

在学习图像分析技术过程中，接触到 Mean-Shift 算法，由于对模式识别相关知识不太了解故理解有点困难，参考了清华《模式识别》和覃剑的《基于边界力的 Mean-Shift 核窗宽自适应算法》初步得到了如下的理解。

1、目标模型，算法采用的是特征值的加权概率分布来描述目标模型。这应该是模式识别中主要描述目标的模型，不同于自动控制理论中采用的状态方程。

目标模型 $q = \{q_u\}$ $u \in [1, \dots, m]$ 共 m 个特征值。

$$q(u) = C \sum_{i=1}^n k \left(\left\| \frac{X_i - X_0}{H} \right\|^2 \right) \delta(b(X_i) - u) \quad i \in [1, \dots, n]$$

其中 X_0 是窗口中心点向量值（可能为 RGB 向量或者灰度值）， X_i 是窗口内第 i 点向量值。

C 为归一化常数，保障 $\sum_{u \in [1, \dots, m]} q_u = 1$ 。 $k(x)$ 为核函数， H 为核函数的带宽向量。 M 为特征值的个数，对应于图像处理可以理解为灰度等级划分的个数，从而特征值 u 为对应的灰度等级。 δ 函数为脉冲函数，保证只有具有 u 特征值的像素才对概率分布作出贡献。从而 k 函数可以理解为 u 灰度值的一个加权频数。

2、匹配对象，也采用特征值加权概率分布

$$p_u(Y) = C_h \sum_{i=1}^{n_h} k \left(\left\| \frac{X_i - Y}{H_h} \right\|^2 \right) \delta(b(X_i) - u) \quad i \in [1, \dots, n_h]$$

其中， Y 为匹配对象的中心， X_i 是匹配窗口内第 i 点向量值， H_h 为匹配窗口的核函数带宽向量。 C_h 为匹配窗口特征向量的归一化常数。

3、匹配对象与目标模型的相似程度，相似函数可采用 Bhattacharyya 函数

$$\rho(p(Y), q) = \sum_{u=1}^m \sqrt{p_u(Y)q_u}$$

4、匹配过程就是寻找相似函数最大值的寻优过程，Mean-Shift 采用的是梯度下降法。首先将 $\rho(Y)$ 在 $p(Y_0)$ 附近进行泰勒级数展开，取前两项。即：

$$\rho(Y) \approx \rho(Y_0) + \frac{d\rho}{dp} (p(Y) - p(Y_0))$$

$$\text{定义 } \rho_u(Y) = \sqrt{p_u(Y)q_u},$$

$$\text{从而 } \rho_u(Y) = \rho_u(Y_0) + \frac{q_u}{2\sqrt{p_u(Y_0)q_u}} (p_u(Y) - p_u(Y_0)) = \frac{1}{2} \rho_u(Y_0) + \frac{1}{2} \sqrt{\frac{p_u(Y)q_u}{p_u(Y_0)q_u}}$$

$$\text{因此, } \rho(Y) = \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(Y_0)q_u} + \frac{1}{2} \sum_{u=1}^m (p_u(Y) \sqrt{\frac{q_u}{p_u(Y_0)q_u}})$$

要使得 $\rho(Y)$ 向最大值迭代，只要 Y 的搜索方向与梯度方向一致即可，通过求导可得到 Y_0 的

梯度方向为：

$$\nabla \rho(Y_0) = \frac{C_h}{H_h^2} \left[\sum_{i=1}^{n_h} w_i g\left(\left\| \frac{Y_0 - X_i}{H_h} \right\|^2\right) \right] \left[\frac{\sum_{i=1}^{n_h} X_i w_i g\left(\left\| \frac{Y_0 - X_i}{H_h} \right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\| \frac{Y_0 - X_i}{H_h} \right\|^2\right)} - Y_0 \right]$$

其中 $w_i = \sum_{u=1}^m \sqrt{\frac{q_u}{p_u(Y_0)}} \delta(b(X_i) - u)$ 为权值。因此如果如下确定 Y_1 ，那么 $Y_1 - Y_0$ 将与梯度方

向一致。

$$Y_1 = \frac{\sum_{i=1}^{n_h} X_i w_i g\left(\left\| \frac{Y_0 - X_i}{H_h} \right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\| \frac{Y_0 - X_i}{H_h} \right\|^2\right)}$$

总结：感觉到从博士或者硕士论文中看到的公式推理过程总是不太连贯，不容易理解，总有点小错误，不知道自己的这番理解是否犯了同样的毛病，一般来说，别人的毛病容易看见，自己的呢，也就是那回事，不见得多好！