CENG 483

Introduction to Computer Vision

Fall 2023-2024

Take Home Exam 1

Instance Recognition with Color Histograms

Full Name: **Meriç Karadayı**

Student ID: **2448553**

1 3D Color Histogram (RGB)

Q. Interval	Query Set		
	Query 1	Query 2	Query 3
128	0.735	1.0	0.05
64	0.995	1.0	0.14
32	1.0	1.0	0.12
16	1.0	1.0	0.145

Table 1: Top-1 accuracy results using 3D color histogram (RGB).

When we compare the accuracy results of queries. We can say that most accurate results are from Query 2 and the worst are from Query 3. When we inspect the images at queries it makes total sense. When we compare the queries with support:

- Query1 is zoomed out version of support.
- Query2 is rotated version of support.
- Query 3 is color changed version of support like some filter is applied to support.

For Query2, since the all images on Query2 and Support have same pixels and only their locations are different. 100% accuracy is result of this condition. Also quantization intervals size does not matter due to this reason.

For Query1, images are very similar to Support when we examine them, but their pixels are not same like Query2 because of the zoom process. And we are making predictions according their RGB values. Therefore, accuracy values are high but not accurate as query2.

For Query3 values accuracy results are way less than other 2 queries.Since colors of the images are changed for Query 3 with respect to Support, their pixel RGB values also different from the Support. Therefore this will cause low accuracy with such method to find image similarity. Images are similar but not the aspect of RGB values.

If we look at the affect the change of quantization intervals on the accuracy, we can see that decreasing interval also results higher accuracies. I believe that the reason for this situation can be explained as follows. When we increase the interval size we lose information about RGB color values. For example let's consider a $pixel_1$ as R = 255,G = 255 and B = 255 and a $pixel_2$ as R = 129,G = 129 and B = 129. When we take the interval size as 128 they both considered as the same color. However, if we decrease the interval size to 64, they are no longer considered as the same color in the histogram. Therefore smaller interval size preserves more information then the bigger interval size. So this increases our accuracy.

2 3D Color Histogram (HSV)

Q. Interval	Query Set		
	Query 1	Query 2	Query 3
128	0.82	1.0	0.18
64	0.99	1.0	0.195
32	1.0	1.0	0.135
16	1.0	1.0	0.125

Table 2: Top-1 accuracy results using 3D color histogram (HSV).

When we compare the 3D color histogram with respect to HSV and RGB values, we can clearly see that HSV have better results than RGB for both Query1 and Query3. Query2 results are 1.0 for both. For this increament at the result,

For Query1, therefore I believe that for this dataset zooming in the image is changing mostly the HSV value rather than RGB value. But, I think making a generalization such as "Zoomed images would give better accuracy with their HSV for image recognition" would not be true. This can vary according to the dataset. In our dataset zoom amount is not too big therefore mostly the tone of the image is changing and this lead smaller differences in HSV. Actually we can also see that for quantization interval = 64 RGB is giving more accurate result which is not fitting into to generalization.

For Query3, again most of the results are getting better. As I mentioned at Question 1 difference between Query3 and and Support is that there is some operation like filtering, so for such result I believe that this filtering is applied on their HSV values mostly. Therefore it is giving better results.

3 Per-Channel Color Histogram (RGB)

Q. Interval	Query Set		
	Query 1	Query 2	Query 3
128	0.555	0.995	0.04
64	0.915	1.0	0.14
32	0.97	1.0	0.17
16	0.97	1.0	0.19
8	0.975	1.0	0.185

Table 3: Top-1 accuracy results using per-channel color histogram (RGB).

When we analyze the results for this section. Again for the accuracies we can see that Query2 >Query1 >Query3. This result is same as the 3D, so I think the reasons for such ordering is same.

When we make the comparison between 3D and Per-Channel for RGB values.

- For Query1, 3D values are giving better results than per-channel. 3D color values are carrying more information than per-channel histogram. This can be explained as follows. For example there are 2 images with 2 pixels as following.

$Image_1$ - $pixel_{11}$ = (255,255,0), $pixel_{12}$ = (0,0,255)

$Image_2$ - $pixel_{21}$ = (255,0,255), $pixel_{22}$ = (0,255,0)

If we make the histogram with per-channel we will get same result for both images, although even none of their pixels are same. But If we make the histogram with 3D we will have 2 different histograms. This is because of 3D is carrying more information.

So since the images on the Query1 are similar as RGB values using less information with per-channel approach decreased accuracy.

- For Query2 unlike 3D for interval = 128 we could not get 1.0 accuracy. I believe that the reason for this is that the information carried is decreased too much such that 2 different images look same on histories although they are not same. I think even changing the image compare order can change the lead to 1.0 accuracy.

- For Query3, we are having a little better result than 3D version. I think the main reason for this is that for some images RGB values of pixels are different mostly for only one channel. For example Support has R = 255, G = 255, B = 255 for a pixel and Query3 has R = 255, G = 255, B = 0 after filter. With per-channel histogram 2/3 of the channels are same but 3D does not evaluate these pixels that similar. So use of per-channel results in a little higher accuracy. I think this could be different for a different filter applied dataset.

4 Per-Channel Color Histogram (HSV)

Q. Interval	Query Set		
	Query 1	Query 2	Query 3
128	0.635	1.0	0.175
64	0.97	1.0	0.285
32	1.0	1.0	0.325
16	1.0	1.0	0.32
8	1.0	1.0	0.335

Table 4: Top-1 accuracy results using per-channel color histogram (HSV).

The comparison result between per-channel(RGB)/per-channel(HSV) and 3D(RGB)/3D(HSV) are so similar. Like the 3D histogram, using HSV for image similarity again increased accuracy for the all queries. The main difference between affect of HSV between per-channel and 3D is increment value at Query3 results. It has a significant rise than the 3D.

For Query1, I think the reason behind the increment is same as the 3D version. For this dataset zooming changing RGB values more than HSV.

For Query2, we can accept the results same as the RGB because I think the accuracy difference between RGB (0.995) is negligible.

For Query3, we see a increase on the accuracies. After we have observed this result, I think we can say that the filter applied for the images for creating Query3 is mainly changing only one channel at HSV. In RGB part I have answered that there are images with only one channel change done for RGB channels, but after this analyze I think mostly the changes are done in HSV channels.

Best Configuration

- Color space: HSV

- Quantization interval for 3D color histogram: 64
- Quantization interval for per-channel color histogram: 8

5 Grid Based Feature Extraction - Query set 1

Histogram Type	Spatial Grid			
	2×2	4×4	6×6	8×8
3D	1.0	1.0	1.0	1.0
Per-Channel	1.0	1.0	1.0	1.0

Table 5: Top-1 accuracy results on query set 1.

5.1 Questions

- There is not any difference between the results when grid is changed.
- For Query1 which is zoomed in version of the Support best accuracies result without use of grid were 0.99 for 3D and 1 per channel with HSV and interval = 64 configuration. These were already high accuracies for our dataset. After we have used grids we also add some coordinate informations of pixels in image. So increasing the information such way made our accuracy 1.0 for also the 3D. However, by adding grids we are increasing the histogram number and this increasing calculation time and also occupying memory. I think the advantage ,the increase of accuracy, does not compensate the time-memory disadvantage for bigger spatial grids.

6 Grid Based Feature Extraction - Query set 2

Histogram Type	Spatial Grid			
	2×2	4×4	6×6	8×8
3D	0.56	0.495	0.425	0.37
Per-Channel	0.695	0.6	0.555	0.55

Table 6: Top-1 accuracy results on query set 2.

6.1 Questions

- Query2 is rotated version of Supports, so using grids are actually decreasing our accuracy because in such scenario adding coordinate information is misleading. Moreover by increasing the grid number we are providing more information about the location of pixels and this is decreasing our results more.
- For this query using spatial grid will not help our program to predict their similarities, however if we use this results for comparing the effects of increasing grid number. If we observe such a significant reduction pattern after comparison, we can interpret that the images may have been rotated.

7 Grid Based Feature Extraction - Query set 3

Histogram Type	Spatial Grid			
	2×2	4×4	6×6	8×8
3D	0.245	0.3	0.33	0.365
Per-Channel	0.53	0.695	0.805	0.835

Table 7: Top-1 accuracy results on query set 3.

7.1 Questions

- Increasing the grid number is increasing our accuracy. Since more grid means more precise location and more information. It is also increasing the accuracy.
- For Query3, without grids our results were 0.195 for 3D interval = 64 and 0.335 for per-channel interval = 8. When we compare the old result with the new ones even using a 2x2 = 4 grid is increasing accuracy remarkably use of 8x8 = 64 grid is improving the result wildly. The filter of Query3 is causing change in color but it does not cause change value on pixels location. Therefore adding this information is slightly balancing the different color values.

8 Additional Comments and References

After this experimental homework, I have learnt that histograms also can be used for calculating image similarities for some cases like slightly zoomed images, rotated images and color filtered applied images. When histograms are being built deciding interval size, use of per-channel or 3D color approach and use of HSV or RGB values have affects on the results. Most of the time using smaller interval sizes improves the accuracy. But I think use of per-channel/3D and HSV/RGB are mostly depends on the image data set. Because even though 3D is containing more information than per-channel, we had better accuracy results for per-channel.

But the biggest disadvantage using histograms is that absence of pixels location information for some cases. But using grid based feature extraction methods can be a solution for this problem for some scenarios as we have seen at Query3. In addition making the grids smaller to increase grids and histograms are also can improve the results if there is not any rotation, mirror like operation. I also think that adding a grid for images where the camera position is different will have a bad effect. Maybe crossing the grid result for calculating similarity can help to improve that scenario.

To sum up, using histograms for image recognition can useful for some cases but I think it has many disadvantages and misleading probabilities when used.