Project 2: Content-based Image Retrieval

by Yixiang Xie for CS5330 CVPR

Project Description

This is the second project for the class CS5300 Pattern Recognition & Computer Vision. The idea of the project is to let students continue the process of learning how to manipulate and analyze images at a pixel level.
The overall task for the project is, given a target image and a database of images, find images in the database with similar content to the target image. Instead of using neural networks or object recognition methods, this project will focus on more generic characteristics of the images such as color, texture, and their spatial layout. This project will allow students to practice with working with different color spaces, histograms, spatial features, and texture features.
Task 1 is to perform baseline matching by using the 9x9 square in the middle of the image as the feature vector and using sum-of-the-differences as the distance metric. This allow the students to set up the pipeline to read the images, generate the feature vectors, do the comparisons and get the top N results.
Task 2 is to perform histogram matching by using a single normalized color histogram as the feature vector and using histogram intersection as the distance metric. Students should write their own code to calculate histogram and distance metric.
Task 3 is to perform multi-histogram matching. This is similar to task 2, but using two or more color histograms and the histograms should represent different spatial parts of the image.
Task 4 is to perform texture and color matching by using a whole image color histogram and a whole image texture histogram as the feature vector. Students should choose a texture metric, e.g. gradient magnitude.
Task 5 asks students to choose a specific type of image to execute CBIR. Students need to select a subset of the data as the training data set and design a feature vector and a distance metric for that type of image. Finally students evaluate the method using the remaining data.

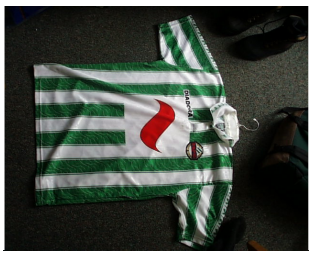
Required Images

1. Baseline Matching

Use the 9x9 square in the middle of the image as the feature vector. Use sum-of-the-differences as the distance metric. Use pic.1016.jpg as the target image and find top three matches. The top three matches here are pic.0986.jpg, pic.0641.jpg, and pic.0233.jpg.


pic.1016.jpg


pic.0986.jpg

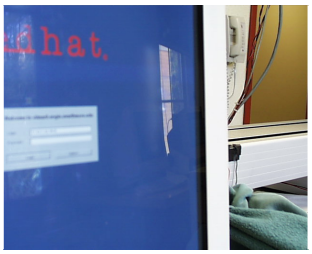

pic.0641.jpg


pic.0233.jpg

2. Histogram Matching

Use a whole image rg chromaticity histogram using 16 bins for each of r and g as the feature vector. Use histogram intersection as the distance metric. Use pic.0164.jpg as the target image and find top three matches. The top three matches here are pic.0080.jpg, pic.0489.jpg, and pic.0461.jpg.


pic.0164.jpg


pic.0080.jpg


pic.0489.jpg


pic.0461.jpg

3. Multi-histogram Matching

Use two RGB histograms using 8 bins for each of RGB as the feature vectors, representing the top and bottom halves of the image. Use histogram intersection as the distance metric. The two distances generated are weighted equally. Use pic.0274.jpg as the target image and find top three matches. The top three matches here are pic.0273.jpg, pic.0409.jpg, and pic.1031.jpg.


pic.0274.jpg


pic.0273.jpg


pic.0409.jpg


pic.1031.jpg

4. Texture and Color Matching

Use the whole image RGB histogram using 8 bins for each of RGB and whole image gradient magnitudes histogram using 8 bins for the magnitudes as the feature vectors. Use histogram intersection as the distance metric. The two distances generated are weighted equally. Use pic.0535.jpg as the target image and find top three matches. Compare this matching methods with task 2 and 3.
Task 4 - whole image RGB histogram and whole image gradient magnitudes histogram top three matching for pic.0535.jpg are pic.0636.jpg, pic.0582.jpg, and pic.0829.jpg.
Task 2 - whole image rg chromaticity histogram top three matching for pic.0535.jpg are pic.0733.jpg, pic.0731.jpg, and pic.0340.jpg.
Task 3 - top and bottom halves RGB histogram top three matching for pic.0535.jpg are pic.0285.jpg, pic.0952.jpg, and pic.1105.jpg.
From the results we can see that, task 2 and task 3 methods mainly focus on matching the overall color, while task 4 methods also takes the texture into consideration (e.g. the horizontal stone tiles in the walls).

pic.0535.jpg pic.0636.jpg pic.0582.jpg pic.0829.jpg

Task 4 target and top three matches

pic.0535.jpg pic.0733.jpg pic.0731.jpg pic.0340.jpg

Task 2 target and top three matches

pic.0535.jpg pic.0285.jpg pic.0952.jpg pic.1105.jpg

Task 3 target and top three matches

5. Finding red furry toys

For task 5, I chose images containing red furry toys to execute CBIR. The subset of data set I chose is from pic.0941.jpg to pic.0980.jpg. The method I designed to find the red furry toys pays extra attention to the center of the images. This method uses a) the whole image RGB histogram, b) the RGB histogram of half-size of the image at the center and c) the gradient magnitude histogram of that, d) the RGB histogram of quarter-size of the image at the center and e) the gradient magnitude histogram of that and f) the RGB histogram of eighth-size of the image at the center and g) the gradient magnitude histogram of that, each using 8 bins for each of RGB or 8 bins for the magnitudes as the feature vectors. Use for each pair of histograms, the method uses histogram intersection as the distance metric. The weight of the distances are 1, 2, 2, 4, 4, 8, 8.
To evaluate this method, I picked two target images pic.0928.jpg and pic.1015.jpg, and got the top ten matching results. For image pic.0928.jpg, the top ten matching results are pic.0937.jpg, pic.0924.jpg, pic.0930.jpg, pic.0933.jpg, pic.1016.jpg, pic.0935.jpg, pic.0399.jpg, pic.1048.jpg, pic.0627.jpg and pic.0547.jpg. For image pic.1015.jpg, the top ten matching results are 1014.jpg, pic.1017.jpg, pic.0328.jpg, pic.1010.jpg, pic.0329.jpg, pic.0221.jpg, pic.0233.jpg, pic.0170.jpg, pic.1019.jpg and pic.1016.jpg.

Target image pic.0928.jpg

pic.0937.jpg pic.0924.jpg pic.0930.jpg pic.0933.jpg pic.1016.jpg

pic.0935.jpg pic.0399.jpg pic.1048.jpg pic.0627.jpg pic.0547.jpg

Top 10 Matching images for pic.0928.jpg, 4/10 correctness

Target image pic.1015.jpg

pic.1014.jpg pic.1017.jpg pic.0328.jpg pic.1010.jpg pic.0329.jpg

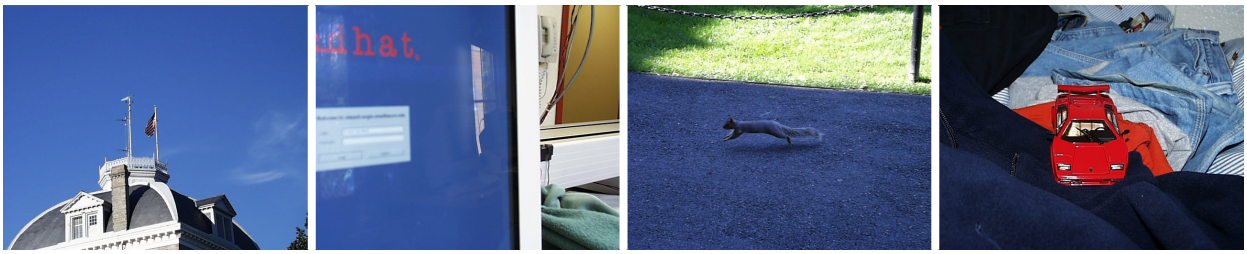
pic.0221.jpg pic.0233.jpg pic.0170.jpg pic.1019.jpg pic.1016.jpg

Top 10 Matching images for pic.1015.jpg, 5/10 correctness

Extensions

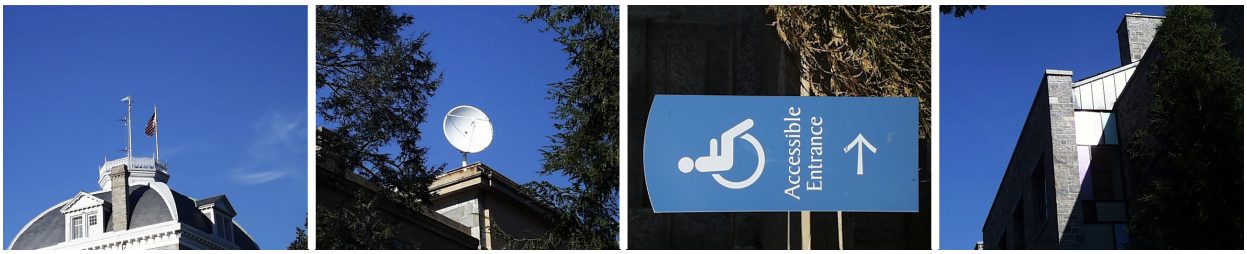
1. Extending task 2 - whole image RGB histogram matching

Besides using the whole image rg chromaticity histogram using 16 bins for each of r and g, I also implemented a histogram matching method using the whole image RGB histogram using 8 bins for each of RGB and histogram intersection as the distance metric. The results are as the following. The whole image RGB histogram seems to look more unified compare to using rg chromaticity histogram.



pic.0164.jpg pic.0080.jpg pic.0489.jpg pic.0461.jpg

Whole image rg chromaticity histogram



pic.0164.jpg pic.0110.jpg pic.0976.jpg pic.1032.jpg

Whole image RGB histogram

2. Extending task 3 - different combinations of image parts

Besides using two RGB histograms representing the top and bottom halves of the image, I also implemented methods using rg chromaticity histograms representing the top and bottom halves of the image, both histograms representing the top and bottom halves of the image and the quarter-size image of the center (weight of the distances are 1, 1, 2), and both histograms representing the whole image and the quarter-size image of the center (weight of the distances are 1, 2). The results are as the next page.

From the results we can see that using RGB histogram or using the center of the image and weighing more on that also has a better change at finding the water tower. Using the top and bottom halves or the whole image combined with the center the image doesn't seem to have a better performance.




pic.0274.jpg pic.0273.jpg pic.0409.jpg pic.1031.jpg

8 bins RGB histograms of the top and bottom halves of the image




pic.0274.jpg pic.0409.jpg pic.1055.jpg pic.0825.jpg

16 bins rg chromaticity histograms of the top and bottom halves of the image




pic.0274.jpg pic.0275.jpg pic.1055.jpg pic.0273.jpg

8 bins RGB histograms of the top and bottom halves and center of the image




pic.0274.jpg pic.1055.jpg pic.0825.jpg pic.0213.jpg

16 bins rg chromaticity histograms of top and bottom halves and center of the image




pic.0274.jpg pic.1055.jpg pic.0275.jpg pic.0209.jpg

8 bins RGB histograms of the whole image and center of the image




pic.0274.jpg pic.0209.jpg pic.0213.jpg pic.0825.jpg

16 bins rg chromaticity histograms of the whole image and center of the image

3. Extending task 4 - magnitude, orientation or 2D gradient, alone, or in combination with color information

Besides using the whole image RGB histogram and the whole image gradient magnitudes histogram as the feature vectors, I also implemented using the gradient magnitudes alone, using the gradient orientations alone and using the 2D histogram of gradient orientation and magnitude. Also these three methods methods combined with either rg chromaticity histogram or RGB histogram. The results are as the following.

From the results we can see that using the gradient alone is not the best idea. It usually finds images that are not relevant. It is best combined with RGB histogram to get the color information and texture information. And between magnitude, orientation and 2D gradient, the magnitude seems to have a better result regarding the wall texture.

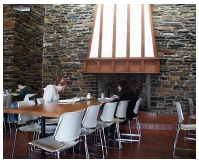

pic.0535.jpg


pic.0262.jpg


pic.0106.jpg


pic.0410.jpg

16 bins histogram of magnitude alone

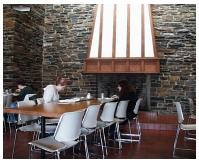

pic.0535.jpg


pic.0636.jpg


pic.0582.jpg


pic.0829.jpg

8 bins histogram of magnitude and RGB histogram (weight 1, 1)

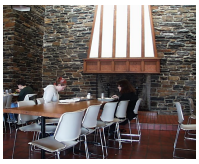

pic.0535.jpg


pic.0733.jpg


pic.0004.jpg


pic.0745.jpg

16 bins histogram of magnitude and rg histogram (weight 1, 1)

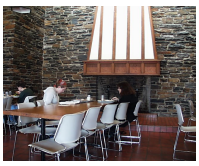

pic.0535.jpg


pic.0394.jpg

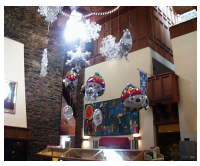

pic.1061.jpg


pic.0419.jpg

16 bins histogram of orientation alone

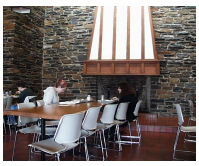

pic.0535.jpg


pic.0582.jpg


pic.0628.jpg


pic.0285.jpg

8 bins histogram of orientation and RGB histogram (weight 1, 1)

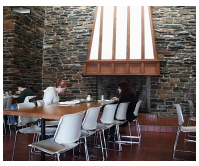

pic.0535.jpg


pic.0733.jpg


pic.0731.jpg


pic.0340.jpg

16 bins histogram of orientation and rg histogram (weight 1, 1)

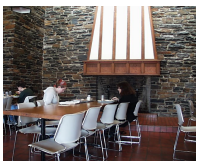

pic.0535.jpg


pic.0568.jpg


pic.0576.jpg


pic.0262.jpg

16 bins histogram of 2D gradient alone

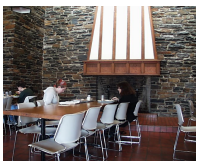

pic.0535.jpg


pic.0582.jpg


pic.0636.jpg


pic.0628.jpg

8 bins histogram of 2D gradient and RGB histogram (weight 1, 1)


pic.0535.jpg


pic.0733.jpg


pic.0745.jpg


pic.0171.jpg

16 bins histogram of 2D gradient and rg histogram (weight 1, 1)

4. Extending task 5 - finding different objects

Based on the method I developed in task 5, I fine tuned the weights for finding objects of different sizes. The methods in task 5 is for finding small objects, and has the weights configured as 1, 2, 2, 4, 4, 8, 8, giving more preference to the smallest size of image at the center. For medium size objects, the weights are configured as 1, 2, 2, 8, 8, 4, 4. For large size objects, the weights are configured as 1, 8, 8, 4, 4, 2, 2.

For finding the blue trash can bins with pic.0287.jpg, using the medium configuration returns 4/11. It also finds 2 post bins that are pretty similar to the blue trash can bins.


pic.0287.jpg


pic.0289.jpg


pic.0665.jpg


pic.0291.jpg


pic.0666.jpg


pic.0214.jpg


pic.0548.jpg


pic.0288.jpg


pic.0132.jpg


pic.0407.jpg


pic.0999.jpg


pic.0920.jpg

For finding the green trash can bins with pic.0746.jpg, using the large configuration returns 7/11. It also finds 4 green bottles.


pic.0746.jpg


pic.0750.jpg


pic.0751.jpg


pic.0904.jpg


pic.0748.jpg


pic.0896.jpg


pic.0747.jpg


pic.0921.jpg


pic.0753.jpg


pic.0752.jpg


pic.0755.jpg


pic.0880.jpg

For finding the bananas with pic.0343.jpg, using the small configuration returns 2/11. It returns three duck toys that has a similar yellow color and smooth texture.


pic.0343.jpg


pic.0347.jpg


pic.0345.jpg


pic.0458.jpg


pic.0279.jpg

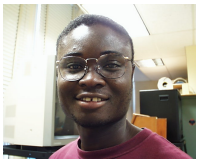

pic.0514.jpg


pic.0464.jpg


pic.0406.jpg


pic.0463.jpg


pic.0469.jpg


pic.0363.jpg

pic.0178.jpg

Reflections

In this project, I learned more about manipulating and analyzing images at pixel levels. I learned how to perform content-based image retrieval using color, texture, and their spatial layout. I developed several matching methods exploring different color spaces, histograms, spatial features, and texture features, such as rg chromaticity histograms, RGB histograms, gradient magnitudes, orientations, 2D gradients. This is a good exercise before learning to use object recognition methods to do the same thing.