Assignment 2 (Due: Dec. 20, 2025)
(Send your solutions to TA: 13893138834@163.com)

1. (**Math**) In the augmented Euclidean plane, there is a line $x - 3y + 4 = 0$, what is the homogeneous coordinate of the infinity point of this line?

2. (**Math**) On the normalized retinal plane, suppose that $\mathbf{p}_n$ is an ideal point of projection without considering distortion. If distortion is considered, $\mathbf{p}_n=(x, y)^T$ is mapped to $\mathbf{p}_d=(x_d, y_d)^T$ which is also on the normalized retinal plane. Their relationship is,

$$\begin{cases} x_d = x\left(1 + k_1 r^2 + k_2 r^4\right) + 2\rho_1 xy + \rho_2 \left(r^2 + 2x^2\right) + xk_3 r^6 \\ y_d = y\left(1 + k_1 r^2 + k_2 r^4\right) + 2\rho_2 xy + \rho_1 \left(r^2 + 2y^2\right) + yk_3 r^6 \end{cases}$$

where $r^2 = x^2 + y^2$

For performing nonlinear optimization in the pipeline of camera calibration, we need to compute the Jacobian matrix of $\mathbf{p}_d$ w.r.t $\mathbf{p}_n$, i.e.,

$$\frac{d\mathbf{p}_d}{d\mathbf{p}_n^T}$$

It should be noted that in this question $\mathbf{p}_d$ is the function of $\mathbf{p}_n$ and all the other parameters can be regarded as constants.

3. (**Math**) In our lecture, we mentioned that for performing nonlinear optimization in the pipeline of camera calibration, we need to compute the Jacobian of the rotation matrix (represented in a vector) w.r.t its axis-angle representation. In this question, your task is to derive the concrete formula of this Jacobian matrix. Suppose that

$$\mathbf{d} = \theta\mathbf{n} \in \mathbb{R}^{3\times 1}, \text{ where } \mathbf{n} = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix} \text{ is a 3D unit vector and } \theta \text{ is a real number denoting the rotation angle.}$$

With Rodrigues formula, $\mathbf{d}$ can be converted to its rotation matrix form,

$$\mathbf{R} = \cos\theta\mathbf{I} + (1 - \cos\theta)\mathbf{n}\mathbf{n}^T + \sin\theta\mathbf{n}^\wedge$$

and obviously $\mathbf{R} \triangleq \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ is a 3×3 matrix.

Denote $\mathbf{r}$ by the vectorized form of $\mathbf{R}$, i.e.,

$$\mathbf{r} \triangleq \left(r_{11}, r_{12}, r_{13}, r_{21}, r_{22}, r_{23}, r_{31}, r_{32}, r_{33}\right)^T$$

Please give the concrete form of Jacobian matrix of $\mathbf{r}$ w.r.t $\mathbf{d}$, i.e., $\dfrac{d\mathbf{r}}{d\mathbf{d}^T} \in \mathbb{R}^{9\times 3}$.

In order to make it easy to check your result, please follow the following notation requirements,

$$\alpha \triangleq \sin\theta, \beta \triangleq \cos\theta, \gamma \triangleq 1 - \cos\theta$$

In other words, the ingredients appearing in your formula are restricted to $\alpha, \beta, \gamma, \theta, n_1, n_2, n_3$.

4. (**Programming**) Bird's-eye-view generation. The geometric transform between the physical plane and its bird's-eye-view image can be simply described by a **similarity transformation** matrix. Bird's-eye-view is very useful in autonomous industrial inspection, ADAS, etc. In this question, your task is to create the bird's-eye-view image of a physical plane, e.g., the wall of your room. For this purpose, you may need to,
   1) make a calibration board with chessboard patterns;
   2) calibrate your camera (the camera mounted on your laptop or the camera of your mobile phone with fixed focal length) to get its intrinsics;
   3) attach regular patterns (e.g., chessboard patterns) to the wall, determine the 2D coordinate system C_W of

the wall, and determine the coordinates $\left\{\mathbf{x}_{Wi}\right\}_{i=1}^{N}$ of the feature points of the regular patterns with respect to C_W;

4) take the image I_d of the wall with regular patterns;
5) undistort image I_d with the camera's intrinsics to get the undistorted image I;
6) For each $\mathbf{x}_{Wi}$, determine its image $\mathbf{x}_{Ii}$ on I;
7) solve the homography matrix $P_{W \to I}$ between the wall and the image I wall using $\left\{\mathbf{x}_{Wi} \leftrightarrow \mathbf{x}_{Ii}\right\}_{i=1}^{N}$;
8) generate the final bird's-eye-view image of the wall using the technique introduced in our lecture.

For submission, you **only** need to submit the following items to TA:
1) the intrinsic parameters of your camera;
2) the original image of the wall (or other physical planes) taken by your camera; make sure that your name is painted or attached on the wall (or the plane); (maybe similar to following image I provide to you)
3) the generated bird's-eye-view image of the wall (or other physical planes).

5. (Programming) I have established a dataset for training models for detecting speed-bumps and persons. This dataset can be downloaded from
https://github.com/csLinZhang/CVBook/tree/main/chapter-15-YOLO/For-yolov4
Using this dataset, please train a speed-bump detection model and test your model on the provided test video (on the course website). For this question, you only need to hand in your video with detected bounding-boxes to the TA. A sample frame of our result video may like the following image.