图像特征 harris 角点检测

我们知道角点检测是图像的一种比较有价值的特征，它在图像识别和匹配，物体跟踪等领域有十分广泛的作用。这篇文章主要说明一下 harris 角点检测的数学原理，我们会给出一系列的数学推导来帮助大家更好地理解 harris 焦点检测。

角点具备什么特性呢，它和图片其它部分有什么不同呢，比如和图片的背景，边缘的区别在哪里。很多教材或者网上的文章都有说明，相信大家也都知道。这里我再简要说明一下，角点在图片沿着X轴和Y轴方向移动的话，会有较明显的灰度值变化，即发生较大的频率波动。而背景沿着X轴和Y轴方向移动的话，并没有明显变化。而边缘部分沿着X轴和Y轴方向移动，会在其中一个方向产生较大的灰度值变化，而在另外一个方向上并不会产生大的变化。我们判断角点就是根据这个原则。那么我们就可以给出判断角点角点检测的数学表达式：

$$\boldsymbol{dv(x,y,\triangle x,\triangle y)=\sum_{u,v\in W(x,y)}W(u,v)\,(I(u,v)-I(u+\triangle x,v+\triangle y))^2} \quad \textbf{(1)}$$

这里$\boldsymbol{I(x,y)}$是指图像在坐标$\boldsymbol{(x,y)}$位置上的像素值。$\boldsymbol{W(u,v)}$是指以点$\boldsymbol{(x,y)}$为中心的滑动窗口的权值，我们一般用一个常量或者高斯加权的方式给出这个权重值。这个窗口和卷积窗口一样我们设置一个滑动窗口，在图像上下左右地滑动，判断每个窗口区域的中心点$\boldsymbol{(x,y)}$是否是角点。判断依据就是在滑动窗口范围内的所有像素点分别减去对应像素点移动$\boldsymbol{(\triangle x,\triangle y)}$后的像素值，即使$\boldsymbol{I(u+\triangle x,v+\triangle y)}$，加上平方是因为可正可负，但我们并不关正负，我们只在乎变化的程度。$\boldsymbol{dv(x,y,\triangle x,\triangle y)}$就是像素点移动后与原来位置像素相比，发生的变化程度。这个值很大的话说明就是角点。

现在我们给出的这个判断角点的计算公式，显然还不足够帮助我们解决问题。这个公式里面的关键是我们并不知道$\boldsymbol{I(u+\triangle x, v+\triangle y)}$是多少。如何解决这个问题呢。精确值我们确实无法获得，那么我们可以想办法求出个近似值。如何求解近似值呢，我们就要用到高等数学里面的泰勒公式了，我们将$\boldsymbol{I(u+\triangle x, v+\triangle y)}$做一阶泰勒展开，得到下面式子：

$$\boldsymbol{I(u+\triangle x, v+\triangle y) = I(u,v) + I_x(u,v)\triangle x + I_y(u,v)\triangle y + O(\triangle x^2, \triangle y^2)}$$

$\boldsymbol{I_x}$和$\boldsymbol{I_y}$是图像在X轴和Y轴方向的导数，我们可以通 sobel 算子求出。$\boldsymbol{O(\triangle x^2, \triangle y^2)}$是一个关于$\boldsymbol{\triangle x^2, \triangle y^2}$的高阶无穷小。所以我们可以将它忽略。那么我们就得到下面这个式子：

$$\boldsymbol{I(u+\triangle x, v+\triangle y) \approx I(u,v) + I_x(u,v)\triangle x + I_y(u,v)\triangle y} \quad \textbf{(2)}$$

这样我们就等到了$\boldsymbol{I(u+\triangle x, v+\triangle y)}$的近似解了。代入公式 1 里面可以得到下面的表达式：

$$\boldsymbol{dv(x,y,\triangle x,\triangle y) \approx \sum_{u,v\in W(x,y)} W(u,v)(I_x(u,v)\triangle x + I_y(u,v)\triangle y)^2} \quad \textbf{（3）}$$

接下来我们把$\boldsymbol{\sum_{u,v\in W(x,y)} W(u,v)(I_x(u,v)\triangle x + I_y(u,v)\triangle y)^2}$这一部分转换成矩阵乘法的形式如下：

$$\boldsymbol{Mat = \sum_{u,v\in W(x,y)} \begin{bmatrix} W(u,v)I_x(u,v)^2 & W(u,v)I_x(u,v)I_y(u,v) \\ W(u,v)I_x(u,v)I_y(u,v) & W(u,v)I_y(u,v)^2 \end{bmatrix}}$$

$$\boldsymbol{= \begin{bmatrix} \sum\limits_{u,v\in W(x,y)} W(u,v)I_x(u,v)^2 & \sum\limits_{u,v\in W(x,y)} W(u,v)I_x(u,v)I_y(u,v) \\ \sum\limits_{u,v\in W(x,y)} W(u,v)I_x(u,v)I_y(u,v) & \sum\limits_{u,v\in W(x,y)} W(u,v)I_y(u,v)^2 \end{bmatrix}}$$

即是：

$$\boldsymbol{dv(x,y,\triangle x,\triangle y)\approx[\triangle x\quad\triangle y]Mat\begin{bmatrix}\triangle x\\\triangle y\end{bmatrix}}$$

我们设$A=\sum_{u,v\in W(x,y)}W(u,v)I_x(u,v)^2$，$B=\sum_{u,v\in W(x,y)}W(u,v)I_y(u,v)^2$，$C=\sum_{u,v\in W(x,y)}W(u,v)I_x(u,v)I_y(u,v)$。这样便于我们表述问题。那么可得到下面化简的式子：

$$\boldsymbol{dv(x,y,\triangle x,\triangle y)\approx[\triangle x\quad\triangle y]\begin{bmatrix}A & C\\C & B\end{bmatrix}\begin{bmatrix}\triangle x\\\triangle y\end{bmatrix}}$$

即是：

$$\boldsymbol{dv(x,y,\triangle x,\triangle y)\approx A\triangle x^2+2C\triangle x\triangle y+B\triangle y^2}$$

这个式子其实是个椭圆方程。只是不是我们高中接触的标准椭圆方程。标准椭圆方程如下：

$$\boldsymbol{\frac{x^2}{a^2}+\frac{y^2}{b^2}=1}$$

标准椭圆方程是二维空间中，如果写成三维空间即是：

$$\boldsymbol{\begin{cases}\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=z\\ z=1\end{cases}}$$

$\boldsymbol{z}$类似于等高线的存在。而我们的$dv(x,y,\triangle x,\triangle y)$可以看做$\boldsymbol{z}$。$\boldsymbol{A\triangle x^2+2C\triangle x\triangle y+B\triangle y^2}$这一部分是椭圆，只是不是标准的以原点为中心的椭圆。它是个在三维空间中倾斜的椭圆。我们可以通过矩阵变换的方式将它转为标准椭圆。而转化为标准椭圆，并不会影响到我们对角点的判断。因为一个角点即使做了变换，它依然还是角点。所谓的通过矩阵转换，其本质就是将原来图像上的点进行平移旋转而已，所以它不会影响到我们最终结果的判断。为什么要转换为标准椭圆呢，其实即使为了把$\boldsymbol{2C\triangle x\triangle y}$这一部分清除掉吗，这样我们会很容易判断$\boldsymbol{dv(x,y,\triangle x,\triangle y)}$在$X$方向和$Y$方向的大小变化程度，那么如何消除掉$2C\triangle x\triangle y$呢，这就需要运动矩阵变换的方式了。

$\begin{bmatrix} A & C \\ C & B \end{bmatrix}$这个矩阵可是个对称矩阵，我们知道对称矩阵是可以对角化的，即可以转化成如下格式：

$$\begin{bmatrix} \boldsymbol{\lambda_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda_2} \end{bmatrix} = \boldsymbol{p^T} \begin{bmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{bmatrix} \boldsymbol{P}$$

其中λ_1，λ_2是Mat的特征值，P是Mat的特征向量组成的正交阵。这段不清楚的可以去查阅线性代数。

那么就说明我们可以实施一个矩阵变换吗，将Mat对角化，从而得到一个标准椭圆方程，该方程如下所示：

$$\boldsymbol{dv(x, y, \triangle x, \triangle y)} \approx [\triangle \boldsymbol{x} \quad \triangle \boldsymbol{y}] \begin{bmatrix} \boldsymbol{\lambda_1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda_2} \end{bmatrix} \begin{bmatrix} \triangle \boldsymbol{x} \\ \triangle \boldsymbol{y} \end{bmatrix}$$

即：

$$\boldsymbol{dv(x, y, \triangle x, \triangle y)} \approx \boldsymbol{\lambda_1} \triangle \boldsymbol{x^2} + \boldsymbol{\lambda_2} \triangle \boldsymbol{y^2}$$

这就得到一个标准的椭圆方程了。$\boldsymbol{dv(x, y, \triangle x, \triangle y)}$这个值越大，则说明这个坐标点$(x, y)$移动后的变化越大。特征值$\lambda_1$，$\lambda_2$代表了这两个标准椭圆长短周的变化方向的程度，即是沿着X轴和Y轴方向的变化程度。所以λ_1，λ_2都很大，并且两个特征值之间差距不大，说明是角点。反之一个很大，一个很小，表明只有一个方向有巨大变化，另外一个方向变化很小，这说明是边缘。而两个特征值都很小的情况，说明是背景或者非边缘和角点的地方。

，