Computer Vision (600.461/600.661)
Homework 5: Two-View Geometry and Optical Flow

Instructor: René Vidal

Due 11/13/2014, 11.59PM Eastern

1. **(20 points) Properties of $so(3)$ and $SO(3)$.** Given a nonzero vector $\boldsymbol{w} = (w_1, w_2, w_3) \in \mathbb{R}^3$, one can build a 3×3 skew-symmetric $\widehat{\boldsymbol{w}} = \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix} \in so(3)$.

   (a) (4 points) Show that $\operatorname{rank}(\widehat{\boldsymbol{w}}) = 2$. What is the vector in the null space of $\widehat{\boldsymbol{w}}$? Give a geometric interpretation to your result.

   (b) (4 points) Show that $\widehat{\boldsymbol{w}}^2 = \boldsymbol{w}\boldsymbol{w}^\top - \|\boldsymbol{w}\|^2 I$ and $\widehat{\boldsymbol{w}}^3 = -\|\boldsymbol{w}\|^2 \widehat{\boldsymbol{w}}$.

   (c) (4 points) Assume that a camera moves with a rotation $R(t) \in SO(3)$ as a function of time $t \in \mathbb{R}$. Use the relationship $R(t)R(t)^\top = I$ to show that there exists a skew symmetric matrix $\widehat{\boldsymbol{w}}(t)$ such that $\dot{R}(t) = \widehat{\boldsymbol{w}}(t)R(t)$. Assume now that $\widehat{\boldsymbol{w}}(t) \equiv \widehat{\boldsymbol{w}}$ is constant, i.e., it doesn't change with time. Show that the solution to the differential equation $\dot{R}(t) = \widehat{\boldsymbol{w}}R(t)$ is given by

   $$R(t) = \exp(\widehat{\boldsymbol{w}}t)R_0, \quad t \geq 0, \tag{1}$$

   where $R_0 \in SO(3)$ is the initial rotation at $t = 0$ and $\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}$ is the matrix exponential.

   (d) (4 points) Let $R = \exp(\widehat{\boldsymbol{w}}\theta)$ for some $\boldsymbol{w} \in \mathbb{R}^3$ such that $\|\boldsymbol{w}\| = 1$. Show that

   $$R = I + \sin(\theta)\widehat{\boldsymbol{w}} + (1 - \cos(\theta))\widehat{\boldsymbol{w}}^2. \tag{2}$$

   (e) (2 points) Show that $\operatorname{trace}(R) = 1 + 2\cos(\theta)$.

   (f) (2 points) Show that $\boldsymbol{w}$ is an eigenvector of R and give a geometric interpretation to your result.

2. **(15 points) Optical flow of a perspective camera.**

   (a) (8 points) Let $\boldsymbol{x}(t) = (x(t), y(t), 1) \in \mathbb{P}^2$ be the projection of a fixed point $\boldsymbol{X} = (X, Y, Z) \in \mathbb{R}^3$ onto a perspective camera moving with rotation $R(t) \in SO(3)$ and translation $T(t) \in \mathbb{R}^3$ at time t, i.e., $\boldsymbol{x}(t) = \pi(R(t)\boldsymbol{X} + T(t))$, where $\pi(\boldsymbol{X}) = \boldsymbol{X}/Z$ denotes perspective projection. Assume that the initial pose of the camera is $R(0) = I$ and $T(0) = \mathbf{0}$. Let $\boldsymbol{w}(t) \in \mathbb{R}^3$ and $\boldsymbol{v}(t) \in \mathbb{R}^3$ be, respectively, the rotational and translational velocities of the camera at time t, and recall that $\dot{R}(t)R(t)^\top = \widehat{\boldsymbol{w}}(t)$ and $\dot{T}(t) = \boldsymbol{v}(t)$. Show that the optical flow induced by the camera at time $t = 0$ is given by

   $$\begin{aligned} \dot{\boldsymbol{x}} &= (-\widehat{\boldsymbol{x}} + \boldsymbol{x}e_3^T\widehat{\boldsymbol{x}})\boldsymbol{w} + \frac{1}{Z}(I - \boldsymbol{x}e_3^T)\boldsymbol{v} \\ \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} &= \begin{bmatrix} -xy & 1 + x^2 & -y \\ -(1 + y^2) & xy & x \end{bmatrix} \boldsymbol{w} + \frac{1}{Z} \begin{bmatrix} 1 & 0 & -x \\ 0 & 1 & -y \end{bmatrix} \boldsymbol{v}, \end{aligned} \tag{3}$$

   where $e_3 = (0, 0, 1) \in \mathbb{R}^3$, and $\boldsymbol{w}$ and $\boldsymbol{v}$ are the rotational and translational velocities at $t = 0$.

   (b) (7 points) Assume that a perspective camera observes a 3-D plane $\boldsymbol{N}^T\mathbf{X} = d$, where $\boldsymbol{N} = (n_1, n_2, n_3)$ is the normal to the plane and d is the distance from the plane to the camera center. Show that the optical flow observed by a camera moving with angular and linear velocity $\boldsymbol{w}$, $\boldsymbol{v}$ is a polynomial of degree 2 in (x, y) with 8 different coefficients. Find explicit formulae for each one of the coefficients as a function of $\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{N}, d$. Under what conditions on $\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{N}, d$ can the optical flow be approximated by an affine flow model $\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{b}$? Find expressions for A and $\boldsymbol{b}$ as a function of $\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{N}, d$ under those conditions.

3. **(15 points) 8-point algorithm.** Implement a function `[R,T,X] = twoviewSFM(x1,x2)` that reconstructs camera motion `(R,T)` $\in SE(3)$ and 3D structure $X \in \mathbb{R}^{3\times P}$ from a set of point correspondences `x1,x2` $\in \mathbb{P}^{2\times P}$ using the 8-point algorithm (Alg. 5.1 of book-MASKS). Test the algorithm on the dataset $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ available here. The given point correspondences are noise free, and so your algorithm should be able to recover the given R_{true}, T_{true} and X_{true}. To evaluate the robustness of your algorithm with noisy data, add Gaussian noise with standard deviation $\sigma \in \{0, 10^{-4}, 2\times 10^{-4}, \ldots, 10^{-3}\}$. Be careful to add noise only to the two pixel coordinates. Repeat the experiment for $N = 100$ trials. For each trial, compute the the error in rotation, $\theta_R = |\operatorname{acos}\big((\operatorname{trace}(R_{est}^\top R_{true}) - 1)/2\big)|$, the error in translation $\theta_T = |\operatorname{acos}(T_{est}^\top T_{true})|$, and the error in 3D reconstruction $\theta_{\boldsymbol{X}} = |\operatorname{acos}(\operatorname{trace}(X_{est}^\top X_{true}))|$ (measured in degrees) between the estimated and the true quantity. Make sure that $\|T\|_2 = 1$ and $\|X\|_F = 1$ before you compute errors.

Submission instructions. Send email to vision14jhu@gmail.com with subject 600.461/600.661:HW5 and attachment `firstname-lastname-hw5-vision14.zip` or `firstname-lastname-hw5-vision14.tar.gz`. The attachment should have the following content:

1. A file called `hw5.pdf` containing your answers to each one of the analytical questions. If at all possible, you should generate this file using the latex template hw1-vision14.tex. If not possible, you may use another editor, or scan your handwritten solutions. But note that you must submit a single PDF file with all your answers.

2. For coding questions, submit a file called `README`, which contains instructions on how to run your code. Use separate directories for each coding problem. Each directory should contain all the functions and scripts you are asked to write in separate files. For example, for HW2 the structure of what you should submit could look like

   (a) `README`

   (b) `hw2.pdf`

   (c) `hw2q3`: `hw2q3c.m`, `hw2q3e.m`

   (d) `hw2q4`: `hw2q4b.m`, `hw2q4c.m`

   The TA will run your scripts to generate the results. Thus, your script should include all needed plotting commands so that figures pop up automatically. Please make sure that the figure numbers match those you describe in `hw2.pdf`. You do not need to submit input or output images. The output images should be automatically generated by your scripts so that the TA can see the results by just running the scripts. In writing your code, you should assume that the TA will place the input images in the directory that is relevant to the question solved by your script. Also, make sure to comment your code properly.