

Building Discriminant Visual Vocabularies

Daru Xu

Media Communication Lab

Introduction

Content-Based Image Retrieval

- Search image using image itself !!

- NO keywords input

Applications:

- Data redundancy
- Copyright infringement
- Landmark recognition ...



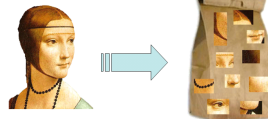
Baseline method

Visual Bag-of-Words

- Represent image as "visual word" histogram

Advantage:

- Compact representation
- Fast image comparison



Limitations

- Visual words are less descriptive



- Abandon spatial information



Semi-supervised clustering

- Aimed for classification
- Impose constraints on unsupervised clustering
- Supervised by "side information":
 - 1) "must-link" constraint ("similar" features)
 - 2) "cannot-link" constraint ("dissimilar" features)

- Based on distance metric learning

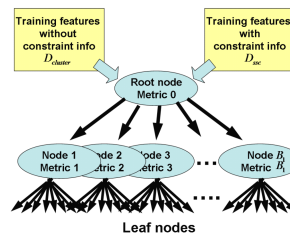
$$d_A(x_1, x_2) = \sqrt{(x_1 - x_2)^T A (x_1 - x_2)}$$

- Goal: learn Mahalanobis matrix A such that

- 1) Similar features are close
- 2) Dissimilar features are far away

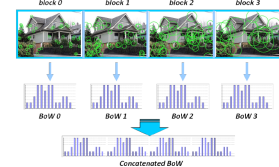
Better Visual Vocabulary

- Generate side information
- Feature correspondences between perspective images
- Hierarchical SSC framework
- Combine metric learning and Hierarchical K-means

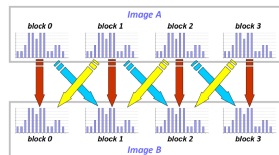


Scale-based Image Pyramid Matching (SIPM)

Define "blocks" based on scale ranges



Ordered pyramid matching – scale invariance



Experiments and Evaluations

Training data

- ImageNet: 26k images from 10k+ categories
- Aimed to be generic
- Non-overlapping training set for K-means and metric learning

Training datasets



Testing data

- Three benchmarks for near-duplicate image retrieval

	images #	Query #	Measure of Accuracy
UKBench	10200	10200	Kentucky Score
Holidays	1491	500	MAP
Oxford5k	5061	55	MAP

UKBench



Holidays



Oxford5k



Testing datasets

Results of image retrieval

- HSSC vocabulary lead to consistently better results
- Image pyramid is generally effective, but not consistent

UKBench			
	Euclidean (Kentucky Score)	HSSC (Kentucky Score)	Performance Gain
BoW	3.010	3.164	3.85%
BoW + SIPM	3.127	3.281	3.85%
Holidays			
	Euclidean (MAP)	HSSC (MAP)	Performance Gain
BoW	0.582	0.632	5.0%
BoW + SIPM	0.573	0.620	4.7%
Oxford5k			
	Euclidean (MAP)	HSSC (MAP)	Performance Gain
BoW	0.241	0.265	2.4%
BoW + SIPM	0.259	0.288	2.9%

Complexity and index storage

Processing time:

- Similar level as the classic BoW model

Storage:

- 16% extra storage for SIPM

Vocabulary training

	Euclidean	HSSC
Processing Time (minute)	48	85

Online retrieval

	Euclidean	HSSC	Euclidean +SIPM	HSSC +SIPM
Processing Time	1 X	1 X	2.7 X	2.7 X