

**著名会议**:IEEE International Conference on Computer Vision (ICCV)、IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)、European Conference on Computer Vision(ECCV)、ACM Multimedia Conference (MM)、IEEE International Conference on Image Processing (ICIP)、IEEE International Conference on Multimedia and Expo (ICME)、International Conference on Pattern Recognition (ICPR)、ACM International Conference on Image and Video Retrieval (CIVR)

**著名期刊**:IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)、IEEE Transactions on Image Processing (TIP)、IEEE Transactions on Circuits and Systems for Video Technology (CSVT)、International Journal of Computer Vision (IJCV)、Pattern Recognition (PR)Image and Vision Computing (IVC)

**国内著名期刊**：、Journal of Computer Science and Technology (JCST)(SCI、CCF-B)、计算机学报、软件学报、计算机研究与发展 、电子学报

**图像的概念**: 图像是对客观存在的一种相似性的生动模仿或描述，是物体的一种不完全、不精确，但在某种意义上是适当的表示。

**图像处理的基本概念**：①将一幅图像变为另一幅经过加工的图像,是图像到图像的过程②将一幅图像转化为一种非图像的表示,如语义信息或决策。

**图像处理难点**: 语义鸿沟、计算能力、数据的获取

**计算机图形学**: 用计算机将由概念所表示的物体 (不是实物) 图像进行处理和显示; 侧重于根据给定的物体描述模型、光照及想象中的摄像机的 成像几何, 生成一幅图像; **图像数字化**: 是指将模拟图像经过离散化之后, 得到用数字表示的图像。图像的数字化包括了空间离散化 (即采样) 和明暗表示数据的离散化 (即量化)。**采样**是指将在空间上连续的图像转换成离散的采样点 (即像素) 集的操作 (在 x 轴和 y 轴两个方向上进行, 一般等间隔)。采样间隔太小, 则增大数据量; 太大, 会发生信息的混叠, 导致细节无法辨认。**分辨率**是指映射到图像平面上的单个像素的景物元素的尺寸; 或者是指要精确测量和再现一定尺寸的图像所必需的像素个数。**量化**是将各个像素所含的明暗信息离散化后, 用数字来表示。一般的量化值为整数 (8bit 量化: 0-255)。(3bit 以下的量化, 会出现伪轮廓现象。) **均匀量化**是简单地在灰度范围内等间隔量化。**非均匀量化**: 对灰度变化比较平缓的部分用比较多的量化级, 在灰度变化比较剧烈的地方用比较高的分辨率。

**图像的质量**: 层次、对比度、亮度、分辨率、细节信息、颜色饱和度

**灰度级**: 表示像素明暗程度的整数量

**层次**: 表示图像实际拥有的灰度级的数量

**对比度**: 是指一幅图像中灰度反差的大小 对比度 = 最大亮度 / 最小亮度

**灰度直方图 图形表示形式**: 横坐标表示灰度级, 纵坐标表示图像中对应某灰度级所出现的像素个数

**数组表示形式**: 数组的下标表示相应的灰度级, 数组的元素表示该灰度级下的像素个数。**性质**: 所有的空间信息全部丢失; 每一灰度级的像素个数可直接得到。**数字化参数**: 直方图给出了一个简单可见的指示, 用来判断一幅图像是否合理的利用了全部被允许的灰度级范围。一幅图像应该利用全部或几乎全部可能的灰度级, 否则等于增加了量化间隔。丢失的信息将不能恢复。**分割阈值选取**: 假设某图像的灰度直方图具有二峰性, 则表明这个图像较亮的区域和较暗的区域可以较好地分离。取二峰间的谷点为阈值点, 可以得到好的二值处理的效果。

**对比度展宽的目的**是, 将人所关心的部分强调出来。**原理**: 进行像素点对点的, 灰度级的映射。**效果**: 新、旧图的灰度级分别为 g 和 f, 要求 g 和 f 均在[0, 255]间变化, 但是 g 的表现效果要优于 f。**公式: 灰度窗**: 只显示指定灰度级范围内的信息。**灰度窗切片**: 只保留感兴趣的部分, 使其为二

$$g(i,j)=\begin{cases} \alpha f(i,j) & 0\leq f(i,j)<a \\ \beta(f(i,j)-a)+g_s & a\leq f(i,j)<b \\ \gamma(f(i,j)-b)+g_s & b\leq f(i,j)<255 \end{cases}$$
  
(i=1,2,...,m; j=1,2,...,n)

值化图像中的白色部分, 其余部分置为 0。  
**幂律变换**: 低灰度值拉伸, 高灰度值压缩, 调整γ进行伽马变换 (图像更接近真实值)。伽马参数降低细节增加对比度降低, 反之亦然。**公式**: s=c<sup>r^γ</sup>  
**动态范围**: 是指图像中从暗到亮的变化范围。**动态范围对人**

**视觉的影响**: 由于人眼所可以分辨的灰度的变化范围是有限的, 所以当动态范围太大时, 很高的亮度值把暗区的信号及所关心部分都掩盖了。**动态范围调整原理**: 通过动态范围的压缩可以将所关心部分的灰度级的变化范围扩大。**公式: 灰度窗公式** (需额外定义映射倍数): **非线性动态范围调整**:

$$h'(i,j)=\begin{cases} 0 & h(i,j)<a \\ \frac{255}{b-a}h(i,j)-\frac{255a}{(b-a)} & h(i,j)\in(a,b) \\ 255 & h(i,j)>=b \end{cases}$$
  
$$g(i,j)=\begin{cases} 0 & 0\leq f(i,j)<a \\ \beta(f(i,j)-a) & a\leq f(i,j)<b \\ 255 & b\leq f(i,j)<255 \end{cases}$$

通常用取对数的方法  
信号的处理是有一个近似对数算子的环节。

**节。公式**: g(i,j)=clg(1+f(i,j))

**直方图均衡化方法: 基本思想**: 对在图像中像素个数多的灰度级进行展览, 而对像素个数少的灰度级进行缩减。从而达到清晰图像的目的。**步骤 1**: 求出图像 f 的总体像素个数 N<sub>t</sub> =长\*宽; **步骤 2**: 计算每个灰度级的像素个数在整个图像中所占的百分比; **步骤 3**: 计算图像各灰度级的累计分布 hp; **步骤 4**: 新图像 g 的灰度值 g(i,j) = 255×hp(k)。

**比较**: 对比度展宽和动态范围调整均需要提前获知感兴趣的灰度范围[a,b]; 在现实生活中, 这个范围因为图像的复杂性较难获取。直方图均衡化不需要提前的假设, 因此在应用中被广泛应用消除光照不均匀等现象。

**图像噪声**: 图像在摄取时或是传输时所受到的随机干扰噪声。常见的有椒盐噪声和高斯噪声。

**特征: 椒盐噪声**: 出现位置随机, 幅值基本相同; **高斯噪声**: 出现位置一定, 幅值随机。

**均值滤波器: 原理**: 在图像上, 对待处理的像素给定一个模板, 该模板包括了其周围的邻近像素。将模板中的全体像素的均值来替代原来的像素值的方法。**均值滤波器的缺点**是, 会使图像变模糊, 原因是它对所有的点都是同等对待, 在将噪声点分摊的同时, 将景物的边界点也分摊了。为了改善效果, 就可采用加权平均的方式来构造滤波器。加权平均滤波器需配平, 权值大说明注意力集中。**中值滤波器**: 如果在某个模板中, 对像素进行由小到大的排列, 最亮或者最暗的点 (噪声) 会被排在两侧。取模板中排在中间位置上的像素的灰度值替代待处理像素的值, 就可以达到滤除噪声的目的

对于椒盐噪声, 中值滤波效果比均值滤波效果好。原因: 1.椒盐噪声是幅值近似相等但随机分布在不同位置上, 图像中有干净点也有污染点。均值将改变干净点的值。2.因为噪声分布的均值不为 0, 所以均值滤波不能很好地去除噪声点。

对于高斯噪声, 均值滤波效果比中值滤波效果好。原因: 1.高斯噪声是幅值近似正态分布, 但分布在每点像素上。2.因为图像中的每点都是污染点, 所以中值滤波选不到合适的干净点。3.因为正态分布的均值为 0, 所以均值滤波可以消除或减弱噪声。

**边界保持类平滑滤波器的提出**: 经过平滑滤波处理之后, 图像就会变得模糊。**原因**: 在图像上的景物之所以可以辨认清楚是因为为目标物之间存在边界。边界点与噪声点有一个共同的特点是, 都具有灰度的跃变特性。所以平滑处理会同时将边界也模糊了。**设计思想**: 自然的想法: 在进行平滑处理时, 首先判别当前像素是否为边界上的点; 如果是, 则不进行平滑处理; 如果不是, 则进行平滑处理。

**K 近邻平滑滤波器: 算法**: 1、以待处理像素为中心, 作一个 m\*m 的作用模板。在模板中, 选择 K 个与待处理像素的灰度差为最小的像素。将这 K 个像素的灰度均值替换掉原来的像素值。**效果分析**: K 近邻滤波器因为有了边界保持的作用, 所以它在去除椒盐以及高斯噪声时, 在图像景物的清晰保持方面的效果非常明显。**代价**: 算法的复杂度增加了。

**Sigma 平滑滤波器: 原理**: 根据统计数学的原理, 属于同一类别的元素的置信区间, 落在均值附近 ±2σ 范围之内。Sigma 滤波器是构造一个模板, 计算模板的标准差σ, 置信区间为当前像素值的 ±2σ 范围。将模板中落在置信范围 (实质将模板内区域划分为两个区域: 边界区与非边界区) 内的像素的均值替换原来的像素值。**核心**: 平滑处理应避免在两个或多个不同区域进行计算。可以采用同类相似的概念进行判别。

**图像锐化: 目的**: 使灰度反差增强, 加强图像中景物的细节边缘和轮廓。因为边缘和轮廓都位于灰度突变的地方, 所以锐化算法的实现基于微分作用。

**单方向一阶锐化: 定义**: 对某个特定方向山轮廓边缘信息进行增强。单方向锐化实际上是包括水平

方向与垂直方向上的锐化。**水平方向的锐化**, 通过可以检测出水平方向上像素值变化的模板来实现。

**垂直锐化算法**的设计思想与水平锐化算法相同。**单方向锐化的后处理**: 解决像素值为负的问题。**方法 1**: 整体加一个正整数, 以保证所有的像素值均为正。可以获得类似浮雕的效果。**方法 2**: 将所有的像素值取绝对值。可以获得对边缘的有方向提取。PS: 有方向一阶锐化对于人工设计制造的具有矩形特征物体 (如楼房、汉字等) 的边 缘的提取很有效。对于不规则形状 (如: 人物) 的边缘提取, 则存在信息的缺损。

**无方向一阶锐化**: 为了解决上述问题, 希望提出对任何方向边缘信息均敏感的锐化算法, 即无方向的锐化算法。1、交叉微分算法 (Roberts 算法) 2、Sobel 锐化 (得到的边缘信息较强, 可用于边缘

$$g(i,j)=|f(i+1,j+1)-f(i,j)|+|f(i+1,j)-f(i,j+1)|$$
 提取) 3、Prewitt 锐化

$$g(i,j)=\{d_x^2(i,j)+d_y^2(i,j)\}^{\frac{1}{2}}$$
 
$$d_x=\begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$
 
$$d_y=\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

Sobel 算法与 Prewitt 算法的思路相同, 属于同一类型, 因此处理效果基本相同。Roberts 算法的模板为 2×2, 提取出的信息较弱。

$$g(i,j)=\{d_x^2(i,j)+d_y^2(i,j)\}^{\frac{1}{2}}$$
 
$$d_x=\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$
 
$$d_y=\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

二阶微分锐化: 从图像景物细节的灰度分布特性可知, 有些图像灰度变化特性的一阶微分描述不是很明确。采用二阶微分能够获得更丰富的景物细节。1、突变形的细节, 通过一阶微分的极大值点, 二阶微分的过零点均可检测到。2、细线形的细节, 通过一阶微分的过零点, 二阶微分的极小值点均可检测出来。3、对于渐变的细节, 一般情况下很难检测, 但二阶微分的信息比一阶微分的信息略多。

**Laplacian 算法**: 
$$H_1=\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$
 **变形算子**: 
$$H_2=\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$
 
$$H_3=\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$
 
$$H_4=\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

H1,H2 的效果基本相同, H3 的效果最不好, H4 最接近原图

**Wallis 算法**: 
$$s=[\log f(i-1,j)+\log f(i+1,j)+\log f(i,j-1)+\log f(i,j+1)]$$
 
$$g(i,j)=\log[f(i,j)]-\frac{1}{4}s$$

为了防止对 0 取对数, 计算时实际上是用 log(f(i,j)+1); 因为对数值很小 log(256)=5.45, 所以计算时用 46×log(f(i,j)+1)。(46=255/Log(256)) **特点**: Wallis 算法考虑了人眼视觉特性, 因此, 与 Laplacian 等其他算法相比, 可以对暗区细节进行比较好的锐化。

**一阶与二阶微分的锐化效果比较**: Sobel 算子获得比较粗略的边界, 反映的边界信息较少, 但是所反映的边界比较清晰; Laplacian 算子获得比较细致的边界。反映的边界信息包括了许多的细节信息, 但是所反映的边界不是太清晰。锐化本质是空间的微分处理, 频域上的高通滤波处理。频率域中高频部分代表着图像的细节部分 (边缘属于其中)。

**边缘检测**: 边缘(edge)是指图像局部强度变化最显著的部分; 边缘主要存在于目标与目标、目标与背景、区域与区域(包括不同色彩)之间; 边缘检测是图像分割、纹理特征和形状特征等图像分析的重要基础。**局部强度变化 (边缘) 分类**: 阶跃不连续, 即图像强度在不连续处的像素灰度值有着显著的差异; 线条不连续, 即图像强度突然从一个值变化到另一个值, 保持一较小行程后又回到原来的值。**术语: 边缘检测器(Edge detector)**: 从图像中抽取边缘集合的算法; **轮廓(Boundary)**: 一条表示边缘轮廓的拟合曲线; **边缘连接(Edge linking)**: 从无序边缘形成有序边缘度的过程。**边缘检测与锐化的关系**: 锐化加强图像边缘, 使得图像看起来轮廓更加明显。边缘检测是检测图像中的边缘 (定位), 可以用于分割等多项应用。二者紧密相关, 因此边缘检测的一阶、二阶算法的基础是锐化的一、二阶算法。**基本步骤: 滤波**: 改善噪声影响; 一般滤波器(均值滤波器)导致边缘的损失; 增强边缘和降低噪声之间需要折衷。**增强(锐化)**: 将邻域强度值有显著变化的点突出出来。**检测**: 最简单的边缘检测判据是梯度幅值阈值。**定位**: 估计边缘的位置和方位。

**一阶边缘检测算子**: 1、Roberts 算子  $G[i,j]=|f[i,j]-f[i+1,j+1]|+|f[i+1,j]-f[i,j+1]|$  2、Sobel 算子 3、Prewitt 算子 (与 Sobel 算子的方程完全一样, 但 c=1)

$$\mathcal{M}=\sqrt{s_x^2+s_y^2}$$
 
$$s_x=\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$
 
$$s_y=\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$
 
$$s_x=\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$
 
$$s_y=\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

**二阶微分算子: 拉普拉斯算子**:

**LoG (Laplacian of Gaussian) 算法: 特征**: 平滑滤波器是高斯滤波器。增强步骤采用二阶导数(二维拉普拉斯函数)。边缘检测判据是二阶导数零交叉点。

$$DoG\triangleq G_{\sigma_1}-G_{\sigma_2}=\frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sigma_1}e^{-(x^2+y^2)/2\sigma_1^2}-\frac{1}{\sigma_2}e^{-(x^2+y^2)/2\sigma_2^2}\right)$$

**Canny 算子**: 使用高斯函数的一阶导数, 是对信噪比与定位之乘积的最优化逼近算子。**步骤**: 1) 用**高斯滤波器**平滑图像; 2) 用**一阶偏导**有限差分计算梯度幅值和方向; 3) 对梯度幅值应用**非极大值抑制**; 4) 用**双阈值算法检测**和**连接边缘**。

**图像的几何变换**: 图像的几何变换包括了图像的形状变换、图像的位置变换以及仿射变换。图像的形状变换是指图像的放大、缩小与错切。图像的位置变换是指图像的平移、镜像与旋转。图像的仿射变换则使用数学映射变换公式, 表示几何变换。

**图像缩小**: 设原图像大小为 M\*N, 缩小为 k<sub>1</sub>M\*k<sub>2</sub>N, (k<sub>1</sub><1, k<sub>2</sub><1)。算法步骤如下: 1) 设原图为 F(i,j), i=1,2,...,M, j=1,2,...,N。压缩后图像是 G(x,y), x=1,2,...,k<sub>1</sub>M,y=1,2,...,k<sub>2</sub>N。2) G(x,y)=F(c<sub>1</sub>\*x,c<sub>2</sub>\*y), c<sub>1</sub>=1/k<sub>1</sub>、c<sub>2</sub>=1/k<sub>2</sub>。

**图像放大**: 设原图像大小为 M\*N,放大为 k<sub>1</sub>M\*k<sub>2</sub>N, (k<sub>1</sub>>1, k<sub>2</sub>>1)。算法步骤如下: 1) 设旧图像是 F(i,j), i=1,2,...,M, j=1,2,...,N。新图像是 G(x,y), x=1,2,...,k<sub>1</sub>M,y=1,2,...,k<sub>2</sub>N。2) G(x,y)=F(c<sub>1</sub>\*i,c<sub>2</sub>\*j), c<sub>1</sub>=1/k<sub>1</sub>、c<sub>2</sub>=1/k<sub>2</sub>。

**基于双线性插值的图像放大方法: 顶点填充**: SrcX=(dstX+0.5)\*(srcWidth/dstWidth)-0.5 SrcY=(dstY+0.5)\*(srcHeight/dstHeight)-0.5。**插值: x: y:**

$$f(R_1)\approx\frac{x_2-x}{x_2-x_1}f(Q_{11})+\frac{x-x_1}{x_2-x_1}f(Q_{21})$$
 Where  $R_1=(x,y_1)$   $f(P)\approx\frac{y_2-y}{y_2-y_1}f(R_1)+\frac{y-y_1}{y_2-y_1}f(R_2)$

$$f(R_2)\approx\frac{x_2-x}{x_2-x_1}f(Q_{12})+\frac{x-x_1}{x_2-x_1}f(Q_{22})$$
 Where  $R_2=(x,y_2)$

**错切**: 
$$\begin{cases} x'=x+d_xy \\ y'=y \end{cases}$$
 (x方向的错切)  $d_x=\tan\theta$  
$$\begin{cases} x'=x \\ y'=y+d_yx \end{cases}$$
 (y方向的错切)  $d_y=\tan\theta$

**镜像**: 
$$\begin{cases} x'=x+\Delta x \\ y'=y+\Delta y \end{cases}$$
 
$$\begin{cases} x'=x \\ y'=-y \end{cases}$$
 (水平镜像) 
$$\begin{cases} x''=x'=x \\ y''=y'+N+1=N+1-y \end{cases}$$
 (坐标平移)

大画布, 计算出 x 和 y' 的最大、最小值, 即 x'min、x'max 和 y'min, y'max。画布大小为: x'max - x'min、y'max -y'min。**旋转后插值**: 找出当前行的最小和最大的非背景点的坐标, 记作: (i,k1)、(i,k2)。在(k1,k2)范围内进行插值, 插值的方法是: 空点的像素值等于前一点的像素值。同样的操作重复到所有行。

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$
 **水平错切**: 
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} 1 & d_x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$
 **垂直错切**: 
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & d_y \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$
 **水平镜像**: 
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$
 **垂直镜像**: 
$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

**图像二值化处理目的**: 图像二值化处理是指通过某种方法, 使得画面场景被分为“目标物”及“非目标物”两类,即将图像的像素变换为黑、白两种。图像的二值化处理属于数字图像分割范畴。**p-参数法**: 对固定分辨率下的目标物, 根据目标物在画面中所占的比例来选择阈值, 进行二值化处理。**步骤**: 1、设图像的大小为 m\*n, 计算得到原图的灰度直方图 h。2、输入目标物所占画面的比



例 p。3、尝试性地给定一个阈值 Th=Th<sub>0</sub>。4、计算在 Th 下判定的目标物的像素点数 N<sub>5</sub>。5、判断 p<sub>5</sub>=N/(m\*n)是否接近 p？ 是，则输出结果； 否则，Th=Th+dT; (if ps<p, 则 dT>0; else dT<0), 转4) 直到满足条件。

**类间类内最大方差比阈值法：思想**：“物以类聚”属于“同一类别”的对象具有较大的一致性。**手段**：以均值与方差作为度量一致性的数字指标。**步骤**：1、给定一个初始阈值 Th=Th<sub>0</sub>。(例如：可以默认为1, 或者是 128 等)，则将原图分为 C1 和 C2 两类、2、分别计算两类的类内方差。分别计算两类像素在图像中的分布概率：选择最佳阈值 Th=Th\*，使得下式成立

$$\sigma_0^2 = P_1(\mu_1 - \mu)^2 + P_2(\mu_2 - \mu)^2 \quad \sigma_m^2 = P_1\sigma_1^2 + P_2\sigma_2^2 \quad \eta|_{Th^*} = \max \left\{ \frac{\sigma_b^2}{\sigma_m^2} \right\} \quad (\text{遍历 } 1-254)$$

**聚类处理**：步骤 12 同上， 步骤 3、进行分类处理：如果{|(x,y0-u1|<{|(x,y1-u2|, 则 f(x,y)属于 C1, 否则 f(x,y)属于 C2。4、对上一步重新分类后得到的 C1 和 C2 中的所有像素，分别重新计算其各自的均值与方差。5、如果下式成立，则输出计算得到的阈值 *Th(t-1)*，算法结束； 否则重转 4)

$$[p_1\sigma_1^2 + p_2\sigma_2^2]_{|Th(t-1)} \leq [p_1\sigma_1^2 + p_2\sigma_2^2]_{|Th(t)}$$

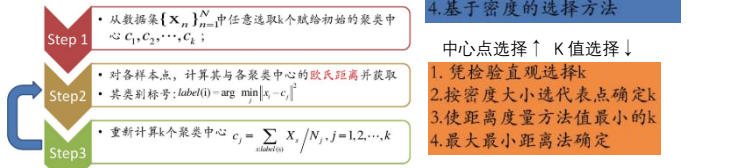
**K-means 聚类算法**：基于**距离**的聚类算法(clustering algorithm), 属于经典的无监督机器学习算法。有监督机器学习算法 (SVM)、半监督机器学习算法(Transductive SVM 转导支持向量机)、无监督机器学习算法(聚类算法)

**K-means 聚类算法的基本思想**：

- 一、指定需要划分的簇的个数k值；
- 二、随机地选择k个初始数据对象点作为初始的聚类中心；
- 三、计算其余的各个数据对象到该k个初始聚类中心的距离，把数据对象划归到距离它最近的那个中心所处的簇类中；
- 四、调整新类并重新计算出新类的中心。
- 五、计算聚类准则函数E，若E不满足收敛条件。重复二、三、四，六、结束

**伪代码**

```
1 选择 K 个点作为初始质心
2 repeat
3     将每个点指派到最近的质心，形成 K 个簇
4     重新计算每个簇的质心
5 until 簇不发生变化或达到最大迭代次数
```



**阈值法：聚类法**：见上

**边界法：原理**：将待分割的区域看做由一个闭合边界包围的像素集合，因此边界法实现分割的前提是边界的提取

**Hough 变换**：将图像由图像空间变换为参数空间。1) 图像空间中的一条线（点）对应 Hough 空间中的一个点（线）。2) Hough 空间中两条线的交点用来表示过点(x0,y0)和点(x1,y1)的直线。

**Hough 变换（极坐标系）**：图像空间中的每个点(x,y)被映射为一个(r,θ)空间中的正弦曲线。2) 图像空间中共线的点所对应的(r,θ)空间中正弦曲线 相交于一点(r',θ')。PS：把在图像空间中检测直线的 问题转化为在极坐标参数空间中找到过点(r,θ)的最多正弦曲线数的问题。 图像空间中圆上的点映射到参数空间中一族圆锥的交点(a0, b0, r)正好对应于圆的圆心坐标和圆的半径。 在参数空间中的交点就代表了图像空间中的某个圆。

**直线 Hough 检测步骤**：1) 构建（参数空间）变换域累加器数组，并将其初始化为 0；2) 读入一幅二值化图像，遍历图像像素点；3) 对每一个 像素点，进行霍夫变换，按照 r 和 θ 的值在变换域累加器数组中的相应位置上加 1；4) 遍历累加器数组，寻找局部极大值。

**圆 Hough 检测步骤**：1) 基于以上原理对参数空间适当量化，得到一个三维的累加器阵列，阵列中的每一个立方小格对应的参数离散值。2) 对图 像空间所存在的圆进行边缘检测，将每个边缘点映射到 abr 三维空间，同时将对应立方小格的累加器累加。3) 当对全部边缘点变换完成后，查找所有累加器的最大值，其对应坐标就是图像空间中圆的圆心和半径。

**区域法：思路**：综合考虑待分割区域内部像素的特征**同质性**和**近邻性**，通过反复地将**相邻且特征相似**的像素或区域纳入同一分割区域，以实现分割。**区域生长法：基本步骤**：1、通过人工指定或设计类似的初始化算法，寻找一组**生长点**（单个像素或某小区域）、2、搜索生长点的领域，比较领域像素与生长点像素的**特征相似性**，若满足合并条件，则合并为同一区域，形成新的生长点 3、重复步骤 2 直到没有新的生长点出现。**关键问题**：初始化：选择一组能正确代表所需区域的生长点像素。生长过程：选择有意义的特征，用于相似性判断。停止条件：确定相似性准则，即获取生长过程停止的准则。

**图论法：基本思想**：将数字图像视为具有拓扑结构（常为矩形）的图(graph)，将像素视为点(vertex)，将相邻像素点间的联系视为边(edge)，定义某种基于类别标记的**目标函数**，通过对目标函数的优化实现分割 **图论分割算法**：最小割算法、马尔可夫随机场、生成树算法

**二值图像处理**：图像二值化处理是指通过某种方法，使得画面场景被分为“**目标物**”及“**非目标物**”两类。二值图像处理在二值图像基础上进行计算。目的：1.区分所提取出的不同目标物。2对不同的目标物特征差异进行描述与计算。**连接：四连接、八连接、连通域。**

**贴标检测步骤**：1、初始化：设标号为 L=0、已贴标像素 N=0，标签矩阵 g 为全 0 阵。2、检查相邻像素的状态：如果扫描过的像素均为 0，则 L=L+1, g(i)=L, N=N+1, 如果扫描过的像素标签号相同，则 g(i)=L, 如果扫描过的像素标签号不相同，例如 L>L<sub>i</sub>, 则 g(i)=L<sub>i</sub>, N=N-1, 修改所有为 L<sub>i</sub> 的像素值，使之变为 L<sub>i</sub>。3、将全部的像素进行步骤 2 的处理，直到所有的像素全部处理完成。4、判断最终的 L 是否满足 L=N, 如果是，则贴标检测完成，如果不是，则表明已贴标存在不连号情况，将进行一次编码整理，消除不连续编号的情况。

**形态学操作：基本思想**：用具有一定形态的**结构元素**找到图像中的对应形状以达到图像分割识别的目的，基本的操作有**膨胀、腐蚀、开和闭运算**。

**腐蚀**：消除连通域的边界点，使边界向**内收缩**的处理。设计一个**结构元素**，结构元素的原点定位在待处理的目标像素上，通过判断是否覆盖，来确定是否该点被腐蚀掉。（全覆盖才不腐蚀）

**膨胀**：将与目标区域的背景点合并到该目标物中，使目标物边界向**外部扩张**的处理。设计一个结构元素，结构元素的原点定位在背景像素上，判断是否覆盖有目标点，来确定是否该点被膨胀为目标点。（有黑点就膨胀）

**开运算**：先腐蚀后膨胀。开运算可以在分离粘连目标物的同时，基本保持原目标物的大小。

**闭运算**：先膨胀后腐蚀。闭运算可以在合并断裂目标物的同时，基本保持原目标物的大小。

**色彩的描述**：通过建立色彩模型来实现的，不同的色彩模型对应于不同的处理目的。**RGB 色系**：CIE 规定了以 700nm(红)、546.1nm(绿)、435.8nm(蓝)三个色光为三基色。自然界的所有颜色都可以过选用这三基色按不同比例混合而成。

**HSI 色系**：RGB 色系颜色的构成与人对颜色的理解方式不同，所以在进行处理与调整时，比较不容易获得准确的参数。HSI 彩色系统格式的设计反映了人类观察色彩的方式。I 表示**光照强度**或称为**亮度**，它确定了像素的整体亮度，而不管其颜色是什么。**H 表示色度**，由角度表示，反映了该颜色最近的光谱波长。0°为红色，120°为绿色，240°为蓝色。**S 表示饱和度**，饱和度参数是色环的原点到彩色点的半径长度。在环的外围圆周是纯的或称饱和的颜色，其饱和度值为 1。在中心是中性（灰）色，即饱和度为 0。**CMYK**：印刷用减色系统。

**RGB 到 HSI 的转换**：
$$I = \frac{1}{3}(R + G + B) \quad S = 1 - \frac{3\min(R,G,B)}{R+G+B} \quad H = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases}$$
**HSI 到 RGB 的转换**：
$$R = \frac{1}{\sqrt{3}}[1 + \frac{S \cos(H)}{\cos(60^\circ - H)}] \quad B = \frac{1}{\sqrt{3}}(1 - S) \quad G = \sqrt{3}I - R - B$$
**H 0-120 120-240 240-300**

$$G = \frac{1}{\sqrt{3}}[1 + \frac{S \cos(H - 120^\circ)}{\cos(180^\circ - H)}] \quad R = \frac{1}{\sqrt{3}}(1 - S) \quad B = \sqrt{3}I - R - G$$
**RGB-CMYK**
$$B = \frac{1}{\sqrt{3}}[1 + \frac{S \cos(H - 240^\circ)}{\cos(300^\circ - H)}] \quad G = \frac{1}{\sqrt{3}}(1 - S) \quad R = \sqrt{3}I - G - B$$
**YUV 表色系统**：**Y：亮度；U，V：色差信号。**
$$\begin{cases} R = Y + V \\ U = B - Y \\ V = R - Y \end{cases} \quad \begin{cases} R = Y - 0.192U - 0.509V \\ B = Y + U \end{cases}$$
**目的**：为了可以使电视节目同时被黑白电视及彩色电视接收。**白平衡算法步骤**：1、计算输入的具有色偏的原图亮度。2、根据计算出的亮度值来寻找图像中的白色点。考虑到实际中，白色的点不一定是理想状态下的白点，因此在这里只是将白色定义为亮度值为最大的点。3、考虑到对环境光照具有一定的适应性，寻找出原图中所有亮度值不小于 0.95 倍最大亮度值的点。令这些点构成白色点集合。4、计算白色点集 Q<sub>white</sub> 中所有像素的 R,G,B 三个颜色分量的均值。5、按照下面的公式计算颜色均衡化的调整参数。K<sub>x</sub>=Y/X6、对整幅图像的 R,G,B 三个颜色分量，进行彩色平衡调整如下。X'=K<sub>x</sub>X。56 步中 x 为 RGB。

**频域分析**反映了信号不同频率组成及其分量成分大小，能够提供比时域信号波形更直观，丰富的信息。**图像中频率代表什么内容？**低频：图像基本外形轮廓特征、总体灰度信息；中低频：图像分块和区域特征；中高频：图像边缘、线条、纹理等细节信息；高频：部分图像高频噪声(白噪声除外)。(变换后的图像，大部分能量分布于低频谱段，这对图像的压缩、传输都比较有利)**傅立叶级数定义**：任何周期函数都可以用正弦函数和余弦函数构成的无穷级数来表示。选择正弦函数与余弦函数作为基函数是因为它们是正交的。欧拉公式将正弦波统一成了简单的指数形式。**连续函数 f(x)的傅立叶变换 F(u)：傅立叶变换 F(u)的反变换：将欧拉公式带入可得离散傅立叶公式**：F(u)由 f(x)与对应频率的正弦和余弦乘积和组成。u 值决定了变换的频率成份，因此，F(u)覆盖的域称为频率域，其中每一项都被称为 FT 的频率分量。频率域及频率分量与 f(x)的时间和空间成分相对应。

**傅立叶变换作用**：傅立叶变换将信号分成不同频率成份。类似光学中的分光棱镜把白光按波长(频率)分成不同颜色，称**数学棱镜**。信号变化的快慢与频率域的频率有关。噪声、边缘、跳跃部分代表图像的高频分量，背景区域和慢变部分代表图像的低频分量。**二维离散 Fourier 变换**：设图像大小为 M\*N，原图为 f(x,y)，其频谱为 F(u,v)，则。**作用**：1、可以得出信号在各个频率点上的强度。2、用于计算卷积。

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \cdot e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} = fT_{列}(fT_{行}(f(x,y)))$$
$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cdot e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} = \frac{1}{MN} fT_{列}^{-1}(fT_{行}^{-1}(f(x,y)))$$

**傅立叶变换作用**：傅立叶变换将信号分成不同频率成份。类似光学中的分光棱镜把白光按波长(频率)分成不同颜色，称**数学棱镜**。信号变化的快慢与频率域的频率有关。噪声、边缘、跳跃部分代表图像的高频分量，背景区域和慢变部分代表图像的低频分量。**二维离散 Fourier 变换**：设图像大小为 M\*N，原图为 f(x,y)，其频谱为 F(u,v)，则。**作用**：1、可以得出信号在各个频率点上的强度。2、用于计算卷积。

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \cdot e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} = fT_{列}(fT_{行}(f(x,y)))$$
$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cdot e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} = \frac{1}{MN} fT_{列}^{-1}(fT_{行}^{-1}(f(x,y)))$$

**傅立叶变换作用**：傅立叶变换将信号分成不同频率成份。类似光学中的分光棱镜把白光按波长(频率)分成不同颜色，称**数学棱镜**。信号变化的快慢与频率域的频率有关。噪声、边缘、跳跃部分代表图像的高频分量，背景区域和慢变部分代表图像的低频分量。**二维离散 Fourier 变换**：设图像大小为 M\*N，原图为 f(x,y)，其频谱为 F(u,v)，则。**作用**：1、可以得出信号在各个频率点上的强度。2、用于计算卷积。

$$f(x,y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \cdot e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} = fT_{列}(fT_{行}(f(x,y)))$$
$$F(u,v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \cdot e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} = \frac{1}{MN} fT_{列}^{-1}(fT_{行}^{-1}(f(x,y)))$$

**模式识别定义：模式识别**(Pattern Recognition)：确定一个样本的类别属性（模式类）的过程，即把某一样本归属于多个类型中的某个类型。**样本** (Sample)：一个具体的研究（客观）对象。如患者，某人写的一个汉字，一幅图片等。**模式**(Pattern)：对客体特征的描述（定量的或结构的描述），是取自客观世界的某一样本的测量值的集合。**特征**(Features)：能描述模式特性的量（测量值）。在统计模式识别方法中，通常用一个向量表示，称之为**特征向量模式类**(Class)：具有某些共同特性的模式的集合。**模式识别系统的主要环节**：特征提取、特征选择、学习和训练、分类识别。**训练集**：是一个已知样本集，在监督学习方法中，用它来开发模式分类器。**测试集**：在设计识别和分类系统时没有用过的独立样本集。**系统评价原则**：为了更好地对模式识别系统性能进行评价，必须使用一组**独立于训练集的测试集**对本系统进行测试。**训练过程**(Training)：信息获取、预处理、特征提取、训练（学习分类规则）。识别过程：信息获取、预处理、特征提取、分类（利用学到的分类规则）。**特征提取和特征选择**：由对象空间到特征空间的变换和选择（降维操作）。**类型判别**：特征空间到类型空间所作的操作。

**图像点**：图像中一区域的灰度幅值与其领域值有着明显的差异。两条直线相交的顶点可看作是角点；物体的几个平面的相交处也可以看作是角点。图像中周围灰度变化较为剧烈的点可看作是角点；图像边界上曲率足够高的点也可看作是角点

**SUSAN 曲率原理**：通过核值相似区实现角点特征的检测。**核心点邻域划分**：一.灰度值相似于核心点灰度值的区域，称为核值相似区(USAN 区域)；二.与核心点像素灰度值相差明显的像素组成的区域。USAN 区域大小反映了图像局部特征的强度。

**纹理**：由某种模式重复排列所形成的结构。图像纹理反映了物体表面颜色和灰度的某种变化，与物体本身的属性相关。从宏观上看，纹理是物体表面拓扑逻辑的一种变化模式；从微观上看，它由具有一定的不变性的视觉单元（通称**纹理基元**）组成。不同物体表面的纹理可作为描述不同区域的一种明显特征。**特征**：某种局部的序列性在该序列更大的区域内不断重复出现。序列由基本的纹理基元非随机排列组成。在纹理区域内各部分具有大致相同的结构和尺寸。**特征描述与提取**：**统计分析法**，基于统计纹理特征的检测方法，主要包括灰度直方图法、灰度共生矩阵法等。**原理**：根据小区域统计特征的统计分布情况，通过计算像素的局部特征分析纹理的灰度级的空间分布。**适用范围**：统计分析法对木纹、沙地、草地这种完全无法判断结构要素和规则的图像的分析很有效。

**优势**：方法简单、易于实现，尤其是灰度共生矩阵法是公认的有效方法**结构分析方法**，纹理基元具有规范的关系，因而假设纹理图像的基元可以分离出来，并以基元的特征和排列规则进行图像纹理特征构造。**原理**：根据空间排列关系以及偏心率、面积、方向等特征，分析图像的纹理基元的形状和排列分布特点，以获取结构特征和描述排列的规则。**适用范围**：应用于已知基元的情况，对纤维、砖墙这种结构要素和规则则明确的图像分析比较有效

**模型分析法**根据每个像素和其领域存在的某种相互关系为图像中各个像素点建立模型，再由不同的模型提取不同特征量。典型的模型分析法有自回归方法、马尔可夫随机场方法和分形方法等。**频谱分析法**将纹理图像从空间域变换到频率域，然后通过计算峰值处的面积、峰值与原点的距离平方、峰值处的相位等，来获得在空间域不易获得的纹理特征，如周期、功率谱信息等。典型的谱分析法有二维傅立叶变换方法、Gabor 变换变换和小波方法等。

**图像形状特征**：拓朴特征，距离、周长和面积的测量，几何特征，形状特征的描述等。**图像形状特征描述**：可由其几何属性（如长短、距离、面积、周长、形状、凹凸等）、统计属性（如不变矩等）、拓朴属性（如孔、连通、欧拉数）等来描述。**矩形体**：指目标区域的面积与其最小外接矩形面积之比，反映了目标对其外接矩形的**充满程度**。

**HOG 特征提取算法**：1、标准化 gamma 空间和颜色空间。2、计算梯度。3、单元格梯度投影 4、块内归一化梯度直方图。5、收集 HOG 特征。

**SIFT 算法**：1、提取关键点。2、对关键点附加详细的信息（局部特征）构建描述器。3、通过两特征点（附带上特征向量的关键点）的两两比较找出相互匹配的若干对特征点，建立景物点的对应关系。

**计算机图形学**：用计算机将由概念所表示的物体（不是实物）图像进行处理和显示；侧重于根据给定的物体描述模型、光照及想象中的摄像机的成像几何，生成一幅图像；目前的一个热门研究方向是虚拟现实技术。

**计算机视觉**：学科内涵可以表述为图像分析、图像理解；目的是发展出能够理解自然景物的系统。在机器人领域中，计算机视觉为机器人提供眼睛的功能；目前，机器视觉领域的研究中，一些热门方向是基于内容的图像检索(content-based image retrieval)、无人驾驶车等。**2012 年前**，计算机视觉依靠“手工特征”+“机器学习分类”来做识别与检测。**2012 年后**，众多学者对深度学习抱迟疑态度，Hinton 为了回应质疑，让其学生 Alex 应用卷积神经网络(CNN)参加 ImageNet 数据大赛，结果大获全胜，从此开创了深度神经网络空前的高潮。

**图像处理---底层视觉、图像分析---中层视觉、图像理解---高层视觉**