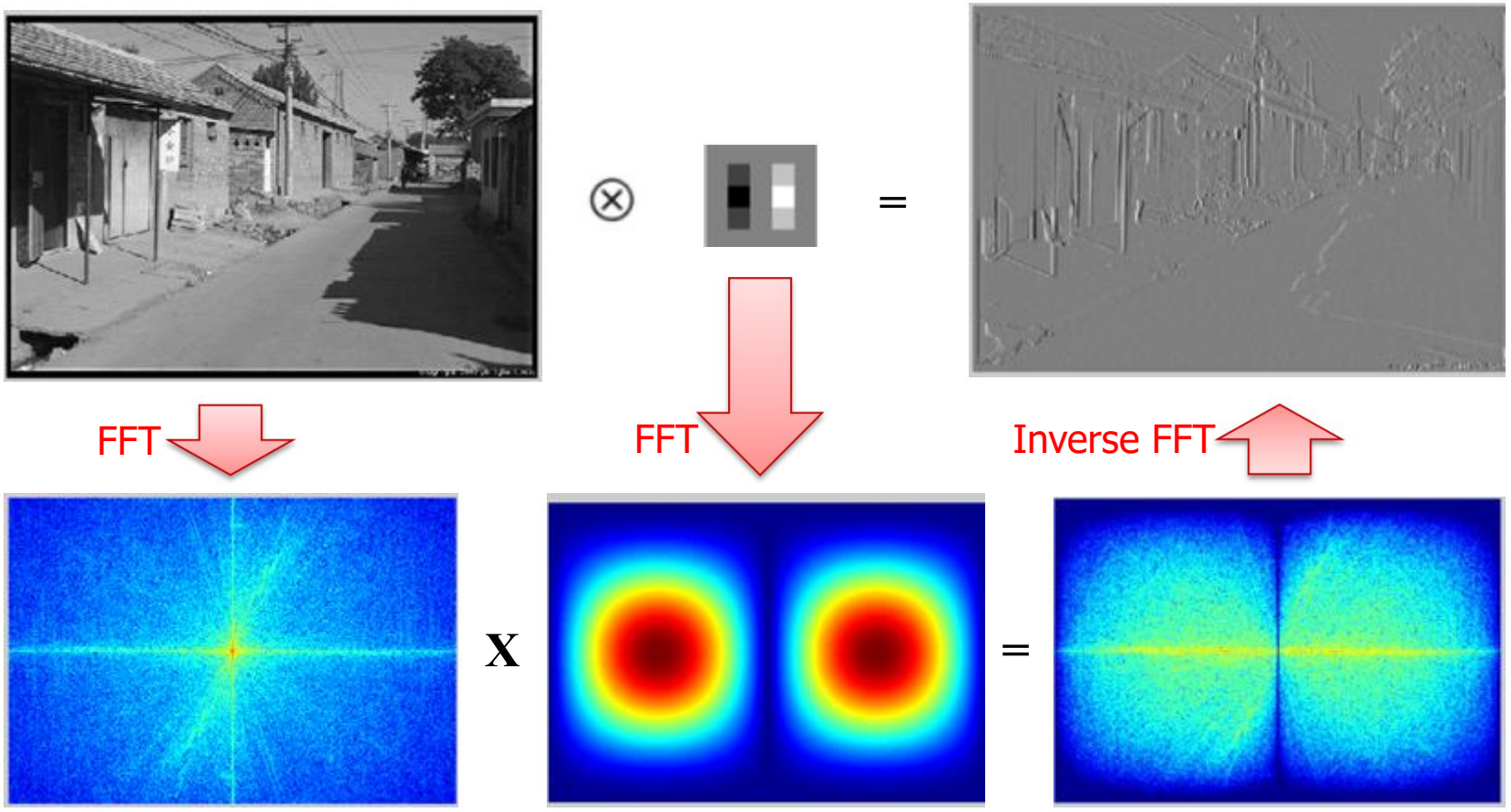


- 锐化



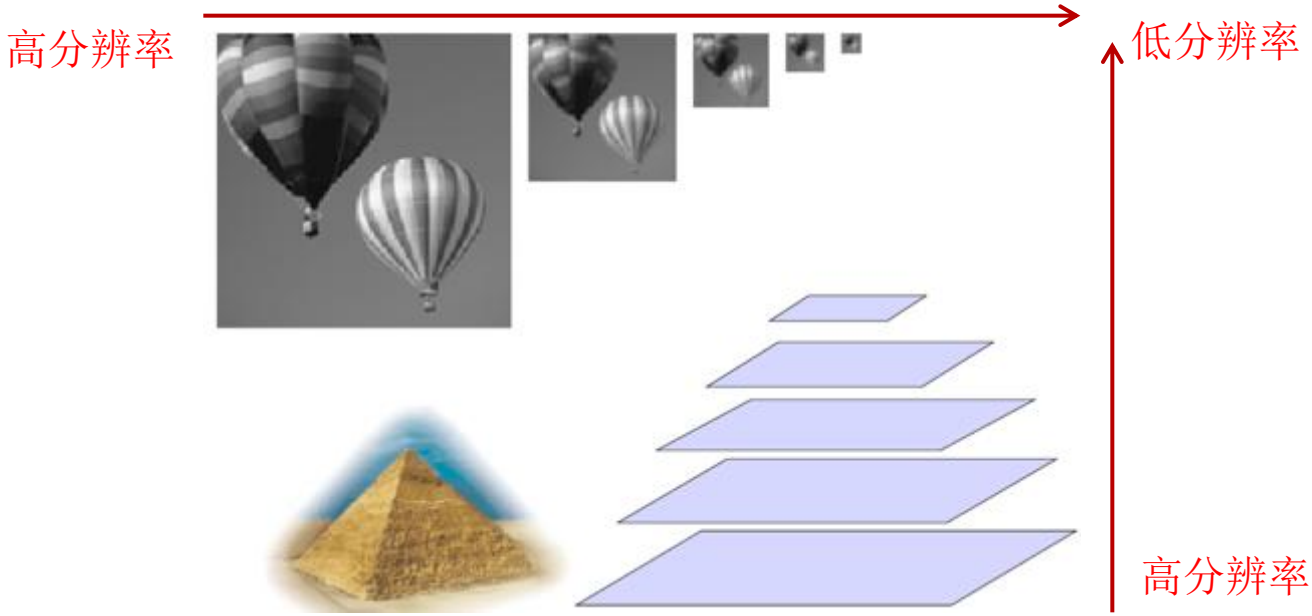
频域分析及变换

如何让卷积更快：空域卷积=频域乘积



高斯金字塔

- 图像金字塔化：先进行图像平滑，再进行降采样，
- 根据降采样率，得到一系列尺寸逐渐减小的图像。
- 操作：n次（高斯卷积→2倍降采样）→n层金字塔
- 目的：捕捉不同尺寸的物体



图像缩放

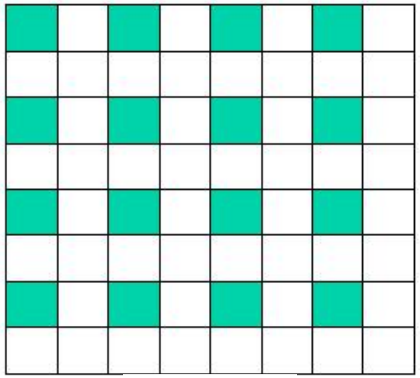
高斯金字塔

- 高斯滤波的必要性：
- 直接降采样损失信息

高斯金字塔本质上为信号的多尺度表示法









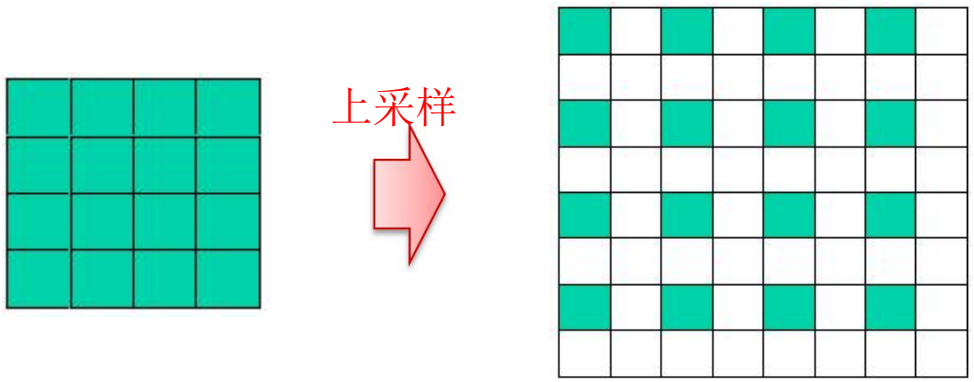
原图　　直接降采样　　高斯+降采样

拉普拉斯金字塔

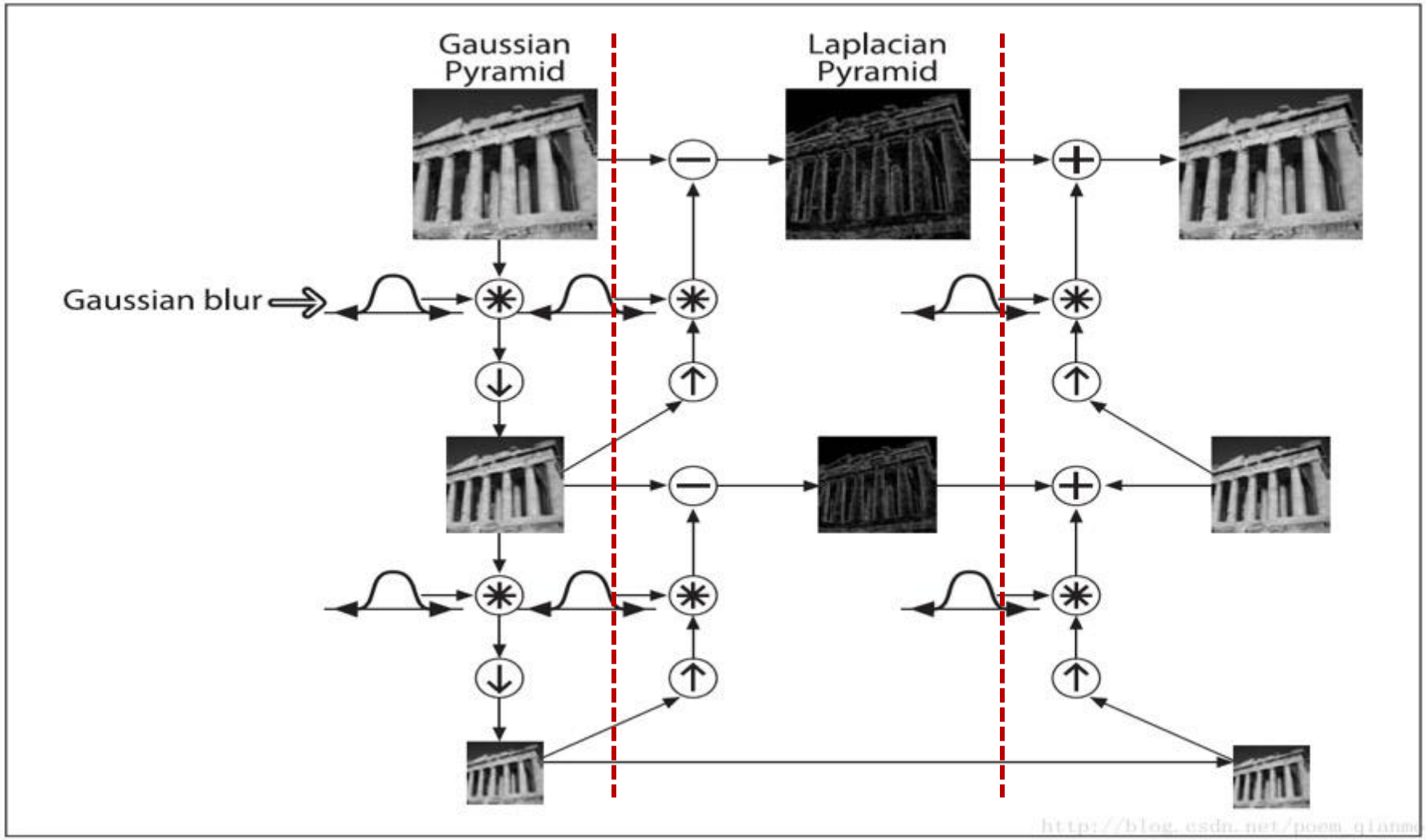
拉普拉斯金字塔（Laplacian）

- 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

$$L_i = G_i - \mathrm{UP}(G_{i+1}) \otimes \mathcal{G}_{5\times5}$$



高斯金字塔与拉普拉斯金字塔



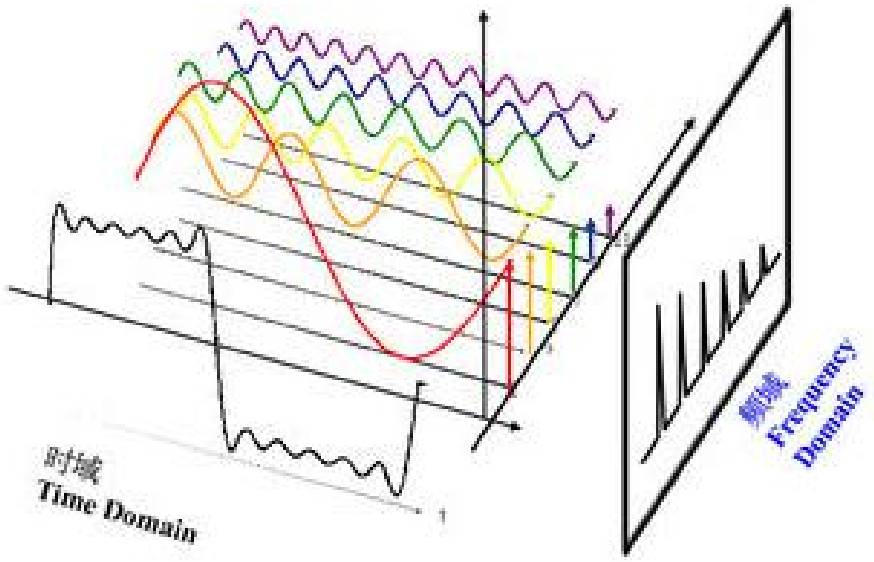
频率域分析——傅里叶变换

傅里叶变换

傅里叶（Baron Jean Baptiste Joseph Fourier，(1768-1830），法国数学家、物理学家。

- 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

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

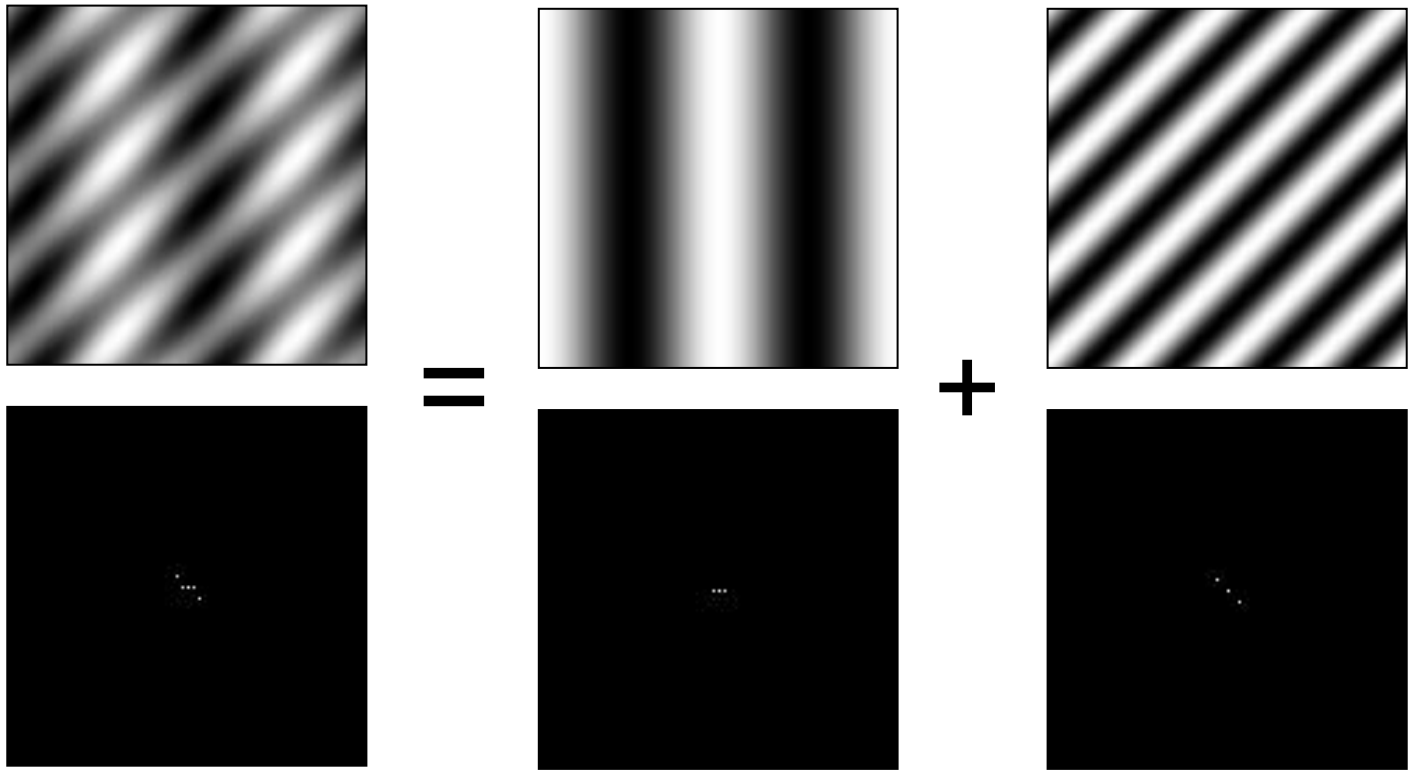




傅里叶变换

傅里叶变换

- 信号分解



傅里叶变换

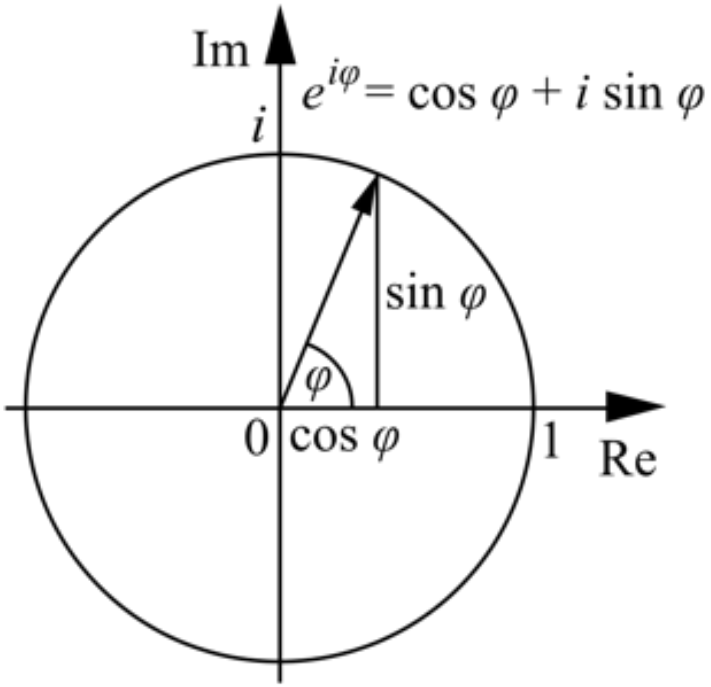
傅里叶变换的数学

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

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \cdots$$



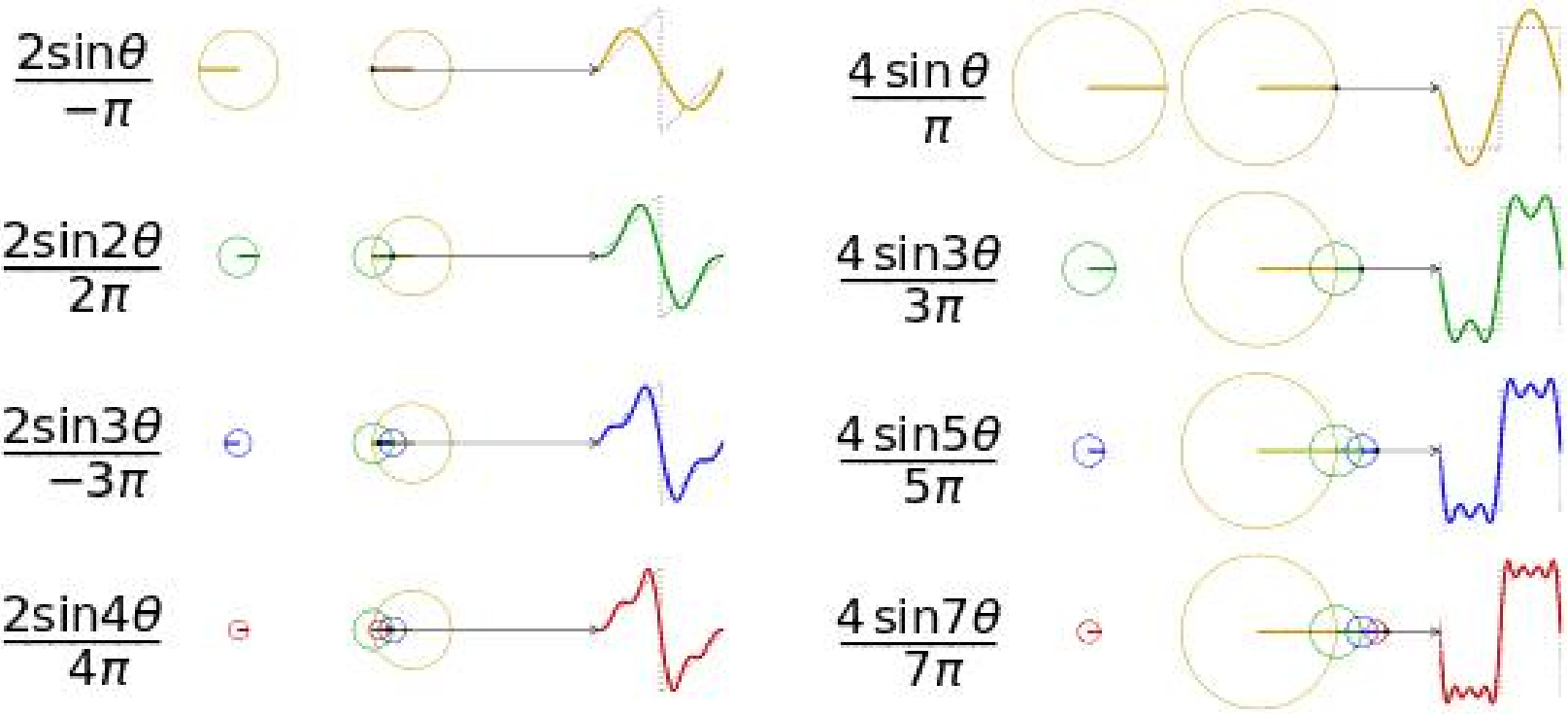
欧拉公式

$e^{ix} = cosx + isinx$

$e^{\pi i} + 1 = 0$

傅里叶变换

欧拉公式描述的是一个随着时间变化，在复平面上做圆周运动的点；傅里叶变换描述的就是一系列这样的点的运动叠加的效应。

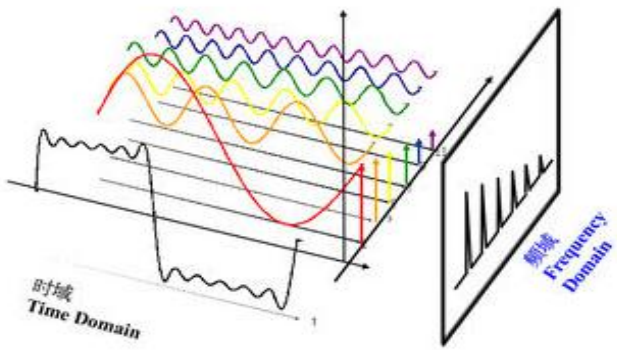


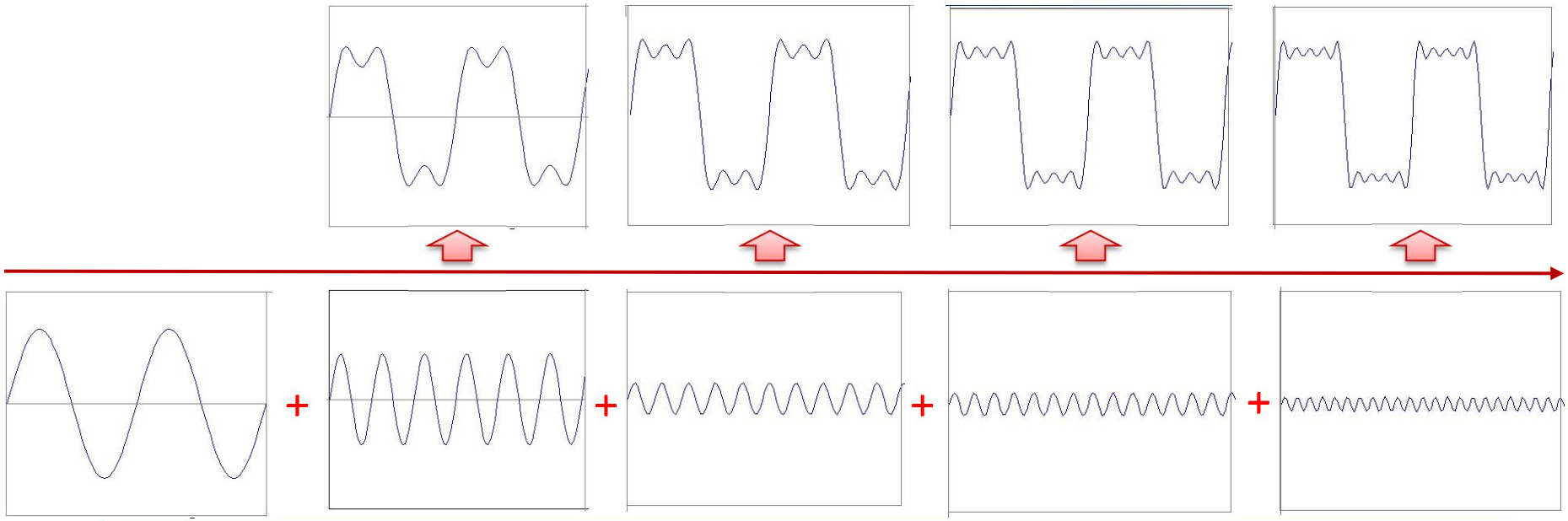
傅里叶变换

再看矩形波：

这里看逆变换就好理解得多！

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{iwt}d\omega$$

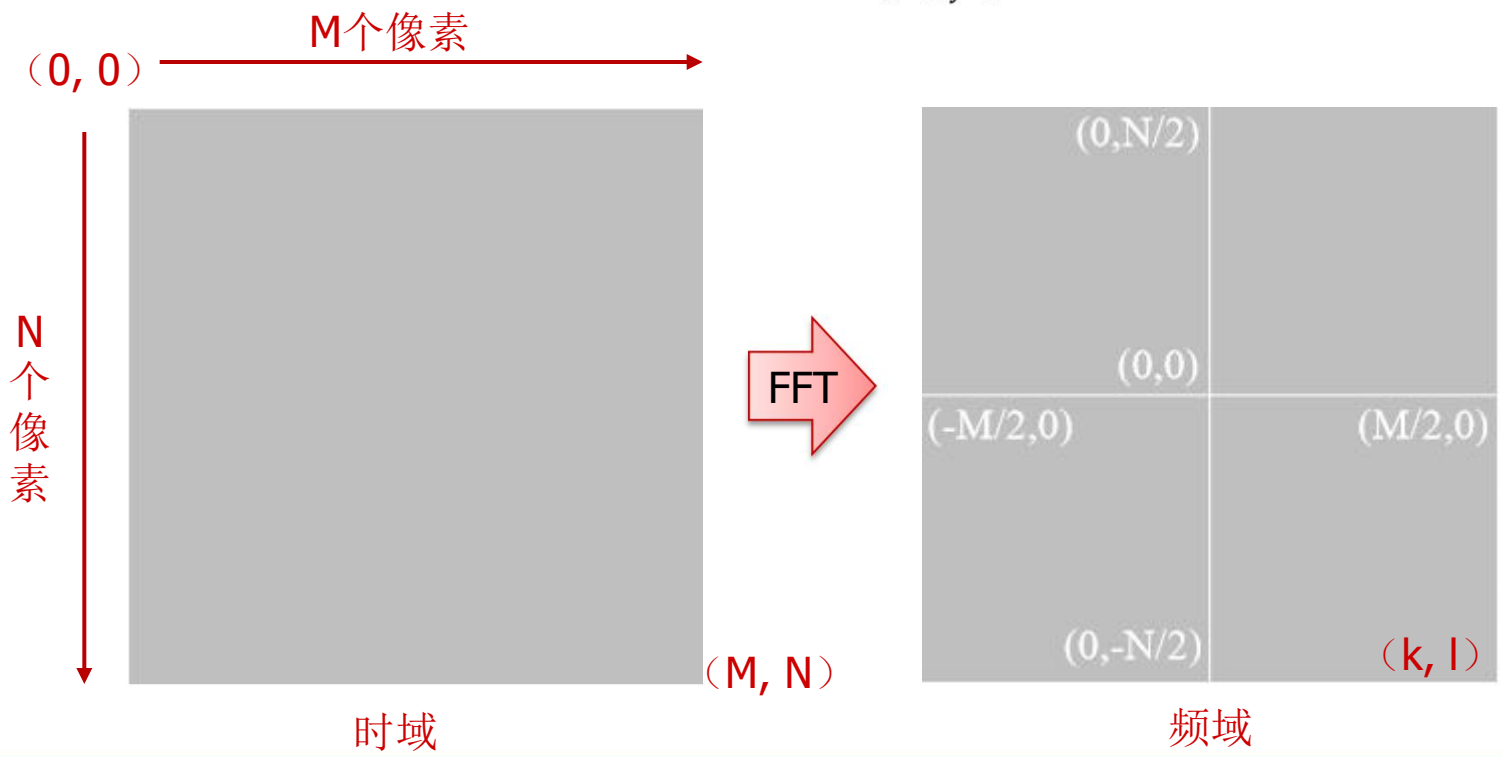




离散傅里叶变换

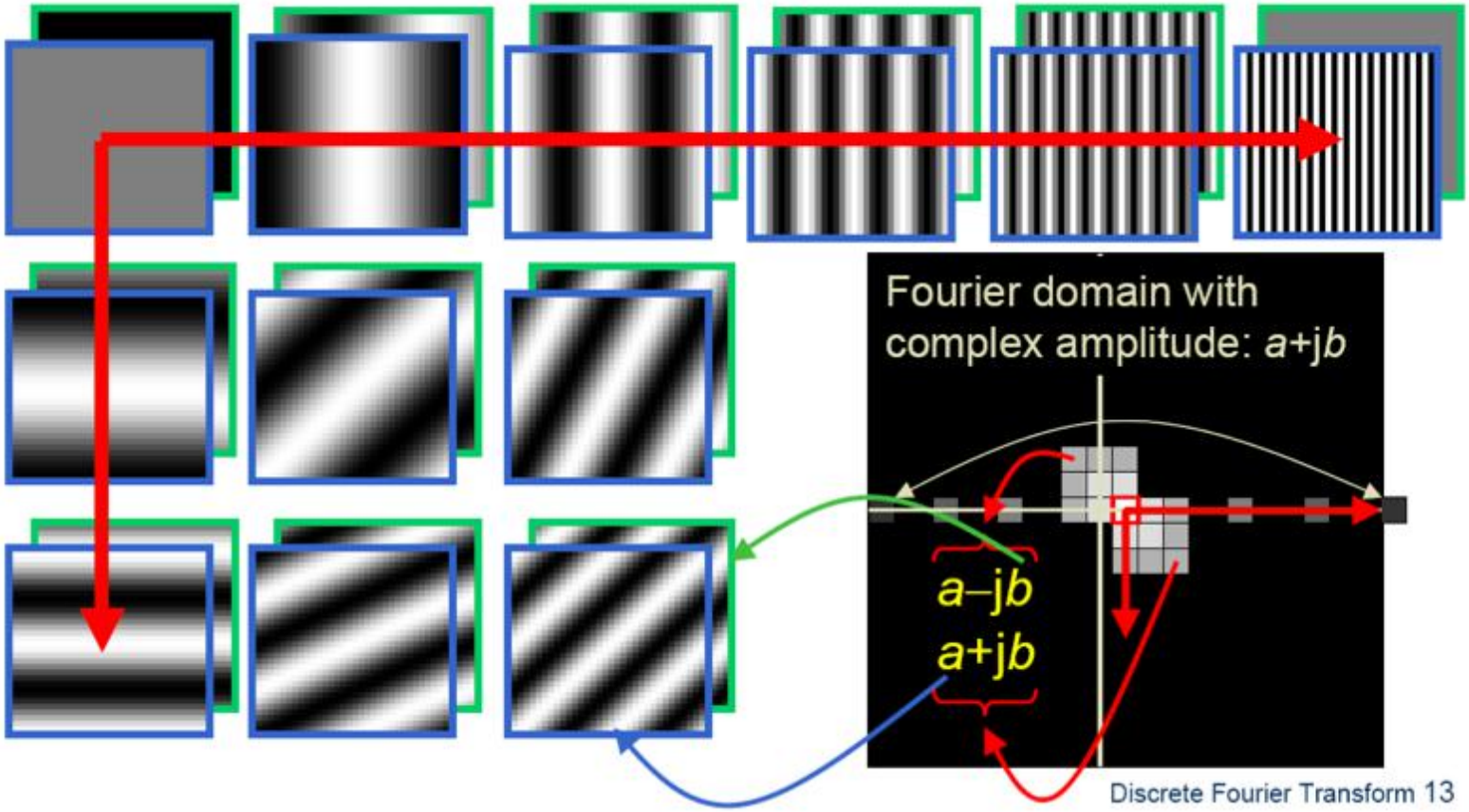
离散傅里叶变换：$X(\omega)=\sum_{n=-\infty}^{\infty}x_n e^{-i\omega n}$

二维离散傅里叶变换：$F(u,v)=\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}f(x,y)e^{-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}$



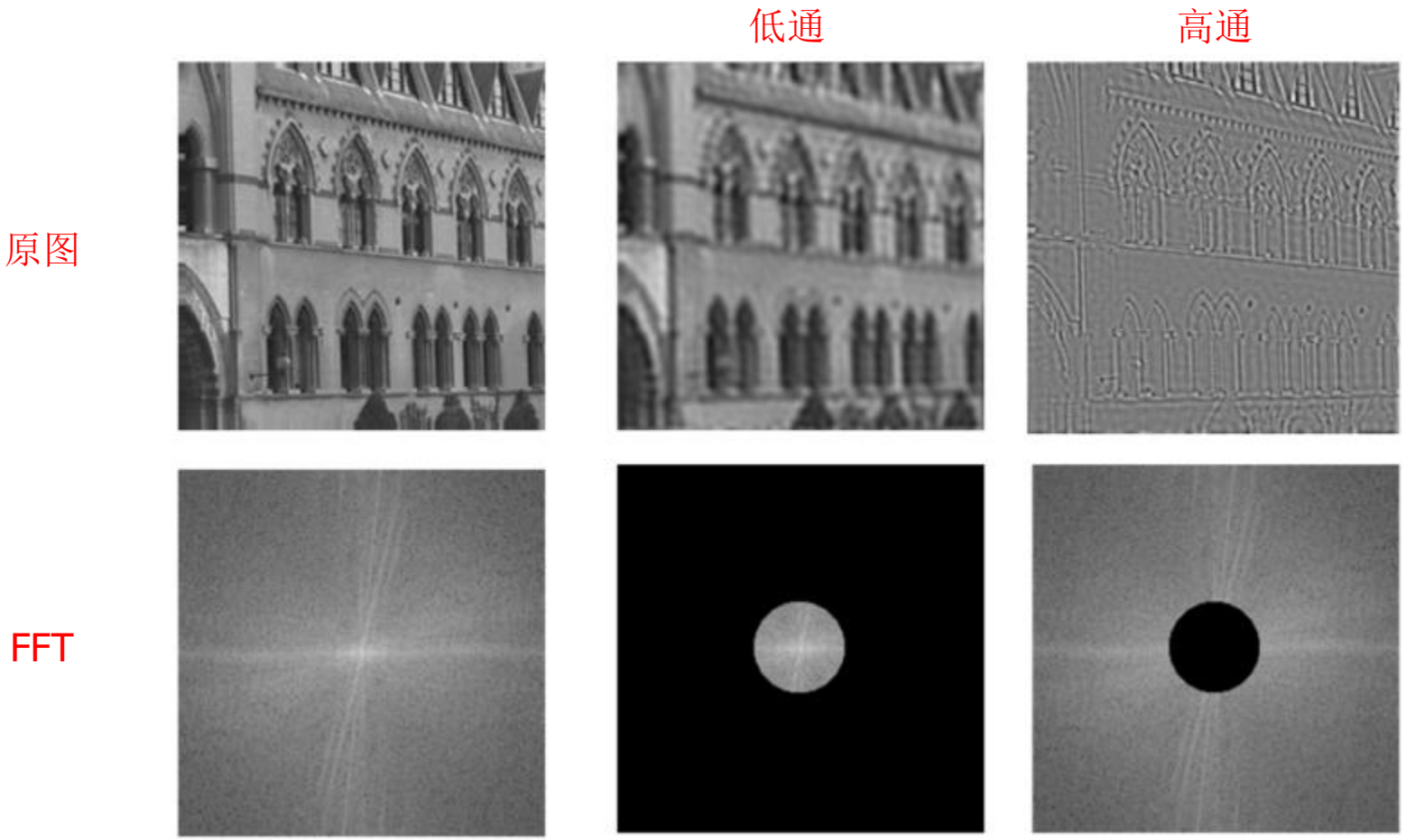
2D离散傅里叶变换

2D傅里叶基图片



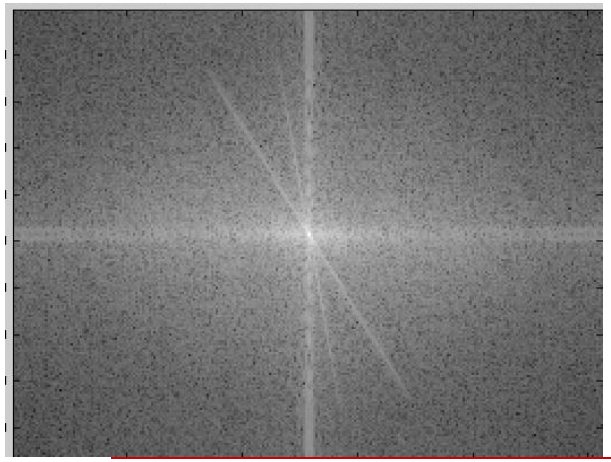
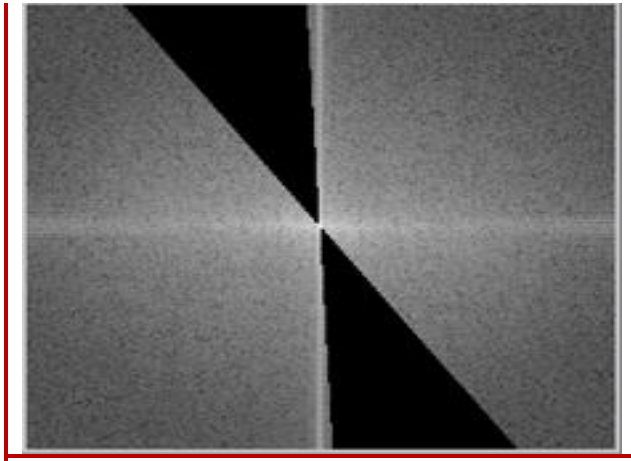
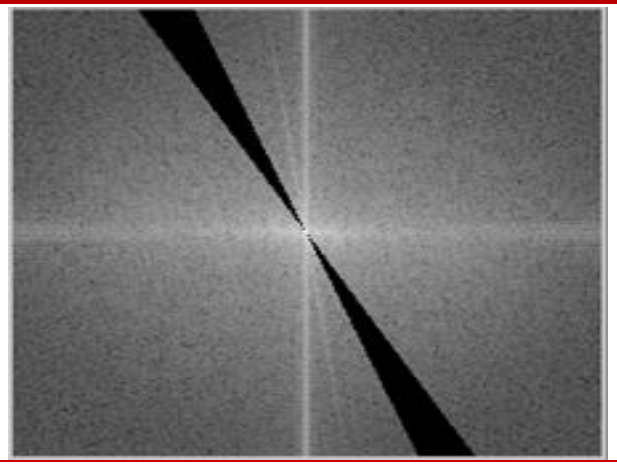
频谱滤波

基于傅里叶变换的滤波：低通滤波



频谱滤波

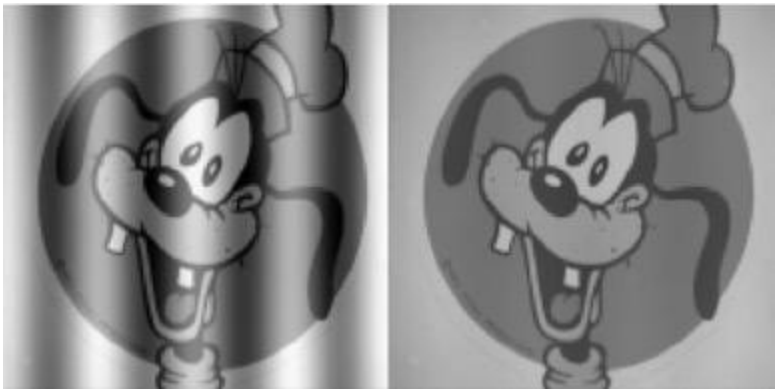
基于傅里叶变换的滤波：相位滤波

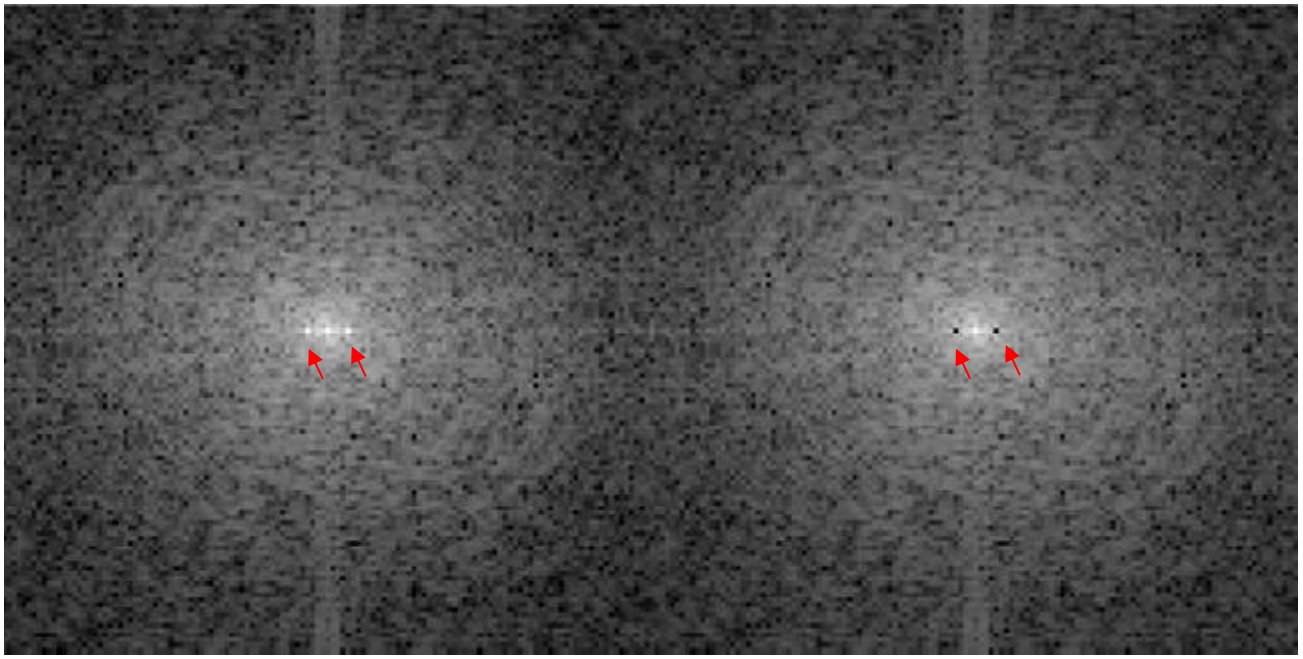







频谱滤波

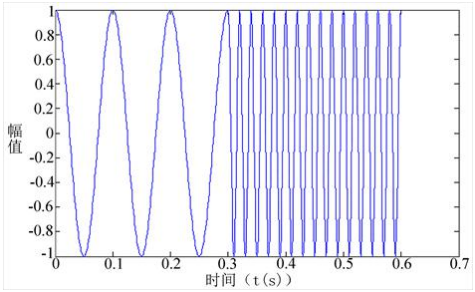
频段滤除

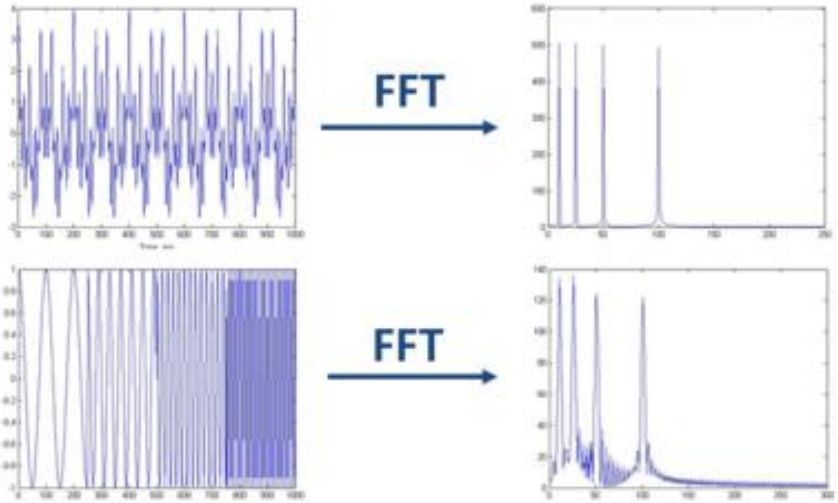




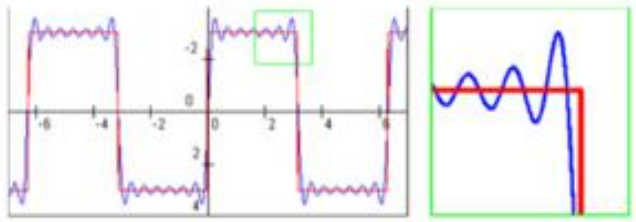
傅里叶变换的不足

- 关键问题——傅里叶变换假设前提为信号平稳，但实际中信号多数为非平稳信号。
  - 缺乏时间和频率的定位功能
  - 对于非平稳信号的局限性
  - 在时间和频率分辨率上的局限性



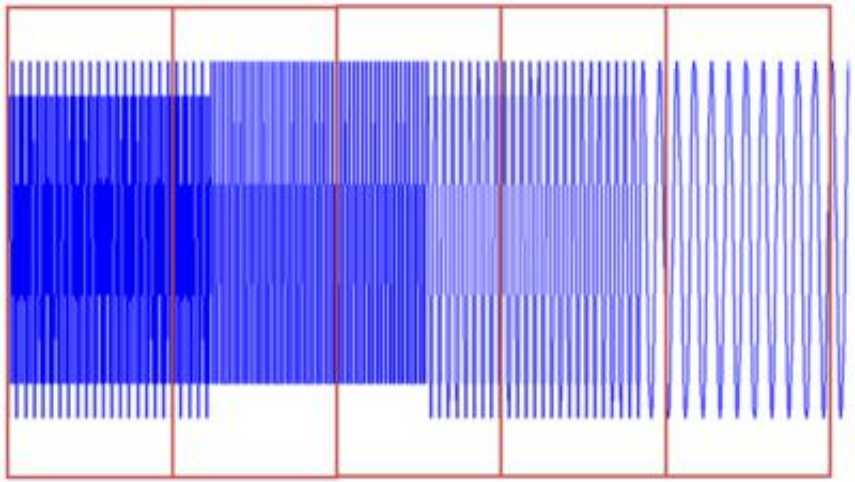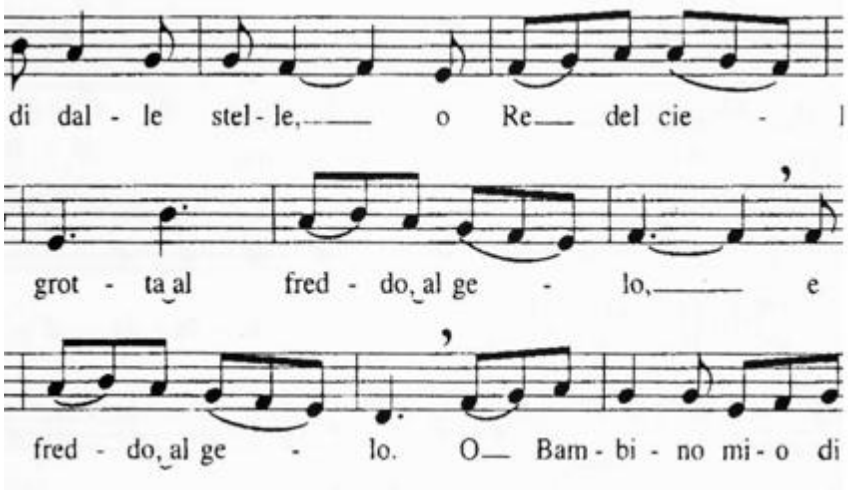


对于突变信号



短时傅里叶变换

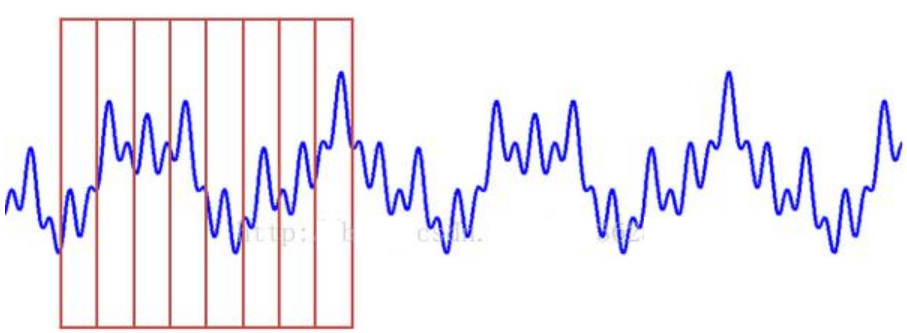
- STFT（短时傅里叶变换）添加时域信息的方法是**设置窗格**，认为窗格内的信号是平稳的；
- 然后对窗格内的信号**分段进行傅里叶变换**。
  - 优点是可以获得频域信息的同时可以获得时域信息。
  - 缺点是窗格大小很难设置。



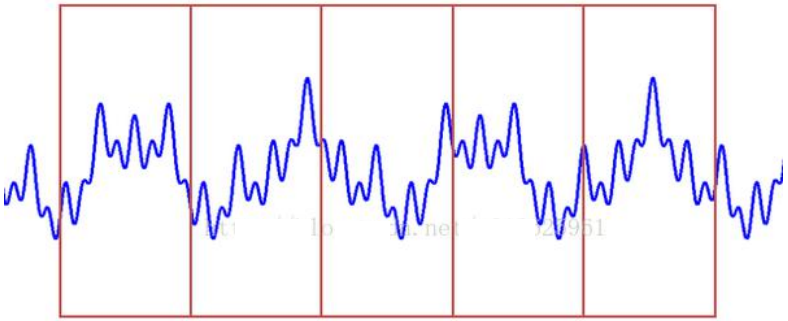


短时傅里叶变换的特点

- 窄窗口时间分辨率高、频率分辨率低；宽窗口时间分辨率低，频率分辨率高。
- 对于时变的非稳态信号，高频适合小窗口，低频适合大窗口。
- 可是STFT的窗口是固定的……



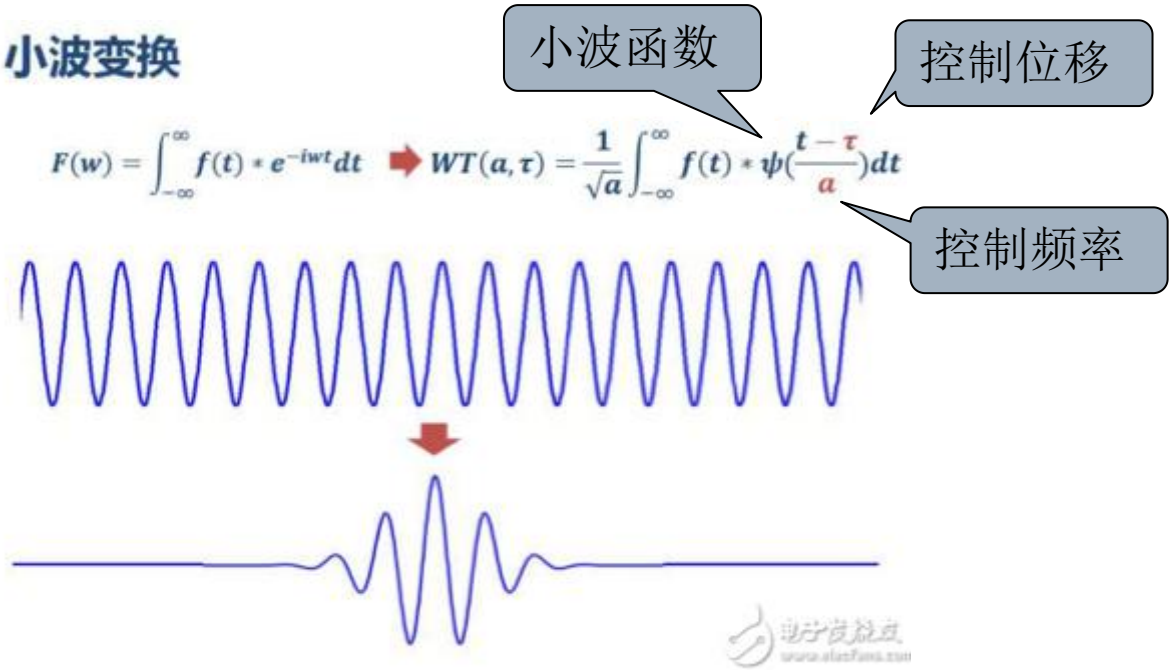
框太窄 → 频率分辨率差



框太宽 → 时间分辨率差

小波变换

- 小波变换与STFT思路接近，但小波变换直接把傅里叶变换的基给换了——将无限长的三角函数基换成了有限长的会衰减的小波基。
- 这样不仅能够获取频率，还可以定位到时间。



频率域分析——小波变换

- 所谓“小波函数”是一族函数，需要满足
  - 1.均值为0；
  - 2.在时域和频域都局部化

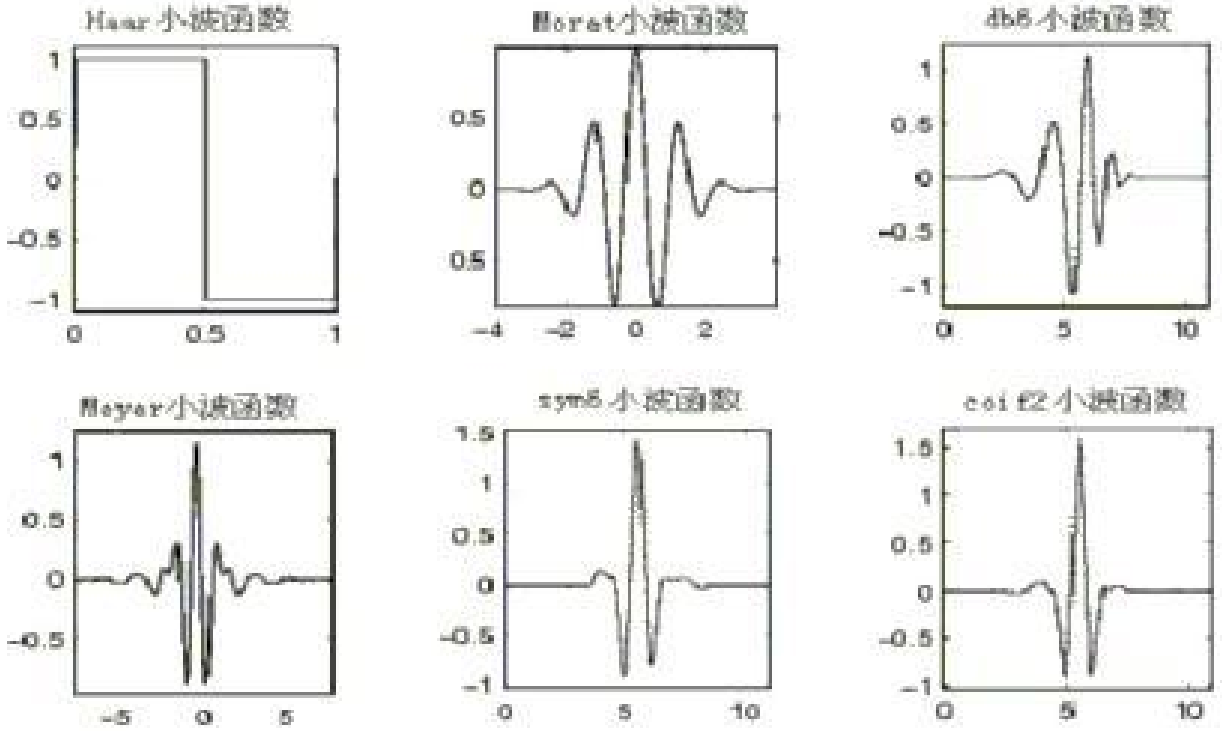
不蔓延到整个坐标轴

- 满足这两条的函数就是小波（Wavelet）函数，有很多种类，最简单的是Haar小波 。

- 小波变换要做的就是将原始信号表示为一组小波基的线性组合，然后通过忽略其中不重要的部分达到数据压缩（即降维）的目的。

常用的小波函数

- 常用的小波函数包括：
  - Haar系列，Daubechies系列，Moret系列
  - Sym系列，Meyer系列，Coif系列。



演示环节

- 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
- 金字塔
  - http://docs.opencv.org/master/dc/dff/tutorial_py_pyramids.html

参考资料

- http://www.360doc.com/content/13/1105/14/10724725_326803150.shtml
- https://blog.csdn.net/xbcreal/article/details/52629465
- https://www.cnblogs.com/jsxyhelu/p/6435601.html?utm_source=debugrun&utm_medium=referral
- https://blog.csdn.net/zhaomengszu/article/details/72628015

问答互动

在所报课的课程页面，

1、点击“全部问题”显示本课程所有学员提问的问题。

2、点击“提问”即可向该课程的老师和助教提问问题。



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：**小象学院**

