Computer Vision homework-1

September 23, 2015

Name: Michael Jaison Gnana Sekar

AndrewId: mgnanase

Date: 9/22/2015

1.0 Properties of filter functions:

Filters 1-5:

Guassian filters with sigma - 1,2,4,8,sqrt(2)*8.

These filters blur the image at various intensities, so that noise is reduced and images are smoothened.

Filters 6-10:

Laplacian of Guassian filter with same scale.

These filters convolute with the second derivative of guassian filter. This helps to pick up and highlight the edges in the images.

Filters 11-15:

x-derivative of Guassian filter with same scale.

These filters convolute with the first x-derivative of guassian filter. It highlights the change in image along the x-direction.

Filters 16-20:

y-derivative of Guassian filter with same scale.

These filters convolute with the first y-derivative of guassian filter. It highlights the change in image along the y-direction.

2.5 Quantitative Evaluation

System parameters:

k-means - number of visual words	100
alpha	50

Confusion matrix:

	airport	auditorium	bedroom	campus	desert	football stadium	landscape	rainforest
airport	12	2	1	2				3
auditorium	4	9	3	1	1	1	1	
bedroom	5	3	7	1	2	1	1	
campus	3	2		6	2	1	3	3
desert		1	1	1	11	2	4	
football stadium		1	1	3	2	9	3	1
landscape	3	1		3	4	2	5	2
rainforest	2	1		1			1	15

Accuracy: $\frac{74}{160} = 46.25\%$

There are 74 images classified correctly out of 160 images. This gives the accuracy of 46.25%.

2.6 General trends in errors

If we analyze each category of images seperately and calculate accuracy for them, we can understand the behaviour of images and its similarity with other images.

Airport images: More than half of the images are classified correctly. But false positives are more than false negatives in this category. A lot of other category images are wrongly guessed as Airport images. 29 images are classified as airport images, out of which only 12 are true airport images. Accuracy is $\frac{12}{20} = 60\%$.

Auditorium images: Most of the wrongly classified images are classified as Airport. This may be due the similariy of lounges in Airports and the auditorium itself. It is difficult for the system to identify patterns between them. False positives and False negatives are more similar for this category. Accuracy is $\frac{9}{20} = 45\%$.

Bedroom images: Most of the images are wrongly classified as Airport or Auditorium. Accuracy is $\frac{7}{20} = 35\%$.

Campus images: Misclassified images are wrongly assiged with Airport, Landscape and Rainforest labels. The system struggles a lot in this category. It may be due to different aspects of campus. A campus picture may show a students lounge which may be similar to airport lounges, may show campus buildings which may be similar to Landscape category, may show trees inside campus which may confused with Rainforest. Accuracy is $\frac{6}{20} = 30\%$.

Desert images: Even though, It is more easier to classify desert images compared with other category, the system finds it difficult to differentiate some deserts with landscape category. Accuracy is $\frac{11}{20} = 55\%$

Football stadium: Some of it are wrongly classified as Campus and Landscape. It might be the football stadium might have the same grass pattern as trained campus images, also some the stadium outer layer may resemble landscape images. Accuracy is $\frac{9}{20} = 45\%$

Landscape: The system failed the most in classifying landscape images. Only 5 images are identified correclty out of 20. True landscape images are wrongly classified as Campus or Desert images. Also most of the Campus, Desert and Football stadium. Accuracy is $\frac{5}{20} = 25\%$

Rainforest: The system attained a great success in identifying rainforest images. It may be easier because of the greenish layout and the patterns of trees in the images. But, some of the campus images which had trees are wrongly assigned with this label. Accuracy is $\frac{15}{20} = 75\%$.

Among the 8 categories of images, Landscape and Campus classes performed poorly in the bags of word approach. The systems uses the pattern in the image to recognize the scene, and it migh

2.7 Custom work on algorithm

Acuracy persective:

- Surprisingly, after increasing the k and alpha, the accuracy decreased by 1%. (with k=150, alpha=250).
- Changing k = 200 & alpha = 150, increased the accuracy by 0.5%.

- Even increasing the number of layers to 3 in SPM did not help much with the accuracy of classification.
- By removing SPM and using histograms of the image direcly, the system was able to classify 70 test images correctly, instead of 74 with SPM.

Increasing the number of visual words (k), or the amount of sampling pixels per image (alpha) supplied to the k-means clustering algorithm does not influence the accuracy of the system drastically.

Performance perspective:

To increase the performance of the system, and to reduce the time elapsed to train the system, I used second level gaussian pyramid approach to scale down the images. This increased the performance drastically as mentioned below.

Modules performance before and after gaussian pyramid:

- createDictionary: 744.8 seconds -> 490.36 seconds. After the change, Efficiency increased by 34.5%.
- buildRecognitionSystem: 10 secondes -> 3.73 seconds. Efficiency increased by 62.7%.
- evaluateRecognitionSystem: 75.4 seconds -> 47.52 seconds. Efficiency increased by 63.02%.

Even though this approach increased the performance drastically, it affects the accuracy in a very big scale. The accuracy of the system came down to 35.62%, compared with 46.25% in original system.