Theory part

Sergio Fernandez

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 require to enhance an image?

We require to enhance an image in order to improve the artificial intelligence model prediction. As a paper with read, in this semester, the enhancement of an image, cleaning the information which is not important and disturbs the model, can make the algorithm way more accurate.

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

When we are working with a grayscale image we have the advantage of working with just one color channel it is better for calculations and for processing time also it points out the edges and corners of the objects so when you try to detect something in grayscale it's even easier. We can mention the example of Huawei cameras which use dual or triple lenses on its devices and have grayscale cameras used just for corner and edge detection so the image can have more detail.

3. Why do we want to create binary images?

It is necessary to create binary images when processing them with morphological operations such as erosion, dilation, opening and closing. As a result of these morphological operations, we can reduce the noise of the image. Not only,the negative side of the image, but also, the positive side.One application which is really useful is for x-ray detection, this morphological operations can help us to eliminate some particles which are useless information and could make the image interpretation less accurate.

4. Why don't printers use 30000 DPI?

It is not necessary because most photography use of printers is maximum 2400 DPI. If someone wanted to buy a 30,000 DPI printer would be extremely expensive and must be extremely accurate in the process of printing. Finally, the benefits of having this printer would not be worth the price you could pay for this printer that, as well, surely must be slower than a normal printer.

5. Why would you use 32 bits of color depth?

At 32 bit of color depth means that you can create Shadows transparency or gradients and almost every color that you can even imagine you have 1 million color combinations and this color depth means an acceptable quality for almost any application. As an example, nowadays, in many games, we have a 32-bit color space and this games are more realistic every time.

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

We don't use 1-bit color-depth images because they can only have two values 1 or 0. If we are talking about pixels it would be black or white and we wouldn't have the option to analyze the image correctly, because of the data loss, especially if the image was colorful. A 1-bit color-depth image is more useful when you are doing some image

processing like morphological operations. It is necessary to have a 1-bit color-depth image.

B) Image filtering and processing

For this section: choose 10 questions randomly, and answer them thoroughly with examples.

1. Explain the reason for filtering images.

When you implement the high pass filter the email shows white spaces between the edges in the images. However, when we implement a low pass filter when we want to reduce the image frequency and blur the image.

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

The big difference between these two types of filters is that linear filters are way easier to work with than non-linear filters. As an example, we can convolve with these filters. However, with no linear filters you cannot make convolution nor a Fourier multiplication.

3. Why do we use the FFT algorithm?

FFT(Fast Fourier Transform) is a digital variation of the DFT(Discrete Fourier Transform) algorithm. In addition, this FFT algorithm it's much faster than the DFT so it is useful because it does not require so much processing time to get images into the frequency domain. In some cases the DFT time process could require years and an FFT just requires a few seconds or minutes.

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

The main difference between these three types of filters is the application. Because we can apply, for example, a medium filter to a different kind of noise from a gaussian filter. Additionally, the gaussian filter works better in images with gaussian noise. Also, the kernel is different and the way the algorithm is modified is different, in each case. As a result, every algorithm reduces different kinds of noises and has different results.

5. Explain the intuition behind the Sobel Operator.

The sobel filter is a filter which has kernels that have elements in the middle which are bigger than the elements in the outside of the kernel. Consequently, we have thicker borders so we are able to detect them easier with the future algorithm such as Candice detection or similar. However this civil filter can just be applied in the x-axis or in the y-axis.

6. Explain how the Canny Edge detector work.

Canny Edge detection is an algorithm which tries to detect edges on an image. First, the image must be converted into a grayscale image so it's easier to work and process. Then. it is necessary to apply a low pass filter in order to reduce some noise on the image. After that, a high pass filter is applied so the edges of the image are highlighted. Finally, the algorithm applies a binary threshold in order to detect the edges and separate them from the rest of the image in a two-bit space.

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

Canny edge detection works better than sobel filters because sobel filters are algorithms based on kernels which can be applied on the x-axis or y-axis. However, Canny edge detection works in both directions at the same time because it uses low and high pass filters additionally to a binary threshold. Even if you apply sober filters in both directions it won't be the same result as Canny edge Detection, due to the data loss generated.

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

By the time I have been working with image processing and opencv, I found rather use HSV instead of RGB useful when I am looking for a mask it's easier because when you change the colors you just have to change almost the v value of HSV. Another application is when you predict something based on the brightness of the image.

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

The hough transform for lines method can fail if you apply a hough transform in images that have multicolored lines. A human being is able to detect multicolor lines because of past experience that improve the ability of detecting multicolor lines. However a hough transform method bases its detection on frequency. In this case, as the frequency changes so much between colors, the algorithm is not able to recognize lines in that situation. As a result, probably a solution would be to apply machine learning so the algorithm could base its detection on colorful lines.

10. Explain what would be the use of high-frequency pass filters.

The high pass filters are really useful when you're trying to find it just on an image. This is why high pass filters are present in canny edge detection just before binary threshold.

C) Visual Information Compression and Analysis

For this section: answer all questions thoroughly with examples.

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

First processing times are related with lower dimension images because lower dimension images require less processing because the operations required for that amount of data are less. As an example when you apply kernels to an image it's easier and faster to apply to an image 600x400 than 1920x1080.

2. Explain the encoder-decoder-called colorful they couple mechanism for compression.

When we compress an image. It is necessary to follow some steps to encode. The three steps are, mapper, quantizer and symbol code. The mapper reduces the spatial or temporal reduced redundancies by grouping similar pixels. After that, we have the Quantizer which reduces the precision of mappers output, this means, reducing the decimal values. Finally, the symbol coder generates variable codes for quantifiers output, as a result, we have an encode format image which is lighter than order image format. If you want the computer to show up the image you have to decode the compressed format so you have to follow the encode steps from the end to the beginning.

3. Why is JPEG commonly used?

Jpeg is commonly used because it has three benefits. The first benefit, as we are talking about encoded compressed images, is the weight in this format is lighter than other image formats. The second benefit is that the format is supported by most of the platforms, so this format is well known. Finally, the third benefit is that JPEG is a lossless compression method; this means that when we are using a JPEG format we don't lose any information.

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

When we convert any machine to jpeg format we have to follow some steps. The first step is the mapper which changes the color space of the image into ycrcb color space. Then, we have the image division which fractionates the image into several blocks. The following step of the mapper is normalized the data in order to make the values more

similar to each other and tend to zero. We continue the algorithm with the discrete cosine transform applied to the data we generate in the previous step. After that, we change the values of the data with the prominence quantization table and the luminance quantization table so the algorithm can have previous known values. Finally, we make a symbol encoder or Huffman coding which will fill the values on the sub Matrix after quantization, completing the data moving into zigzag.

5. Why do lossless compression techniques not achieve higher compression rates than lossy compression techniques?

Because if we want to preserve the information into our computers or cloud storage we do have to sacrifice some space this is why Lucy compression methods don't have a limit we can use a Lucy compression method until the image information is completely vanish.

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

When you implement a very high compression rate on an image with a low pass effect you can lose a lot of information. In some cases, the image could lose almost all the information,, and even the image could seem to be black.

7. Explain the main differences between JPEG and JPEG2000.

Jpeg 2000 format is the upgrade of jpeg format, it has improvements on compression performance and image quality. Jpeg format and jpeg 2000 format are almost the same format. You must have an extremely trained eye in that area so you can recognize and know the difference between formats even if you see two images side to side it is really difficult to know which format is it. This is why. as long as the difference is almost imperceptible, the moment jpeg 2000 format showed up, was not accepted immediately.

8. Why is compression so important for transmission?

Compression is really important for transmission because it saves money and time for a company and also improves the speed of transferring files. For example, if you want to transmit an image which is not compressed and the weight is 8 MB, with the bandwidth of two megabytes per second, the file will arrive in 4 seconds. In another situation, if you compress the image and pass it, the image will arrive in 1 second this means a quarter of the time spent on sending.

9. Explain the role of features when recognizing objects.

When we are implementing an algorithm for recognizing objects we have to take into consideration three types of features. First, we have the edges which are lines, circles, and complex shapes. Then, we have the corners which are intersections between the edges and this feature is the most important because it gives us more relevant information from the processed image. Finally, we have to consider the blobs which are the least important, and when we are talking awbout blobs, we are referring to the textures of the image.

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

PCA is an algorithm that is based on feature extraction and orthogonal projections for data transformation. This algorithm tries to preserve most of the data and it's really useful because when you are applying PCA, you canize the data and work wir and faster.

11. Why can PCA be used as an image denoising algorithm?
PCA can reduce the noise of an image because as it takes the principal component of the data and reduces the dimensionality this, sometimes, could make these noisy data disappear.
12. Explain how PCA is related to the dimensionality reduction task.
PCA is an algorithm that reduces the dimension of a data set you can reduce and dimensions, that is the way the algorithm works. PCA reduces the dimensions and based on that reduction generates axis that can make predictions based on which section is placed the value.