Question 1:

A calibration pattern is a set of 3D points at known positions. Such a pattern is used to calibrate a stereo rig (determine the transformation between the two cameras).

Suppose we take two images of a calibration pattern with the cameras at unknown positions and orientations (so a point $P$ from the pattern is expressed as $R_1 P + t_1$ in the coordinate system of image 1; and $P \rightarrow R_2 P + t_2$ in image 2). To calibrate the cameras in this case we need a more general expression for the essential matrix.

Express the essential matrix between the two images in terms of $R_1, t_1, R_2, t_2$.

Question 2:

A camera is imaging an object at two time instances. Let $P_i = (X_i, Y_i, Z_i)^T$, $i = 1, \ldots, N$ be a set of 3D points in space at the first time instance, and $p_i = (x_i, y_i, f)^T = \frac{1}{f} P_i$. At the second time instance the points are displaced relative to the camera by $P'_i = R P_i + t$, where $R$ is the rotational component and $t$ is the translational component of either the camera or object motion. In the second image therefore the points are projected to $p'_i = \frac{1}{Z'_i} P'_i$. Assume point correspondences are given across the two images.

Can depth ($Z_i$) be recovered if between the two images:

1. There is only a camera rotation ($t = 0$)?
2. There is only an object rotation (about its center of mass)?

In case depth cannot be recovered show explicitly that it can be eliminated from the equations relating the two images. (Hint: show the existence of a mapping from $p_i$ to $p'_i$ that does not include 3D structure (i.e., depth $Z_i$). Explain why the existence of such a mapping proves the claim.)

Question 3:

1. Show that the right epipole (the epipole in the first image) is given by $v = \alpha R^T t$ (for some $\alpha \neq 0$). Hint: To show this show that $Ev = 0$.

   Explain why this implies that $v$ is the epipole (the intersection of all epipolar lines).

2. Derive an expression for the left epipole (in other words, find $u$ such that $u^T E = 0$). Hint: invert the rigid transformation that relates $P$ and $Q$ and use the expression derived for the right epipole.
Question 4:
Let $H$ be a non trivial homography ($H \neq \alpha I$). A point is called fixed if it does not change location following the application of $H$. A line is called fixed if each point on the line is mapped by $H$ to some point along the same line. A line is called pointwise fixed if each point along the line is fixed.

1. Derive an expression for a fixed point; what should such a point satisfy? How many isolated points can be fixed under $H$.
2. In the case of maximal number of isolated fixed points, identify the fixed lines.
3. How many lines can be pointwise fixed?

Question 5:
Two images of the same scene are related by general rotation and translation. Corresponding points $p_i$ and $p'_i$ in the two images satisfy the fundamental relation $p'_iT F p_i = 0$. Suppose that all the points $P_i$ lie on some plane, and so corresponding points in the two images are related by a homography $p'_i \propto Hp_i$ (i.e., $p'_i = \alpha_i Hp_i$ for some scalar $\alpha_i \neq 0$). SHOW that in this case there exist infinitely many solutions to the fundamental matrix $F$.

Hint: A matrix $M$ is called skew-symmetric if it has the form:

$$M = \begin{pmatrix} 0 & m_{12} & m_{13} \\ -m_{12} & 0 & m_{23} \\ -m_{13} & -m_{23} & 0 \end{pmatrix}.$$

Note that a skew-symmetric matrix $M$ has the following properties:

1. For all $x$: $x^T M x = 0$.
2. $M + M^T = 0$.

(a) Show that $H^TF$ is a skew-symmetric matrix.
(b) Show that this fact induces only 6 linearly independent equations for $F$, instead of 8 (use hints above).
(c) If the points $P_i$ come from more than just one plane in the scene, how many planes are needed in order to uniquely recover $F$?

Question 6:
In class we saw that the disparity map can be computed by minimizing an energy function of the general form

$$\min_d \sum_p D_p(d_p) + \sum_{p,q} V(d_p, d_q),$$

where $p, q$ are points in the left image and $d$ denotes disparity values.
1. Assume that in addition to the pair of stereo images, you also have a laser scan of the scene: for each pixel \( p \) it gives a rough estimation of the depth, \( Z_p \), at the pixel. Write a new energy function that incorporates this additional information from the laser scan. You may assume that that the focal length \( (f) \) and the distance between the two cameras \( (b) \) are known, and \( d_p = (fb)/Z_p \).

2. Depth discontinuities are more likely to happen along edges in the image (e.g., pixels with strong intensity gradient). Write a new energy function that incorporates this knowledge.

Question 7:

SVD factorization can be used to recover shape and motion of \( p \) points in \( f \) frames under orthographic projection. Modify the factorization algorithm to handle the following cases:

1. Scaled orthographic projection (each frame is subject to a different unknown scale factor that multiplies its rotation).
2. Unknown translation, in addition to rotation (i.e., instead of eliminating translation before producing the measurement matrix \( M \) include it in the algorithm).

In each case show what the rank of \( M \) is and how the constraints to resolve the ambiguities should be applied to determine a unique solution.

Question 8:

RANSAC is a method for finding a transformation between images from point matches when some (possibly many) of the matches are erroneous. The objective of RANSAC is to identify the transformation that is supported by the largest number of matches.

Given two images, suppose we use a feature detector to identify \( m \) points in the first image and \( n \) points in the second image, and suppose we do not know in advance which of the \( m \) points in the first image matches which of the \( n \) points in the second image. Write an expression for the number of transformations that may be computed with RANSAC in the worst case and the total run-time complexity for the following cases:

1. The transformation relating the two images is a homography.
2. The two images are related by a general fundamental matrix and we use the 8-point algorithm to compute the transformation.