深蓝学院

shenlanxueyuan.com

视觉SLAM：从理论到实践
第六次课 光流法与直接法

主讲人 高翔

清华大学 自动控制与工程 博士
慕尼黑工业大学计算机视觉组 博士后
Email: gao.xiang.thu@gmail.com

第六讲 光流法与直接法

深蓝学院
shenlanxueyuan.com

1. 光流法
2. 直接法

1. 光流法

1. 光流法

- 特征点法流程：
  1. 在图像中提取特征点并计算特征描述　　非常耗时 ~10+ms in ORB
  2. 在不同图像中寻找特征匹配　　非常耗时 O(n^2) in brute force matching
  3. 利用匹配点信息计算相机位姿　　比较快速 < 1ms

- 能否不使用特征匹配计算VO？

1. 光流法

- 不使用特征匹配的思路：
  - 通过其他方式寻找配对点：光流
  - 不需要配对点：直接法

1. 光流法

- 光流：追踪源图像某个点在其他图像中的运动
- 一般分为稀疏光流和稠密光流
  - 稀疏以Lucas-Kanade（LK）光流为代表
  - 稠密以Horn－Schunck（HS）光流为代表
  - 本质上是估计像素在不同时刻图像中的运动

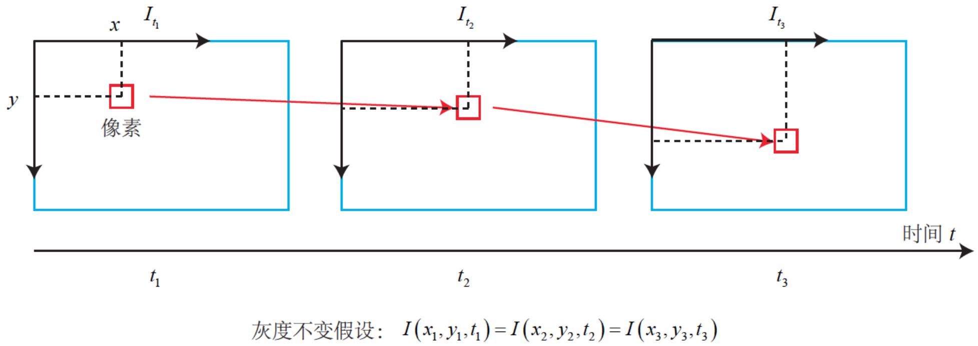


图 8-1　LK 光流法示意图。

1. 光流法

- 设 t 时刻位于 x, y 处像素点的灰度值为 $\boldsymbol{I}(x, y, t)$.
- 在 t+dt 时刻，该像素运动到了 $\boldsymbol{I}(x+\mathrm{d}x, y+\mathrm{d}y, t+\mathrm{d}t)$
- 希望计算运动 dx， dy

- 灰度不变假设：$\boldsymbol{I}(x+\mathrm{d}x, y+\mathrm{d}y, t+\mathrm{d}t)=\boldsymbol{I}(x, y, t)$.
- 注意：灰度不变是一种理想的假设，实际当中由于高光/阴影/材质/曝光等不同，很可能不成立。

1. 光流法

深蓝学院
shenlanxueyuan.com

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

- 对 t+dt 时刻的灰度进行Taylor展开并保留一阶项：

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

- 由于灰度不变，所以

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

因此 =>

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

x方向梯度　　　随时间变化

y方向梯度

- 希望求解dx/dt，dy/dt

1. 光流法

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

- 但本式是一个二元一次线性方程，欠定
  - 需要引用额外的约束
  - 假定一个窗口（ w´ w）内光度不变：

$$\begin{bmatrix} \boldsymbol{I}_x & \boldsymbol{I}_y \end{bmatrix}_k \begin{bmatrix} u \\ v \end{bmatrix} = -\boldsymbol{I}_{tk}, \quad k=1,\ldots,w^2.$$

  - 通过超定最小二乘解求得运动 u,v

$$\boldsymbol{A}=\begin{bmatrix} [\boldsymbol{I}_x,\boldsymbol{I}_y]_1 \\ \vdots \\ [\boldsymbol{I}_x,\boldsymbol{I}_y]_k \end{bmatrix}, \boldsymbol{b}=\begin{bmatrix} \boldsymbol{I}_{t1} \\ \vdots \\ \boldsymbol{I}_{tk} \end{bmatrix}.$$

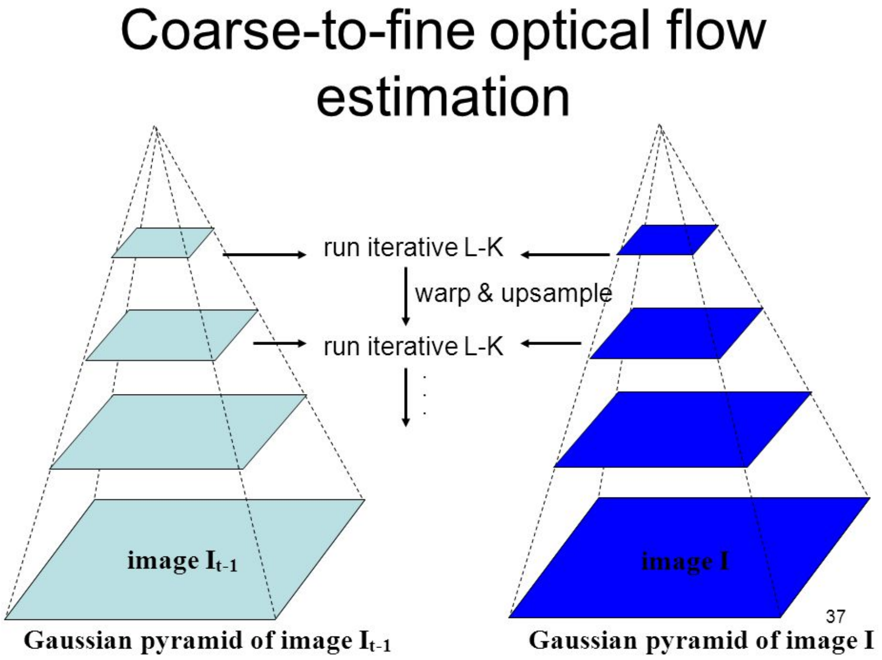
$$\begin{bmatrix} u \\ v \end{bmatrix}^* = -\left(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}.$$

1. 光流法

LK光流的结果依赖于图像梯度

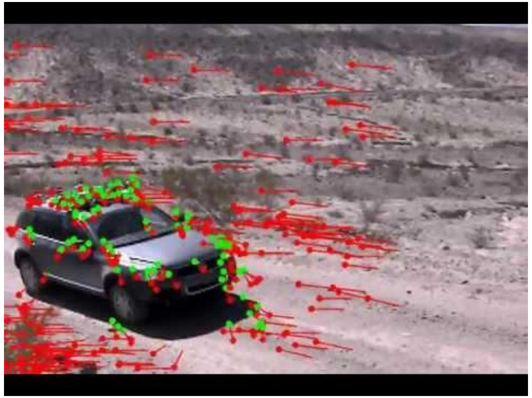
- 但梯度不够平滑，可能剧烈变化
- 局部的梯度不能用于预测长期图像走向

- 解决方式：多层光流



1. 光流法

- 注解：
  - 可以看成最小化像素误差的非线性优化
  - 每次使用了Taylor一阶近似，在离优化点较远时效果不佳，往往需要迭代多次
  - 运动较大时要使用金字塔
  - 可以用于跟踪图像中的稀疏关键点的运动轨迹
  - 得到配对点后，后续计算与特征法VO中相同
  - 按方法可分为正向/反向+平移/组合的方式，具体在习题中介绍



2. 实践：LK光流

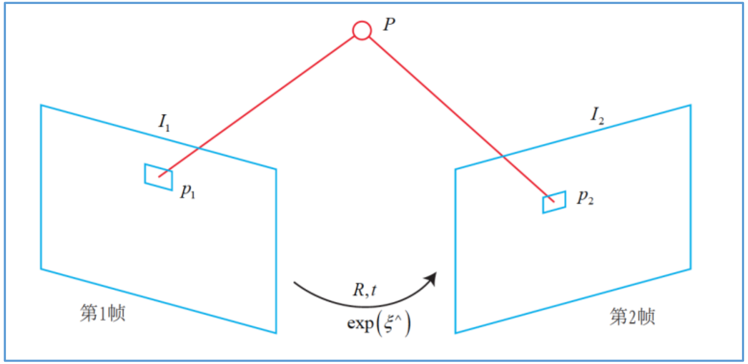
3. 直接法

3. 直接法

- 光流仅估计了像素间的平移，但
  - 没有用到相机本身的几何结构
  - 没有考虑到相机的旋转和图像的缩放
  - 对于边界上的点，光流不好追踪
  - 直接法则考虑了这些信息

3. 直接法

- 直接法的推导
  - 假设有两个帧，运动未知，但有初始估计 R,t
  - 第1帧上看到了点P，投影为p1
  - 按照初始估计，P在第2帧上投影为p2



投影关系：

$$\boldsymbol{p}_1 = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_1 = \frac{1}{Z_1}\boldsymbol{K}\boldsymbol{P},$$

$$\boldsymbol{p}_2 = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_2 = \frac{1}{Z_2}\boldsymbol{K}\left(\boldsymbol{R}\boldsymbol{P} + \boldsymbol{t}\right) = \frac{1}{Z_2}\boldsymbol{K}\left(\exp\left(\boldsymbol{\xi}^\wedge\right)\boldsymbol{P}\right)_{1:3}.$$

3. 直接法

- 为了估计相机的运动，建立最小化问题

最小化光度误差：$e = \boldsymbol{I}_1(\boldsymbol{p}_1) - \boldsymbol{I}_2(\boldsymbol{p}_2)$. $\min_{\boldsymbol{\xi}} J(\boldsymbol{\xi}) = \sum_{i=1}^{N} e_i^{\mathrm{T}} e_i, \quad e_i = \boldsymbol{I}_1(\boldsymbol{p}_{1,i}) - \boldsymbol{I}_2(\boldsymbol{p}_{2,i})$.

- 待估计的量为相机运动 x ，我们关心误差相对于相机的导数

3. 直接法

深蓝学院 shenlanxueyuan.com

$$e(\delta T \oplus T) = I_1(p_1) - I_2\left(\frac{1}{Z_2}K\exp(\delta\xi^\wedge)TP\right)$$

$$\approx I_1(p_1) - I_2\left(\frac{1}{Z_2}K(1+\delta\xi^\wedge)TP\right)$$

$$= I_1(p_1) - I_2\left(p_2 + \frac{1}{Z_2}K\delta\xi^\wedge TP\right)$$

$$\boldsymbol{q} = \delta\boldsymbol{\xi}^\wedge \exp\left(\boldsymbol{\xi}^\wedge\right)\boldsymbol{P},$$

$$\boldsymbol{u} = \frac{1}{Z_2}\boldsymbol{K}\boldsymbol{q}.$$

$$e(\delta T \oplus T) = I_1(p_1) - I_2(p_2 + u)$$

$$\approx I_1(p_1) - I_2(p_2) - \frac{\partial I_2}{\partial u}\frac{\partial u}{\partial q}\frac{\partial q}{\partial \delta\xi}\delta\xi$$

$$= e - \frac{\partial I_2}{\partial u}\frac{\partial u}{\partial q}\frac{\partial q}{\partial \delta\xi}\delta\xi$$

3. 直接法

$$e(\boldsymbol{\xi}) - \frac{\partial \boldsymbol{I}_2}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{q}} \frac{\partial \boldsymbol{q}}{\partial \delta \boldsymbol{\xi}} \delta \boldsymbol{\xi}.$$

- 第一部分：图像梯度
- 第二部分：像素对投影点导数（见上一章）

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

- 第三部分：投影点对位姿导数： $\frac{\partial \boldsymbol{q}}{\partial \delta \boldsymbol{\xi}} = [\boldsymbol{I}, -\boldsymbol{q}^\wedge].$
- 综上：

$$\boldsymbol{J} = -\frac{\partial \boldsymbol{I}_2}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \delta \boldsymbol{\xi}}. \qquad \frac{\partial \boldsymbol{u}}{\partial \delta \boldsymbol{\xi}} = \begin{bmatrix} \frac{f_x}{Z} & 0 & -\frac{f_x X}{Z^2} & -\frac{f_x XY}{Z^2} & f_x + \frac{f_x X^2}{Z^2} & -\frac{f_x Y}{Z} \\ 0 & \frac{f_y}{Z} & -\frac{f_y Y}{Z^2} & -f_y - \frac{f_y Y^2}{Z^2} & \frac{f_y XY}{Z^2} & \frac{f_y X}{Z} \end{bmatrix}.$$

3. 直接法

$$e(\boldsymbol{\xi}) - \frac{\partial \boldsymbol{I}_2}{\partial \boldsymbol{u}} \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{q}} \frac{\partial \boldsymbol{q}}{\partial \delta \boldsymbol{\xi}} \delta \boldsymbol{\xi}.$$

- 可以看到，直接法的雅可比项有一个图像梯度因子
  - 因此，在图像梯度不明显的地方，对相机运动估计的贡献就小
- 根据使用的图像信息不同，可分为：
  - 稀疏直接法：只处理稀疏角点或关键点
  - 稠密直接法：使用所有像素
  - 半稠密直接法：使用部分梯度明显的像素

3. 直接法

- 直接法的直观解释
  - 像素灰度引导着优化的方向
  - 要使优化成立，必须保证从初始估计到最优估计中间的梯度一直下降
  - 这很容易受到图像非凸性的影响（可部分地由金字塔减轻）

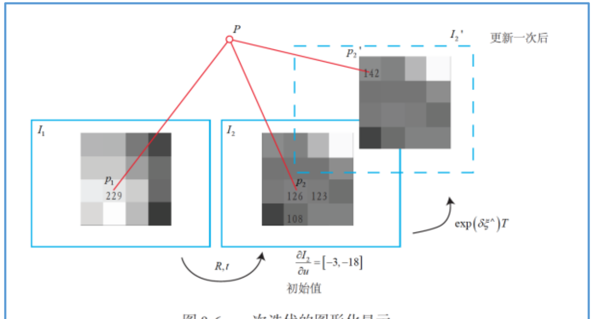
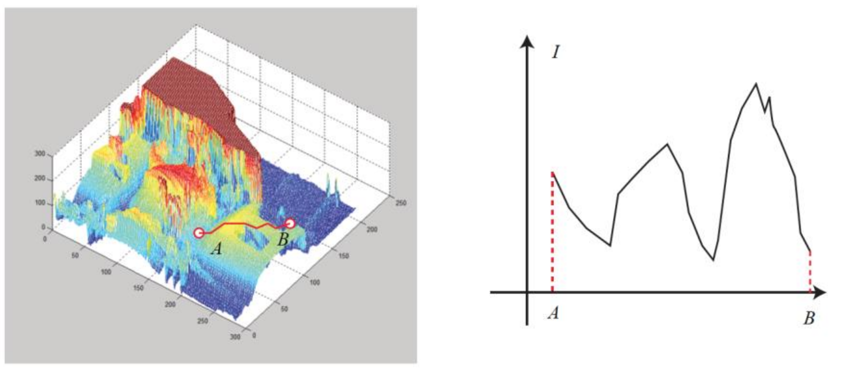


图 8-6 一次迭代的图形化显示。



3. 直接法

- 优缺点小结
- 优势
  - 省略特征提取的时间
  - 只需有像素梯度而不必是角点（对白墙等地方有较好效果）
  - 可稠密或半稠密
- 劣势
  - 灰度不变难以满足（易受曝光和模糊影响）
  - 单像素区分性差
  - 图像非凸性

4. 实践：直接法