UID: u6656110

Name: Zhiyuan Huang

**Q1**

1. Low-level task: Image filtering

Mid-level task: Image segmentation

1. H stands for Hue, S stands for Saturation, V stands for Value.

H represents a color, with a degree from 0 to 360.

1. Outliers are data that have significant different values compared to others in the dataset. RANSAC and IRLS are two algorithms that robust to outliers.
2. A can be factorized into UΣVT, U is a k\*k matrix, Σ is a k\*n matrix, V is an n\*n matrix.

The rank of A is determined by Σ if the diagonal elements in Σ are in a descending order.

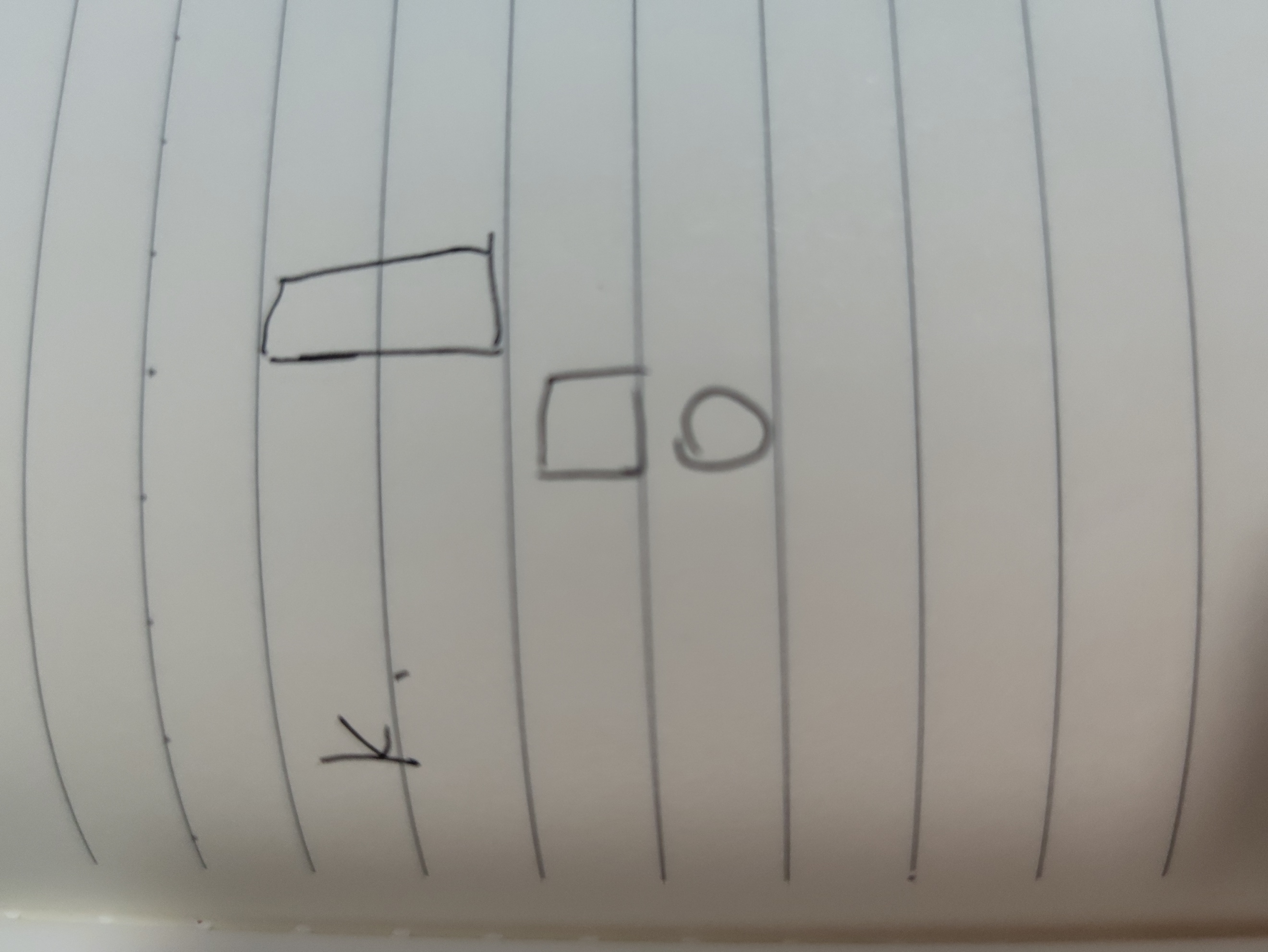
1. In a binary classifier, suppose we have N data in total, R of them are labelled True and thus (N-R) are labelled False. The classifier predicts T data, among of which P data are really True. The precision is P/T. The recall is P/R. Average precision is the averaged value of precision in multiple data retrieval processes.
2. Generative model: generates distributions for all classes, determine the scope of each class.

Discriminative model: find the boundary between classes.

Generative models can be used in super-resolution.

Q2





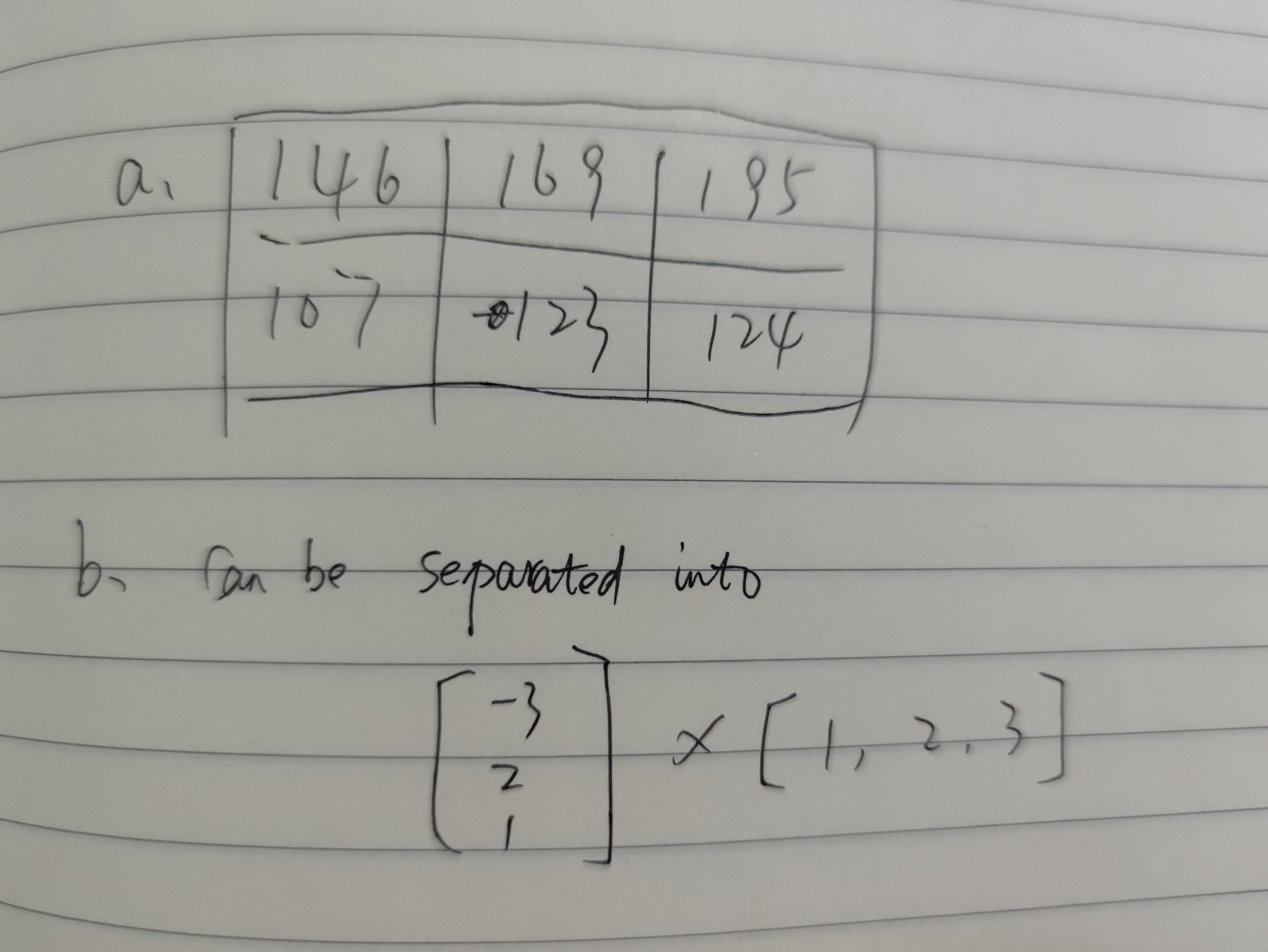
Q3

1. When gamma is 0.4, lower level grayscale pixels will be magnified and high level grayscale pixels are suppressed. When gamma is 2.5, lower level grayscale pixels will be suppressed whereas high level grayscale pixels are magnified.
2. First, calculate the probability of grayscale x in the histogram P(x=i)= ni/n, n is total number of pixels and ni is the pixel number with grayscale i.

Then, the transformed grayscale values are based on the CDF function:

1. The result applying equalization twice is the same as only apply once. Because the probability distribution does not change, though their values may be different.

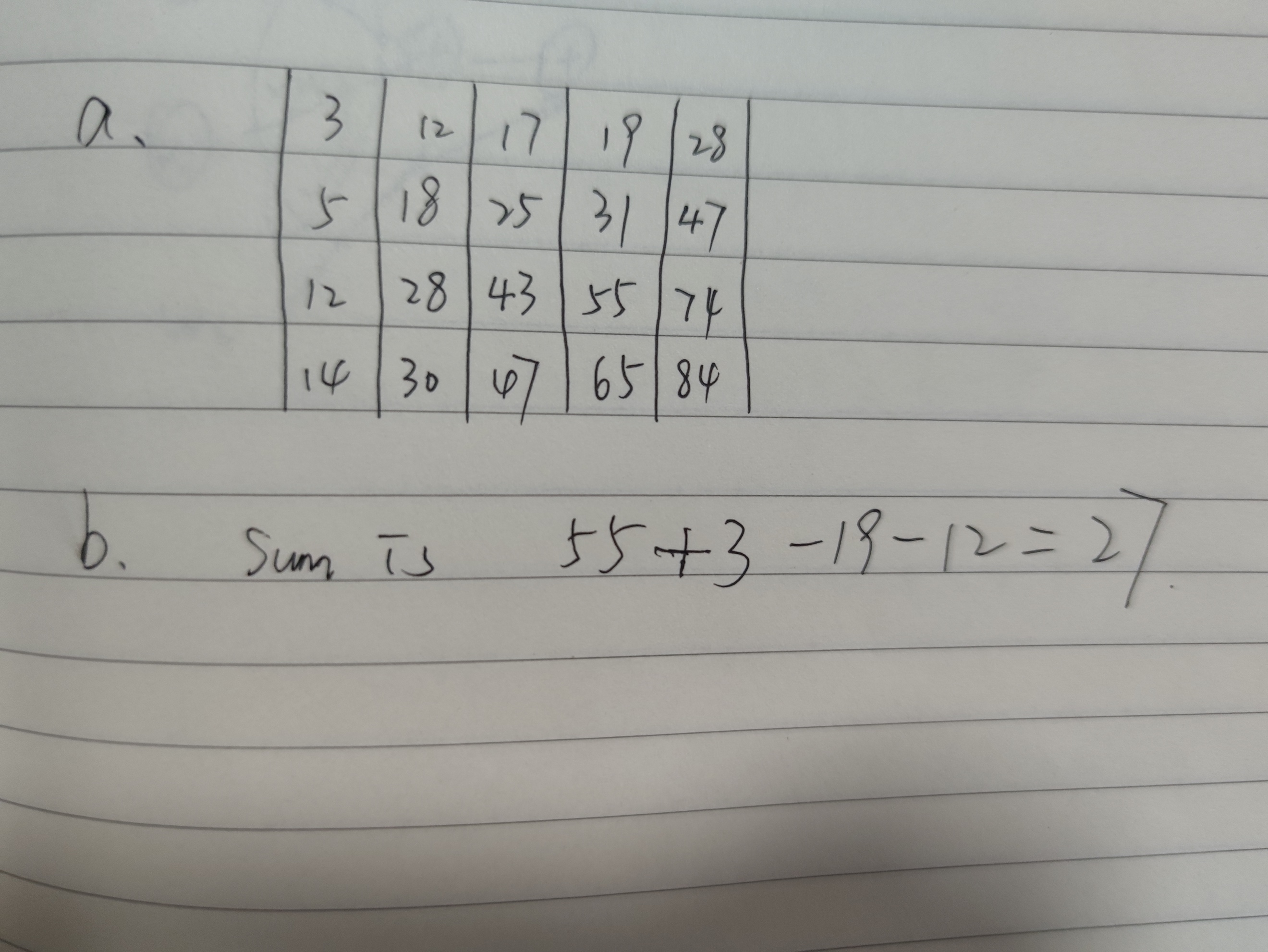
Q4.



c.﷒ f1 filter magnify edge pixel on the diagonal directions.

f2 filter smooth pixels according to the neighbors in one row, reduce variations of the pixel.

Q5



c. The histogram of the pixel gradient, including orientations and magnitude.

d. First, use SIFT to detect features in the image, store the SIFT descriptor of each feature (probably 128 dimension space) in a list

e. 3

f. EM iterations vary at each step. First, it determines expectation function of likelihood. Second, it calculates the maximum parameters. Then it goes to the first step again, and then the second step then the first step and so on until it converges.

K means iterations are the same, it does not change.

Q6

a.

b and c



1. Residual blocks are information from previous layers.

Q7.

1. We can use bag-of-words for visual features.

First, for all images in dataset, we can extract their features based on the SIFT method, store their SIFT descriptors and corresponding images in a library.

Then, we extract feature using SIFT in the query images, iterate through the feature library to check if there is one similar in the library. If the difference is less than a threshold, we can see these two images match.

1. Each SIFT is a 128 dimension vector space. The time complexity is O(n) since the searching is linear.

Q8

Using SIFT algorithm. For each room, take several photos, store their SIFT descriptors and their position in the museum in a library. If a tourist were to find the position, take a photo of where he is and upload the image to the system. The system extract features using SIFT and compare the SIFT descriptors with the descriptors in the library. If two SIFT descriptors match, then the position can be found from the library.