CENG 483

Introduction to Computer Vision

Spring 2018-2019

Take Home Exam 1

Content Based Image Retrieval

1 Grayscale Histogram

Number of bins used and mAP results obtained for the grid level 1, where the star (*) indicates the highest mAP value:

- Bins: 5 - mAP: 0.11224
- Bins: 12 - mAP: 0.18897*
- Bins: 32 - mAP: 0.18448
- Bins: 36 - mAP: 0.17981
- Bins: 256 - mAP: 0.18069

Quantization in grayscale histogram essentially means dividing the range of intensity values into gray levels. If a few bins are used, pixels are distinguished by large intensity differences. On the other hand, if too many bins are used, even small changes in intensity values become very distinctive. While, in the former, many pixels can be placed in the same bin (in this sense, using bin size of 1 is meaningless since all pixels will be eventually placed in this bin), in the latter, so many pixels are distributed in separate bins. Therefore, a balance between these two possibilities must be redressed to obtain a proper distribution. The experiments show that larger bin numbers generally result in better results than smaller ones. However, using much higher number of bins will likely to reduce precision after some point due to over quantization. Moreover, although the large number of bins seemingly resulted in better results in general, the optimal value is obtained when bin size is around 12.

2 3D RGB Histogram

Number of bins used and mAP results obtained for the grid level 1, where the star (*) indicates the highest mAP value:

- Bins: 5 - 0.35172
- Bins: 9 - 0.40424*
- Bins: 11 - 0.40084

- Bins: 15 - 0.38521
- Bins: 20 - 0.36876

The arguments stated for the grayscale histogram is valid in this case as well. Here, the major difference is when the number of bins is increased, since the same quantization is applied to all of the 3 color channels, the resulting histogram includes the cube of the number of bins used. Quantizing 3 channels, the RGB color histogram essentially uses more information and, in the experiments, it is observed that color histogram generally provides better results than grayscale histogram when the same number of bins are used per channel. The experiments also show that, again, larger bin numbers often yield better results than smaller ones. However, the increase in precision tends to stop and the precision starts to go down after certain number of bins due to over quantization. The optimal number of bins observed is 9 (729 bins in total).

3 Gradient Histogram

The steps taken in the histogram of oriented gradients method:

1. As a first step, two different filtered versions of the image in grayscale format are obtained by convolving horizontal and vertical filters with the image. I used Prewitt operator in the centered format for filtering operation.

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

Figure 1: Horizontal (G_x) & Vertical (G_y) filters

2. After filters are applied, gradient magnitude and angle values of each gradient vector are calculated. In the calculation, unsigned angle values are used. Therefore, all of the resulting angle values lie within $[0, 180)$ range.

3. Then, dividing 180 degrees into equally-spaced bins with respect to the number of bins inputted, a histogram is generated.

4. By using the angle value (which indicates the direction), the corresponding bin in the histogram is determined for each magnitude and angle value pairs. If the angle value does not match exactly with the value of a bin, depending on the distance to the right and left bin, weights are calculated and the magnitude value is distributed to both bins based on the weights.

Figure 2: Original "AiRTnhTcVx.jpg" Image

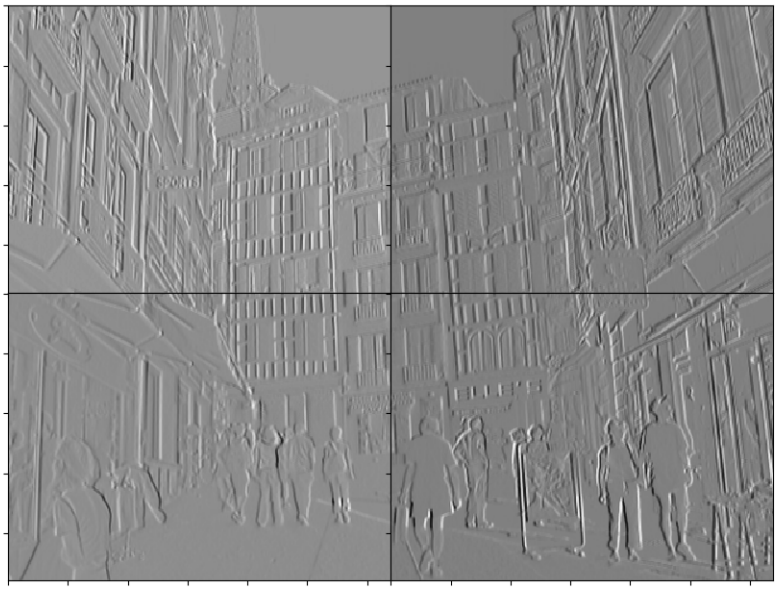

Figure 3: Result of Applying Horizontal Filter on "AiRTnhTcVx.jpg" with Grid Level 2

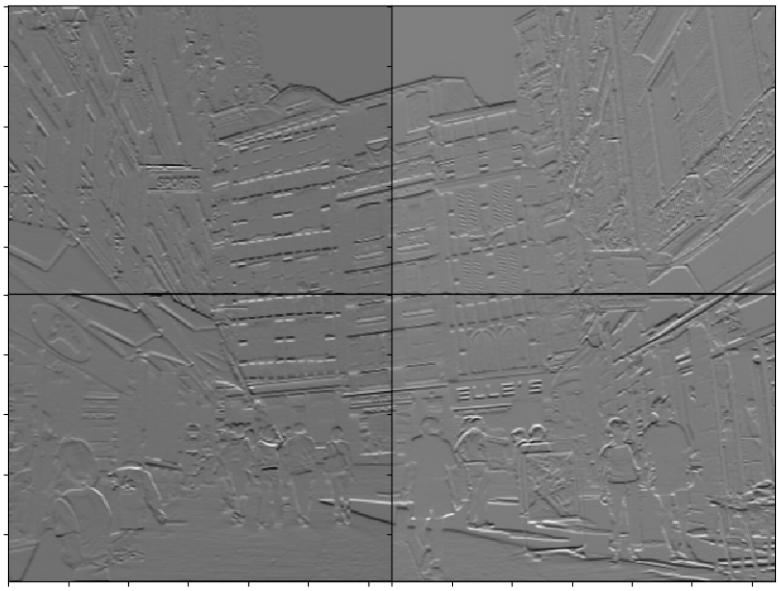

Figure 4: Result of Applying Vertical Filter on "AiRTnhTcVx.jpg" with Grid Level 2

Number of bins used and mAP results obtained for the grid level 1, where the star (*) indicates the highest mAP value:

- Bins: 30 - 0.22787
- Bins: 90 - 0.25496
- Bins: 180 - 0.28131
- Bins: 360 - 0.29167*
- Bins: 450 - 0.29144

The explanation for the grayscale histogram is valid here as well. The difference here is that the angles of gradients (directions) of the pixels are quantized instead of intensity values. Since the domain of angles used in the implementation (0-180) is narrower than that of intensity values (0-255), it is more likely for gradient histogram to use values that are much closer to each other than it is for grayscale histogram. Therefore, it makes sense that higher quantization levels are needed to distinguish pixels with different orientation in gradient histogram. Again, however, it is observed that after certain number of bins, the precision tends to decrease due to over quantization. The optimal mAP value is obtained when 360 bins are used.

4 Grid Based Feature Extraction

mAP configurations for the configurations as follows (the star (*) indicates the highest mAP value):

4.1 Level 1

- Grayscale Histogram: 0.18897 (Bins: 12)
- 3D RGB Histogram: 0.40424* (Bins: 9)
- Gradient Histogram: 0.29167 (Bins: 360)

4.2 Level 2

- Grayscale histogram: 0.26393 (Bins: 12)
- 3D RGB Histogram: 0.40110 (Bins: 9)
- Gradient histogram: 0.30284 (Bins: 360)

4.3 Level 3

- Grayscale histogram: 0.26930 (Bins: 12)
- 3D RGB Histogram: 0.32847 (Bins: 9)
- Gradient Histogram: 0.27630 (Bins: 360)

4.4 Answers to the Questions

1. From grid level 1 to 2 mAP score of grayscale histogram increased significantly while others almost remained constant. From grid level 2 to 3, mAP score for grayscale histogram remained almost constant while others decreased. These results indicate that each grid level has its own optimal bin values. The possible reason is as follows, the number of bins used for the entire image (i.e. the level 1) has to be large to distinguish values properly because there can be many values separated by a small amount in the overall image and these minor differences can only be properly differentiated by using large number of bins. However, the same number of bins are not that useful in distinguishing possibly larger changes in the values that exist in smaller patches (in levels greater than 1) because the use of large number of bins can result in over quantization. This argument can be supported by the additional tests performed on each method with different quantization and grid level values. These tests will be discussed in the last section.

2. In the grid level 2 & 3 tests, the histograms of patches are combined with concatenation. Adding up the histogram would simply result in the histograms obtained from level 1 (if the same number of bins are used for the histograms of all patches. In my tests, this is the case.), therefore no improvement in results can be obtained by adding the histograms. On the other hand, concatenation allows treating smaller patches of an image as separate images and compare each patch in the query image with its corresponding patch in the other image. Since this approach concentrates on local features by eliminating possible outliers coming from other patches, descriptors depending on local features will work better with concatenated histograms of patches.

5 Your Best Configuration

- 89 different experiments are conducted. These experiments test the performance of each method separately with different number of bins and with different grid levels. Unfortunately, the combination of all three methods could not be tested. The best mAP observed in all tests is 0.40424 and this result was obtained by using 3D Color Histogram with 9 bins in grid level 1.
- The exact command used for this test is as follows:

```
$ python3 the1.py --pipemode full --extmode color --bins 9 --level 1
```

- In general, it is possible to run the script with the following command template:

```
$ python3 the1.py --pipemode [full|ext|query]\
                  --extmode [gray|color|grad]\
                  --bins [N] --level [N]
```

–*pipemode* flag simply specifies which part of CBIR pipeline will be executed.

1. *ext*: Extraction only
2. *query*: Query (distance calculation) only
3. $full$: Extract first and then perform queries

–*extmode* flag specifies which extraction method will be used.

1. *gray*: Extract using grayscale
2. *color*: Extract using 3D Color Histogram
3. *grad*: Extract using Histogram of Oriented Gradients

–*bins* flag specifies the number of bins to in the histogram to be built. (Note that for 3D Color Histogram this value is used for each axis, resulting in a histogram in the size which is **the cube of the inputted bin size**.)

–*levels* flag specifies the level of the grid following the convention given in the homework text.

Notes about the usage of the script:

- By default script takes images from a director named *dataset/* which is located in the same directory as the script. To read from an arbitrary directory use –$drfolder$ flag additionally.
- By default script takes image names from *images.dat* file located in the same directory as the script. To read from an arbitrary file use –*imagedb* flag additionally.
- By default script takes queries from *validation_queries.dat* file located in the same directory as the script. To read from an arbitrary file use –*querydb* flag.
- By default script creates two separate directories for extraction and distance measurement in the same directory as the script. The directory names are generated automatically from the extraction method, bin and level parameters used.
- In all pipeline modes (including querying) –*extmode*, –*bins*, –*levels* flags must be provided.

- Some of the matches obtained by using the best configuration given above:

Figure 5: QVULMICOPi.jpg - mUjbfEDiJR.jpg: 0.062

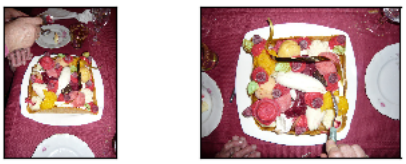

Figure 6: WpoNzXoDGr.jpg - wWsgkPdSqR.jpg: 0.087

Figure 7: TsdbuHwrZH.jpg - rKREsHIylG.jpg: 0.306

- Average precision (AP) generates a single score metric which is used to determine the precision of the method in retrieving the image when a single query is made and the mean average precision (mAP) generates a single score metric which gives the mean of the average precision values obtained from all queries made for the image retrieval. The higher the precision is on the average, the higher the mAP. Thus, high mAP values indicate a better performance.

6 Additional Comments and References

- I conducted many other tests to have a better understanding for the relationship between the performance and the use of different levels and bins.
Some additional results from grayscale histogram:

- Bin: 9 Level-1: 0.15633 (Second worst result for level-1)
- Bin: 9 Level-2: 0.25669 (Second best result for level-2, after 12 bins)
- Bin: 9 Level-3: 0.27162 (Second best result for level-3, better than 12 bins)

Some extra results from gradient histogram:

- Bin: 10 Level-1: 0.18439 (Second worst result for level-1)
- Bin: 10 Level-2: 0.32862 (Better than 360 bins for level-2)
- Bin: 10 Level-3: 0.35350 (Best result for level-3)

These results show that it is possible to obtain a better performance from grayscale and gradient histograms with lower number of bins in higher levels. The tests for other bin values of RGB histogram did not yield a significantly different observation from the one made in the section 4. Despite 22 other tests with different combinations, it seems that 9-bin & level-1 combination yields the best result for the RGB histogram. It can be concluded that while it is possible to increase the performance of grayscale and gradient histograms by increasing the number of levels and decreasing the number of bins, 3D color histogram provides enough quantization levels in level-1 and it performs better when working on the image itself rather than its patches.

- Note that all of the results provided in this report are obtained by using the *validation_queries.dat* file.
- Results might have been affected by the image size inconsistency discussed in the forum. Until I see the related post, I have already done almost all of the tests. There was not enough time to conduct all the experiments again and, after having already done many experiments, I decided not to add resizing functionality in order to conduct all experiments consistently. In brief, the current implementation uses image sizes as they are.