Notes on Visual SLAM 14 lectures – Chapter 8

1 Motivation of Direct Method

A. extraction of key points and the calculation of the descriptor are very time-consuming. (SIFT cannot be done in real-time CPU and ORB always requires 20ms of calculation)

B. not all the information is used in the feature method

C. camera sometimes moves to featureless places

solutions:

A. keep feature points but discard their descriptors

B. only calculate key points not descriptors

2 2D Optical Flow

Dense optical flow: calculation of all pixels in an image Sparse optical flow: Lucas-Kanade optical flow

2.1 Lucas-Kanade Optical Flow

image can be regarded as a function of time $\mathbf{I}(t)$, for a pixel (x, y) at time t, its grayscale can be written as:

$$\mathbf{I}(x, y, t) \tag{1}$$

Constant grayscale assumption: the pixel's grayscale is constant in each image

suppose pixel at (x, y) at time t moves to $(x + dx, y + dy)$ at time $t + dt$:

$$\mathbf{I}(x + dx, y + dy, t + dt) = \mathbf{I}(x, y, t) \tag{2}$$

the assumption of constant brightness is not true in practice. Carry out Taylor expansion on the left side:

$$\mathbf{I}(x + dx, y + dy, t + dt) \approx \mathbf{I}(x, y, t) + \frac{\partial \mathbf{I}}{\partial x}dx + \frac{\partial \mathbf{I}}{\partial y}dy + \frac{\partial \mathbf{I}}{\partial t}dt \tag{3}$$

since we assume constant brightness, we have:

$$\frac{\partial \mathbf{I}}{\partial x}dx + \frac{\partial \mathbf{I}}{\partial y}dy + \frac{\partial \mathbf{I}}{\partial t}dt = 0 \tag{4}$$

divide both sides by dt:

$$\frac{\partial \mathbf{I}}{\partial x}\frac{dx}{dt} + \frac{\partial \mathbf{I}}{\partial y}\frac{dy}{dt} = -\frac{\partial \mathbf{I}}{\partial t} \tag{5}$$

denote dx/dt as u, and dy/dt as v, $\partial\mathbf{I}/\partial x$ as $\mathbf{I}_x$, $\partial\mathbf{I}/\partial y$ as $\mathbf{I}_y$, and $\partial\mathbf{I}/\partial t$ as $\mathbf{I}_t$:

$$\begin{bmatrix} \mathbf{I}_x & \mathbf{I}_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -\mathbf{I}_t \tag{6}$$

Consider a window of size w x w, which contains w^2 pixels, since the pixels in this window are assumed to have the same motion, we have a total of w^2 equations:

$$\begin{bmatrix}\mathbf{I}_x & \mathbf{I}_y\end{bmatrix}_k \begin{bmatrix}u \\ v\end{bmatrix} = -\mathbf{I}_{tk}, \quad k = 1, ..., w^2 \tag{7}$$

stack them:

$$\mathbf{A} = \begin{bmatrix}\begin{bmatrix}\mathbf{I}_x & \mathbf{I}_y\end{bmatrix}_1 \\ \begin{bmatrix}\mathbf{I}_x & \mathbf{I}_y\end{bmatrix}_2 \\ \vdots \\ \begin{bmatrix}\mathbf{I}_x & \mathbf{I}_y\end{bmatrix}_k\end{bmatrix}, \mathbf{b} = \begin{bmatrix}\mathbf{I}_{t1} \\ \mathbf{I}_{t2} \\ \vdots \\ \mathbf{I}_{tk}\end{bmatrix} \tag{8}$$

then we have:

$$\mathbf{A}\begin{bmatrix}u \\ v\end{bmatrix} = -\mathbf{b} \tag{9}$$

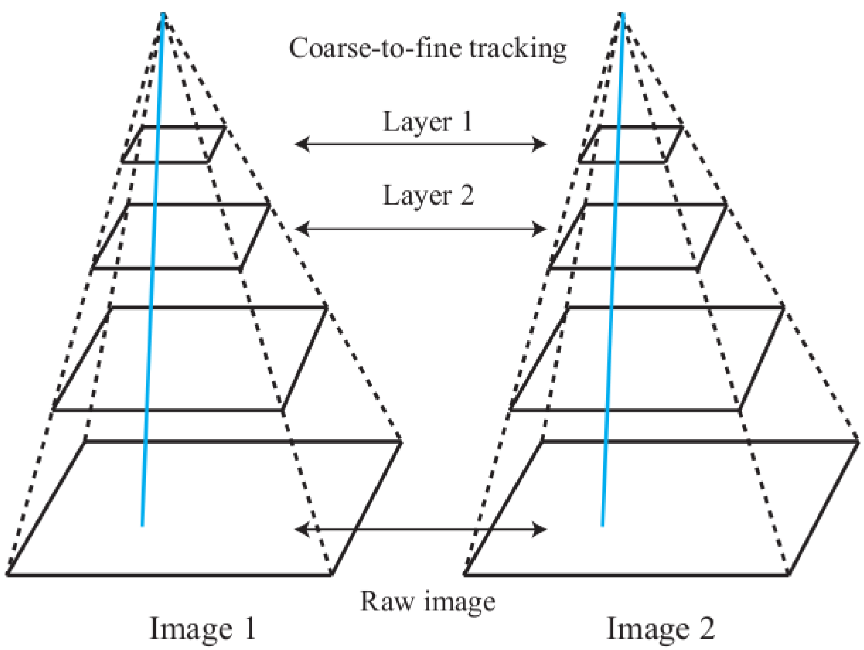
we can find its least-square solution:

$$\begin{bmatrix}u \\ v\end{bmatrix}^* = -(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \tag{10}$$

2.1.1 Single-layer optical flow

$$\min_{\Delta x, \Delta y} ||\mathbf{I}_1(x, y) - \mathbf{I}_2(x + \Delta x, y + \Delta y)||_2^2 \tag{11}$$

residual is inside the brackets, and the corresponding Jacobian is the gradient of the second image at



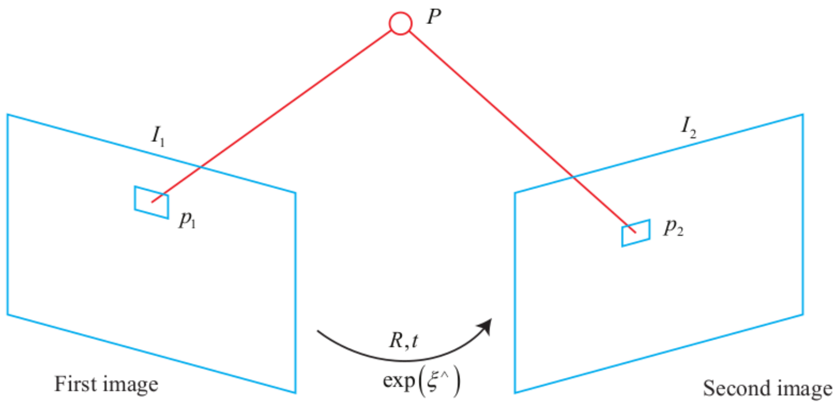
$x + \Delta x, y + \Delta y$. The gradient can also be replaced by the gradient $\mathbf{I}_1(x, y)$ of the first image.

Inverse gradient method: gradient of $\mathbf{I}_1(x, y)$ remains unchanged, so we can use the result calculated in the first iteration in the subsequent iterations.

2.1.2 Multi-layer optical flow

If camera moves fast and difference between the two images is obvious, the single-layer image optical flow can be easily stuck at a local minimum, but it can be resolved by image pyramids. start from the top image, and go down one by one using the previous layer's tracking result as the initial value of the next layer's optical flow (coarse-to-fine optical flow)

2.2 Direct Method



we can get two projection equation:

$$\mathbf{p}_1 = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_1 = \frac{1}{Z_1}\mathbf{KP} \tag{12}$$

$$\mathbf{p}_2 = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}_2 = \frac{1}{Z_2}\mathbf{K}(\mathbf{RP}+\mathbf{t}) = \frac{1}{Z_2}\mathbf{K}(\mathbf{TP})_{1:3} \tag{13}$$

minimize the photometric error:

$$e = \mathbf{I}_1(\mathbf{p}_1) - \mathbf{I}_2(\mathbf{p}_2) \tag{14}$$

the objective then becomes:

$$\min_{\mathbf{T}} J(\mathbf{T}) = ||e||^2 \tag{15}$$

which still requires constant brightness assumption. For N space points P_i, the whole camera pose estimation problem becomes:

$$\min_{\mathbf{T}} J(\mathbf{T}) = \sum_{i=1}^{N} e_i^T e_i, \quad e_i = \mathbf{I}_1(\mathbf{p}_{1,i}) - \mathbf{I}_2(\mathbf{p}_{2,i}) \tag{16}$$

define two intermediate variables:

$$\mathbf{q} = \mathbf{TP}$$
$$\mathbf{u} = \frac{1}{Z_2}\mathbf{Kq}$$

in which $\mathbf{q}$ is the coordinate of P in the second camera coordinate system, and $\mathbf{u}$ is its pixel coordinates. Consider the left perturbation model in Lie algebra:

$$e_i = \mathbf{I}_1(\mathbf{p}_1) - \mathbf{I}_2(\mathbf{u}) \tag{17}$$

then we get:

$$\frac{\partial e}{\partial \mathbf{T}} = \frac{\partial \mathbf{I}_2}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \delta\xi}\delta\xi \tag{18}$$

A. The first term is the grayscale gradient at pixel $\mathbf{u}$

B. the second term is the derivative of the projection equation with respect to the 3D point in the camera frame:

$$\frac{\partial \mathbf{u}}{\partial \mathbf{q}} = \begin{bmatrix} \frac{\partial u}{\partial X} & \frac{\partial u}{\partial Y} & \frac{\partial u}{\partial Z} \\ \frac{\partial v}{\partial X} & \frac{\partial v}{\partial Y} & \frac{\partial v}{\partial Z} \end{bmatrix} = \begin{bmatrix} \frac{f_x}{Z} & 0 & -\frac{f_x X}{Z^2} \\ 0 & \frac{f_y}{Z} & -\frac{f_y Y}{Z^2} \end{bmatrix} \tag{19}$$

C. the third term is the derivative of the transformed 3D point with respect to the transformation:

$$\frac{\partial \mathbf{q}}{\partial \delta \xi} = \begin{bmatrix} \mathbf{I} & -\mathbf{q}^\wedge \end{bmatrix} \tag{20}$$

we then get:

$$\frac{\mathbf{u}}{\partial \delta \xi} = \begin{bmatrix} \frac{f_x}{Z} & 0 & -\frac{f_x X}{Z^2} & -\frac{f_x XY}{Z^2} & f_x + \frac{f_x X^2}{Z^2} & -\frac{f_x Y}{Z} \\ 0 & \frac{f_y}{Z} & -\frac{f_y Y}{Z^2} & -f_y - \frac{f_y Y^2}{Z^2} & \frac{f_y XY}{Z^2} & \frac{f_y X}{Z} \end{bmatrix} \tag{21}$$

and we have:

$$\mathbf{J} = -\frac{\partial \mathbf{I}_2}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \delta \xi} \tag{22}$$