# Homework 4 Questions

## Instructions

- 4 questions.
- Write code where appropriate.
- Feel free to include images or equations.
- **Please use only the space provided and keep the page breaks.** Please do not make new pages, nor remove pages. The document is a template to help grading.
- If you really need extra space, please use new pages at the end of the document and refer us to it in your answers.

## Questions

**Q1:** Imagine we were tasked with designing a feature point which could match all of the following three pairs of images. Which real world phenomena and camera effects might cause us problems? Use the OpenCV function $cornerHarris$ to investigate.

**A1:** The problem of finding a feature point which matches all the three pairs of images, is that each image pair has different 'problems' to deal with.
one of the Chase images is very blurry and it is hard to find corners at all (compared to the not blurry one).

The Observatory images come with very different viewpoint angles and different zooms. This makes it harder to find good matching features later on. Also on the one picture are many kind of noisy environment changes, that lead to corner defections that couldn't be matched to the other picture.

The Library images come with a very repetitive pattern across the whole images. In general it is hard to match a images that have allot of the same patterns on the image. One point that makes it even harder to find a proper feature, it that both images have different objects in the focus of the image and also very different contrasts and illumination.

**Q2:** In designing our feature point, what characteristics might we wish it to have? Describe the fundamental trade-off between feature point invariance and discriminative power. How should we design for this trade-off?

**A2:** By designing features points, it is desired to find features that are repeatable and distinctive. These points usually come with a high appearance- or geometric-variation. Also it is desired to have less feature points (compared to the total amount of image-pixel data). To reach such a good feature, we try to reach a high invariance (regarding to scale, rotation and illumination). So using the feature, it should result into the same features if they only differ in scale, rotation and illumination.

$$features(transformation(image)) = features(image)$$

On the other side there is the discriminative power, which ensures that each feature has a distinctiveness and is robust to deformations.
Both desired properties exclude each other, because having similar features to small changes and return different features on small changes results into a contradiction. The best way to perform this issue is to view the image and choose the feature based on the image it is applied on (are there repetitive patterns). If this is not possible, it is essential to get further information (metadata) that may indicate if there is a high invariance.

**Q3:** In the Harris corner detector, what do the eigenvalues of the 'M' second moment matrix represent? Discuss both how they relate to image intensity and how we can interpret them geometrically.

**A3:** The eigenvalues of the second moment matrix represent the axis length of the ellipse ($\lambda_{max}$ for the small axis and $\lambda_{min}$ for the big axis).
Interpreting the values of both of the eigenvalues, it is possible to make a guess of the shape in that location.
$\lambda_1, \lambda_2 \approx 0 \longrightarrow$ flat region
$\lambda 1 \approx 0$ and $\lambda_2 >> \lambda_1 \longrightarrow$ edge
$\lambda 2 \approx 0$ and $\lambda_1 >> \lambda_2 \longrightarrow$ edge
$\lambda_1, \lambda_2 >> 0 \longrightarrow$ corner

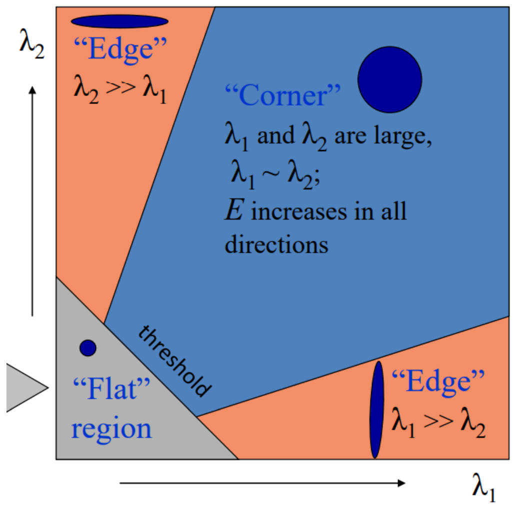


Figure 1: Geometric interpretation of the eigenvalues

Regarding the image intensity, the eigenvalues could be used as a metric for the intensity change in a patch.

**Q4:** Explain the difference between the Euclidean distance and the cosine similarity metrics between descriptors. What might their geometric interpretations reveal about when each should be used? Given a distance metric, what is a good method for feature descriptor matching and why?

**A4:** The Euclidean distance is defined as:

$$d_1(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$$

It represents the length of the line between $x$ and $y$ in a $n$-dimensional space. So it is influenced by the direction (angle) and magnitude (length) of the vector.
The Cosine similarity is defined as:

$$d_2(x, y) = \frac{\sum_{i=1}^{n} x_i * y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} * \sqrt{\sum_{i=1}^{n} y_i^2}}$$

It measures the cosine between two vectors and so is a good measure for the direction (angle) but not for the magnitude (length). In general the Cosine similarity is invariant to magnitude changes of vectors.
So, the Euclidean distance should be used, when there is a focus on discriminative power and the Cosine similarity should be chosen if the focus is on invariance. Also due to the fact that the magnitude is invariant in the Cosine similarity, you should use this measure if you just need to consider the angle between the vectors.

Given a distance metric, the euclidean distance should be used for feature description matching. As pointed out the Cosine similarity is invariant to changes in length of a vector. So a distance metric would always score the same value. This is strongly not desired and the reason why the euclidean distance should be used. It is also influenced by the angle between the vectors (which is irrelevant in this scenario), but also uses the length differences to match the features.