

Midterm Exam

IMT-344 Computer Vision 1 - 2021

| Alejandro Núñez Arroyo

3. INQUIRIES

Specifications: This inquiries' list contains all the questions for the midterm. They are categorized according to sections of what was covered during this first part of the course.

A) Computer vision and image processing foundations

For this section: choose 6 questions randomly, and answer them thoroughly with examples.

1 Why do we use an RGB based color system for image processing?

The RGB system is not the only system with which you can represent images, this is widely used because only with the combination of three colors: red, green and blue we can represent different colors.

People do not have the same range of colors as the RGB system as this is often used for electronic devices, the advantages of this system is that by having a depth of three dimensions we can work independently and join them to have the same result.

3 Why do we include a pre-processing stage for CV based systems?

The pre-processing phase is necessary in a CV system because at this stage noise reduction, color correction and scaling are performed according to the needs of the image.

If I have an algorithm that recognizes vehicles in the pre-processing phase I must reduce the noise that may exist by the time factors (reduce the brightness of the image if it is daytime) or if it is night (increase the brightness of the image) and

remove the background, so the image is ready for the stage of segmentation of characteristics.

5 Why do we require to enhance an image?

We need to improve an image as these may have noise or imperfections in them that may be detrimental to us when using a CV technique, not getting the best results no matter how much we use the best techniques.

If we want to detect the outline of a person in an image the correct thing to do is to remove the background of the image to avoid errors and depending on the noise within them apply a filter, convert the image to grayscale and apply a grayscale equalization to reduce the computation to do the whole process if we enter the original image to the system.

6 Explain what advantages we have when dealing with gray images.

A great advantage we have when using grayscale images is the processing of the image which is much less if we compare it with one that is in RGB because of the size of the image.

In addition to having only one working channel we can use some techniques such as histograms that allow us to better understand the image.

In addition, in a grayscale image we only work on a numerical scale in the range of 0 - 255 for each pixel.

11. Explain why we don't use 1 bit color-depth images.

We do not use 1 bit of depth because with the following formula in relation of bits and depth we obtain:

$$2^b = levels$$

$$b : bits$$

$$2^1 = 2$$

The result of 2 indicates that the images should only have 2 color scales, being a high or low value, this would not be useful in any CV process, since we do not have a scale to distinguish image features.

12 In uniform quantization, why do we have 256 levels when using 8 bits?

This data relation is linked to the conversion that is made between analog and digital since to make this conversion the following formula is used that represents

the levels (values) that can be stored of an image, the range is 256 for the following:

$$2^b = \text{levels}$$

$$b : \text{bits}$$

$$2^8 = 256$$

So an 8-bit image can represent 256 values in one channel.

B) Image filtering and processing

2 Explain the reason for filtering images.

Filters are used to filter out unwanted information in the image and amplify the image features of interest.

In addition, this helps our CV system to recognize the features and process the image in a better way.

3. Explain the difference between linear and non-linear filters.

The difference is that linear filters are based on convolutions, which go through all the pixels of the image obtaining a result, these filters are the High and Low pass filters and non-linear filters use other methods to reach this result.

6. Why do we use the FFT algorithm?

7. Explain the differences between the Median, bilateral, and Gaussian filters.

The median filter is a probabilistic filter, which selects a series of points, and according to the kernel size orders the elements and the value of the center is the one that is replaced.

Gaussian filter is based on the normal distribution, it takes into account a central pixel and a pixel that is around the image, the coordinates are operated and multiplied with the Gaussian function, this process is performed between all pixels.

Bilateral it is based on the Gaussian filter and adds another Gaussian function of intensity, calculating the coordinate and intensity in parallel.

8. Explain the intuition behind the Sobel Operator.

The sobel filter is very similar to the Prewitt filter with the difference that it gives us the possibility to modify the center of our kernel which allows us to have thicker borders in the vertical or horizontal edges.

9. Explain how does the Canny Edge detector work.

The Canny edge detector is a combination of low-pass filter and high-pass filter among other things. It is divided into four main parts.

1. The images are filtered through a Gaussian filter, which is used to remove noise
2. Sobel filters are implemented on the x-axis and y-axis to obtain all the edges.
3. Both filters are added together and the resulting image is subjected to non-maximum suppression, which consists of reducing the thickest edges.
4. Hysteresis is implemented, which consists in evaluating which edges remain connected and which are not eliminated.

10. Why would you rather use the Canny Edge Detector instead of Sobel filters?

It is more preferable to use Canny Edge Detector instead of the sobel filters because a part of the Canny filter is the implementation of the sobel filter on the "x" and "y" axis, that means that the sobel filters would not have the full features (use of the Gaussian filter, use of non-maximum suppression and implementation of hysteresis) and would be only a part of the Canny Edge Detector.

11. Why would you rather use HSV instead of RGB?

It is preferable to use the HSV system when we have the presence of brightness or shadow in our image and we do not want a lot of variability where we need to focus more on color.

In addition, it allows us to isolate the values of Hue, Saturation, Value of each pixel of the image, which allows us to have more control when studying the image.

13. Explain under what conditions the Hough Transform for Lines method can fail.

The method fails when the recognition is applied to situations where there is not a constant straight line, for example to detect circular edges there is a specific method, in case we want to recognize the silhouette of a person with this method

we will not have the best results because it can not detect the set of lines that form the silhouette.

C) Visual Information Compression and Analysis

1 Why lower dimension images are related to fast processing times?

They are related because they do not need much storage, the smaller the dimensions the more images we can store, another important area is the transmission of images because we work in environments where we have to load our images and if the weight of these would be much, it would take a long time. These features will help us to have a faster computation and image processing.

2. Explain the encoder - decoder couple mechanism for compression.

The encoder is responsible for grouping similar characters from the image by grouping pixels together, it also reduces the precision of the output by specifying a specific symbol that has the same properties as the decoder.

The decoder is similar to the encoder in reverse, following the same process.

3. Why is JPEG commonly used?

JPEG is one of the most used compressions since it has a compatibility with different electronic formats, since it has a lower weight compared to other formats, which helps it to have a better transmission of information, it is also a standard that supports large amounts data (24 bits)

4. Explain how JPEG can be used to compress without losses.

Lossless compression needs several components to have an efficient compression, we have a first layer where you enter the data in RGB and a color transformation is applied to it, it goes to a block piece called pixel blocks, then it goes through a DCT and finally to a quantizer which, according to the size of the image, will encode it with its coefficients.

5. Why does the lossless compression techniques cannot achieve higher compression rates than the lossy compression techniques?

6. Explain the low-pass effect that JPEG produces when aiming at very high compression rates.

When you want to achieve very high compression rates in JPEG images they tend to remain blurred at higher compression ratios, and the filter itself acts as a trigger for this effect to occur, since the image measurements will be converted to neutral values, eliminating the image characteristics.

7. Explain the main differences between JPEG and JPEG2000.

The biggest difference between the two is that JPEG is a compression mostly used, as it is a compression with loss on the other hand, JPEG2000 is a compression that seeks to have the original capture resolution, a disadvantage so it is still not used enough at present is the high computational needs as it uses the wavelet compression algorithm, it is estimated that in the not too distant future will be the new standard,

8. Why is compression so important for transmission?

Among the most important advantages when performing the compression is higher speed of data transmission, since at the time of loading data to the colab environment we are making use of the internet and having very heavy files does not help to have accessible data, another advantage is the computational capacity, when we work in environments with more data, we need to find ways to optimize the reading of data because we may not have enough computational power to train or verify the models.

9. Explain the role of features when recognizing objects.

The features are very important when recognizing an object, in addition, it is important to determine what specific feature will help us to make this recognition, since we can extract many features from an image but only some are relevant to their study, among the most common features to make a classification of images is the color, size, the presence of contours, the appearance of edges.

10. Why do we require to use PCA for feature extraction?

A great feature of PCA is the decrease of the image dimensionality, on the other hand, compressing the images, this helps us to keep the main characteristics of an image, and at the same time decrease the space needed to process the image.

11. Why PCA can be used as an image denoising algorithm?

It can be used as a filter when we keep the main characteristics of the image, because when it is grouped and its dimensionality is reduced, it will take into

account values that are grouped, since what we are looking for is a grouping of pixels, so it can work as a filter.

12. Explain how PCA is related to the dimensionality reduction task.

The PCA is related to dimensional reduction by its operation as an algorithm, since it obtains a collection of points in a space of real coordinates, which can be grouped as a sequence of unit vectors, so that we can obtain a dimensional reduction.