MMAI HW2 R02922083 邵元輔

HW2-(1)

File Briefing

The main code for HW2-(1) is in hw2_main.m with Matlab. I have implement regional hsv color histogram with L1, L2 and L6. The color feature is extracted with function ColorVector(img, grid_x, grid_y) located in ColorVector.m. grid_x, grid_y is the image slicing number of x and y. The similarity calculateion function is ColorSimilarity(vector_a, vector_b, type) in ColorSimilarity.m. The L6 covariance similarity is pre-trained by L6_learning.m and paste the matrix into ColorSimilarity.m.

Algorithm and Discussion:

The hsv is used for color histogram. The bin selection is based on [1]. I choose 2 grey bin and 2 color bin with 2 saturation bin and 10 hue bin, total 2x10x2+2=42 bin. This bin selection is considered the discrimination of human recognition and also to lower the bin for minimize the possibility of curse of high dimension and calculate faster. With this bin selection I then test with different regional split for feature input and also varies the Lx similarity to build a MAP matrix. The result is shown the Table 1. The result showed that 3x3 regional color histogram with L1 similarity could retrieve best MAP = 0.1739. The result is reasonable with L1, L2 with a similar MAP. However, L6 is not as good as shown in the paper and also degrade with high regional slicing. The explanation is suggested that the training set is too small. When doing the inverse matrix in L6-training, it is observed that the condition converges very slowly and generate warning about the singularity of the matrix. So, with a smaller data set, the matrix learning might have a worse result compared with L1 and L2. The MAP of different query of 3x3-L1 color is shown in figure 1. The best MAP is bracelet. The result could be inferred from directly observe the images of query and database that most of bracelet are close to red and located at center. And the worst case gge snack could also clearly observe from image comparison. The query of gge snack is much darker compared with other images in database.

color-histogram MAP	1x1	2x2	3x3
L1	0.1509	0.1556	0.1739
L2	0.1492	0.1436	0.1728
L6	0.1412	0.1182	0.0807

Table 1 Color histogram MAP matrix

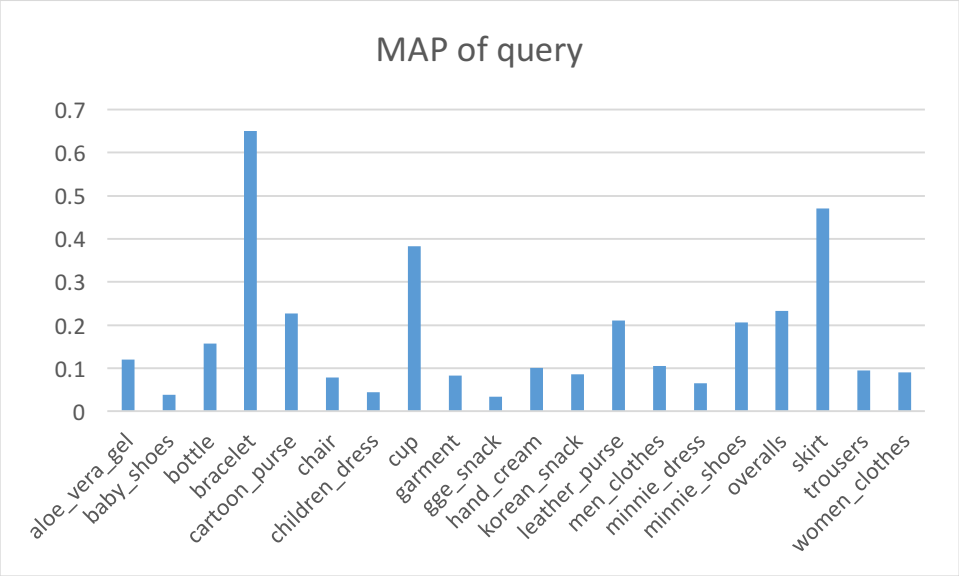

Figure 1. 3x3 regional color histogram with L1 test result of query

HW2-(2)

File Briefing

The main code for HW2-(2) is in hw2_texture.m. I have used regionial Fourier features and phog for texture feature. The regional Fourier feature is implemented by myself in FFTVector.m with function FFTVector(img, grid_x, grid_y). The PHOG feature is tested and used of function in annl_phog.m[2] with anna_phog(img, bin number, 360, pyramid layers, region of interest). 2 different texture features were tested with different grid or layer and L1 or L2 similarity.

Algorithm and discussion

Regional Fourier feature is implemented by separate the image into grids. For each region, calculate the absolute value of Fourier transform of the region and use a 5 bin of radius and 8 bin of angle to summarize the regional image. The result of Fourier feature is in figure 2. It is shown that the best of regional Fourier feature is 3x3 grid with L1 similarity. Noted that the both L1 L2 are not normalized because of a slightly performance degradation after normalization. The MAP is 0.12 which is the worst of the three different features (color, Fourier, phog). However, the MAP of query shows a totally different behavior compared with color histogram. The best query Minnie shoes are very bad in color histogram. These type of queries do have a clear texture in a large enough area. In the other hand, the performance of cup is very bad because it is dominated by the texture of background.

PHOG feature is also test with different layers and L1 L2 similarity. For more layer, the features will growth dramatically because it considers all layer information below. For example layer=1 with 50

features if we set the split bin to 10. But layer=3 with 850 features with the same settings. The result is in table 3. It is shown that the best of PHOG feature is 2 layers with L1 similarity. The MAP is 0.1917, the best performance among all three different features. The distribution of MAP of query of PHOG in figure 3 looks very similar to regional Fourier but with a higher MAP value. This is because both Fourier and PHOG are texture related feature. A similar distribution is just a reflection of two group of related features.

Fourier(5r,8a) MAP	1x1	2x2	3x3
L1(not normalized)	0.1026	0.0996	0.1205
L2 (not normalized)	0.105	0.0976	0.114

Table 2 The regional Fourier feature test result

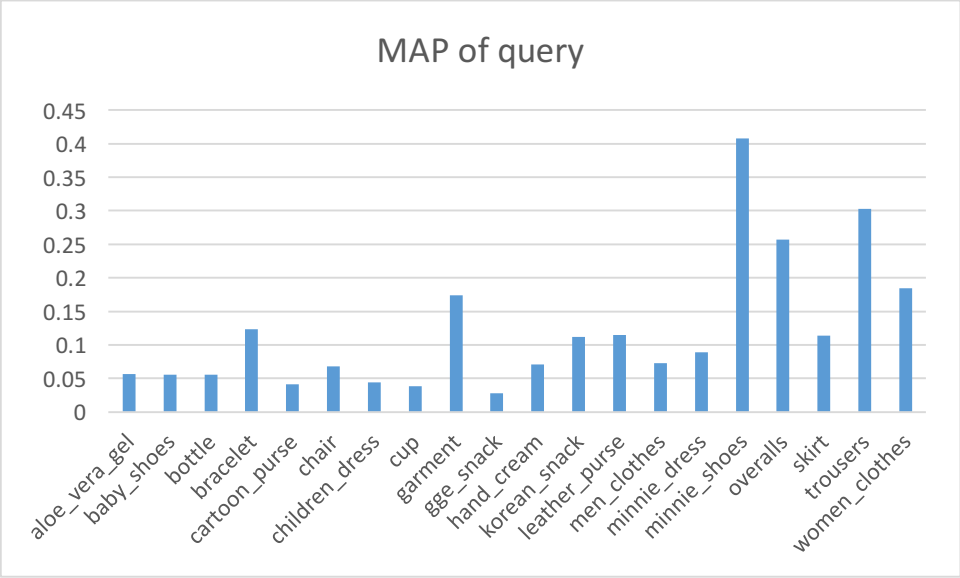

Figure 2. The 3x3 regional Fourier feature with L1 test result of query

PHOG(3-p) MAP	L1	L2	L3
L1(not normalized)	0.1791	0.1917	0.1907
L2 (not normalized)	0.1777	0.18	0.1798

Table 3 The PHOG feature test result

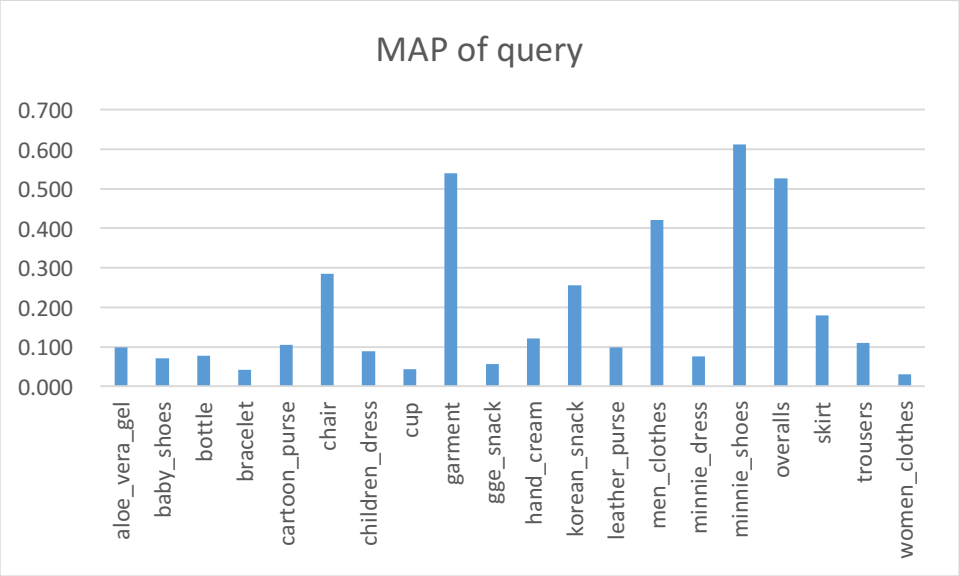

Figure 3. The layer 2 PHOG feature with L1 test result of query

HW2-(3)

File Briefing

The main code for HW2-(3) is in LMNN_learning.m. I have used LMNN[3] library for matrix learning. The function for LMNN is lmnnCG(vector, lable, k_nearest_neighbor) in lmnnCG.m. The color and texture feature is selected with the different combination and combined together for training.

Algorithm and discussion

I have test LMNN learning with different combination of features and different k nearest neighbor settings. The result is showned in table 4. The best performance is not the combination of highest features with the highest kn. The best is 2x2 color+ L2 phog +kn=14 or 1x1 color+L1phog +kn=19 instead. The result also shows that LMNN with 2 combination feature performed worse than any single feature with L1 similarity. This might also be a result of small training data set. The LMNN MAP of query is shown in figure 4. The distribution is different from color or texture feature only.

LMNN	kn=3	5	7	10	14	max kn=19
3x3 color + L3 phog	0.0848	0.068	0.0653	0.0643	0.058	0.0548
2x2 color + L2 phog	0.097	0.0891	0.0895	0.0871	0.102	0.1011
1x1 color + L1 phog	0.0684	0.076	0.0744	0.0726	0.0851	0.1022

Table 4 The LMNN with color + phog feature test result

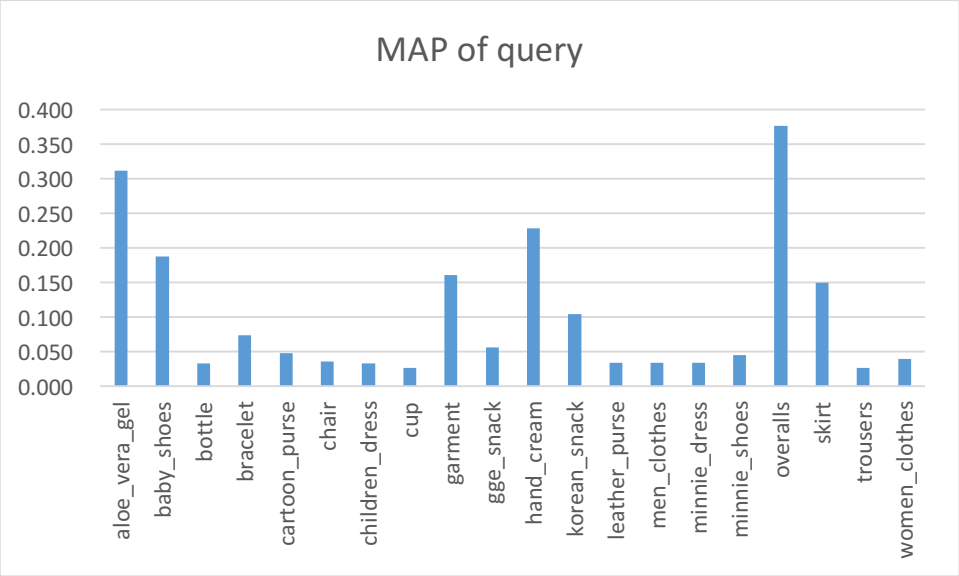

Figure 4. The layer 2 PHOG + 2x2 regional color feature LMNN learning test result of query

[Reference]

[1] *Image Databases: Search and Retrieval of Digital Imagery* Castelli'01

[2] Pyramid Histogram of Oriented Gradients (PHOG) Anna Bosch and Andrew Zisserman: http://www.robots.ox.ac.uk/~vgg/research/caltech/phog.html

[3] LMNN (large margin nearest neighbor): http://www.cs.cornell.edu/~kilian/code/lmnn/lmnn.html