CENG 483

Introduction to Computer Vision

Spring 2018-2019

Take Home Exam 1

Content Based Image Retrieval

Student ID: 2171379

In this Take Home Exam we study on searching a large database for finding images that matches a query image.To do that , I have built a Content Based Image Retrieval(CBIR) which starts with feature extraction of the query iamge and all other images in the database.In the CBIR system a similarity test is applied to the features of the images both in database and the query.After the result of this set, basically the most similar images are identified.

For the feature extraction we use different kind of histograms such as grayscale intensity, color and edge orientation . A histogram is a vector that counts how many instances of a given property exist in the image. First of all we define ranges to determine bins, then each instance is assigned into one of these bins. Counting how many instances occur for each bin in the image gives us the histogram of that image.

Also we are going to use to 3 different spatial levels.Level 1 is the image itself, Level 2 corresponds to constructing a 2×2 grid and Level 3 corresponds to constructing a 4×4 grid.

1 Grayscale Histogram

I start with grayscale histograms of images for Level 1 by quantizing the pixels into histogram bins based on their intensity level.
First step is choosing different quantization levels in order to create different size of histograms . I apply l_1 *normalization* to the histograms such that the total count of each histogram sums up to 1. After obtaining the histogram representations, euclidean distance metric is utilized for similarity test.
Table below gives the mAP results for different quantization levels(bins).

Bin	mAP
2	0.02277
8	0.16206
10	0.17126
16	0.18616
32	0.18016

When the number of bins in a histogram increases, there will be more information we can get from the grayscale histogram ie. the intensity levels of pixels are distributed better when the bin numbers get higher.But this is not true for all the results because when we use bin 16, the mAP result is greater than the mAP result of bin 32.This is possibly histogram vectors are more similar in bin 16 in this homework's dataset.

2 3D RGB Histogram

The color channel histogram is defined as quantizing pixels at each color channel separately. After that we assign each pixel into combination of bins of these three histograms.
To do that I create a 3D numpy array whose dimensions are bin x bin x bin and put pixels' RGB values into array and applied l_1 *normalization*.
The mAP results of histograms with different bins:

Bin	mAP
2	0.152246
4	0.32950
5	0.31485
8	0.35030
10	0.39889

When we look at the table above , we see that when the bin siz increases the mAP results also increases , there is one exception for this idea bin 4 and bin 5. The mAP result of bin size 10 is the greatest one.There is a better color distribution in the histogram when there is more bins , so this could make the difference between mAP results
Also color histograms shows us that for similarity test , they are better than the grayscale histograms because color is a more distinctive feature than the intensity.

3 Gradient Histogram

What about the gradient histograms?
An Edge in an image is a sharp variation of the intensity function. In grayscale images this applies to the intensity or brightness of pixels. In color images it can also refer to sharp variations of color. Properties of edges include gradient and orientation. This part we construct gradient histograms for grayscale images where each bin value is the total magnitude of the pixels whose orientations are quantized into that bin. If we define A as the source image, and Mx and My are two filters which at each point contain the vertical and horizontal derivative approximations respectively, the computations are as follows:

$$M_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \text{A}$$

$$M_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * \text{A}$$

where $*$ here denotes the 2-dimensional signal processing convolution operation. This is called Sobel Kernel operation.

Sobel kernels can be decomposed as the products of an averaging and a differentiation kernel, they compute the gradient with smoothing. For example, Mx can be written as

$$M_x = \left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} * \begin{bmatrix} -1 & 0 & +1 \end{bmatrix} \right) * \mathrm{A}$$

$$M_y = \left(\begin{bmatrix} -1 \\ 0 \\ +1 \end{bmatrix} * \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \right) * \mathrm{A}$$

$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$ denotes Gaussian bluring and $\begin{bmatrix} -1 & 0 & +1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 0 \\ +1 \end{bmatrix}$ denotes vertical and horizantal derivatives respectively.
After defining our kernels, with using scipy's ndimage.convolve() function we apply our filters to images(images in Level1).For the boundaries of the images we use mirror reflection of the pixels.
To calculate gradient magnitude and gradient's direction we use

M $=\sqrt{{M_x}^2 + {M_y}^2}$ and $\theta = atan(M_y/M_x)$ respectively.

And in order to find the indexes of pixels, used $find\_index\_for\_orientation(value, bin)$ function which takes the pixel's value and the number bins as parameters.
After getting gradient histograms for the images in the whole dataset and validation queries,and apply l_1 normalization, $euclidean\_distance(hist1, hist2)$ function calculates the euclidean distance between each histogram and then we calculated mAP for the bins 4,8 and 16:

Bin	mAP
4	0.07450
8	0.11255
16	0.14571

As we see in the table mAP results grater when the bin size increases .This could be because of the fact that larger bin sizes makes comparision of images more accurate since there is more variety of feutures in the histogram bins.

Figure 1: Original image

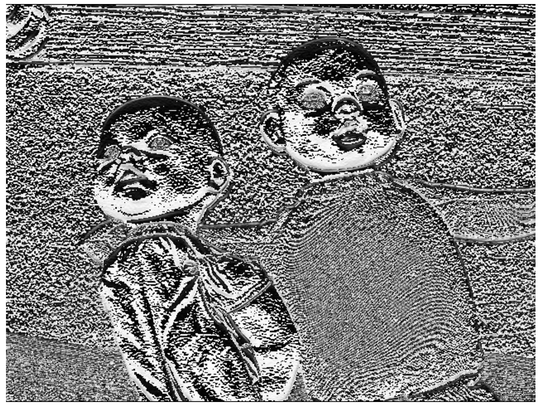

Figure 2: Horizantal filter

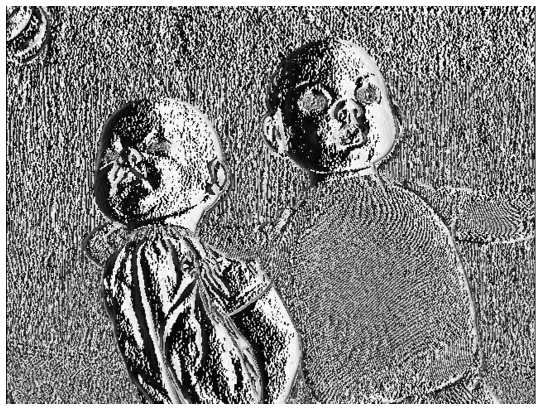

Figure 3: Vertical filter

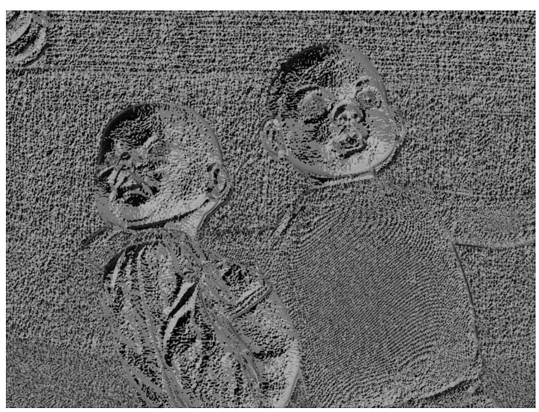

Figure 4: Gradient magnitude image

4 Grid Based Feature Extraction

The best calculation for grayscale and gradient histograms is mAP with the bin 16 . The best configuration for RGB color histogram is the result with bin size 10 but lets choose bin 4 for time efficiency.

4.1 level 1

- grayscale histogram: For grayscale histogram choose bin 16 and mAp 0.18016 which was calculated in section(1)

- 3d rgb histogram: The best color histogram mAP result is 0.32950 with bin size 4 which was calculated in section(2).
- gradient histogram: For best configuration to the gradient histogram is the histogram whose bin size is 16 and mAP is 0.14571 which was calculated in section(3).

4.2 level 2

- grayscale histogram:To obtain grayscale histogram in level2 , we crop the image to 4 equal areas, compute the each part's histogram seperately with bin 16 , apply the l_1 normalization and concatenata 4 histograms ,use euclidean distance formula. And calculate the mAP result for this operation which ,is 0.13282.

- 3d rgb histogram:The mAP result for Level2 RGB histogram is 0.30105 with bin size 4.
- gradient histogram: For the gradient histogram we used the same function for calculating gradient histograms in Level1 , after croping image 4 equal areas . Quantization level is 16 and mAP result for similarity of images is 0.00538.

4.3 level 3

In this section the mAP results was calculated with cropping the image 16 equal areas and appling the histogram functions for each part of the image separately.

- grayscale histogram: The mAP result is 0.22538 with bin 16
- 3d rgb histogram:The mAP result is 0.32044 for bin 4
- gradient histogram:The mAP result is 0.00548 with bin 16

4.4 questions

- What do you think cause the difference between the results? Parsing the image can divide the objects inside the image. For gradient results, it is normal to have lower mAP results in level2 and level3 comparing to level1 because edges are powerful features for similarty and when divide the image we divide the objects at the same time and we lost information about the whole image.Also, gradient results can be easily affected by the shape of the objects, angles, scales of object in the image.

  For the grayscale and color results , we can see that there is no big difference between levels(Especially in color mAP result for level1,level2 and level3). This mAP result can be affected by the size of objects, lighted and shadowed areas in the image and mostly background color/intensity.

- How did you combine the histograms in level 2 and 3? What would you think the difference between to simply sum them and to concatenate them?

I have concatenated the histograms. The difference between concatenation and summing histograms is that when we concatenate , we can compare each divided parts of the image with each other separately .If we sum up the histograms , there will be no reason to divide the image because the summed histogram will be same with level1 histogram.(if the bin sizes are equal)

5 Your Best Configuration

$my\_best(img\_color, img\_grayscale, bin)$ function takes three argument , an color image , a grayscale image (for the same image) and bin size respectively.It produces 4 numpy arrays , 3 of them are for color histograms and one of them is for grayscale histogram. First I calculate the grayscale histogram with Myhistogram_grayscale1() function . Then I calculate color histograms separately using Myhistogram_0 which is for red, Myhistogram_1 which is for green,Myhistogram_2 which is for blue. Finally I concatenate red,green,blue and grayscale histograms respectively.

- Explain mean average precision in your own words:

What I get from mAP is that there is a groundtruth which shows our true results. Then we calculate mAP to understand how our prediction for similarity is correct. When the mAP result increases , our prediction gets better.

6 Additional Comments and References

Although the histogram based similarity tests are used in common ,I think more powerful tools for similarty test are object detection methods. Because in two different image there can be lots of objects similar to each other but using grayscale, color and gradient histograms ,for similarty test, will not be able give efficient result because the objects could be different sizes , different directions or could be appeared in the image in different colors, which can be caused low/high light , noise or quality of the image. Also the background of the image is very effective for the histograms. Background color or intensity can change the similarty results powerfuly.