Lab 4: Bonus Questions

- There are two sets of bonus questions for this lab, each worth 3 points. In total, you can gain up to 18 points. Obtaining 12 out of the 18 points will improve your grade by one if you fulfill all requirements to pass the course.
- You have to work on the questions on your own and have to hand in your solutions individually.
- Your answers will be graded on a continuous scale, i.e., we try to reward partially correct answers.
- Please hand in your solutions as a PDF through Canvas.

1 Camera Geometry, 3 points

(a) Consider the following situation: You have a set of 2D-3D correspondences between features in an image and 3D points on a plane for a camera with an unknown focal length f. The plane is parallel to the X-Y plane of the local coordinate system of the camera and the Z-axis of the local camera coordinate system intersects the plane at the point $(0, 0, z_{\text{plane}})^T$, i.e., the camera has a fronto-parallel view of the plane. However, the distance z_{plane} of the plane to the origin of the local camera coordinate system is unknown.
Derive an algorithm that computes both the unknown focal length f and the distance z_{plane} from a given set of 2D-3D correspondences or prove that no such algorithm exists. You can assume that all matches are correct.

(b) An explicit representation of a line in 3D is given by

$$\mathbf{x}(\lambda) = \mathbf{o} + \lambda \mathbf{d} \ ,$$

where $\mathbf{o} \in \mathbb{R}^3$ is some point on the line, $\mathbf{d} \in \mathbb{R}^3$ is the direction of the line, and $\lambda \in \mathbb{R}$ is a scaling factor.
Consider the set of all parallel lines in a given direction $\mathbf{d}$ in the local coordinate system of the camera. Assume that the direction $\mathbf{d}$ is not parallel to the X- or Y-axis of the camera. The *vanishing point* corresponding to this set of lines is the point in the image in which the projections of all these lines intersect (i.e., the point in which all lines intersect after they have been projected into the image). A classical example for a vanishing point is a camera looking down a railroad track, leading to a point in the image where the two lines defining the track intersect.
Show that the vanishing point corresponding to parallel lines in direction $\mathbf{d}$ (in the local camera coordinate system) corresponds to the intersection of the line $(0, 0, 0)^T + \lambda \mathbf{d}$ with the image plane. For simplicity, assume that the focal length f is equal to 1.

Hint: Consider the projection of points that are infinitively far away along the lines.

(c) Assume that you know the *absolute* poses $[\mathtt{R}_1|\mathbf{t}_1]$ and $[\mathtt{R}_2|\mathbf{t}_2]$ of two images, i.e., $[\mathtt{R}_i|\mathbf{t}_i]$ transforms a point from a global coordinate system into the local camera coordinate system of the i-th image. For a third image, you only know the *relative* poses $[\mathtt{R}_{31}|\mathbf{t}_{31}]$ and $[\mathtt{R}_{32}|\mathbf{t}_{32}]$, where $[\mathtt{R}_{3i}|\mathbf{t}_{3i}]$ transforms a point from the local coordinate system of image i to the local coordinate system of the third image. Since you obtained the relative poses by decomposing essential matrices, $\mathbf{t}_{31}$ and $\mathbf{t}_{32}$ are only known up to a scaling factor, i.e., assume that $||\mathbf{t}_{31}|| = ||\mathbf{t}_{32}|| = 1$. Describe how you can compute the absolute pose $[\mathtt{R}_3|\mathbf{t}_3]$ of the third image from this given information.

2 Generative Neural Networks, 3 points

(a) Using the notation introduced in the lecture, describe how a Variational Autoencoder (VAE) can be used to generate meaningful images.

(b) What is the purpose of the discriminator in a Generative Adversarial Network (GAN)? Could it also be used as part of a VAE? If yes, how would it be integrated?

(c) Which problem can arise in the context of unpaired image-to-image translation when using only discriminators in the loss function? How can a cycle consistency constraint be used to prevent this problem?