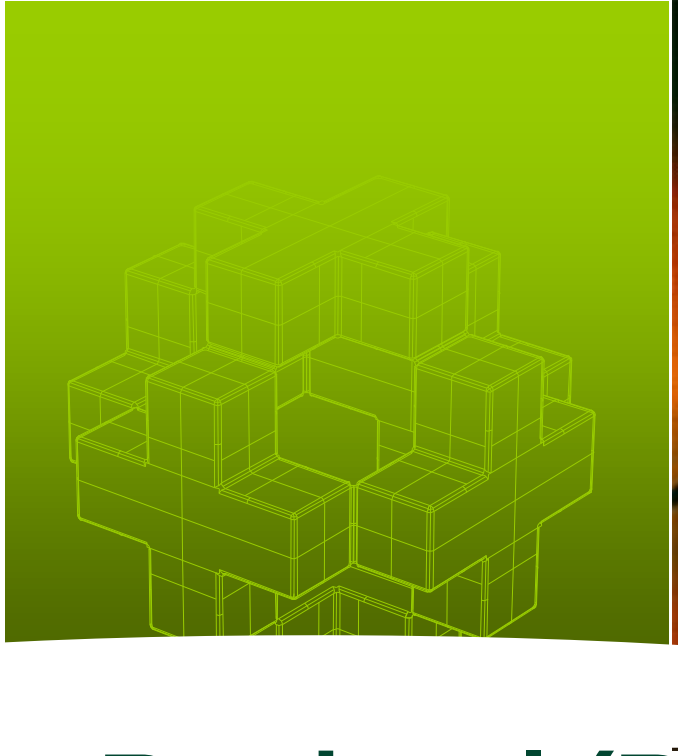
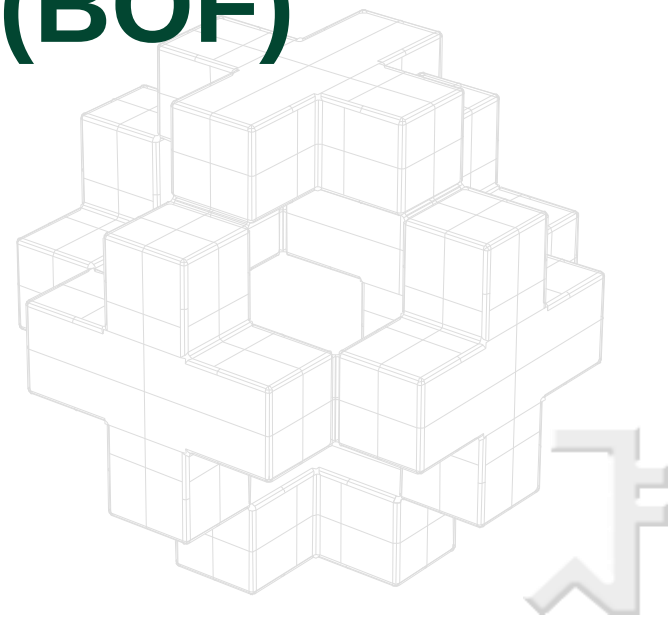




Image Retrieval (BOF)

Lecturer: Sang Hwa Lee

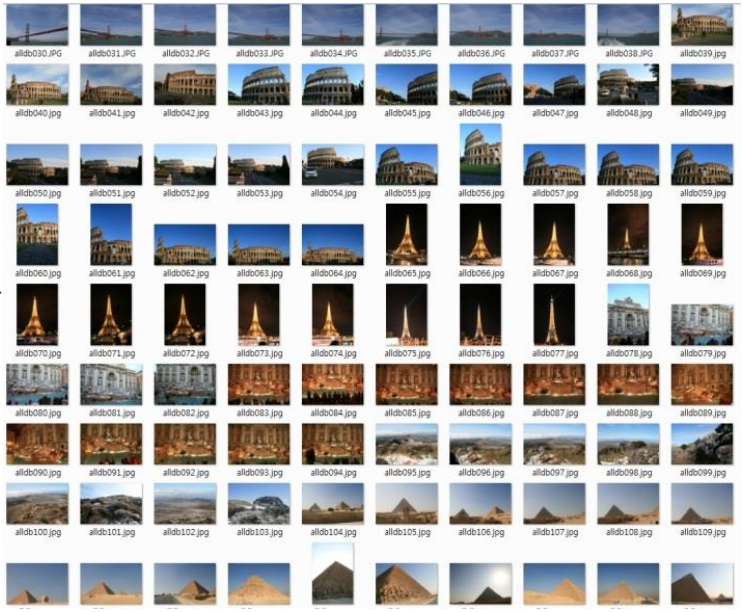


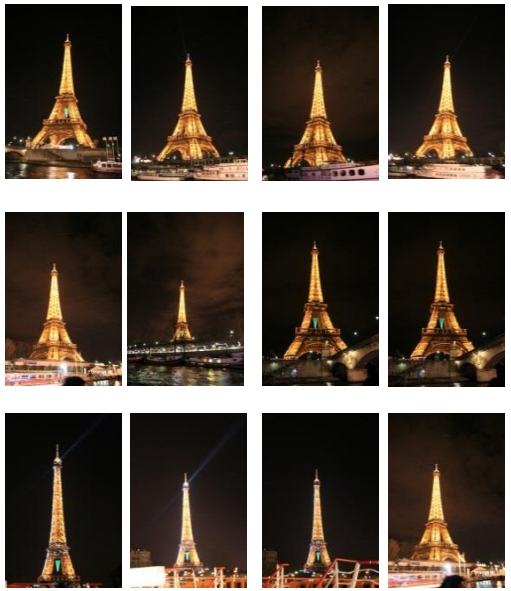
Introduction

❖ Image Retrieval?

Color features, BOF, Learning model

Database Search, Image matching





Query image
- User taken images
- original image

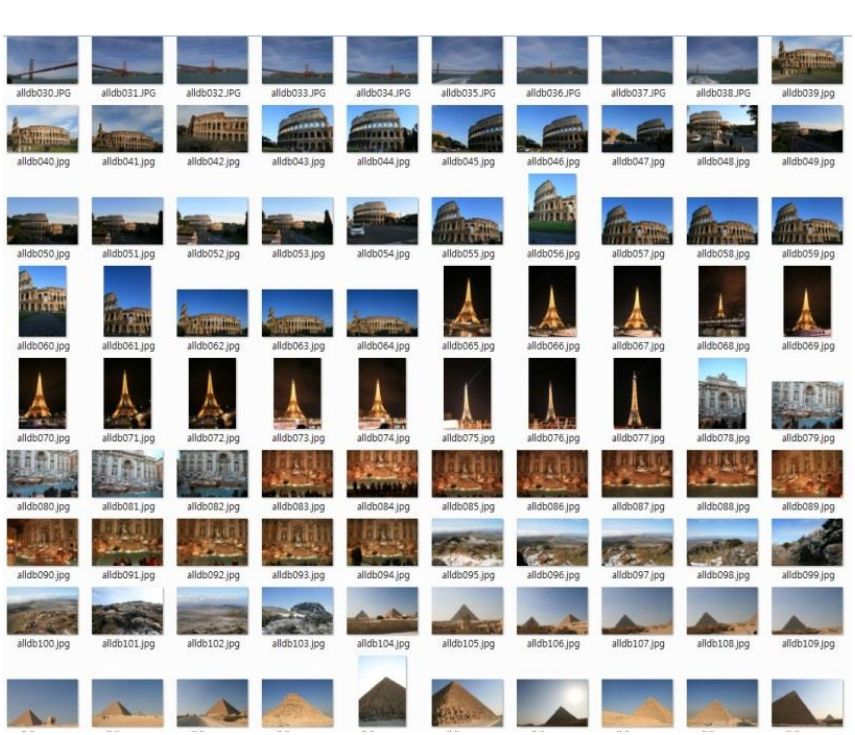
Database images
- Internet images
- Individual DB

Retrieval Results
- Same objects
- Similar scenes

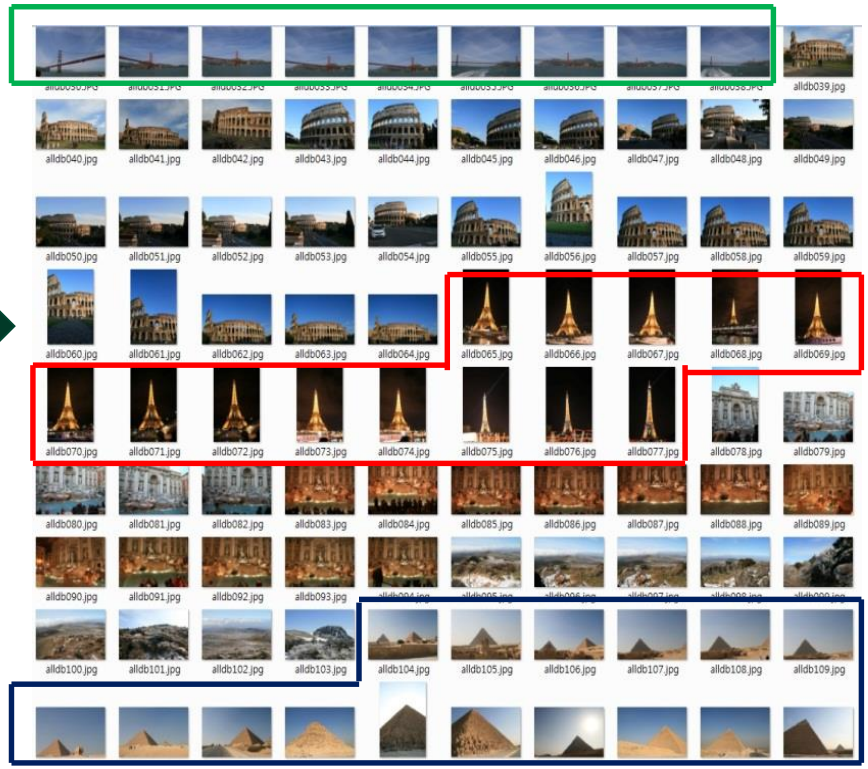
Introduction

❖ Image Categorization?

Content-based classifier models,
Non-supervised clustering



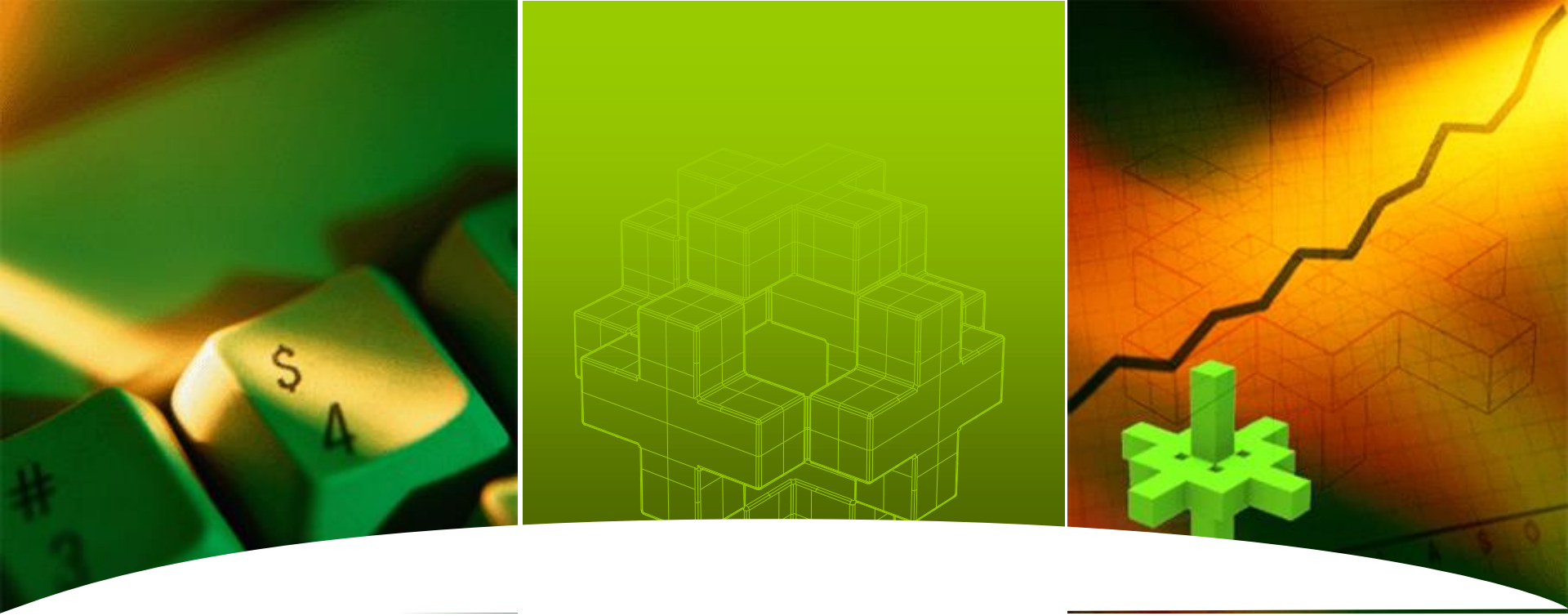
Database or Query



Classification Result

Introduction

- **Main approaches of image retrieval**
  - MPEG-7: color histogram
  - Video google: SIFT/SURF features
- **Image retrieval vs Text retrieval**
  - Get inspiration from text retrieval method
  - Video Google: A text retrieval approach to object matching in videos, ICCV, 2003.



Color Histogram Based Image Retrieval

Color based Image Retrieval

- **MPEG-7 image retrieval system**
  - Using various color histogram
  - Histogram distance measure
- **Reference**
  - MPEG-7 draft
  - Histogram-Based Color Image Retrieval, 2001, http://scien.stanford.edu/pages/labsite/2002/psych221/projects/02/sojeong/

Color Histogram Discrimination

1. Histogram Euclidean distance

$$d^2(h,g) = \sum_A \sum_B \sum_C (h(a,b,c) - g(a,b,c))^2$$

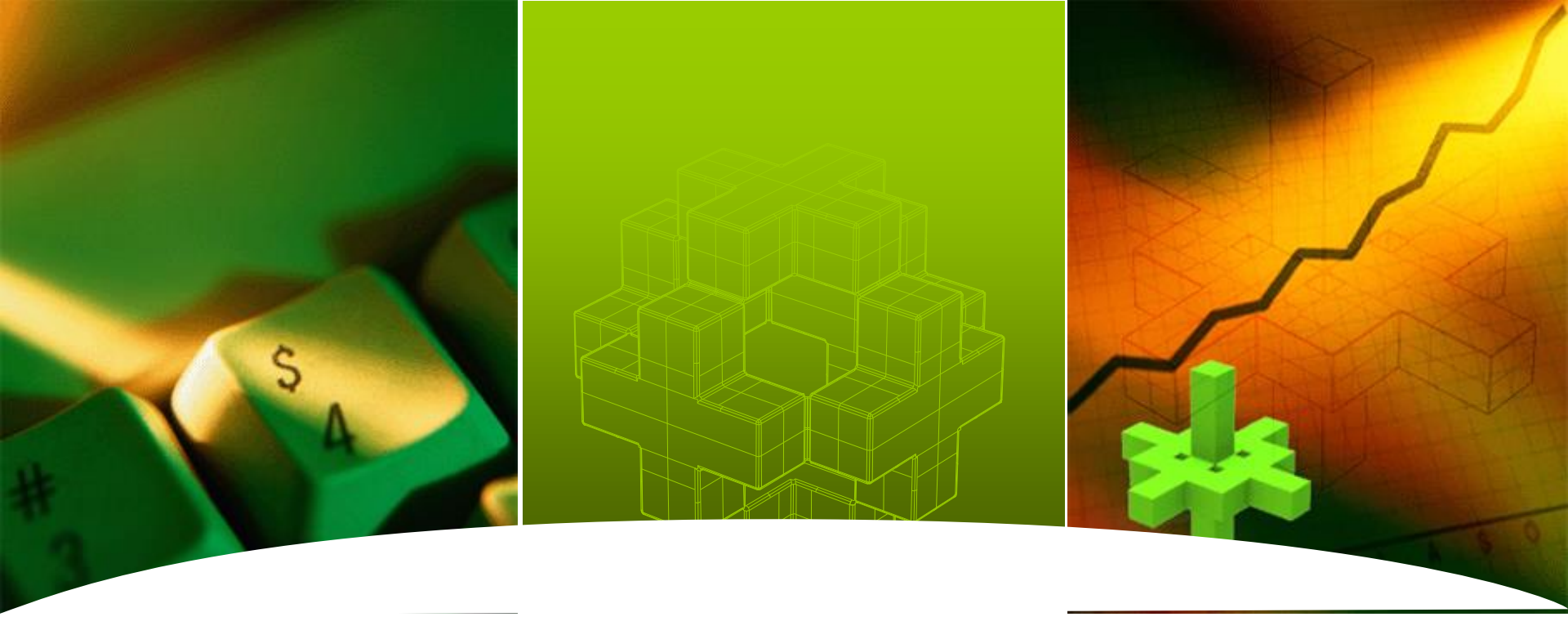
- h 와 g 는 두 이미지의 컬러 히스토그램
- Threshold 를 어떻게 정해줄 것인가?

Color Histogram Discrimination

2. Histogram Intersection Distance

$$d(h,g) = \frac{\sum_{A}\sum_{B}\sum_{C} \min(h(a,b,c), g(a,b,c))}{\min(|h|,|g|)}$$

- |h| 와 |g| 는 각 히스토그램의 크기 (sample 수)
- 이 값이 어느 값(Threshold) 이상이면 Similar 이미지라고 검색한다



Bag-of-Feature (BOF) Based Image Retrieval

Text Retrieval

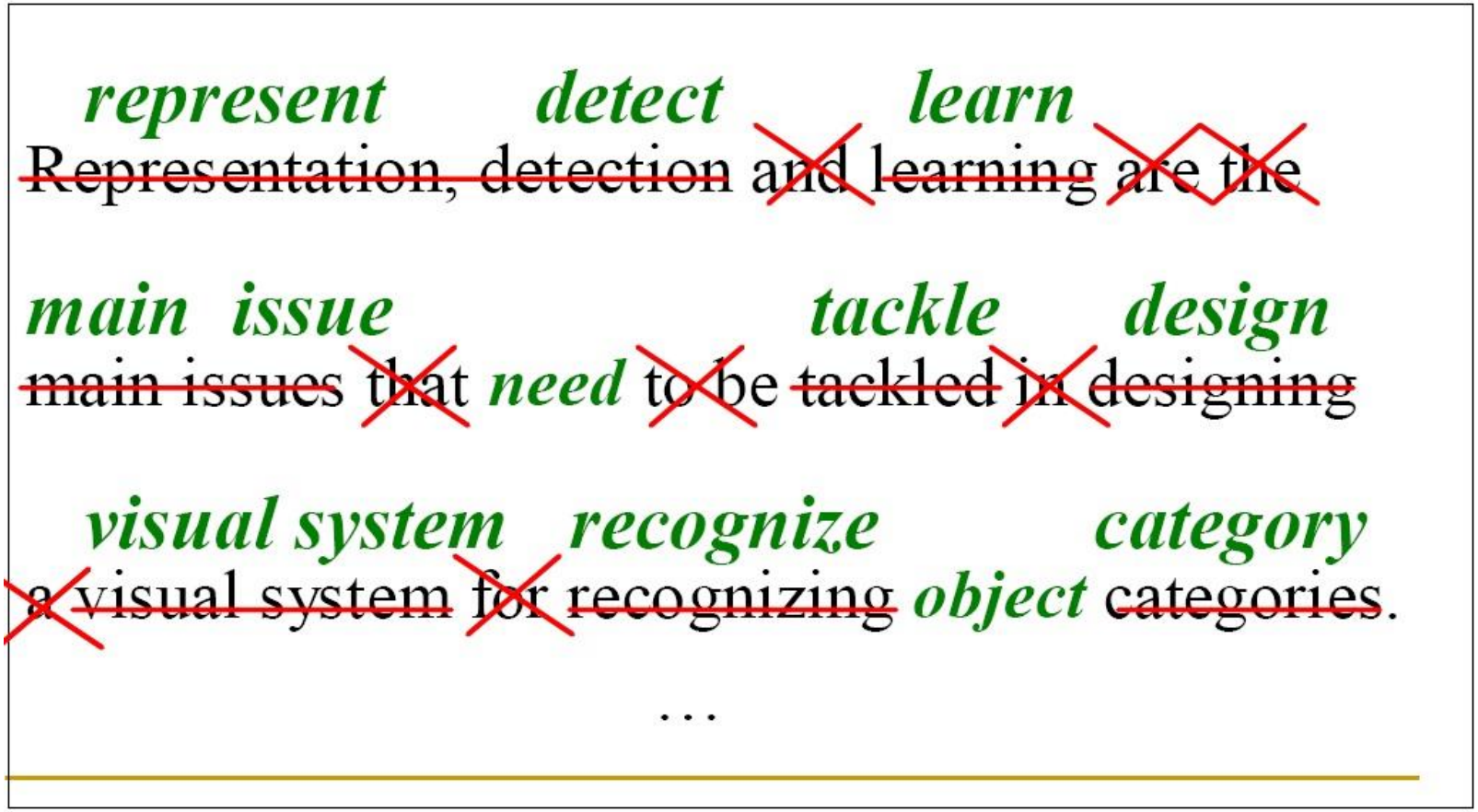
- **Text retrieval procedure**
- **BOW (bag-of-word)**
  - Document parsing → words
  - Word stems

    Ex:‘walk’, ‘walking’, ‘walks’ → ‘walk’
  - Common words: stop list

    Ex: the, a, be, …
  - Each word → unique identifier
  - Document → vector of words

Text Retrieval

- **Text retrieval procedure example**
  - Words stem, stop list



Text Retrieval

❖ Weighting

- Tf-idf weighting
  (term frequency–inverse document frequency)

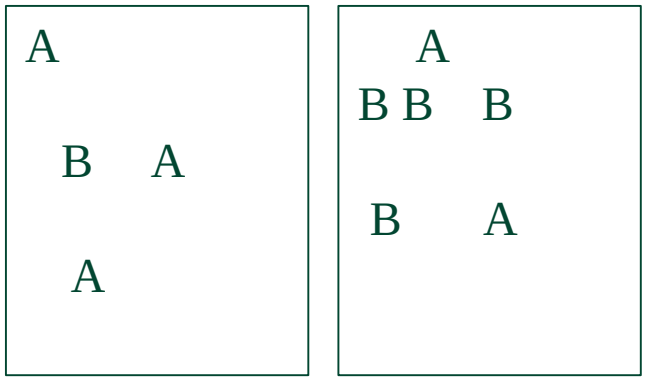
$$t_i = \frac{n_{id}}{n_d} \log \frac{N}{n_i}$$

$n_{id} = \#\ of\ ith\ words\ in\ document$

$n_d = \#\ of\ total\ words\ in\ document$

$n_i = \#\ of\ ith\ words\ in\ total\ documents$

$N = \#$ of total words in total documents



A word 중요 B word 중요

Text Retrieval

- **Inverted file**
  - For retrieval efficiency
  - Key values & position pointers

certified:	1, 4, 56, ...
document:	4, 6, 7, ...
fading:	1, 2, 43, ...
hallow:	2, 12, 34, ...
legal:	4, 5, 17, ...
Potter:	4, 42, 121, ...
reach:	1, 3, 4, ...
sniff:	2, 5, 59, ...
software:	1, 12, 13, ...
witch:	2, 3, 90, ...
wizard:	3, 102, 105, ...
...	

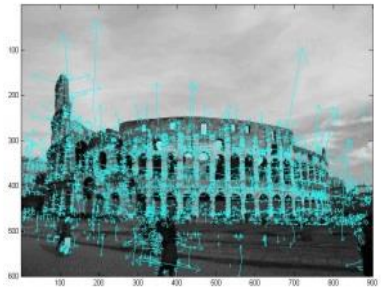
Application to Image Retrieval

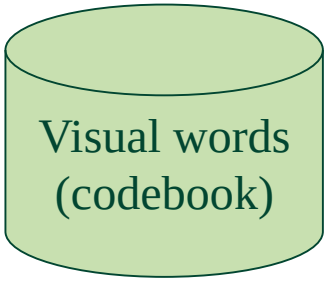
- **BOF (Bag-of-Feature)**
- **Document = image**
- **Words = features & descriptions**
- **Stem = visual words**
  - similar descriptor vectors
  - Codebook (VQ, Mean shift…)
- **Image → vector**
  - Histogram of visual words

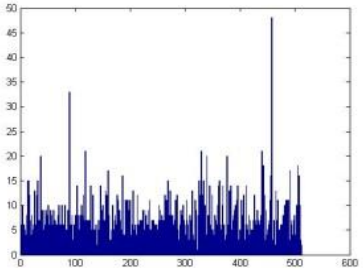
Application to Image Retrieval

❖Bag-Of-Feature 기반 영상 검색

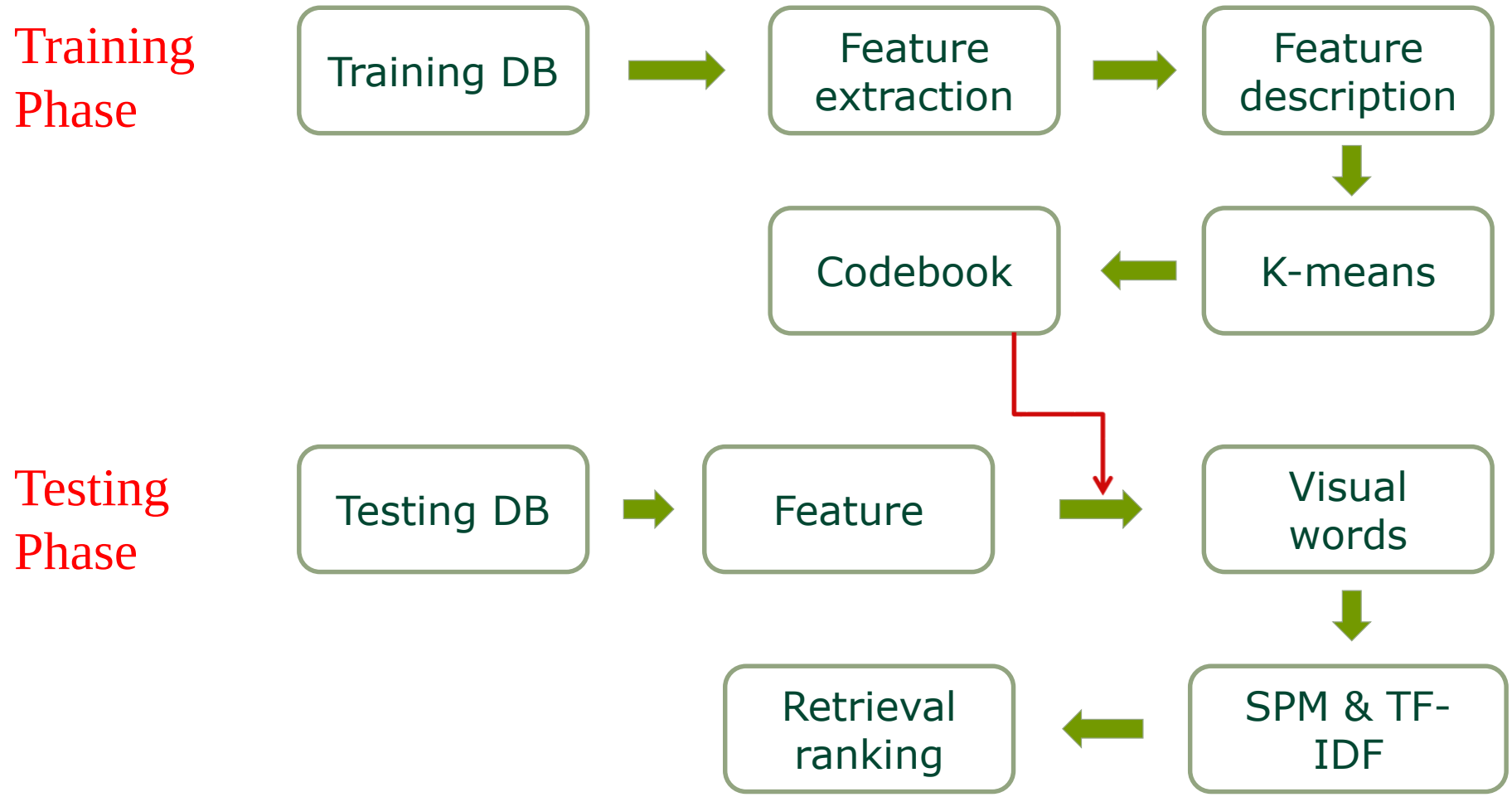
Input Image → Feature extraction → BOF modeling → Histogram matching







Application to Image Retrieval



Visual Words Design

- **Classification of features**
  - Finding a small number of representative feature vectors: 500~5000
  - Vector quantization (coding)

- **Classification Methods**
  - K-means clustering:
    - Vector-distance based
  - Mean shift clustering
    - Distribution based

Detection of Features

- **Local features**
  - Harris affine detector
  - Hessian affine detector
  - MSER
  - DOG, LOG, Fast Hessian detector
  - Regular grid, random spot

- **Global features**
  - Color histogram
  - Spectral analysis

Description of features

- **Local descriptors**
  - SIFT descriptor
  - SURF descriptor
  - PCA-SIFT

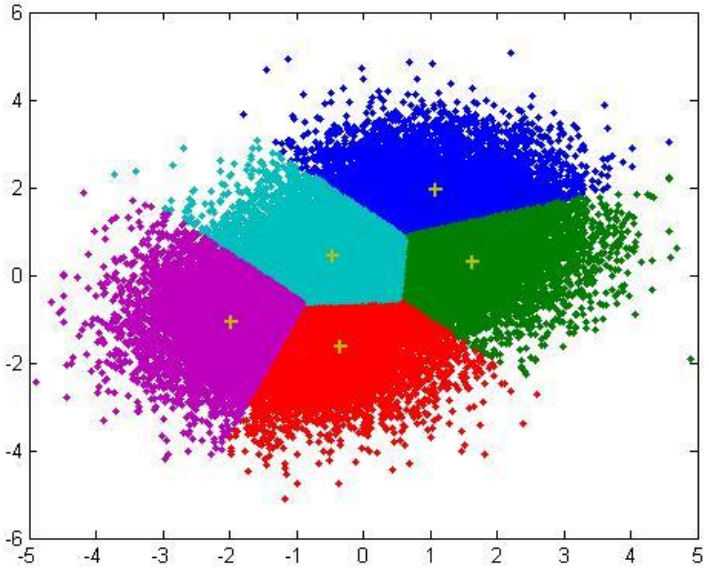
Generation of Visual Words

- **Visual vocabulary classification**
  - K-means clustering
  - Mean-shift clustering
  - Gaussian Mixture Models(GMM)

- **Improvement**
  - Tf-idf weighting, Stop list
  - Binary-BOF, mini-BOF
  - Spatial Pyramid Matching

Generation of Visual Words

- **K-means clustering**
  - Equivalent-distance line (plane) boundary between two centroids.
  - K is a main parameter.
  - Video Google: A text retrieval approach to object matching in videos, ICCV, 2003.



Generation of Visual Words

- **Mean-shift clustering**
  - Moving local means to a convergent point
  - Non-linear boundary according to the convergent points.
  - Local area (radius) is a main parameter.
  - Creating efficient codebooks for visual recognition, ICCV, 2005.

Generation of Visual Words

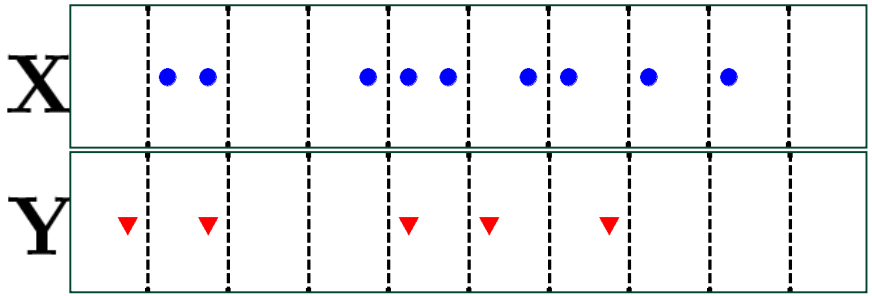
- **GMM representation**
  - Generative approach that estimates the probability distribution.
  - Unknown the number of models(means)
  - Some assumptions
    - Independence samples
    - Each model is uncorrelated

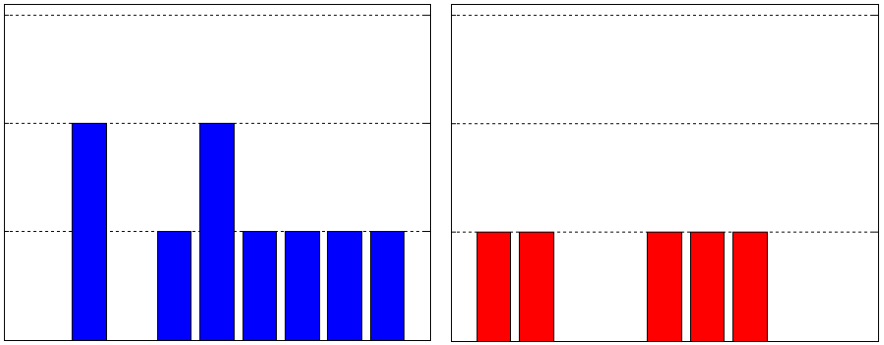
Comparing distances

- **Distance measure**
  - **Euclidean distance**
  - Manhattan distance
  - Earth Mover's Distance(EMD)
  - Bhattacharyya distance
  - Chi-square test
  - Mahalanobis distance
  - **Histogram intersection**
  - Hamming distance

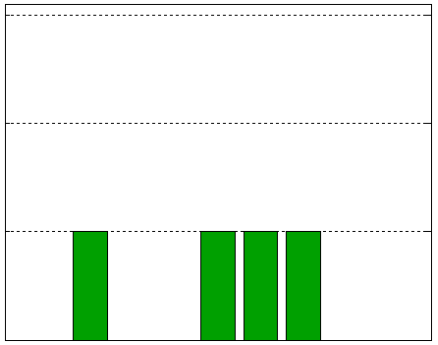
Histogram Intersection

$$\mathcal{I}\left(H(\mathbf{X}), H(\mathbf{Y})\right) = \sum_{j=1}^{r} \min\left(H(\mathbf{X})_j, H(\mathbf{Y})_j\right)$$





Histogram of X and Y



$$\mathcal{I}\left(H(\mathbf{X}), H(\mathbf{Y})\right) = 4$$

Experiment – TF-IDF

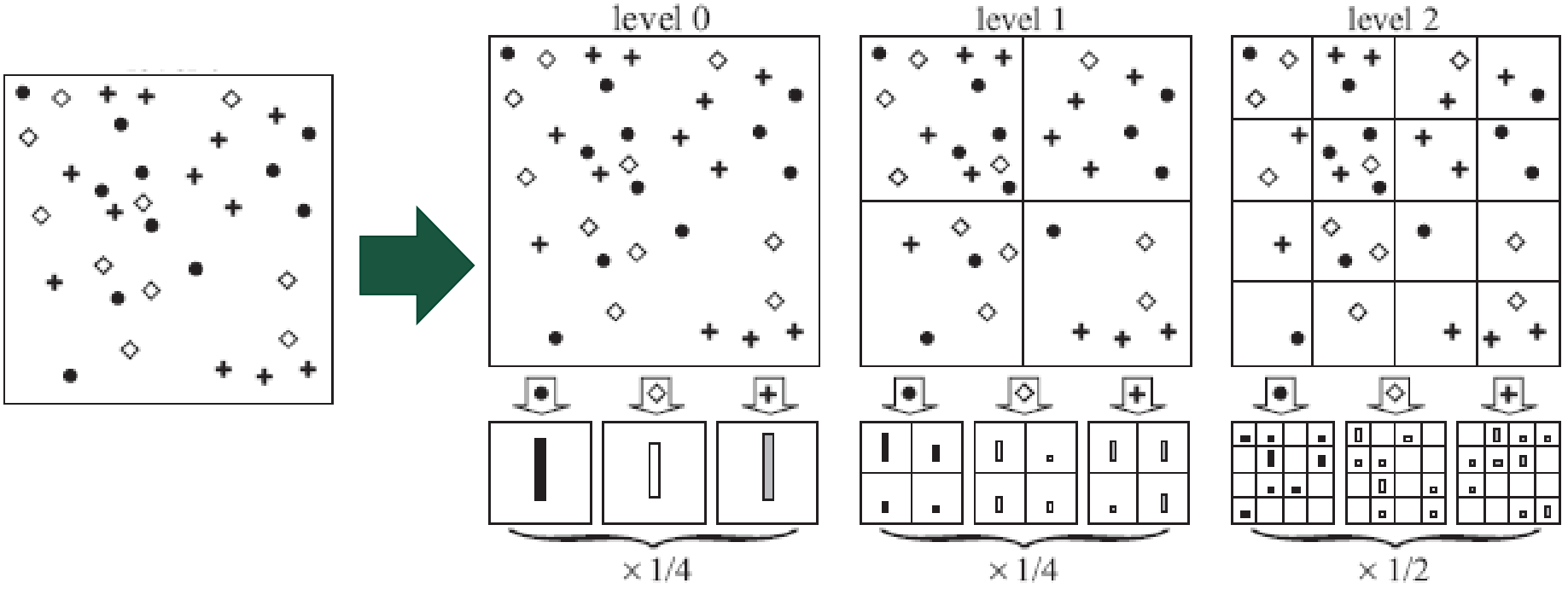
- **Experiment parameters**
  - SIFT detector(DOG)
  - SIFT descriptor
  - K-means clustering (k=512)
  - Tf-idf weighting
  - Euclidean distance

Experiment - SPM

- **Database**
  - Training images: 900x600, 318장
  - Test images: 68
- **BOF 모델링**
  - SIFT 특징 추출 기술 이용: 40만개 특징점
  - K-means clustering: distance 기반 분할
    - 512개 visual words
- **Spatial pyramid matching**
  - 0,1 2 levels
  - Weighted Histogram intersection

Spatial Pyramid Matching (1)

- **Spatial histograms & Intersection**

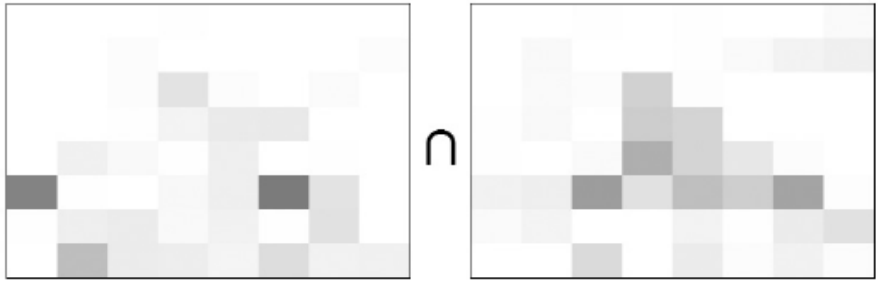


Spatial Pyramid Matching (2)

❖ Spatial histograms & Intersection

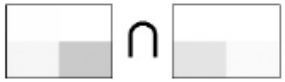
Feature histograms:

Level 3



Level 2

Level 1



Level 0

Total weight (value of *pyramid match kernel*): $\mathcal{I}_3 + \frac{1}{2}(\mathcal{I}_2 - \mathcal{I}_3) + \frac{1}{4}(\mathcal{I}_1 - \mathcal{I}_2) + \frac{1}{8}(\mathcal{I}_0 - \mathcal{I}_1)$

dbvo010.jpg

dbvo078.jpg

dbvo271.jpg

dbvo005.jpg

dbvo006.jpg

dbvo294.jpg

dbvo262.jpg

dbvo056.jpg

dbvo263.jpg

dbvo008.jpg

dbvo010.jpg

dbvo008.jpg

dbvo005.jpg

dbvo203.JPG

dbvo271.jpg

dbvo252.jpg

dbvo003.JPG

dbvo009.jpg

dbvo006.jpg

dbvo294.jpg

SPM

TF-IDF

dbvo013.JPG	dbvo016.JPG
	
dbvo019.JPG	dbvo012.JPG
	
dbvo011.JPG	dbvo020.JPG
	
dbvo015.JPG	dbvo302.jpg
	
dbvo301.jpg	dbvo017.JPG
	

SPM

dbvo013.JPG	dbvo012.JPG
	
dbvo020.JPG	dbvo016.JPG
	
dbvo019.JPG	dbvo302.jpg
	
dbvo015.JPG	dbvo301.jpg
	
dbvo018.JPG	dbvo278.jpg
	

TF-IDF

dbvo055.jpg


dbvo059.jpg


dbvo058.jpg


dbvo057.jpg


dbvo061.jpg


dbvo040.jpg


dbvo078.jpg


dbvo044.jpg


dbvo060.jpg


dbvo230.jpg


SPM

dbvo055.jpg


dbvo059.jpg


dbvo058.jpg


dbvo061.jpg


dbvo057.jpg


dbvo060.jpg


dbvo044.jpg


dbvo056.jpg


dbvo039.jpg


dbvo043.jpg


TF-IDF

dbvo087.jpg

dbvo088.jpg

dbvo083.jpg

dbvo091.jpg

dbvo084.jpg

dbvo094.jpg

dbvo086.jpg

dbvo092.jpg

dbvo085.jpg

dbvo248.jpg

SPM

dbvo088.jpg

dbvo087.jpg

dbvo083.jpg

dbvo086.jpg

dbvo084.jpg

dbvo085.jpg

dbvo091.jpg

dbvo092.jpg

dbvo094.jpg

dbvo093.jpg

TF-IDF

dbvo122.jpg

dbvo121.jpg

dbvo123.jpg

dbvo138.jpg

dbvo113.jpg

dbvo115.jpg

dbvo140.jpg

dbvo118.jpg

dbvo111.jpg

dbvo096.jpg

SPM

dbvo122.jpg

dbvo123.jpg

dbvo121.jpg

dbvo113.jpg

dbvo138.jpg

dbvo115.jpg

dbvo111.jpg

dbvo137.jpg

dbvo118.jpg

dbvo124.jpg

TF-IDF