

MMAI - Content-based Image Retrieval

姓名：江東峻

學號：B01902032 資工四

a. Result

topics	color MAP	shape MAP	metric learning MAP
aloe_vera_gel	0.2059	0.5085	0.0689
baby_shoes	0.3475	0.1605	0.2879
bottle	0.1480	0.2133	0.0914
bracelet	0.2124	0.1294	0.0775
cartoon_purse	0.2851	0.2106	0.1257
chair	0.0991	0.0963	0.0407
children_dress	0.3248	0.2009	0.2391
cup	0.3654	0.1297	0.0937
garment	0.4381	0.2452	0.3897
gge_snack	0.6488	0.3355	0.4265
hand_cream	0.2924	0.0980	0.0484
korean_snack	0.6670	0.2856	0.3852
leather_purse	0.1811	0.3173	0.1090
men_clothes	0.4124	0.2102	0.3496
minnie_dress	0.5979	0.1304	0.6244
minnie_shoes	0.2112	0.1002	0.0511
overalls	0.1888	0.1396	0.1824
skirt	0.4324	0.4294	0.2564
trousers	0.1140	0.2787	0.1236
women_clothes	0.3564	0.1255	0.4382
average MAP	0.3264	0.2172	0.2205

b. Methods

(1) color features

We choose HSV color space and segment the image into $4 * 4$ region. For each region, compute the $3 * 8 * 8$ (192) bins histogram and normalize it to keep the scale invariance. We use l1 distance to be the metric function. (dimension: 3072)

(2) shape features

We use OpenCV¹ to capture the HOG feature. First, we resize the image to $128 * 128$ and use HOGDescriptor(winsize=(64, 64), blocksize=(16, 16), cellsize=(8, 8), nbins=9). We also normalize the features and choose l1 distance to be the metric. (dimension: 1764)

(3) metric learning

Because we need to evaluate with the leave-one-out manner but it spends much time to train the other 399 images (1 for query), we use additional dataset (training.zip) to be the training data of LMNN model (toolkit: metric-learn²). To make dataset more balance, we pick up 20 images for each topic and duplicate images for those topics with less than 20 images. We choose color feature (region histogram) to be the feature of this model and use Mahalanobis metric to compute the distance.

c. Result Analysis

(1) No feature method is absolutely the best. We can see MAPs in “hand-cream” topic, color feature method is apparently better than shape feature method, but in “aloe_vera_gel” topic, the shape feature is better. In “chair” topic, the results in both color and shape feature method are bad, we may need texture feature method to capture the latticed pattern.

(2) Metric learning method takes much time, and l1 distance metric is a faster method and its performance is usually in a certain level.

(3) We can see the performance of LMNN metric learning method is not better than color feature method. There are some possible reason:

¹ OpenCV: a c++ and python CV toolkit, <http://opencv.org>

² Metric-learn: a python based metric learning toolkit, <http://all-umass.github.io/metric-learn/>
2/3

1. Because there are few images of some topics in the additional dataset, we choose to duplicate the images, in “bracelet” topic, there are originally 3 images and we duplicate them to 20 images. But in fact, there are only 3 images information in these 20 images.
2. The times of training iterations are not enough, or learning rate is too large. To save training time, we set a small iteration number and a large learning rate. These may cause the bad result.

But in the topics with originally lots of images, such as “women_clothes” and “minnie_dress”, we can see the MAP scores are higher than color feature method. It means the learning method works, and because the metric can be trained offline, with more data and training time, the metric learning method is a better choice.

d. Some other works

```
=====women_clothes=====
(0.3461859224220864, 'women_clothes_5.jpg')
(0.8522725898588619, 'women_clothes_18.jpg')
(0.8663253106901458, 'women_clothes_20.jpg')
(0.9206101829805302, 'women_clothes_10.jpg')
(1.0756142877876869, 'women_clothes_4.jpg')
(1.1336410344790633, 'women_clothes_16.jpg')
(2.052092016165151, 'women_clothes_17.jpg')
(2.9273638128935544, 'women_clothes_15.jpg')
(4.035930637416495, 'women_clothes_8.jpg')
(4.638335583550181, 'gge_snack_14.jpg')
(4.677005975648134, 'women_clothes_13.jpg')
(4.9076137886252456, 'garment_20.jpg')
(4.9601370748970846, 'baby_shoes_13.jpg')
(4.9646711962506425, 'women_clothes_19.jpg')
(5.062617182698148, 'minnie_dress_4.jpg')
(5.0832813880238215, 'gge_snack_15.jpg')
(5.410146808593408, 'gge_snack_2.jpg')
(5.41531632041151, 'women_clothes_1.jpg')
(5.431384062058074, 'children_dress_13.jpg')
```

visualize the query results for
topic “women_clothes”

the PR curve for color feature method

