

# Review

---

## Chapter 1

---

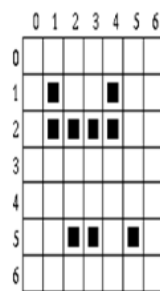
- 成像方式：可见光成像、X射线成像、紫外波段成像、微波波段成像、无线电波段成像
- encoding method 编码方式：无压缩、无损压缩、有损压缩
- bmp：
  - 结构：位图文件头，位图信息头，调色板，位图数据
  - 对齐：每一行的字节数必须是4的整倍数，如果不是，则需要补齐。
  - 存储：bMP文件的数据从下到上，从左到右的。也就是说，从文件中最先读到的是图象最下面一行的左边第一个像素，然后是左边第二个像素.....接下来是倒数第二行左边第一个像素，左边第二个像素.....依次类推，最后得到的是最上面一行的最右一个像素。
- jpeg：压缩比从高频到低频削减。适用互联网
  - 减少高频，压缩比大。
  - 低频保留轮廓色彩
- 景深：镜头前沿能够取得清晰图像的成像所测定的被摄物体前后距离范围突出主题，也是弥散圆间距离。
  - 景深大，整幅画面；景深小，聚焦单个物体
  - 焦距长，景深小
  - 光圈大，景深小
  - 物距大，景深大
- **数码相机成像原理：**
  - (1)当使用数码相机拍
  - (2)当CCD曝光后，光电二极管受到光线的激发而释放出电荷，生成感光元件的电信号。摄景物时，景物反射的光线通过数码相机的镜头透射到CCD上。
  - (3) CCD控制芯片利用感光元件中的控制信号线路对发光二极管产生的电流进行控制，由电流传输电路输出，CCD会将一次成像产生的电信号收集起来，统一输出到放大器。
  - (4)经过放大和滤波后的电信号被传送到ADC，由ADC将电信号（模拟信号）转换为数字信号，数值的大小和电信号的强度与电压的高低成正比，这些数值其实也就是图像的数据。
  - (5)此时这些图像数据还不能直接生成图像，还要输出到DSP（数字信号处理器）中，在DSP中，将会对这些图像数据进行色彩校正、白平衡处理，并编码为数码相机所支持的图像格式、分辨率，然后才会被存储为图像文件。
  - (6)当完成上述步骤后，图像文件就会被保存到存储器上,我们就可以欣赏了。
- 感知颜色
  - priority：色调H > 饱和度S > 亮度V
  - sensitivity：对亮度变化最敏感。与人眼高动态能力匹配。

- 锥形,  $360^\circ H$ , 半径S, 高V
- 视网膜细胞:
  - 杆状: 多, 对光灵敏, 但是不能区分色彩
  - 锥状: 少, 强光才能激发, 用来区分色彩
- 与设备有关的颜色模型:
 

RGB, CMY, HSV
- 与设备无关的颜色模型:
 

CIE XYZ, CIE Lab和 CIE YUV
- RLE压缩编码

一个图像是由一个列表表示, 其中每一行又是一个子列表。子列表的第一个元素表示目前是第几行, 随后的数字成对出现。每对表示一段连续的序列的开头和结尾位置。

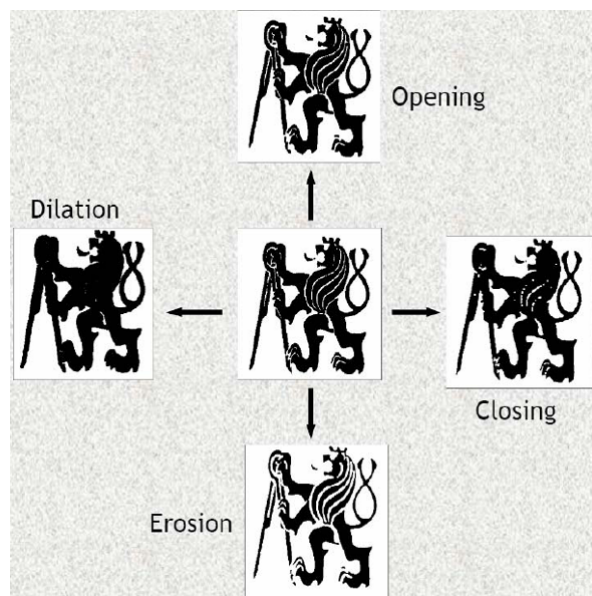


Run length coding; the code is  $\{(11144)(214)(52355)\}$ .

## Chapter 2

### 二值图像

- 大津算法OSTU
  - Step 1: 确定原始图像中像素的最大值和最小值;
  - Step 2: 最小值加1作为threshold对原始图像进行二值化操作;
  - Step 3: 根据对应关系确定前景和背景, 分别计算当前threshold下的内部协方差和外部协方差;
  - Step 4: 回到Step 2直达到像素最大值;
  - Step 5: 找到最大外部和最小内部协方差对应的threshold.
  - 注: 最大外部协方差和最小内部协方差是等价的。
- 光照不均匀 -- 局部自适应操作
  - 设定一个局部窗口, 在整个图像上滑动该窗口sliding window;
  - 对于每一窗口位置, 确定针对该窗口的threshold。
- 腐蚀erosion, 膨胀dilation
- 开运算先e后d, 闭运算先d后e



## Chapter 3

- 可视化增强：对数操作为例
  - - In order to enhance the image's visibility, adjust the pixel value by a logarithmic operator.
    - $L_d = \frac{\log(L_w+1)}{\log(L_{max}+1)}$
    - $L_d$  is display luminance,  $L_w$  is the real luminance,  $L_{max}$  is the maximal luminance value in the image.
    - This mapping function make sure that: no matter the dynamic range of the scene, the maximal luminance value will be 1 (white), and other values changes smoothly.

- 直方图均衡化原理

$$\int_0^r P(r)dr = \int_0^s P(s)ds = \int_0^s 1 \cdot ds = s = T(r)$$

$$s = T(r) = \int_0^r P(r)dr$$

- 为何不能做到绝对平均

所得的sk不可能正好等于8级灰度值中的某一级，因此需要就近归入某一个灰度级中。这样，相邻的多个sk就可能落入同一个灰度级，需要在步骤3时将处于同一个灰度级的像素个数累加。因此，离散灰度直方图均衡化操作以后，每个灰度级处的概率密度（或像素个数）并不完全一样。

- 直方图匹配

所谓直方图匹配，就是修改一幅图像的直方图，使得它与另一幅图像的直方图匹配或具有一种预先规定的函数形状。

直方图匹配的目标，是突出我们感兴趣的灰度范围，使图像质量改善。

- 会考直方图均衡化的过程和直方图匹配的过程，参考ppt即可。

## Chapter 4

- 平移旋转放缩镜像错切
- 旋转后的插值：
  - 行插值：同行前一个元素
  - 最近邻插值：逆映射回原图后距离最近的像素
  - 线性插值，**双线性插值**， $g(x,y) = ax+by+cxy+d$ .
- morph原理：
  - 不同于一般的几何变换warp
  - 从一个关键帧（起始图像）到另一个关键帧（结束图像），自动生成中间帧
  - 图A中每个点像素逐渐变成图B。根据变换快慢，设置步长。可选择等比、等差
  - 需要选择两幅结构相似、大小相同的画面，以保证自然连续
- 表情比例图--解决morph保留不了的表情

假设空间中  $m$  点光源，曲面亮度如下： $I = \rho \sum_{1 \leq i \leq m} S_i I_i n_i \cdot l_i \equiv \rho E(n)$  where

$$S_i = \begin{cases} 0, & \text{cannot be seen} \\ 1, & \text{otherwise} \end{cases}$$

其中  $l_i$  第  $i$  个光源照过来的角度， $n$  表示当前这个曲面的法向量，所以  $n \cdot l_i$  相当于余弦关系； $I_i$  表示第  $i$  个光源的亮度； $S_i$  表示曲面能不能被第  $i$  个点光源照到； $\rho$  表示曲面的材质

由于同一对象前后变换材质相同， $\rho$ 相同，故expression ratio imageERI有  $\frac{I'}{I} = \frac{E(n')}{E(n)}$

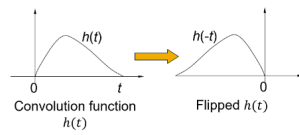
	Neutral	Expression
Person A:	$I_a$	$I'_a$
Person B:	$I_b$	$I'_b = \frac{I'_a}{I_a} I_b$

- 同理转换至表情：输入A, A', B
  - 标记特征点feature points
  - 对每个特征点 $v_b$ ，做warp:  $v_b' = v_b + v_a' - v_a$  得到图像Bg
  - 用Bg对齐A与A'，直到B中每个像素都能找到A与A'上的对应点（用warp方法）
  - 计算比例  $r = A' / A$ 。
  - $B' = B * r$

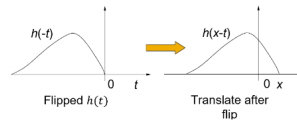
# Chapter 5

已知两个函数，要求计算它们的卷积在任意点 $x$ 处的值。

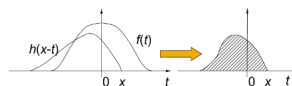
- 将卷积函数  $h(t)$  反折



- 将反折后的卷积函数向右移动距离  $x$



- 计算  $f$  和  $h$  在任意  $t$  的乘积，并积分



- 一维函数卷积：
$$g(x) = f(x) * h(x) = \frac{1}{M} \sum_{t=0}^{M-1} f(t)h(x-t)$$
    - 卷积符合交换律、结合律、分配律。
  - 二维卷积：相当于图像像素加权求和（对应相乘再归一化）
    - 滤波器/掩膜/窗口
    - 平滑空间滤波器spacial filtering for smoothing
      - 效果：**模糊、降低噪声**。掩膜尺寸越大模糊效果越大。
      - 低通滤波
    - 上述滤波器分为：
      - 平滑线性滤波linear smoothing filtering
        - **均值滤波**：计算加权平均值
        - 较好的，和尺寸相关的模糊能力
      - 统计排序滤波statistical sorting filtering
        - **中值滤波**：排序后取中间值为像素值
        - 模糊较均值弱，但可有效去除椒盐（脉冲）噪声（极大/极小白突出点）
      - 高斯滤波gaussian filtering
        - 采用高斯核，按照空间域距离关系乘比例。
- $$G(p) = \frac{1}{2\pi\sigma^2} e^{-\frac{\|p-q\|^2}{2\sigma^2}}$$
- 锐化空间滤波器spacial filtering for sharpening
    - 拉普拉斯变换laplacian operator
    - 效果：**锐化，边缘增强**

$$g(x, y) = \begin{cases} f(x, y) - \nabla^2 f(x, y) & \text{如果拉普拉斯掩模中心系数为负} \\ f(x, y) + \nabla^2 f(x, y) & \text{如果拉普拉斯掩模中心系数为正} \end{cases}$$

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

- 考虑对角线与否，中心系数可正可负
- 边缘保护滤波

#### ■ 双边滤波 Bilateral Filtering

$$BF = \frac{1}{W_q} \sum_{p \in S} G_s(p) G_r(p) * I_p$$

$$= \frac{1}{W_q} \sum_{p \in S} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{\|I_p - I_q\|^2}{2\sigma_r^2}\right) * I_p$$

其中  $W_q$  为滤波窗口内每个像素值的权重和，用于权重的归一化；

$$W_q = \sum_{p \in S} G_s(p) G_r(p) = \sum_{p \in S} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) \exp\left(-\frac{\|I_p - I_q\|^2}{2\sigma_r^2}\right)$$

- q: 原图像素点（对应掩膜中心像素点）；Wq: 该点归一化系数；Gs: space weight空间域权重系数；Gr: intensity weight强度域权重系数；Iq: 该点像素值
- 参数选择：与应用有关； $\sigma_s$ 一般为图像对角线2%（经验参数）； $\sigma_r$ 可取曝光均值/中值等，根据噪声水平调整。
- 参数影响： $\sigma_s$ 越大，模糊效果越大。 $\sigma_r$ 越大，模糊效果越大，且趋于 $\infty$ 时收敛于高斯滤波

#### ■ 导向滤波 Guided Filtering

- 此处略过原理 来不及了
- 与双边滤波比，保边保梯度；非迭代；O(1)速度快；更通用

#### ■ SNF sparse norm filtering 稀疏范数滤波

- 当p = 2时，为均值滤波：求导证明
- 当p = 1时，为中值滤波：贪心证明
- lp范式可改，故可：halo free；去噪；HDR；去抖；引导；分割；分割上色；无缝衔接

### FFT 推导 傅里叶变换消除噪声

不会啊，这是啥啊.....

#### • 相位与幅值

- 幅值表征信号强度
- 相位表征信号结构

- 故而保留相位更重要，以保存图形基本结构

## Chapter 6

$$\begin{aligned}
 E(u, v) &= \sum_{(x, y) \in W} [I(x + u, y + v) - I(x, y)]^2 \\
 &\approx \sum_{(x, y) \in W} [I(x, y) + [I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix} - I(x, y)]^2 \\
 &\approx \sum_{(x, y) \in W} \left[ [I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix} \right]^2
 \end{aligned}$$

This can be rewritten:

$$E(u, v) = \sum_{(x, y) \in W} [u \ v] \underbrace{\begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}}_H \begin{bmatrix} u \\ v \end{bmatrix}$$

Since  $H$  is symmetric, we have  $H = R^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R$

将  $H$  看作椭圆， $R$  决定椭圆的方向，特征值决定椭圆的长轴和短轴，分别表示最慢和最快的变化方向；在角点处，向任何方向移动亮度都要产生较大变化，即  $E(u, v)$  会很大，因此  $E(u, v)$  的最小值也应该变得很大，而  $E(u, v)$  的最小值则是由较小的  $\lambda$  决定，小特征值较大时便很可能是角点。

$$\begin{aligned}
 f &= \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \\
 &= \frac{\text{determinant}(H)}{\text{trace}(H)}
 \end{aligned}$$

(小特征值是如图所示的特征检测方法的一个变种)

**不变性 Invariance:** 旋转和亮度变化（一定范围内）具有不变性；尺寸大小发生改变则不具有不变性。

**Harris-Laplace:** 将原图像缩放成不同的尺度，均进行角点检测，某一点在所有尺度上均检测为角点后，在每个尺度上均计算该角点的拉普拉斯值，选择拉普拉斯的极值所对应的尺度，得到此时的角点坐标与对应尺度，具有一定的尺度不变性。

**SIFT:** 将  $16 \times 16$  的窗口划分成  $4 \times 4$  的单元，计算每个单元的方向直方图，16 个单元格 \* 8 个方向得到 128 维的描述符，因为需要计算取向 Orientation，因此具有旋转不变性。

**SURF 检测:** 计算积分图，算二阶导，非极大值抑制，二次插值

**SURF 描述符:** 将窗口分解为  $4 \times 4$  的单元，计算哈尔小波输出，每个窗口计算

$$v_{\text{subregion}} = \left[ \sum dx, \sum dy, \sum |dx|, \sum |dy| \right] \text{ 得到 64 维描述符。}$$

上图取自：图像信息处理急急急 群友：123 的复习笔记，跪谢大佬

- SIFT 128 特征的得到方法 旋转不变性的原因

旋转不变性：通过各个区块的方向梯度统计确定主方向。

尺度不变性：通过检测空间极值保证尺度不变性

光强不变性：归一化处理

平移不变性：由坐标表征

SIFT 算法通过差分高斯函数检测尺度空间的极值点，关键点的选取是在该点的尺度空间（也就是在不同程度模糊的图像中）都达到极值的情况下进行的

- SURF: 计算积分图快速像素求和。
  - 计算二阶导
  - 二阶导  $H$  矩阵的行列式找局部最大值：每个检测到的关键点都在其周围的尺度和空间邻域内形成一个候选关键点集。然后，这个点的 Hessian 矩阵的行列式值会与其邻居进行比较，只有当此值比所有邻居都大时，这个点才被认为是一个关键点。

- H矩阵行列式：二次多项式，泰勒展开找极值位置，来对关键点的位置和尺度进行调整，得到更精确的关键点位置和尺度，提高匹配的精度。
- 对每一个小方块计算它内部的像素点的梯度方向和强度。将这些信息用一种叫做Haar波的特殊测量方法进行汇总，把每个方块的梯度方向信息和强度信息汇集成一个四元素向量（包括水平和垂直方向的强度，以及它们的绝对值）。对所以所有方块的结果串联起来就形成了最后的特征向量。
- surf 优于 sift:
  - 维度低, 只用 64 维特征, 速度快
  - 在均匀、渐变、只有一条边的图像上 SIFT 无法分辨, 但 SURF 可以
  - 带噪声会使 SIFT 特征凌乱, 对 SURF 几乎没有影响

//

### 拼图：ransac原理与值计算

1. Randomly select a seed group of points on which to base transformation estimate (e.g., a group of matches)
2. Compute transformation from seed group
3. Find inliers to this transformation
4. If the number of inliers is sufficiently large, re-compute least-squares estimate of transformation on all of the inliers
5. Keep the transformation with the largest number of inliers

### ■ How many samples are needed?

- Suppose  $w$  is fraction of inliers (points from line).
- $n$  points needed to define hypothesis (2 for lines)
- $k$  samples chosen.

■ Prob. that a single sample of  $n$  points is correct:  $w^n$

■ Prob. that all  $k$  samples fail is:  $(1 - w^n)^k$

⇒ Choose  $k$  high enough to keep this below desired failure rate.

### pros and cons:

- Pros:
  - General method suited for a wide range of model fitting problems
  - Easy to implement and easy to calculate its failure rate
- Cons:
  - Only handles a moderate percentage of outliers without cost blowing up
  - Many real problems have high rate of outliers (but sometimes selective choice of random subsets can help)



- A voting strategy, the Hough transform, can handle high percentage of outliers

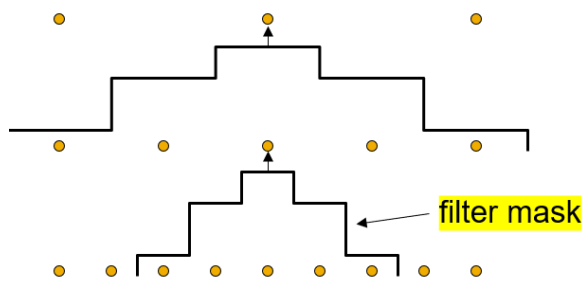
## 拼图总步骤

1. Detect key points
2. Build the SIFT descriptors
3. Match SIFT descriptors
4. Fitting the transformation
5. RANSAC

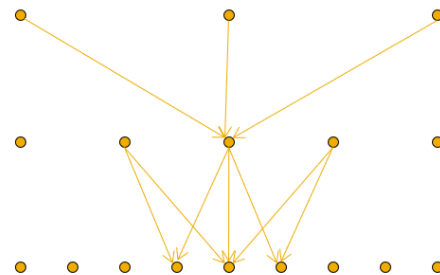
//

image-blending: 高斯/拉普拉斯金字塔

## Pyramid Creation



- “Gaussian” Pyramid
  - “Laplacian” Pyramid
    - Created from Gaussian pyramid by subtraction
- $$L_i = G_i - \text{expand}(G_{i+1})$$



## Chapter 8

人工神经元:

卷积算法

加入卷积核: 减少计算量 数量固定, 全局变局部

pooling池化

卷积过程中尺寸变化 输入输出

## 传统图像分类与深度学习分类区别！

### 交叉熵与交叉熵损失

#### 1. 特征提取：

- 传统图像分类：需要人工设计和选择特征，这可能包括颜色、纹理、形状等特征，特征的选择与设计对分类结果有决定性影响。
- 深度学习：通过神经网络自动从原始输入中学习和提取有效的特征，大大减轻了人工特征工程的负担。

#### 2. 模型构建：

- 传统图像分类：在特征提取后，使用机器学习模型（如SVM、KNN等）进行分类，模型结构相对简单。
- 深度学习：使用深度神经网络，如卷积神经网络（CNN）等进行特征提取与分类，一站式完成，模型结构复杂，能够处理更复杂的图像任务。

#### 3. 性能：

- 传统图像分类：在复杂的图像分类任务上，特别是需要大量数据和高级别抽象特征的任务上，性能往往不如深度学习方法。
- 深度学习：在拥有大量标注数据的情况下，深度学习方法一般能够达到更高的分类精度。