视觉SLAM十四讲|第8讲 视觉里程计2

笨丫头

已关注

3 人赞同了该文章

终于来到第8讲了。

第七讲中的特征点法需要计算特征点，计算量大；如果找不到特征点方法就失效了；忽略了特征点以外的信息。

第八讲介绍光流法和直接法，避免使用特征点。

1. 光流

> 光流是一种描述像素随时间在图像之间运动的方法。

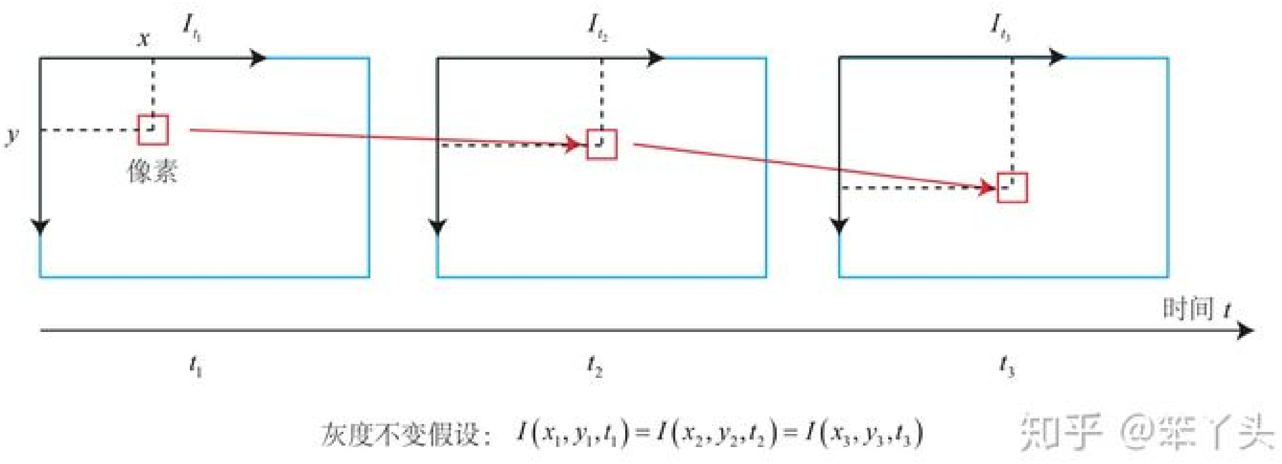

图8-1

如图8-1所示一个像素会随着时间的流逝进行运动，每一时刻的位置都有所变化，光流法就是去追踪像素每一时刻的位置变化，来还原运动过程。

光流法中有一条基本假设：

> 灰度不变假设：同一个空间点的像素灰度值，在各个图像中是固定不变的。

意思是说图8-1中红色框框的像素点，在 t_1, t_2, t_3 时刻的灰度值是一样的，只是每一个时刻在图像中的位置不一样，光流法就是通过这个假设来计算像素点的位置。灰度不变假设很理想，通常都不成立，但仍然有理论贡献。

计算部分像素运动称为悉数光流，以Lucas-Kanade法为代表；计算所有像素运动的称为稠密光流。下面介绍的是LK光流法。

根据前面的描述，可以把图像看成是时间的函数，把一个t时刻处于(x,y)位置像素的灰度记为I(x,y,t)，经过dt时间后它运动到(x+dx, y+dy)处。根据灰度不变假设有：$\boldsymbol{I(x+dy,y+dy,t+dt)=I(x,y,t)}$ 。在这个问题中相当于前后的灰度值已知，求解dx,dy。

对变化后的灰度值进行一阶泰勒展开，有：

$$\boldsymbol{I}(x+\mathrm{d}x,y+\mathrm{d}y,t+\mathrm{d}t)\approx\boldsymbol{I}(x,y,t)+\frac{\partial\boldsymbol{I}}{\partial x}\mathrm{d}x+\frac{\partial\boldsymbol{I}}{\partial y}\mathrm{d}y+\frac{\partial\boldsymbol{I}}{\partial t}\mathrm{d}t.$$

根据灰度不变假设可以得到：

$$\frac{\partial\boldsymbol{I}}{\partial x}\mathrm{d}x+\frac{\partial\boldsymbol{I}}{\partial y}\mathrm{d}y+\frac{\partial\boldsymbol{I}}{\partial t}\mathrm{d}t=0.$$

整理得到：

$$\frac{\partial\boldsymbol{I}}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial\boldsymbol{I}}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}=-\frac{\partial\boldsymbol{I}}{\partial t}.$$

其中，$\boldsymbol{\frac{\partial I}{\partial x},\frac{\partial I}{\partial y}}$ 分别为图像灰度在该点处x方向和y方向的梯度，记为 $\boldsymbol{I_x,I_y}$ ，$\boldsymbol{\frac{\partial I}{\partial t}}$ 是图像灰度对时间的变化量，记为 $\boldsymbol{I_t}$,这三个量已知。$\boldsymbol{\frac{dx}{dt},\frac{dy}{dt}}$ 分别表示像素在x和y轴上的运动速度，记为 $\boldsymbol{u,v}$ 未知。对于图像窗口中的每个像素都可以列出这样的一个方程，如果认为窗口内的像素具有相同的运动，那么每个像素的u,v相同，可以写成一个方程组，形如 $\boldsymbol{A_{N\times2}\cdot\begin{bmatrix}u\\v\end{bmatrix}_{2\times1}=-b_{N\times1}}$ ，可以通过最小二乘法求解，得到：

$$\begin{bmatrix}u\\v\end{bmatrix}^{*}=-\left(\boldsymbol{A}^{T}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{T}\boldsymbol{b}.$$

2. 直接法

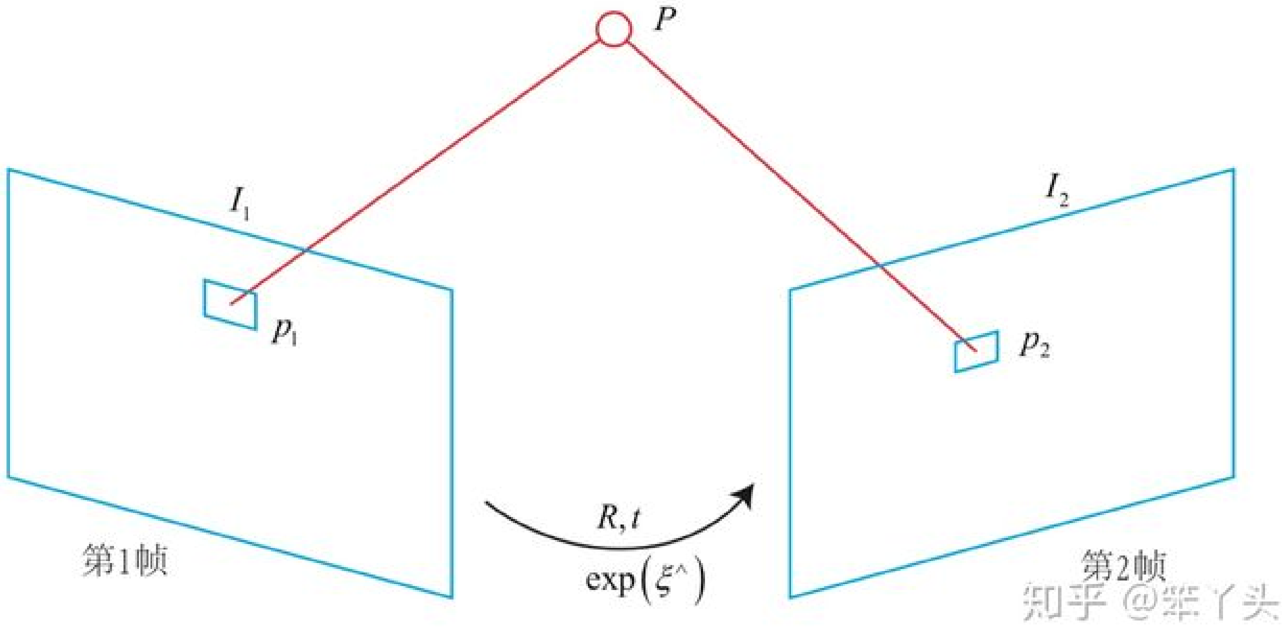

图8-3

如图8-3所示，已知相机内参K，物体P的3D坐标，分别知道第1帧和第2帧上的投影坐标 p_1, p_2 ，但由于没有特征匹配的过程，所以不知道 p_1 和 p_2 是否对应同一个物体P。

直接法的思路是：根据灰度不变假设，对于第1帧中的 p_1 能够在第2帧中摸索到一个点 p_2 ，使得这个点的灰度和 p_1 的灰度误差最小。摸索的过程是：通过相机模型计算出 p_2 的坐标，然后通过调整相机位姿最小化光度误差 $e_1 = I_1(p_1) - I_2(p_2)$ 。

根据目前的已知量，可以计算第1帧的像素坐标为 $p_1 = \frac{1}{Z_1}KP$ ，第2帧的像素坐标为 $p_2 = \frac{1}{Z_2}K(RP+t) = \frac{1}{Z_2}K(exp(\xi^\wedge)P)_{1:3}$ 。

我们的目标函数为：

$$\min_{\boldsymbol{\xi}} J(\boldsymbol{\xi}) = \sum_{i=1}^{N} e_i^T e_i, \quad e_i = \boldsymbol{I}_1(\boldsymbol{p}_{1,i}) - \boldsymbol{I}_2(\boldsymbol{p}_{2,i}).$$

扰动模型求导：

$$\begin{aligned} e(\boldsymbol{\xi} \oplus \delta\boldsymbol{\xi}) &= \boldsymbol{I}_1\left(\frac{1}{Z_1}\boldsymbol{KP}\right) - \boldsymbol{I}_2\left(\frac{1}{Z_2}\boldsymbol{K}\exp(\delta\boldsymbol{\xi}^\wedge)\exp(\boldsymbol{\xi}^\wedge)\boldsymbol{P}\right) \\ &\approx \boldsymbol{I}_1\left(\frac{1}{Z_1}\boldsymbol{KP}\right) - \boldsymbol{I}_2\left(\frac{1}{Z_2}\boldsymbol{K}(1+\delta\boldsymbol{\xi}^\wedge)\exp(\boldsymbol{\xi}^\wedge)\boldsymbol{P}\right) \\ &= \boldsymbol{I}_1\left(\frac{1}{Z_1}\boldsymbol{KP}\right) - \boldsymbol{I}_2\left(\frac{1}{Z_2}\boldsymbol{K}\exp(\boldsymbol{\xi}^\wedge)\boldsymbol{P} + \frac{1}{Z_2}\boldsymbol{K}\delta\boldsymbol{\xi}^\wedge\exp(\boldsymbol{\xi}^\wedge)\boldsymbol{P}\right). \end{aligned}$$

定义中间变量： $q=\delta\xi^{\wedge}exp(\xi^{\wedge})P, u=\frac{1}{Z_2}Kq$ ，有

$$\begin{aligned}
e(\boldsymbol{\xi}\oplus\delta\boldsymbol{\xi}) &= \boldsymbol{I}_1\left(\frac{1}{Z_1}\boldsymbol{K}\boldsymbol{P}\right)-\boldsymbol{I}_2\left(\frac{1}{Z_2}\boldsymbol{K}\exp\left(\boldsymbol{\xi}^{\wedge}\right)\boldsymbol{P}+\boldsymbol{u}\right) \\
&\approx \boldsymbol{I}_1\left(\frac{1}{Z_1}\boldsymbol{K}\boldsymbol{P}\right)-\boldsymbol{I}_2\left(\frac{1}{Z_2}\boldsymbol{K}\exp\left(\boldsymbol{\xi}^{\wedge}\right)\boldsymbol{P}\right)-\frac{\partial \boldsymbol{I}_2}{\partial \boldsymbol{u}}\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{q}}\frac{\partial \boldsymbol{q}}{\partial \delta\boldsymbol{\xi}}\delta\boldsymbol{\xi} \\
&= e(\boldsymbol{\xi})-\frac{\partial \boldsymbol{I}_2}{\partial \boldsymbol{u}}\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{q}}\frac{\partial \boldsymbol{q}}{\partial \delta\boldsymbol{\xi}}\delta\boldsymbol{\xi}.
\end{aligned}$$

其中， $\frac{\partial I_2}{\partial u}$ 是图像的像素梯度，另外两个偏到前面已经求解过：

$$\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{q}}=\begin{bmatrix}\frac{\partial u}{\partial X} & \frac{\partial u}{\partial Y} & \frac{\partial u}{\partial Z} \\ \frac{\partial v}{\partial X} & \frac{\partial v}{\partial Y} & \frac{\partial v}{\partial Z}\end{bmatrix}=\begin{bmatrix}\frac{f_x}{Z} & 0 & -\frac{f_x X}{Z^2} \\ 0 & \frac{f_y}{Z} & -\frac{f_y Y}{Z^2}\end{bmatrix}.$$

$$\frac{\partial \boldsymbol{q}}{\partial \delta\boldsymbol{\xi}}=\left[\boldsymbol{I},-\boldsymbol{q}^{\wedge}\right].$$

由此求出了一阶导数也就是雅各比矩阵，根据雅各比矩阵进行迭代求解。

关于P，如果来自于稀疏关键点，称为稀疏直接法；如果来自于部分梯度明显的像素，称为半稠密的直接法（如果梯度不明显，雅各比矩阵为0）；如果是所有像素，称为稠密直接法。

3. 直接法讨论

直接法是根据图像灰度的变化（梯度）来寻找对应点的，比如图8-6，图像1中P1点的灰度值为229，目标是找到图像2中灰度值和p1相近的点。在P1周围沿着梯度方向搜索，第一次找到P2，继续搜索找到P2' 发现很合适，进行更新，如此迭代。

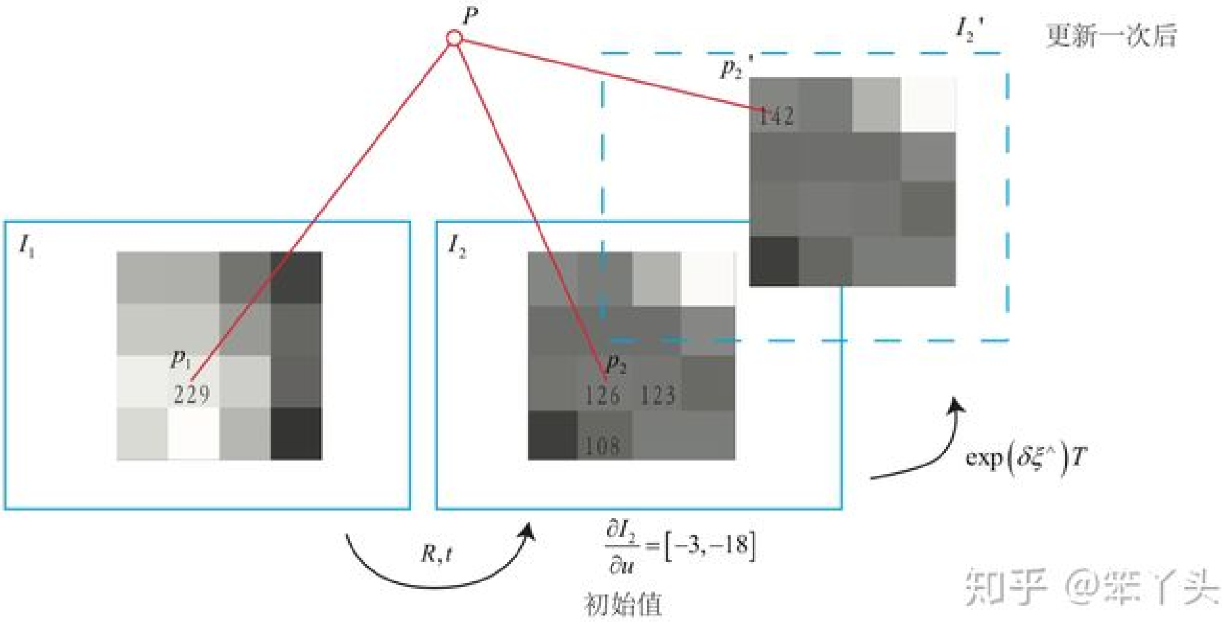

图8-6

直接法的问题和梯度下降法类似，可能会陷入局部最优解。如图8-7所示，图像本身通常是一个强烈的非凸函数，可能在搜索最优点时会找到附近的一个局部最优解。

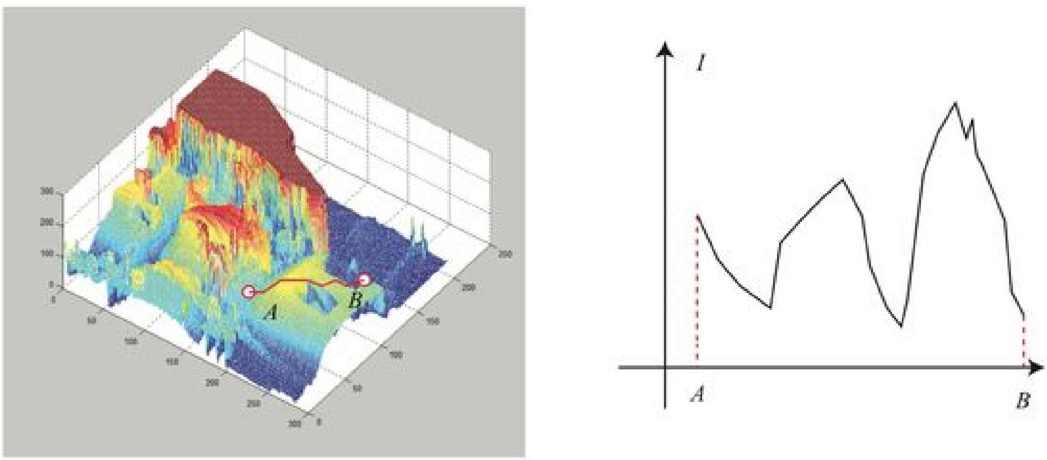

图 8-7　一张图像的三维化显示。从图像中的一个点运动到另一个点的路径不见得是“笔直的下坡路”，而需要经常的“翻山越岭”。这体现了图像本身的非凸性。

为了避免这个问题，可以缩小范围，采用图像块，使用更复杂的差异度量方式等。

直接法的优点是：快，只要求像素梯度，能构建半稠密或稠密地图；缺点是：图像具有非凸性，单个像素区分度低需要计算图像块，灰度值不变假设可能不成立。

[1] 《视觉SLAM十四讲从理论到实践》 高翔，张涛