

Image Retrieval 报告

- ▶ sift bow + 余弦
- ▶ sift bow + svm
- ▶ color feature + 余弦
- ▶ cnn feature + 余弦
- ▶ cnn feature + svm
- ▶ cnn feature + 余弦 + rerank
- ▶ cnn feature + svd + svm
- ▶ Scsm
- ▶ 纯sift匹配

数据说明

- ▶ 图片经过裁剪，尽可能只保留了物品。
- ▶ 对每张图片提取sift时，将图片按长宽等比放缩到总像素数不超过 $500 * 800$ 。
- ▶ 每类物体用20张图片作为训练集合，4张图片做测试集合。
总共训练图片1040张，测试图片208张

Sift Bow + 余弦相似度

使用最简单sift bow特征，利用特征间向量的余弦计算图片相似度

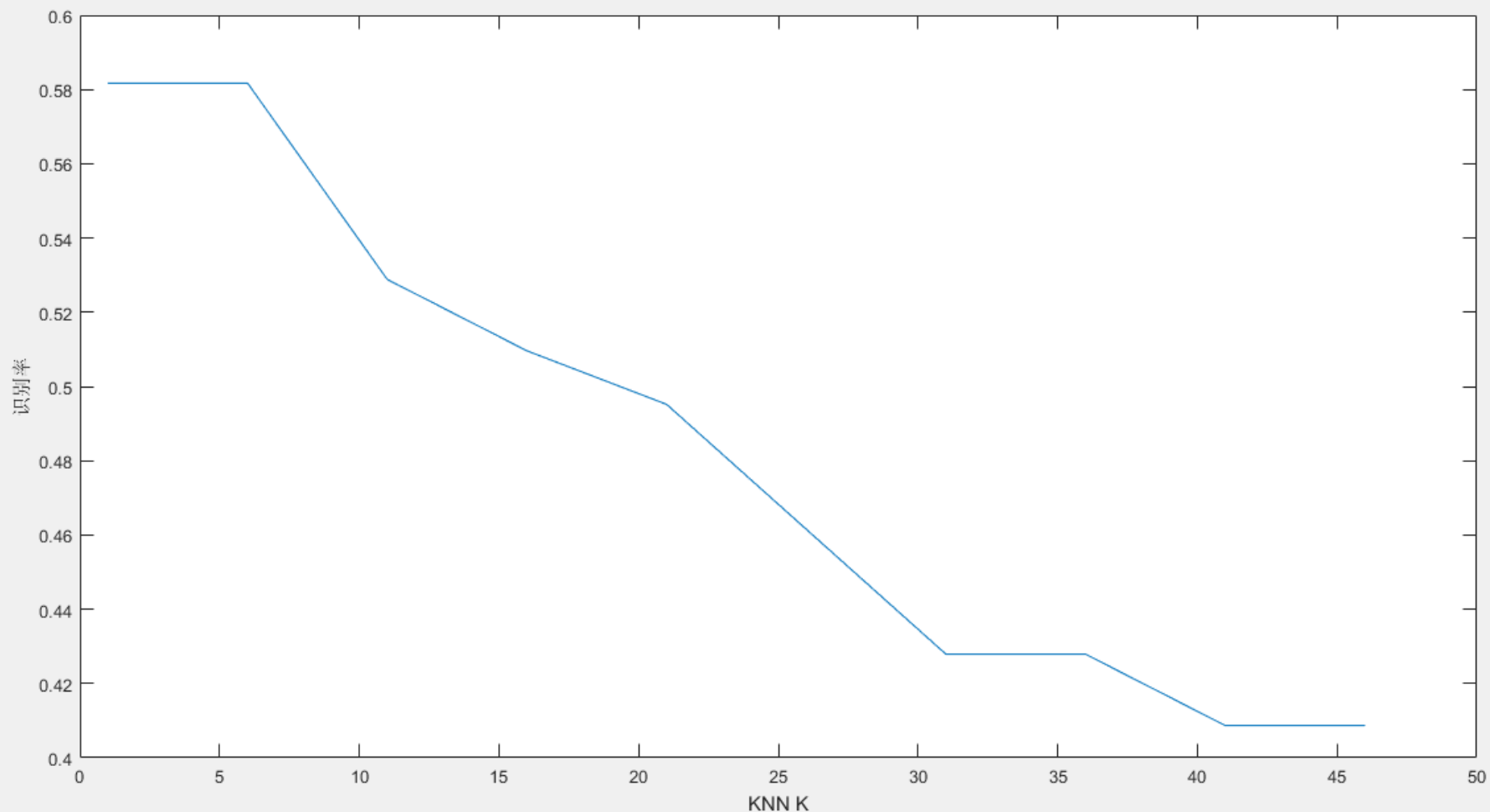
最大识别率为0.5817

Kmeans 聚类 $k = 200$

源码参考 `cv/cv_bow/bow_query`
和 `cv/cv_bow`

Sift Bow + 余弦相似度

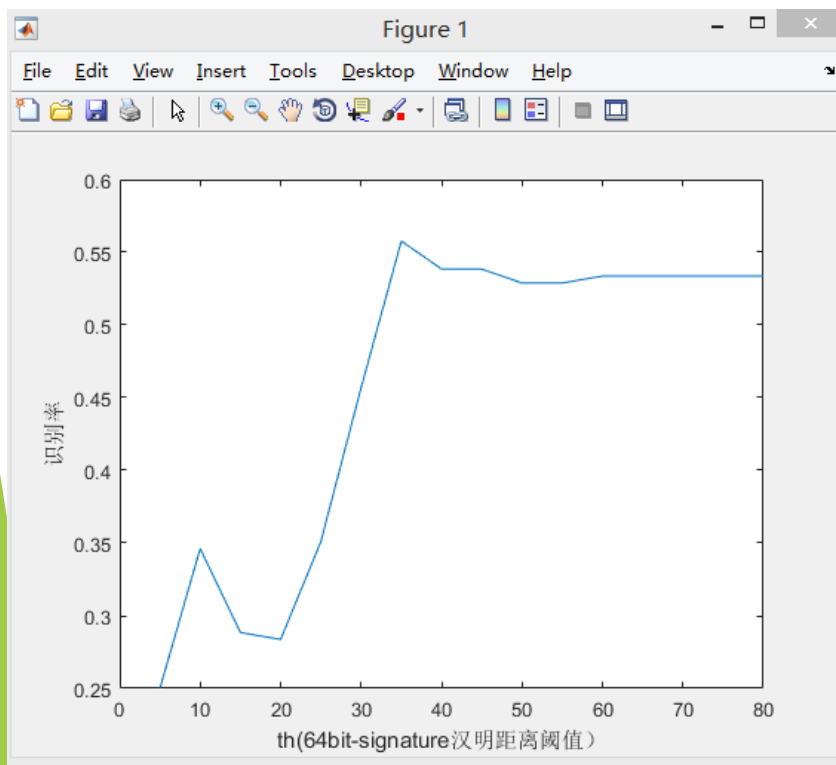
KNN: 基本上随着K的增大, 识别率逐渐下降



Sift Bow + 余弦相似度

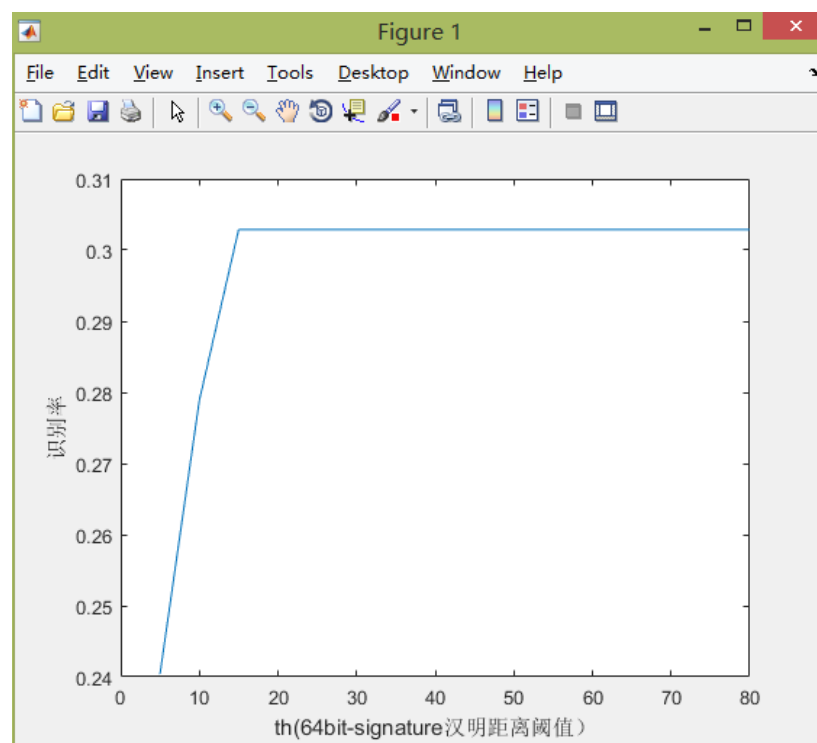
64bit-signature

Score = 1



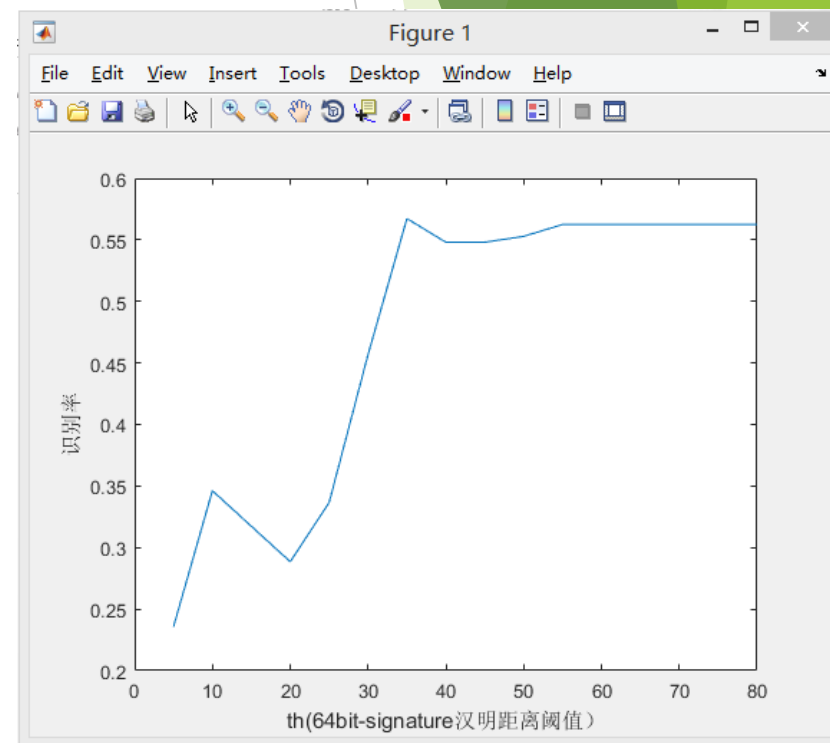
最大识别率: 0.5577

Score = $\exp(-\text{dis}^2/\sigma^2)$



最大识别率: 0.3029

Score = $\log(65 - \text{dis})$



最大识别率: 0.5673

Sift Bow + SVM

- ▶ 将提取到的sift bow特征作为SVM的输入，类别作为输出
- ▶ svm是二分类器，为每一个类建立一个svm分类器进行多分类
- ▶ 最大识别率：0.7308（我测试的所有方法中最高的）

```
Elapsed time is 119.174261 seconds.
```

```
ans =
```

```
    right: [1x208 uint16]
```

```
    test: [1x208 uint16]
```

```
rightRate: 0.7308
```

源码参考 [cv/cv_bow/bow_svm_query](#)
和 [cv/cv_bow](#)

Color feature + 余弦

- ▶ 对以sift点为中心的周围 $\text{block_size} \times \text{block_size}$ (奇数) 区域内的点的颜色值做一个均值统计, 把0-255的区间等间距分成若干段, 每段长度为 bin_size 。统计均值落在每个区间的sift点数, 获得颜色的词频, 作为图片的特征
- ▶ 但是识别率只有0.2873

cnn feature + 余弦

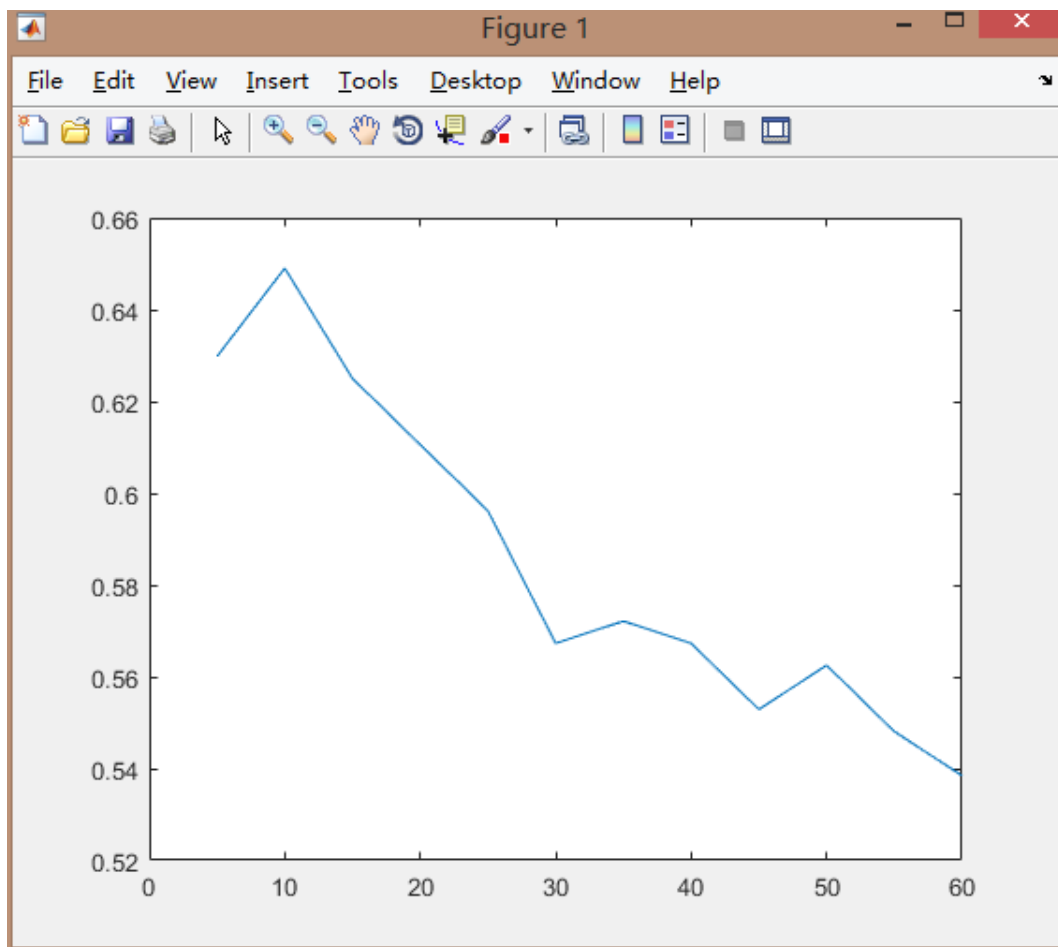
- ▶ 利用CNN提取图片的特征，把最后一层输出作为特征使用
- ▶ cnn model使用的是 imagenet-vgg-f.mat
(<http://www.vlfeat.org/matconvnet/pretrained/>)
该网站应该任何一个模型都可以，对结果影响不大
- ▶ 对特征利用余弦比较相似度

cnn feature + 余弦

最佳识别率: 0.6490

X: KNN 的 K
Y: 识别率

源码参考:
`cv/cnn/cnn_query`



cnn feature + svm

- ▶ 利用cnn提出的feature作为svm的输入，类别作为label
- ▶ 识别率：0.7067

```
Elapsed time is 0.10016 seconds.
```

```
recognize: 9, right: 9
```

```
Elapsed time is 0.139678 seconds.
```

```
recognize: 17, right: 9
```

```
rightRate =
```

```
0.7067
```

cnn feature svd + svm

- ▶ cnn提出的feature是十分稀疏的，而数据的维度1000维相对于数据的数量1040，还是很高的维度，因此想用svd对数据降维
- ▶ 随着选择的维度占总能量的比重的增大，识别率增大，但当占比 $\geq 75\%$ 后，识别率稳定在0.6298，不再提升。
占比为0.75,0.80,0.90,0.95,0.99时，识别率都是0.6298

Cnn feature + 余弦 + rerank

- ▶ 在cnn+余弦计算出的相似度后，采用论文Object retrieval and localization with spatially-constrained similarity measure提出的rerank方法，重新计算相似度。
- ▶ 识别率 0.6779 (KNN K=5)

```
res =  
  
    right: [1x208 double]  
    query: [1x208 double]  
        K: 5  
rightRight: 0.6779
```

scsm

- ▶ 识别率不超过0.3
- ▶ 猜想原因：论文描述的场景物体不存在3D上的旋转，只是相对在相机平面发生位移旋转之类的

纯sift 匹配

- ▶ 按照sift点匹配算法，匹配图片，用匹配的点对数表示相似度
- ▶ 识别率：0.6731

```
Elapsed time is 42.197125 seconds.  
recognize: 8, right: 8 |  
Elapsed time is 33.453567 seconds.  
recognize: 49, right: 9  
Elapsed time is 42.115247 seconds.  
Elapsed time is 60.991467 seconds.  
recognize: 9, right: 9  
recognize: 2, right: 9  
Elapsed time is 56.277105 seconds.  
recognize: 9, right: 9  
right rate : 0.673077
```