

Homework 3 Questions

Instructions

- 4 questions.
- Write code where appropriate.
- Feel free to include images or equations.
- Please make this document anonymous.
- **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 MATLAB function `corner` to investigate. `corner(I, 1000)`.

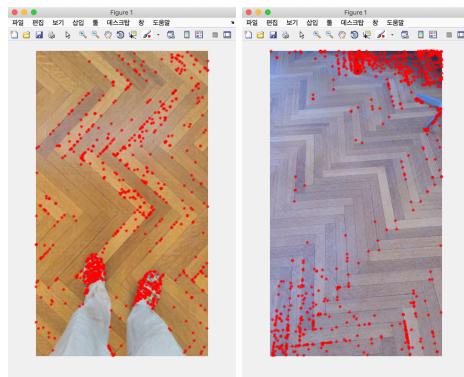
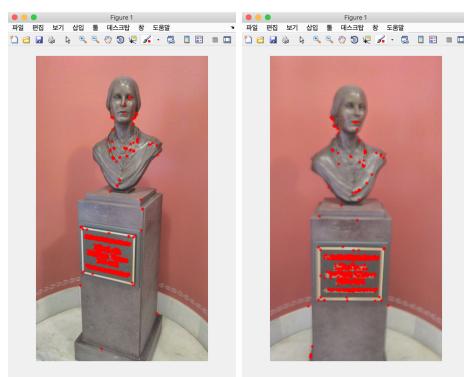
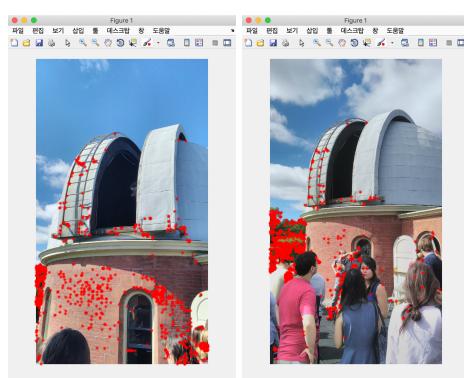
RISHLibrary — Chase — LaddObservatory

A1: Your answer here.

In the case of *RISHLibrary*, we can see that the feature detected in the first picture follows the pattern of the floor, but there cannot be observed much pattern-following feature points in second picture. It seems that the tone of light could affect the feature detection.

In the case of *Chase*, we can see that the second picture cannot be detected the eye of the object, while first picture could. This result is thought to be occurred because of the camera shivering.

In the case of *LaddObservatory*, the features detected at the first picture show us the pattern of the bricks. But in the second picture, we could not find the brick features that much. The scale of the objects described in the picture may affect the feature detection.

Figure 1: Result of *RISHLibrary*Figure 2: Result of *Chase*Figure 3: Result of *LaddObservatory*

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: The feature point should describe the significant point of the image. Invariance means which does not change during the transformation of the image. For example, rotation does not change the number of corners of the image. Discriminative power means that how much pictures we can determine as different. So if we want strong invariance, the discriminative power will be decreased.

We should take care of those two terms carefully. We should keep our invariance in some necessary transformations – such as tones of colors, rotations, and scaling.

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 both eigenvalues represent the ellipses' two radius. The ellipses describe the gradient of the image. If both values are big enough and have similar values, it means that the difference in pixels are huge in both directions. So it should be determined as corner. If one value is much smaller than the another, it means that the difference occurs in only one direction – it's edge. If two values are not big enough, the point could be treated as flat.

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: Euclidean distance is the norm value between two vectors: $\|A - B\|$.

However, cosine similarity uses cosine value between two vectors to define how much closer two vectors are. It is described as $\frac{A \cdot B}{\|A\| \|B\|}$.

Using the NNDR (Nearest Neighbor Distance Ratio) method would be useful when a distance metric is given. NNDR method calculated the ratio between first-neighbor distance and the second-neighbor distance. If the ratio is small enough, we could get more confidence for feature descriptor matching.