A pure translation

$W(x;p) = x + p.$

Starting from a rectangular neighborhood of pixels $N \in \{x_d\}_{d=1}^{D}$ on frame I_t, the LK tracker aims to move it by an offset $p = [p_x, p_y]^T$ to obtain another rectangle on the next frame I_{t+1}. To minimize the pixel square difference:

$$p^* = \arg\min_p = \sum_{x \in N} \| I_{t+1}(x+p) - I_t(x) \|^2$$

Starting with an initial guess of p ($p = [0,0]^T$), we can compute the optimal p^* iteratively.

$$I_{t+1}(x' + \Delta p) = I_{t+1}(x') + \frac{\partial I_{t+1}(x')}{\partial x'^T} \frac{\partial W(x;p)}{\partial p^T} \Delta p$$

where $\Delta p = [\Delta p_x, \Delta p_y]^T$ is the change in template offset. $x' = W(x;p) = x + p$. and $\frac{\partial I(x')}{\partial x'^T}$ is a vector of x- and y- image gradients at pixel coordinate x'. Incorporate:

$$\arg\min_{\Delta p} \| A \Delta p - b \|_2^2$$

①

$$W(x;p) = x + p = \begin{bmatrix} x + p_1 \\ y + p_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & p_1 \\ 0 & 1 & p_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\frac{\partial W(x;p)}{\partial p^T} = \begin{bmatrix} \frac{\partial W_x}{\partial p_1} & \cdots & \frac{\partial W_x}{\partial p_n} \\ \frac{\partial W_y}{\partial p_1} & \cdots & \frac{\partial W_y}{\partial p_n} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

② $$p^* = \sum_{x \in N} \| I_{t+1}(x' + \Delta p) - I_t(x) \|_2^2$$

$$= \sum_{x \in N} \left\| \frac{\partial I_{t+1}(x')}{\partial x'^T} \overbrace{\cancel{\frac{\partial W(x;p)}{\partial p^T}}}^{\text{Identity Matrix}} \Delta p - (I_t(x) - I_{t+1}(x')) \right\|_2^2$$

So: $$A = \begin{bmatrix} \frac{\partial I_{t+1}(x'_1)}{\partial x'^T} \\ \vdots \\ \frac{\partial I_{t+1}(x'_D)}{\partial x'^T} \end{bmatrix} \qquad b = \begin{bmatrix} I_t(x_1) - I_{t+1}(x'_1) \\ \vdots \\ I_t(x_D) - I_{t+1}(x'_D) \end{bmatrix}$$

③ $$\underset{\Delta p}{\arg\min} \| A \Delta p - b \|_2^2 \Rightarrow \Delta p = \underbrace{(A^T A)}_{\text{Hessian Matrix}}{}^{-1} A^T b$$

Affine

$$\Delta p = [p_1, \cdots, p_6]^T$$

$$x' = W(x;p) = \begin{bmatrix} 1+p_1 & p_2 \\ p_4 & 1+p_5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p_3 \\ p_6 \end{bmatrix}$$

Then we can represent this affine warp in homogenous coordinates as:

$$\overset{\frown}{x}' = M \overset{\frown}{x}$$

where $M = \begin{bmatrix} 1+p_1 & p_2 & p_3 \\ p_4 & 1+p_5 & p_6 \\ 0 & 0 & 1 \end{bmatrix}$ M will differ between successive image pairs

$$p^* = \sum_{x \in N} \| I_{t+1}(x' + \Delta p) - \overset{p_n}{I_t}(x) \|_2^2$$

$$= \sum_{x \in N} \left\| \frac{\partial I_{t+1}(x')}{\partial x'^T} \frac{\partial W(x;p)}{\partial p^T} \Delta p - (I_t(x) - I_{t+1}(x')) \right\|_2^2$$

$$W(x;p) = \begin{bmatrix} 1+p_1 & p_2 & p_3 \\ p_4 & 1+p_5 & p_6 \\ 0 & 0 & 1 \end{bmatrix} \quad \frac{\partial W(x;p)}{\partial p^T} = \begin{bmatrix} \frac{\partial W_x}{\partial p_1} & \cdots & \frac{\partial W_x}{\partial p_n} \\ \frac{\partial W_y}{\partial p_1} & \cdots & \frac{\partial W_y}{\partial p_n} \end{bmatrix}$$

$$= \begin{bmatrix} x & y & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 \end{bmatrix}$$

Efficient tracking — Inverse Composition LK

In each iteration, try to find the inverse warping $W(x;\Delta p)^{-1}$ that best align our current tracking rectangle $I_{t+1}(W(x;p))$ in the next frame with the template in the current frame $I_t(x)$. This is equivalent to finding the warping on the template $I_t(W(x;\Delta p))$ that best aligns it with $I_{t+1}(W(x;p))$.

$$p^* = \arg\min_p = \sum_{x \in N} \| I_t(W(x;\Delta p)) - I_{t+1}(W(x;p)) \|_2^2$$

where we linearize just around the current frame I_t as:

$$I_t(W(x;0+\Delta p)) \simeq I_t(x) + \underbrace{\frac{\partial I_t(x)}{\partial x^T}}_{\nabla T} \frac{\partial W(x;0)}{\partial p^T} \Delta p$$

$$\Rightarrow \arg\min_{\Delta p} \| A'\Delta p - b' \|_2^2$$

Update: $W(x;p) \leftarrow W(W(x;\Delta p)^{-1}, p)$

using affine warp where $p \Leftrightarrow M$ and $\Delta p \Leftrightarrow \Delta M$, results in $M \leftarrow M(\Delta M)^{-1}$.

Then Hessian Matrix is:

$$H = A'^T \cdot A' \quad \sum_{x \in N} \left[\nabla T \frac{\partial W}{\partial p}\right]^T \left[\nabla T \frac{\partial W}{\partial p}\right],$$

The Jacobian $\frac{\partial W}{\partial p}$ is evaluated at $(x;0)$, so H can be pre-computed.