Multiple View Geometry: Solution Sheet 5

Prof. Dr. Daniel Cremers,
Shenhan Qian, Simon Weber, Anna Ribic, and Tarun Yenamandra
Computer Vision Group, TU Munich
Wednesdays 16:15-18:15 at Hörsaal 2, "Interims I" (5620.01.102), and on RBG Live

Exercise: June 4th, 2025

1. **The Lucas-Kanade method**

   (a) Prove that the minimizer $\hat{\mathbf{v}}$ of $E(\mathbf{v})$ can be written as

$$\hat{\mathbf{v}} = -M^{-1}\mathbf{q}$$

where the entries of M and $\mathbf{q}$ are given by

$$m_{ij} = G * \left(I_{x_i} \cdot I_{x_j}\right) \quad \text{and} \quad q_i = G * (I_{x_i} \cdot I_t)$$

In general holds for real valued vectors

$$\|A\mathbf{x} + \mathbf{b}\|^2 = \langle A\mathbf{x} + \mathbf{b}, A\mathbf{x} + \mathbf{b}\rangle = \langle A\mathbf{x}, A\mathbf{x}\rangle + 2\langle A\mathbf{x}, \mathbf{b}\rangle + \langle \mathbf{b}, \mathbf{b}\rangle$$

Now, for each term in the sum we take the derivative (gradient) with respect to $\mathbf{x}$:

$$\nabla_{\mathbf{x}}\|A\mathbf{x} + \mathbf{b}\|^2 = 2A^T A\mathbf{x} + 2A^T\mathbf{b}$$

We then set the gradient to zero to find the minimizer

$$0 = 2A^T A\mathbf{x} + 2A^T\mathbf{b}$$
$$\Rightarrow \mathbf{x} = -(A^T A)^{-1}A^T\mathbf{b}$$

Now we can apply this to the given energy function

$$E(\mathbf{v}) = G * (A\mathbf{v} + b)$$

with
$A = \nabla I(\mathbf{x}', t)^\top$, and $\mathbf{b} = \partial_t I(\mathbf{x}', t)$. Note that the derivative is linear and thus a linear combination of derivatives can be written as a derivative of a linear combination. This also holds for the convolution with the kernel G:, as convolution is a linear operation and G does not depend on $\mathbf{v}$. Thus, we can write

$$\begin{aligned}\nabla_{\mathbf{v}} E(\mathbf{v}) &= \int_{W(\mathbf{x})} G(\mathbf{x} - \mathbf{x}') 2\nabla I(\mathbf{x}', t)\nabla I(\mathbf{x}', t)^\top \mathbf{v} d\mathbf{x}' + \\ &\quad + \int_{W(\mathbf{x})} G(\mathbf{x} - \mathbf{x}') 2\nabla I(\mathbf{x}', t)\partial_t I(\mathbf{x}', t) d\mathbf{x}' = \\ &= 2\left(G * \left(\nabla I \nabla I^\top\right)\right)\mathbf{v} + 2\left(G * (\nabla I \partial_t I)\right) =: 2M\mathbf{v} + 2\mathbf{q}\end{aligned}$$

where M is defined as $G * \left(\nabla I \nabla I^\top\right)$ and $\mathbf{q}$ as $G * (\nabla I \partial_t I)$. We further know $\nabla I \nabla I^\top$ and $\nabla I \partial_t I$:

$$\nabla I \nabla I^\top = \begin{pmatrix} I_x \\ I_y \end{pmatrix}\begin{pmatrix} I_x & I_y \end{pmatrix} = \begin{pmatrix} (I_x)^2 & I_x I_y \\ I_x I_y & (I_y)^2 \end{pmatrix} \quad \text{and} \quad \nabla I \partial_t I = \begin{pmatrix} I_x I_t \\ I_y I_t \end{pmatrix}$$

which proves that the entries of M and $\mathbf{q}$ are as stated. Since we want to find a minimizer $\hat{\mathbf{v}}$ of $E(\mathbf{v})$, we require

$$\left.\frac{\mathrm{d}E(\mathbf{v})}{\mathrm{d}\mathbf{v}}\right|_{\mathbf{v}=\hat{\mathbf{v}}} = 0 \quad \Rightarrow \quad 2M\hat{\mathbf{v}} + 2\mathbf{q} = 0 \quad \Rightarrow \quad \hat{\mathbf{v}} = -M^{-1}\mathbf{q}$$

(b) Show that if the gradient direction is constant in $W(\mathbf{x})$, i.e. $\nabla I(\mathbf{x}', t) = \alpha(\mathbf{x}', t)\mathbf{u}$ for a scalar function α and a 2D vector $\mathbf{u}$, M is not invertible.

$\mathbf{u}$ does not depend on $\mathbf{x}'$, so it can be pulled out of the convolution integral. Thus,

$$M = G*\left(\nabla I \nabla I^\top\right) = \left(G * \alpha^2\right)\mathbf{u}\mathbf{u}^\top \quad \Rightarrow \quad \mathrm{det}M = \left(G*\alpha^2\right)^2\left(u_1^2 u_2^2 - (u_1 u_2)^2\right) = 0.$$

Explain how this observation is related to the aperture problem.
The aperture problem states that it is impossible to determine the motion orthogonal to the gradient direction in regions with constant gradient direction,. M not being invertible means that there is no unique solution $\hat{\mathbf{v}}$, which is the mathematical formulation of "the motion cannot be determined".

(c) Write down explicit expressions for the two components $\hat{v}_1$ and $\hat{v}_2$ of the minimizer in terms of m_{ij} and q_i.

$$\hat{\mathbf{v}} = -M^{-1}\mathbf{q} \quad \text{where} \quad M^{-1} = \frac{1}{\mathrm{det}M}\begin{pmatrix} m_{22} & -m_{12} \\ -m_{12} & m_{11} \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \hat{v}_1 \\ \hat{v}_2 \end{pmatrix} = \frac{-1}{m_{11}m_{22} - m_{12}^2}\begin{pmatrix} m_{22} & -m_{12} \\ -m_{12} & m_{11} \end{pmatrix}\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} \frac{m_{22}q_1 - m_{12}q_2}{m_{12}^2 - m_{11}m_{22}} \\ \\ \frac{m_{11}q_2 - m_{12}q_1}{m_{12}^2 - m_{11}m_{22}} \end{pmatrix}$$

2. **The Reconstruction Problem**

The bundle adjustment (re-)projection error for N points $\mathbf{X}_1$,...,$\mathbf{X}_N$ is

$$E(R, \mathbf{T}, \mathbf{X}_1, ..., \mathbf{X}_N) = \sum_{j=1}^{N}\left(\|\mathbf{x}_1^j - \pi(\mathbf{X}_j)\|^2 + \|\mathbf{x}_2^j - \pi(R\mathbf{X}_j + \mathbf{T})\|^2\right)$$

(a) What dimension does the space of unknown variables have if ...

- ... R is restricted to a rotation about the camera's y-axis? $4 + 3N$
- ... the camera is only rotated, not translated? $3 + 3N$.
- ... the points $\mathbf{X}_j$ are known to all lie on one plane? $9 + 2N$.