Practice 2: A basic content-based image search system

Image Understanding

Academic year: 2021/22

GROUP 13:

Víctor Morcuende Castell

Antonio Ruiz García

TABLE OF CONTENTS

1. Questionnaire

1.1. What considerations did you have when implementing the distance function based on cosine similarity?

First of all, we used the formula provided in the documentation paper. This formula gives us the angle between two or more vectors described between the values 0 to 1. Specifically, this angle describes the similarity of two vectors/images, being 1 a matched pair of images and the value 0 meaning that there is no match between the images. However, we are required to give the cosine distance, which means that for a matched pair of images the distance would logically be 0 and for a no match set of images the similarity would be the maximum (in this case 1). Therefore, we had to obtain the cosine similarity and subtract this similarity to 1, in order to get the cosine distance.

1.2. What differences do you find between the different experiments performed? Discuss the results obtained in terms of how they are affected by the characteristics extracted and the distances used.

As it can be seen in the images below, using the raw pixels as features, the labels predicted are not the best. This can be seen for the 18 images retrieved, as only 7 correspond to the right label, in this case a horse. However, if we take into account the overall results, we consider them to be pretty good as most images correspond to white animals that look similar to the query image or white objects. Moreover, these images possess a dark background, as the test image one. This is because, the technique used (raw pixels) compares the test image with the rest pixel by pixel considering the color value. Taking this into consideration, the results are significantly satisfactory.

Results with raw pixels using the Euclidean distance:

Results with raw pixels using the Cosine distance:

On the other hand, if we evaluate the results for the histogram feature extraction technique, it can be observed that the results are significantly worse than the previous one, getting only 1 out of 18 retrieved images correctly. The cause of these results is that, as the test image is made of mostly black and white pixels, the values of the histograms from the images retrieved are made up of very low (dark tones) and high (light tones) values. Therefore, the images obtained possess abundant black or dark pixels and light colors. These last colors are usually white but sometimes there are colors with high intensity and brightness.

Results with histogram using the Euclidean distance:

Results with histogram using the Cosine distance:

As it can be appreciated in the images below, using the CNN as features extraction, the labels predicted are better than the previous ones, having the same success rate as the raw pixels features technique, 7 out of 18. However, thanks to the ability of the CNN of extracting features from regions of the image, the results obtained belong to animals with clearly shaped body parts, such as legs, head, sometimes tail, etc.

Results with CNN using the Euclidean distance:

Results with CNN using the Cosine distance:

As it can be seen in the images above, it is clear that the best results are given by the raw pixels features using the cosine distance, as it gets 6 correct horses out of 9 retrieved images. However, evaluating the results by feature extraction, the best one is the CNN, since it gets 3 correct images with the Euclidean distance and 4 with the Cosine one.