1. How can you use the Harris corner detector as an edge detector? How can you get the edge orientation?

2. Each of the following filters is either a low-pass or a high pass filter. Say which is which:
   (a) \[
   \begin{bmatrix}
   1 & 1 & 1 \\
   1 & 1 & 1 \\
   1 & 1 & 1
   \end{bmatrix}
   \]
   (b) \[
   \begin{bmatrix}
   1 & 2 & 1 \\
   2 & 4 & 2 \\
   1 & 2 & 1
   \end{bmatrix}
   \]
   (c) \[
   \begin{bmatrix}
   1 & 0 & -1
   \end{bmatrix}
   \]
   (d) \[
   \begin{bmatrix}
   -1 & 2 & -1
   \end{bmatrix}
   \]

3. Suppose we are interested in image classification. In the table below, each column is a particular kind of feature vector, and each row is a particular transformation. In the cells of the table, write down if the feature vector on the column is invariant to the transformation on the row.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Bag of words</th>
<th>SIFT</th>
<th>Image flattened as a vector</th>
<th>Image gradients flattened as a vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illumination changes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small translations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large translations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Consider a bag-of-words based feature vector and a linear classifier. Recall that for a bag-of-words feature vector, we run k-means on SIFT feature vectors to get cluster centers that form our visual words. How does the \( k \) in this clustering step affect the dimensionality of the feature vector? How does increasing \( k \) affect overfitting?

5. How do the epipoles relate to the fundamental matrix \( F \)? How do they relate to the locations of the camera?

6. Pixels at a larger depth have a _____________ disparity.

7. Object detection is often considered a harder problem than image classification. Write down 3 ways in which this is true.

8. A plane in 3D is given by the equation \( \mathbf{N}^T \mathbf{x} = d \). Under a rotation and translation, the point changes as \( \mathbf{x}' = R \mathbf{x} + t \). How does the equation of the plane change? How does the vanishing line change?

9. Given a set of 3D points and the corresponding 2D image locations, we can get a set of linear equations in the entries of the camera projection matrix \( P \) of the form \( \mathbf{A} \mathbf{p} = 0 \). What additional constraint can be imposed on \( \mathbf{p} \) and why? How is this equation solved? How many equations do we need to solve for \( \mathbf{p} \)?
10. Given multiple images of the same scene with different light sources, which of the following can you estimate?
   (a) Albedo
   (b) Normals
   (c) Absolute depth
   (d) Relative depth

11. Consider a pixel for which the BRDF $\rho(\theta_i, \theta_r) = \cos \theta_i$. Is the color of this pixel dependent on the viewing direction? What is the color of this pixel in terms of the lighting direction $\mathbf{L}$ and the surface normal $\mathbf{N}$?

12. Which of the following can prevent overfitting? Which can reduce training error?
   (a) Increasing the number of output channels for a convolutional layer.
   (b) Converting a convolutional layer to a fully connected layer.
   (c) Reducing the kernel size of a convolutional layer.
   (d) Replacing a $k \times k$ convolutional layer (with identical number of input and output channels) with two convolutional layers: a $k \times 1$ and a $1 \times k$ respectively.

13. What is the computational cost of running a $k \times k$ convolutional layer with $c_{out}$ filters on a collection of $c_{in}$ feature maps of size $h \times w$? How many feature maps get produced as output?

14. The R-CNN system uses segmentation to generate proposals. Why is this better than simply picking random bounding boxes?