Lucas-Kanade 算法

王 琳 (wanglin193@hotmail.com)

一 预备知识

1 泰勒展开 Taylor expansion

(标量)函数 $f(X)$在 $X=X_0$ 附近的一阶泰勒展开式为

$$f(X_0+\Delta X)\approx f(X_0)+\Delta X\left.\frac{df}{dX}\right|_{X=X_0}$$

变量 X 可以是 n 维矢量 $X=(x_0, x_1,\ldots x_n)$，则 df/dX 为梯度矢量

$$\frac{df}{dX}=\left(\frac{df}{dx_0},\frac{df}{dx_1},\ldots\frac{df}{dx_n}\right),$$

另外，如果 $f(X)$是 m 维矢量函数，则得到所谓 Jacobian 矩阵

$$\frac{df}{dX}=\begin{pmatrix}\frac{df_0}{dx_0} & \cdots & \frac{df_0}{dx_n}\\ \cdots & \cdots & \cdots\\ \frac{df_m}{dx_0} & \cdots & \frac{df_m}{dx_n}\end{pmatrix}$$

2 复合函数链式求导法则 Chain rule

在函数 f 对变量 t 求导数的时候，有时 f 和 t 关系是通过中间变量 x,y 联系起来的,$f=f(x,y)$, $x=x(t)$,$y=y(t)$.，那么 df/dt 由下式表示（<,>表示内积）：

$$\frac{df}{dt}=\frac{\partial f}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial t}=\left\langle\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right),\left(\frac{\partial x}{\partial t},\frac{\partial y}{\partial t}\right)\right\rangle$$

二 L-K 算法

目的是用一组参数 **p** 从一个图像区域（patch）采样 $I(x)$，和模板 $T(x)$进行匹配，使得误差最小

$$E=\sum_{patch}\left[I(W(X;p))-T(x)\right]^2$$

其中 W(X:p)是以矢量 **p** 为参数表达的 image warping，实际上可以理解为坐标变换，以决定在图像 I 上采样点坐标。比如仿射变换(Affine transform)，这个 W(X:p)得到新的坐标点(x’,y’)表达为

$$\begin{aligned}W(X;p)&=\begin{bmatrix}x'\\y'\end{bmatrix}=\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}xp_1+yp_3+p_5\\xp_2+yp_4+p_6\end{bmatrix}\\&=\begin{bmatrix}1+p_1 & p_3 & p_5\\p_2 & 1+p_4 & p_6\end{bmatrix}\begin{bmatrix}x\\y\\1\end{bmatrix}\end{aligned},$$

关于这个公式需要说明：

6 个参数的仿射变换，需要至少 3 个对应点。

4 个参数的相似变换只表达变比和旋转，需要至少 2 个对应点：

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} xp_1 - yp_2 + p_3 \\ xp_2 + yp_1 + p_4 \end{bmatrix}$$
$$= \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} xs\cos\theta - \sin\theta + p_3 \\ x\sin\theta + ys\cos\theta + p_4 \end{bmatrix}$$

而如果只用 2 个参数表示位移，则 $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x + p_1 \\ y + p_2 \end{bmatrix}$，就是一般的**光流(Optical flow)法**。

p 也可以是 8 个参数表达的透射变换。

令 E 最小化的过程是一个优化过程，用梯度下降求解。先对 E 的表达使用泰勒展开线性近似

$$E = \sum_{patch} \left[I(W(X;p+\Delta p)) - T(x) \right]^2 \approx \sum_{patch} \left[I(W(X;p)) + \frac{\partial I}{\partial p}\Delta p - T(x) \right]^2$$

令 $\dfrac{\partial E}{\Delta p} = 0$ 有

$$\sum_{patch} \frac{\partial I}{\partial p}\Delta p + (I(W(X;p)) - T(x)) = 0\,,$$

$\partial I / \partial p$ 的求解用到前面提到的 chain rule：

$$\frac{\partial I}{\partial p} = \frac{\partial I}{\partial W}\frac{\partial W}{\partial p} = \left\langle \nabla I, \left[\frac{\partial W_x}{\partial p} \frac{\partial W_y}{\partial p} \right] \right\rangle,$$

其中 ∇I 表示图像梯度。$\partial W / \partial p$ 可表达为如下雅可比矩阵：

$$\begin{bmatrix} \partial W_x / \partial p \\ \partial W_y / \partial p \end{bmatrix} = \begin{bmatrix} x & 0 & y & 0 & 1 & 0 \\ 0 & x & 0 & y & 0 & 1 \end{bmatrix}.$$

这样 $\partial I / \partial p$ 是维数为 1×2 和 2×6 两矩阵相乘，得到 1×6 的矢量。这样对于 patch 上的每个点

$$\frac{\partial I}{\partial p}\Delta p = T(x) - I(W(X;p))$$

上边这个式子可以这样理解：我们要求的梯度下降方向 Δp 是个 6 维矢量(affine 的情况)，它和 $\partial I / \partial p$ (图像函数关于参数向量 **p** 的梯度)求内积的结果，应该等于图像与模板的差。假设拿来计算的**像素个数为 N**，那么我们有 N 个这样的方程构成方程组，求 6 个参数：

$$J\Delta p = E$$

J 是[N*6]矩阵，它的每一行都是一个像素点处求 $\partial I / \partial p$ 的值。E 是[N*1]矢量，表示每个点处的 *T-I* 值。用最小二乘法解(求 J 的伪逆，要对 6*6 矩阵求逆)方程组得到：

$$\Delta p = \left(J^T J\right)^{-1} J^T E = dP * E$$

关于 L-K 算法这个方程组的说明：

1 图像和模板的差是图像匹配的原动力，它分别和 *dP* 矩阵中的的 6 个 N 维矢量做内积得到参数 p 的更新值 p=p+Δp，从而驱动模板动起来。这个求解过程是个迭代过程，无法保证收敛到全局最优解。用多分辨率法求解是个好的选择，可参见 OpenCV 里的光流法文档。

2 用做运动估计的光流法是用前一帧的图像块作为模板，用当前帧图像进行匹配的，这个过程要实时计算二阶逆矩阵（p 的维数是 2）。而在有些应用中，*dP* 的步骤可以只在模板 T 上离线进行，可称作**训练**。模板也可以扩展为一个包含更多参数的模型，比如著名的 Active Appearance Model。有时 warp 过程无法显式表达，计算 *dP* 可以用给 p 加扰动(perturbation)的方法。CMU 的文章把 Cootes 的那套 AAM 用 L-K 光流的框架来描述，据说电影 Avatar 就是用他们的方法做人脸特征跟踪的。

3 文章[1]中把 $H=J^TJ$ 叫作 Hessian 矩阵，它是对 J 求伪逆过程的一个副产品，和表达二阶偏导的 Hessian 阵不同。不过 H 的作用很重要，它能反映图像数据的统计特性。如果不幸 H 是奇异的，也就是说无法求逆矩阵，那么说明这个图像模板包含的各个梯度方向的信息不足以进行跟踪应用。对于光流应用（p 的维数是 2，只有平移运动），H 的对应的特征值(eigenvalue)都要远大于 0，才是“好的特征”（角点），这个特性又被用来做 corner detection。有著名的文章 Good features to track 专门讨论此事。

参考文献:

[1] Simon Baker and Iain Matthews. Lucas-Kanade 20 years on: A unifying framework.

[2] Jianbo Shi and Carlo Tomasi. Good features to track.
