

第一章绪论：

- 像素在数字图像上具有固定的 (x,y) 空域坐标以及该点的值 $f(x,y)$, 其特征为: $x,y,f(x,y)$ 均为离散值和有限值。 $(f(x,y))$ 常代表图像的 intensity 或 gray 等级 (空域坐标系下 y 右 x)
- 数字图像的形式: 单通道 (B&W 或者灰度等级) 三通道 (RGB) 四通道 (RGB+Alpha) 深度
- (光杆) 颜锥 锥状体 (少) 集中在中央凹附近, 对颜色极其敏感; 杆状体 (多) 的分布较为分散, 对低亮度的光照敏感。(两者均不在盲点上有分布) (锥状体 (cones), 杆状体 (rods))
- 图像不仅仅由电磁波生成, 超声波等都可以 (B 超)

5.DIP 与计算机视觉、机器人视觉、机器视觉的区别: **DIP:** 对图像的基础操作, 如几何变换, 增强恢复, 锐化滤波等, 结果仍为图像。计算机视觉: 对给定的图像进行信息提取, 如运动检测物体识别, 获得具体数据。机器视觉: 用于自动化生产, 由机器自主获取图像等信息来获得具体数据, 需要多传感器结合。机器人视觉: 对象是机器人, 使机器人具备视觉能力完成各项任务。

第二章数字图像基础:

- 采样量化: 数字图像的生成需要经过采样和量化的过程。坐标值 (x,y) 的数字化就是采样, 值 $f(x,y)$ 的数字化就是量化。
- 灰度分辨率: 2^k . 大小为 $M \times N$ 的数字图像一般用矩阵表示, 其占用的储存空间为: $M \times N \times k$ (比特)。空域分辨率为 $M \times N$.
- 图片的 Zooming 和 Shrinking 可用最邻近插值或者双线性插值, 前者的缺点是精度低, 可能会存在灰度上的不连续, 在变化的地方出现明显的锯齿状。双线性: $g(E) = (x' - i)[g(B) - g(A)] + g(A)$

$$g(F) = (x' - i)[g(D) - g(C)] + g(C)$$

$$g(E) = (x', y')[g(F) - g(E)] + g(E)$$

(无论上、降采样, 图片所包含内容不变, 只是空间分辨率变了)

- 像素 p 的 4 邻域 (十字形 N_4)、D 邻域 (4 个对角 D) 和 8 邻域 (一圈 N_8)。两个像素 p 和 q 之间的 4 邻接 (**p** 在 **q** 的 N_4)、8 邻接 (**p** 在 **q** 的 N_8) 和 m 邻接 (**p** 在 **q** 的 N_D 且 **p** 的 N_4 与 **q** 的 N_4 相交为空; 或 **p** 在 **q** 的 N_4)

连通 (所有连通的像素构成通路): 邻接 + 像素值都属于集合 V。连通分量: 对于 S 中的任何像素 p, S 中连通到该像素的像素集称为 S 的连通分量。连通集: 只有一个连通分量的集合。

- 欧几里得距离 (De)、街区距离 ($D_4 |x - s| + |y - t|$) 和棋盘距离 ($D_8 \max(|x - s|, |y - t|)$) 的概念。(注: D_4 为菱形, D_8 为方形)

第三章: 灰度变换和空间滤波 (图像增强) 高光 降噪 视觉感染力

- 对于空域, 我们的操作一般是针对像素的邻居。若直接对像素本身进行操作, 则称为灰度变换 $s = T(r)$ 。否则为空间滤波 $g(x,y) = T(f(x,y))$, 多使用掩模。
- 如果想将一个物体从背景分离, 可以使用阈值变换 $s = 1.0(r > threshold) + 0.0(r \leq threshold)$ 。若输入的图像灰度分辨率很大, 可使用对数变换 $s = c * \log(l + r)$ 。

幂律变化 $s = c * r^y$ 则可以将一个较窄范围的灰度等级映射到较大范围的灰度等级 (注: 小凸大凹) (整体变暗, $y > 1$) 或将图像整体变亮 ($y < 1$ 凸显脊柱改善欠曝光); 由于显示器、打印机等对不同亮度的响应非线性, 而是指数 $s = r^y$, 所以使用 y 校正 $s = r^{\frac{1}{y}}$ 来处理; 灰度切片 (变换函数 T 为分段函数) 类似于阈值变换, 对于突出图像中某些特征起作用。

比特平面分层中高阶比特平面包含了最重要的视觉数据, 低阶比特平面贡献了更精细的灰度细节, 存储 4 个高阶比特平面即可重建原图像 (在可接受范围内)。图像相减时存在-255 255 的灰度等级, 所以需归一化或者加 255 除以 2 将其重新变为 0 255, 该方法可用于检测运动的物体或者进行 change detection; 同一场景多张图取平均噪声水平不变。

3. 直方图均衡化: (满足单调递增和区间条件), 可以使输出图像的灰度分布更加均衡, 提高对比度。(均衡化过程较为简单) 直方图: 显示了像素值等级分布情况。 $R_r(w)$ 是概率分布函数。

$$s = T(r) = (L - 1) \int_0^r P_r(\omega) d\omega = (L - 1) \sum_{j=0}^k \frac{n_j}{n}$$

4. 平滑线性滤波器: (均值滤波器) 可用于滤除噪声 (也可以滤除不必要的细节) 或者突出总体特征, 但是会模糊边缘。统计排序滤波器: (中值滤波器) 有时表现更佳 (相对 average), 尤其是滤除椒盐噪声。自适应中值滤波器 (补) 可以滤除空间密度更大的椒盐噪声, 平滑其他噪声并减小失真, 没有边缘效应。自适应中值滤波: if $z_{med} - z_{min} > 0$, $z_{med} - z_{max} < 0$,

$$\{if z_{xy} - z_{min} > 0, z_{xy} - z_{max} < 0$$

$$z_{out} = z_{xy} else z_{out} = z_{med}\} else$$
 扩大所取集合的 size

5. 锐化空间滤波器: (减少模糊部分并突出边缘) 其效果与平滑空间滤波器相反 (积分与微分之区别) 一阶微分 $f(x+1) - f(x)$, 二阶微分 $f(x+1) + f(x-1) - 2f(x)$. 一阶微分非 0 值存在于 step 和 ramp 的起点以及 ramp 沿线; 二阶微分非 0 值存在于 step 和 ramp 的起 (终) 点; 一阶微分产生较粗的边缘, 对 gray level step 有更佳的响应; 二阶微分对细节 (细线、孤立点和噪声) 有更佳的响应, 而对 gray level step 有双响应 (双边缘, 更加明显), 因此二阶微分在增强细节方面更强。

6. 使用拉普拉斯算子锐化图像 (中心-4, 十字 1)。考虑对角项, 则掩模系数变为 (中心-8, 外圈 1)。使用拉普拉斯算子得到的并不是最终图像, 还要根据中心系数的正负, 用原图像 ± 拉普拉斯图像。最终 (中心 5, 十字-1) (中心 9, 外圈-1)

7. 高提升滤波: $g(x,y) = Af(x,y) \pm \nabla^2 f(x,y)$, 当 A 越大, 则越忽略锐化的作用。

8. 使用梯度 (一阶微分) 锐化图像: 我们直接给出 $g_x = \frac{\partial f}{\partial x}, g_y = \frac{\partial f}{\partial y}$ 这两个模板称为 Sobel 算子, 此时的操作为线性操作。再对其求向量 ∇f 的幅值 $M(x,y) = \sqrt{g_x^2 + g_y^2}$ 或者近似化 $M(x,y) = |gx| + |gy|$, 为非线性。

Sobel 算子 X: -1 0 1;-2 0 2;-1 0 1 Y:-1 -2 -1;0 0 0; 1 2 1;

Roberts 交叉梯度算子 X:0 -1;1 0 Y:-1 0;0 1;

第四章: 频率域滤波:

1. 图像 $f(x,y)$ 乘以指数项 $(-1)^{x+y}$ 再做傅里叶变换, 可将 DFT 的原点移到 $F(u - \frac{M}{2}, v - \frac{N}{2})$ 不影响幅度谱。若空间域和频域都用极坐标表示, 空间域与频域的旋转等价。 $F(0,0)$ = 图像的平均灰度。

2.DFT 的幅度谱中, 低频分量 (幅度谱原点周围) 反映了灰度变化缓慢的区域, 是图像在平滑区域上的外观。高频分量反映了灰度剧变的区域 (如边缘噪声), 是图像中精细部分。

- 一般滤波过程: 1. 原图乘以 $(-1)^{x+y}$ 2.DFT 得到 $F(u,v)$ 3. $F(u,v)$ 乘以滤波器 $H(u,v)$ 4.IDFT 之后再乘以 $(-1)^{x+y}$ 得到输出。

4. 陷波滤波器: $H(u,v) = 0 if (u,v) = (\frac{M}{2}, \frac{N}{2}), else H(u,v) = 1$

1. 可以使图像的平均灰度为 0, 而不影响图像的整体外观和细节。

5. 理想低通滤波器: $H(u,v) = 1 if [D(u,v)] < D_0, else H(u,v) = 0$ 其中 $D(u,v)$ 定义为 (u,v) 与 $(\frac{M}{2}, \frac{N}{2})$ 的欧式距离。(注: 由于 $H(u,v)$ 的急剧变化会产生振铃现象)

6. 高斯低通滤波器: $H(u,v) = e^{-\frac{D^2(u,v)}{2D_0^2}}$ 作用: 连接断裂处, PS 人脸图片上的疤痕或者皱纹。

7. 巴特沃斯低通滤波器: $H(u,v) = \frac{1}{1 + [\frac{D(u,v)}{D_0}]^{2n}}$ 截止频率定义为 $H(u,v)$ 下降为 50%。一阶二阶巴特沃斯滤波器几乎观察不到振铃现象。越高阶越接近理想低通。

8. 理想高通滤波器: $H(u,v) = 0 if [D(u,v)] < D_0, else H(u,v) = 1$

9. 巴特沃斯高通滤波器: $H(u,v) = \frac{1}{1 + [\frac{D_0}{D(u,v)}]^{2n}}$ 越高阶越接近理想高通。

10. 高斯高通滤波器: $H(u,v) = 1 - e^{-\frac{D^2(u,v)}{2D_0^2}}$

额外: 连续 $F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-j2\pi x u} e^{-j2\pi y v} dx dy$

$f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x,y) e^{j2\pi x u} e^{j2\pi y v} du dv$

离散 (dft) $F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi \frac{xu}{M}} e^{-j2\pi \frac{yu}{N}}$

$f(x,y) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} F(x,y) e^{j2\pi \frac{xu}{M}} e^{j2\pi \frac{yu}{N}}$

第五章: 图像恢复

1. 空域上: $g(x,y) = f(x,y) * h(x,y) + n(x,y)$

频域上: $G(u,v) = F(u,v) \cdot H(u,v) + N(u,v)$

图像恢复是由 g 估计原图像 f 的过程。

2. 噪声分布: 高斯 (常见) $p(z) = \frac{e^{-\frac{(z-u)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma}}$ 瑞利 (无人驾驶)、爱尔兰、Gamma、均值、脉冲、椒盐

3. 傅里叶谱图有 8 个点说明原图像有横向纵向对角线的周期性噪

$\sin(2\pi u_0 x + 2\pi v_0 y) \leftrightarrow j \frac{1}{2} [\delta(u+u_0, v+v_0) - \delta(u-u_0, v-v_0)]$

4. 带阻滤波器: $H(u,v) = \frac{1}{1 + (\frac{D(u,v)W}{(D(u,v))^2 - D_0^2})^{2n}} H(u,v) = 1 - e^{-0.5[\frac{D^2(u,v) - D_0^2}{D(u,v)W}]^2}$

6. 退化模型: $H(u,v) = e^{-k(u^2 + v^2)^{5/6}}$

5. 陷波滤波器: notch reject: $H(u, v) = \frac{1}{1 + (\frac{D1(u, v)D2(u, v)}{D0^2})^n}$

$H(u, v) = 1 - e^{-0.5[\frac{D1(u, v)D2(u, v)}{D0^2}]}$ 用以去除噪点, 如: 横向纵向噪声

7. 逆滤波: $\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} = F(u, v) + \frac{N(u, v)}{H(u, v)}$ 则即使我们直到退化过程, 由于噪声影响也无法复原。同时必须 H 不能趋于 0, 否则会放大噪声。解决办法: 只取 H 原点附近的频谱, 以外一概不考虑, 一般来说能量集中于原点, 频域幅度在这较高。

8. 维纳滤波器: 使用最小均方误差准则设计

$$\hat{F} = \left[\frac{H^*(u, v)}{|H(u, v)|^2 + S_n(u, v)/S_f(u, v)} \right] G(u, v)$$

第九章: 形态学图像处理 (对于二值图)

1. 平移: $(B)_z = \{c|c = b + z, b \in B\}$ (如平移到 (z_1, z_2))

反射: $\hat{B} = \{w|w = -b, b \in B\}$ (如 $(-x, -y)$ 反射到 (x, y))

Fit: 完全正确。Hit: 部分正确。Miss: 全部错误

2. 膨胀 (Dilation): $A \oplus B = \{z|(\hat{B}) \cap A \neq \emptyset\}$ 或者定义为 A 与 B 的镜像存在 Hit 关系; 膨胀会粗化或者增长物体; 膨胀可以修复断裂处, 修复表面坑洼。(注: 膨胀不一定能粗化, 和结构元有关。)

3. 腐蚀 (Erosion): $A \ominus B = \{z|B_z \subseteq A\}$ 或者定义为 A 与 B 存在 Fit 关系; 腐蚀缩小或细化了物体, 可视为形态学滤波操作, 小于结构元的物体都将滤除。腐蚀可以分开已连接的物体, 可以将物体表面的突出部分剥离。

4. 开运算: $A \circ B = (A \ominus B) \oplus B$ 即对物体先腐蚀再膨胀, 其几何解释为: 球形结构元 B 沿物体 A 的内部边界滑动, 其并集即为开操作 $A \circ B = \bigcup (B_z)|(B_z) \subseteq A$ 可以平滑图像轮廓, 打断物体间的连接部分, 清除突出物。

5. 闭运算: $A \bullet B = (A \circ B) \ominus B$ 即对物体先膨胀再腐蚀, 其几何解释为: 球形结构元 B 沿物体 A 的边界滑动, 其并集即为闭操作 $A \bullet B = \bigcap (B_z)|(B_z) \subseteq A$ 可以平滑轮廓上的缺口, 填充洞口, 连接间隙和断裂部分。

6. Hit or Miss: $A \circledast B = (A \ominus X) \cap [A^C \ominus (W - X)]$ 设感兴趣物体形状为 X, B 为 X 及其背景组成的集合, W 为比 X 大一点的小窗, 用 X 腐蚀 A 产生的集合与用 (W-X) 腐蚀 A 的补集产生的集合的交集就是击中或击不中变换。

7. 边缘提取: $\beta(A) = A - (A \ominus B)$ 即可提取边界, B 是结构元。

8. 区域填充: 需设置一个初始点, 过程可表示为

$X_k = (X_{k-1} \oplus B) \cap A^C$, 直到 $X_k = X_{k-1}$, 完成区域填充, 最终结果图应为 $X_k \cup A$

9. 连通分量的提取: 过程与上面类似, 只是将 A^C 替换成 A, 最终结果图应为 X_k 。

第十章: 图像分割 (基于不连续性与相似性)

1. 灰度变化不连续性是进行图像分割的基础之一, 其处理的图像特征为孤立点、线、边缘。对于相似的灰度, 根据一组预定义的准则将图像分割为相似的区域。(阈值处理, 区域生长、分裂聚合)

2. 点检测: 使用二阶微分 (实际就是进行空间滤波)

如: 拉普拉斯算子 (添加对角项) 的模板系数之和为 0, 表明在恒定灰度区域模板响应为 0。设置阈值 T, 若模板响应 $R(x, y) >= T$, 则输出图像在该点的值 $g(x, y) = 1$, 检测到孤立点。

3. 线检测: 同样可用拉普拉斯算子。由于其产生负值, 我们一般进行正阈值处理, 仅适用拉普拉斯图像的正值。(注: 当线宽比模板尺寸大时, 会被一个零值分开)。若我们对特定方向的线感兴趣, 可将模板上对应方向的系数全换成 2 (最后同样要进行阈值处理)

4. Ramp 模型: 线的厚度与斜率成反比, 斜率与模糊程度成反比。

5. 微弱的可见噪声也严重影响边缘检测所用的两个关键导数, 所以应先进行平滑处理。另一方法是对梯度图像进行阈值处理 (可能会使部分边缘断开)。若要突出主要边缘并尽可能保持连接时, 实践中通常又做平滑处理又做阈值处理。

6. LoG 算子: $\nabla^2 G(x, y) = [\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4}] e^{-\frac{(x^2 + y^2)}{2\sigma^2}}$ ($G(x, y)$ 为二维高斯函数)。根据 LoG 算子生成的模板需要满足系数之和为 0 (以便模板响应在恒定灰度区域为 0)。先高斯模糊后拉普拉斯算子 $\frac{\partial^2(h * f)}{\partial x^2}$, 可以简化成 $\frac{\partial^2 h}{\partial x^2} * f$, 即用特定的模板直接卷积减小计算量。

7. Canny 算法: 1. 用高斯滤波器平滑图像。2. 用梯度算子 (Sobel, Prewitt) 计算梯度方向和幅值。3. 对梯度幅值进行非极大值抑制。4. 双阈值处理减少伪边缘点。(高阈值: 低阈值 = 2:1 或 3:1)

8. 边缘连接: 1. 局部处理, 若边缘点 q 的某个邻居 p (3X3, 5X5,...) 满足条件, 则两者是连接的。2. 使用霍夫变换的全局处理: 预定义一些全局性质, 使我们得以筛选边缘点从而得到指定形状的曲线。

9. 霍夫变换: 线变换: 空间一条直线: $r = x \cos \theta + y \sin \theta$, 在参数空间对应一个点, 而该直线上的每一点在参数空间对应一条正弦曲线, 各正弦曲线的交点即为直线对应的参数空间的点。

算法实现: 1. 获取二值化边缘图。2. 对 (r, θ) 离散化, 给出 (r_{min}, r_{max}) 与 $(\theta_{min}, \theta_{max})$, 划分有限个等间距的离散值, 是参数空间化为一个个方格。3. 设置累加器 A, 每个网格初始为 0, 计算参数空间的曲线方程, 曲线穿过的格子累加器加一。4. 遍历累加器, 找到最大的格子, 其坐标即为直线参数, 转为直线空间显示出来。

圆变换: 霍夫梯度法: 1. Sobel 获取二值化边缘图, 同时储存梯度向量。2. 初始化圆心空间 N(a,b), 并置 0。3. 遍历边缘像素点, 沿梯度方向画线, 线段经过的累加器加一。4. 统计排序得到可能的圆心 N(a,b) 越大越可能为圆心。5. 针对圆心估算半径, 计算所有边缘点到圆心的距离对距离从小到大排序, 取合适阈值选合适半径。初始化半径空间 N(r) 为 0, 遍历边缘点, N(距离)+=1, 取最大 N(r) 为半径画圆。

普通的霍夫变换: 设置三维的累加器, 确定三个参数时可一个一个确定。

8. 基本的全局阈值处理: 1. 设置全局阈值 T。2. 用 T 分割图像

得到两组像素。3. 分别计算两组像素的灰度平均值 m1 和 m2。4. 得到新阈值 $T = 0.5 * (m1 + m2)$ 。5. 重复步骤至最新两次 T 值的差小于预设值。

9. Ostu 方法 (最佳全局阈值处理): 计算快速简单, 不受图像亮度和对比度影响。但是对噪声敏感, 而且只能对单一目标分割; 当目标和背景大小相差悬殊时, 效果不好。

10. 区域生长: 1. 设置种子 (种子生长顺序会影响最终结果) 2. 确定生长准则 (例如灰度差) (注: 考虑像素相似性和连接性)

11. 区域分割: (注: 二维可用四叉树三维用八叉树)

第六章: 彩色图像处理:

1. 品红 $M = R + B$, 青色 $C = G + B$, 黄色 $Y = R + G$ 。

2. 全彩色图像有 24 比特的深度, RGB 分别为 8 比特图像。(RGB 归一化) 注: RGB 不能表征所有颜色 HSI 比 RGB 更易表征颜色。

3. CMY 模型 (面向应用) 注: CIE 国际标准来看显示器和打印机能实现颜色不一样。

4. HSI 模型: H 为色调, S 为饱和度, I 为强度: 对锥体模型: 离原点在 z 轴方向的距离表示 I, 当前圆面内用极坐标表示, 离圆心的距离 r 表示 S, 夹角 θ 表示 H。

5. RGB 到 HSI: $I = \frac{R+G+B}{3}$ $S = 1 - \frac{3\min(R, G, B)}{R+G+B}$ $\theta = \arccos(\frac{0.5 \times [(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}})$ 注: B>G 时, H 为 $360^\circ - \theta$

7. 灰度分层: 类似于灰度切片, 但是将不同分段的像素集合赋予不同的颜色, 有所区别。(伪彩色)

8. 彩色图像处理方法: 1. 对 RGB 三分量图像分别进行处理。2. 对彩色图像的每个像素点上的 Vector (三或四个分量) 进行向量处理。注意: HSI 处理一般在 I 分量处理, 同时改 H 会影响 RGB 三个分量, 故对三通道处理和单独对向量处理有区别 (彩色图像噪声会严重影响 H, 公式里含 cos)

9. 对彩色图像的滤波: 在 HIS 模型一般只能对 I 处理

10. 彩色分层: 设感兴趣的色彩被包围 W、中心在原型 (平均) 颜色点的立方体所包围, 立方体外的颜色被强制置为中性 (对于 RGB 彩色空间, 为 (0.5, 0.5, 0.5))。也可用球体、椭球体所包围。

11. 彩色轮廓提取: 1. RGB 转 HSI 2. 取对应颜色的 his, 在对应阈值内置白色, 阈值外置黑色 3. 滤波 4. findContours 寻找白色区域轮廓 5. drawContours 绘制轮廓。

第七章: 图像特征

链码: 表示边缘的数组, 数字表示方向。

多边形近似: (MPP: 最小周长) 矩形度: 可判断形状与矩形差异。欧拉数: E=C-H。C 为联通分量, H 为洞个数。

矩: $m_{bp} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$

K-means 算法: 1. 随机选择 k 个起始中心点。2. 对每个点标记最近的中心点。4. 计算新的中心点

优点: 简单, 收敛快。缺点: 对初始值敏感。k 值要预设。