Image Search and Classification

Sung-Eui Yoon
(윤성의)

Course URL:
http://sglab.kaist.ac.kr/~sungeui

KAIST

Acknowledgements

- **Collaborators**
  - **My students, YuWing Tai, Pierre-Yves Laffont, Shih-Fu Chang, Junfeng He, Zhe Lin**
- **Funding sources**
  - **Korea Research Foundation**
  - **Ministry of Knowledge Economy**
  - **Samsung**
  - **Microsoft Research Asia**
  - **Adobe**

About the Instructor

- Joined KAIST at 2007

- Main research focus
  - Handling of massive geometric data for various computer graphics and geometric problems

- Research for the topic
  - Studied on nearest neighbor search about 10 years
  - Moved to image search around 5 years ago

KAIST

Main Research Focus

- Handle massive data for various computer graphics and geometric problems
- Paper and video
  - http://sglab.kaist.ac.kr/papers.htm
- YouTube videos
  - http://www.youtube.com/user/sglabkaist

Web-Scale Visual Data and Novel Applications

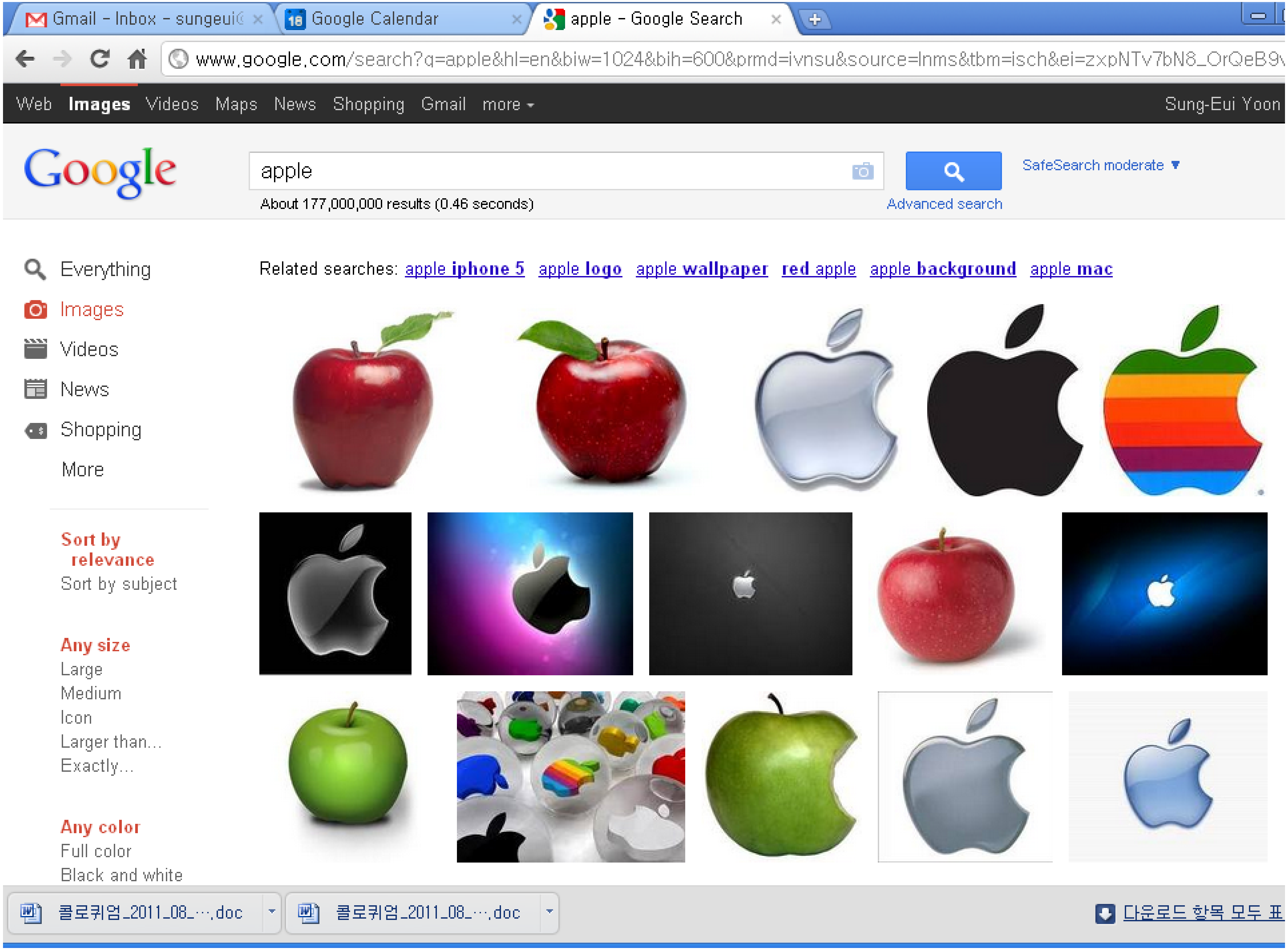
- Visual data are widely used for various communication and, and are more widely consumed at Web and mobile devices
  - YouTube, Facebook, Flickr, etc.
- Processing them requires scalable algorithms
- Web-scale visual data can enable new applications
  - Photo tourism,
  - Scene completion, etc.

About the Course

- We will focus on the following thing:
  - Broad understanding on image retrieval techniques and classification

Content-Based Image Retrieval (CBIR)

- Identify similar images given a user-specified image or other types of inputs


Gmail - Inbox - sungeui
Google Calendar
apple - Google Search
www.google.com/search?q=apple&hl=en&biw=1024&bih=600&prmd=ivnsu&source=lnms&tbm=isch&ei=zxpNTv7bN8_OrQeB9
Web
Images
Videos
Maps
News
Shopping
Gmail
more
Sung-Eui Yoon
Google
apple
About 177,000,000 results (0.46 seconds)
Advanced search
SafeSearch moderate
Everything
Images
Videos
News
Shopping
More
Sort by relevance
Sort by subject
Any size
Large
Medium
Icon
Larger than...
Exactly...
Any color
Full color
Black and white
Related searches: apple iphone 5 apple logo apple wallpaper red apple apple background apple mac
콜로퀴엄_2011_08_....doc
콜로퀴엄_2011_08_....doc
다운로드 항목 모두 표

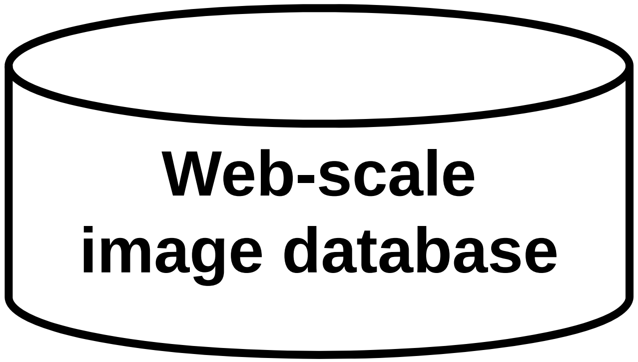

Gmail - Inbox - sungeui
Google Calendar
Google Search
www.google.com/search?tbs=sbi:AMhZZiuyKQN8VbdVKrIWLNXXzF8NwQ3U6nZjHoGicvPiP_1LltJ1Ts9EF9aMlIlZ52wdbI9cGW
Web Images Videos Maps News Shopping Gmail more
Sung-Eui Yoon
Google
sungeui.jpg
describe image here
About 4 results (0.29 seconds)
Advanced search
Everything
Images
Videos
News
Shopping
More
Image size:
200 × 272
Find other sizes of this image:
All sizes - Small
Pages that include matching images
Sungeui Yoon (성의,윤성의)
sglab.kaist.ac.kr/~sungeui/ - Cached
Sung-Eui Yoon (윤 성의) Assistant professor. Scalable Graphics/Geometric Algorithm Lab. Dept. of Computer Science · KAIST ...
200 × 272
صفحه این آدرس - 웹사이언스 공학 WebST ::::::::::::::::::
- [Translate this page]
webst.kaist.ac.kr/content.php?db=professor - Cached
이름Cha, Meeyoung (차미영) 조교수; 연구분야Social Computing, Data-Driven Social Science; 학위PhD, KAIST, 2008; 전화번호+82-42-350-2922; 이 메일meeyoungcha ...
120 × 140
2010.09.13 - KGC 2011 - [Translate this page]
www.kgconf.com/kor/html/conference_c_view.html?cate3... - Cached
Kristian Segerstrale Playfish, 소셜게임의 미래 현재 소셜게임의 현주소와 빠르게 성장하는 소셜게임의 미래를 예리한 견식으로 소개 ...
100 × 100
sungeui.jpg
콜로퀴엄_2011_08_...,doc
콜로퀴엄_2011_08_...,doc
다운로드 항목 모두 표

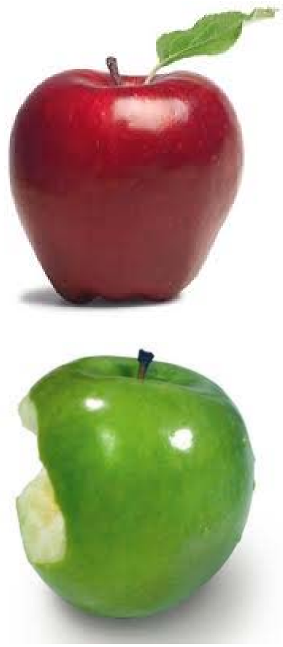
Content-Based Image Retrieval (CBIR)

- **Identify similar images given a user-specified image or other types of inputs**

Extract image descriptors (e.g., SIFT)

Input





Output

Applications

- Search
- Image stitching
- Object/scene/location recognitions
- Robot motion planning
- Copyright detection

Panorama Stitching

(a) Matier data set (7 images)

iPhone version available

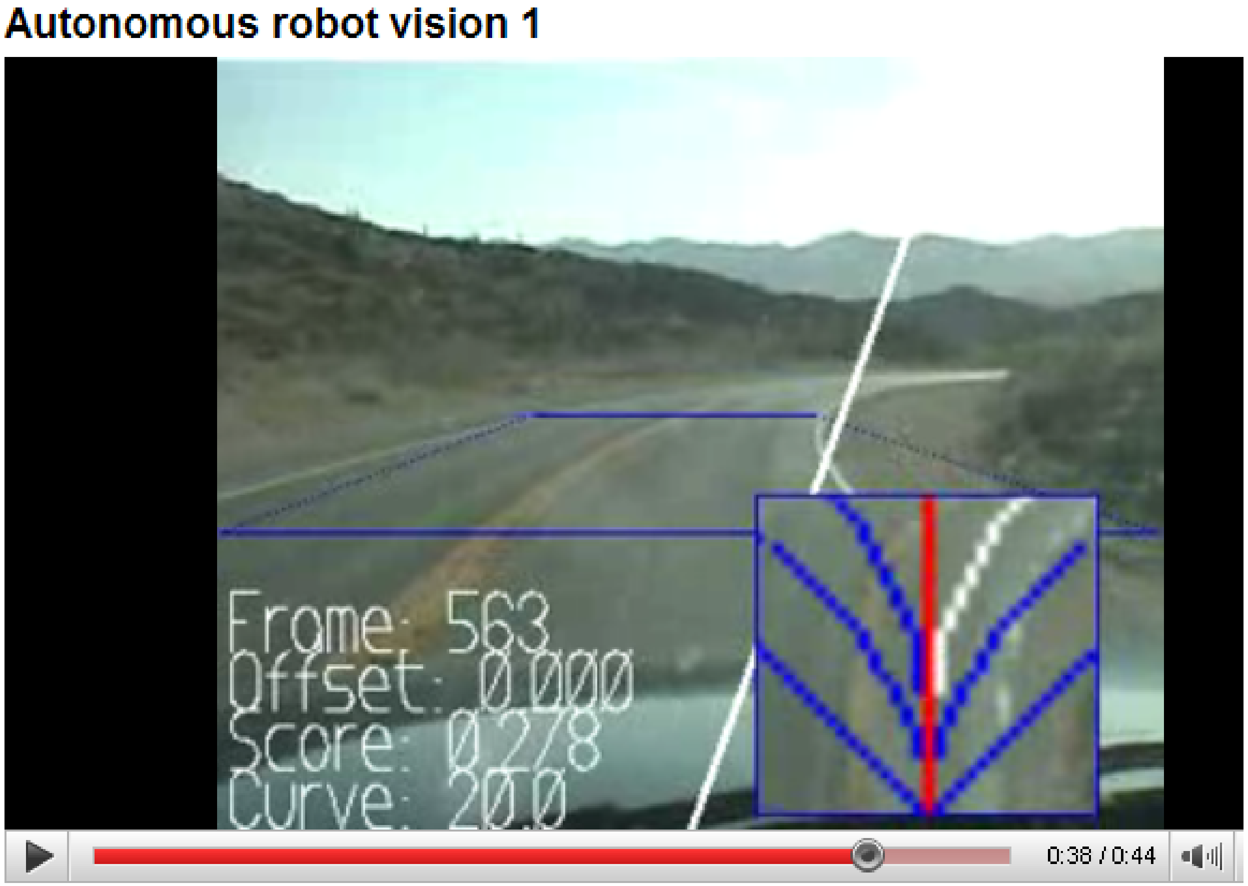
(b) Matier final stitch

[Brown, Szeliski, and Winder, 2005]

http://www.cs.ubc.ca/~mbrown/autostitch/autostitch.html

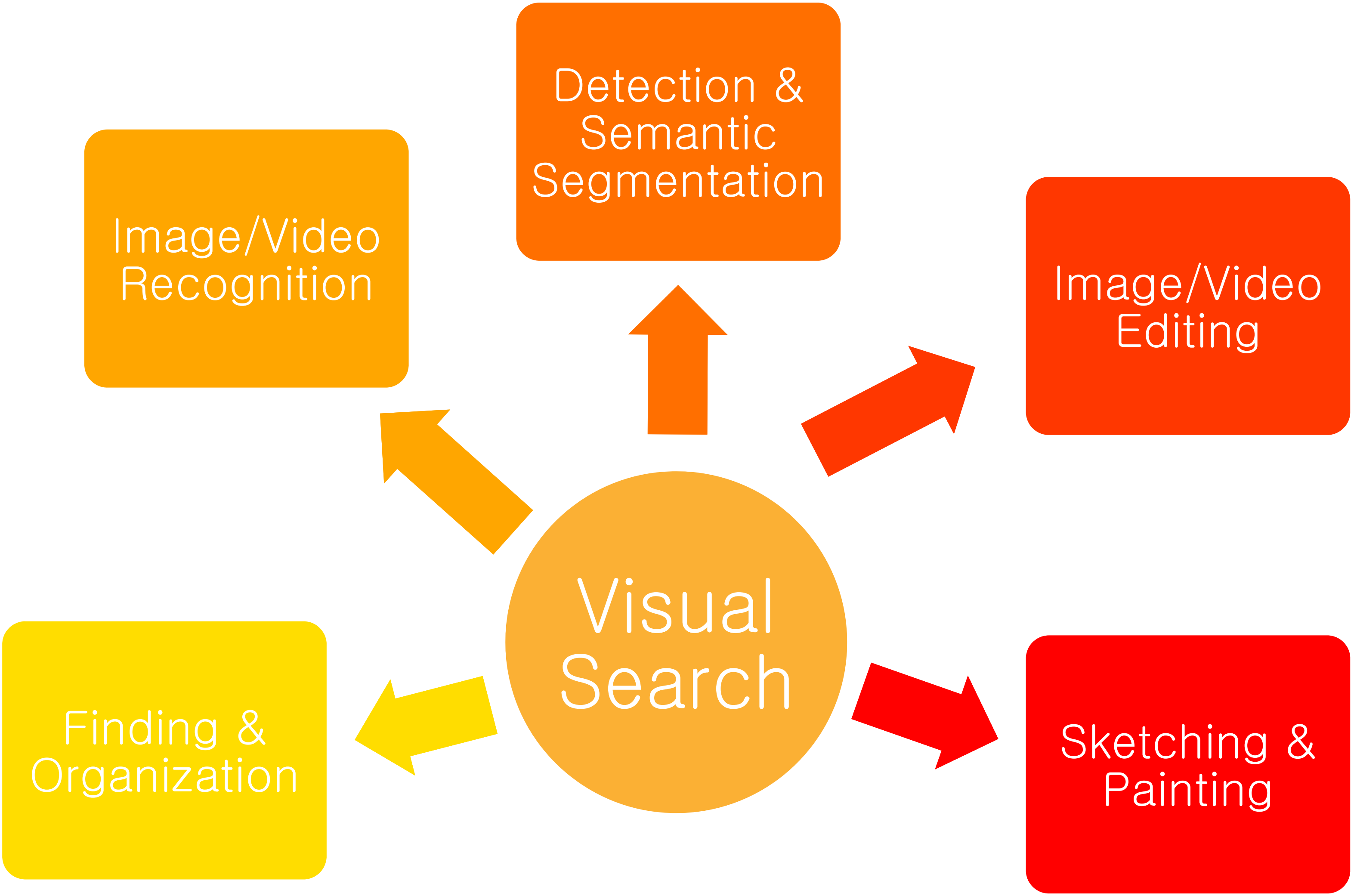
Object Detection

Robot Motion Planning



Autonomous robot
http://www.youtube.com/watch?v=3SQiow-X3ko

Possible Application Domains



Issues for Web-Scale Multimedia Search

- Too many multimedia data and frequent updates

- Accuracy?
- Performance?
- Novel applications?

Gmail - Inbox - sungeui
Google Calendar
Google Search
www.google.com/search?tbs=sbi:AMhZZiuyKQN8VbdVKrlWLNXXzF8NwQ3U6nZjHoGicvPiP_1LltJ1Ts9EF9aMliLZ52wdbl9cGWl
Images Videos Maps News Shopping Gmail more
Sung-Eui Yoon
sungeui.jpg
describe image here
About 4 results (0.29 seconds)
Advanced search
Everything
Images
Videos
News
Shopping
More
- [Translate this page]
www.kgconf.com/kor/html/conference_c_view.html?cate3... - Cached
Kristian Segerstrale Playfish, 소셜게임의 미래 현재 소셜게임의 현주소와 빠르게 성장하는 소셜게임의 미래를 예리한 견식으로 소개 ...
100 × 100
Welcome to ISAC2009!! - [Translate this page]
isac2009.or.kr/isac2009/speakers/domestic_bio.php - Cached
Yoo Mi Choi. 소속: 디자인여성학회 회장 한국디자인 학회 이사 한국애니메이션학회 부회장 인포디자인학회 이사 한국 애니메이션 필림협회 이사 ...
200 × 272
Visually similar images - Report images
Search Help
Give us feedback
sungeui.jpg
콜로퀴엄_2011_08_....doc
콜로퀴엄_2011_08_....doc
다운로드 항목 모두 표시...

It took a few seconds to get this result on my desktop computer.

Key Components

- Image representations
- Indexing algorithms
- Matching methods
- Classification, Localization, etc.
  - Apply image search (or nearest neighbor search)
  - Data-driven approach

Image Representations

- SIFT, GIST, etc.
- Invariant to different transformations

Image Retrieval

- At pre-processing, build an database for efficient retrieval at runtime

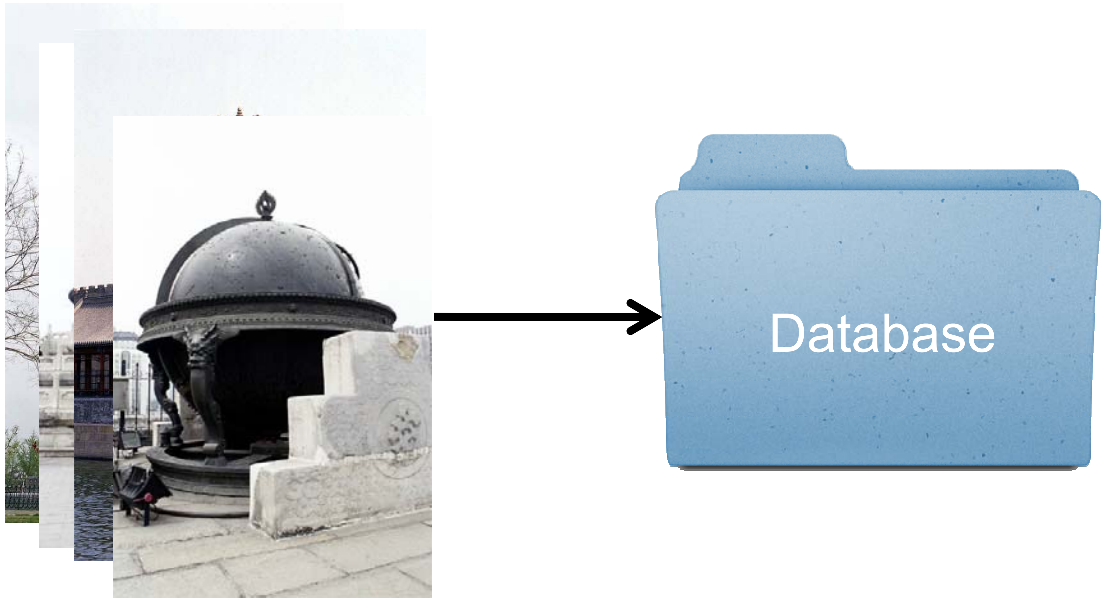
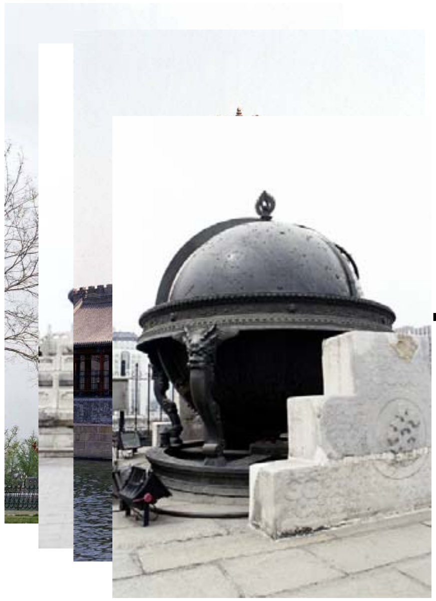


Image Retrieval

- **At pre-processing, build an database for efficient retrieval at runtime**



vocabulary trees, hashing, and inverted files

Image Retrieval: Runtime Procedure

Query image

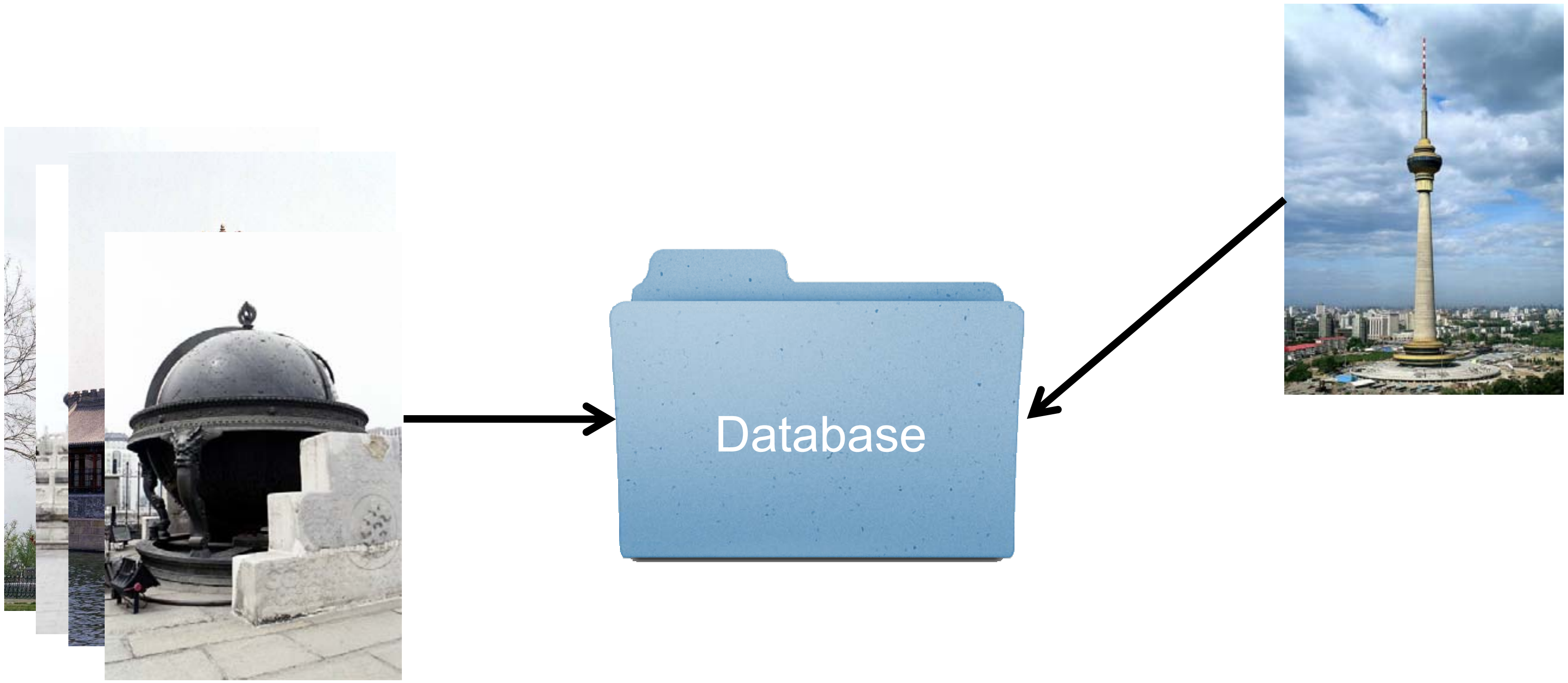


Image Retrieval: Runtime Procedure

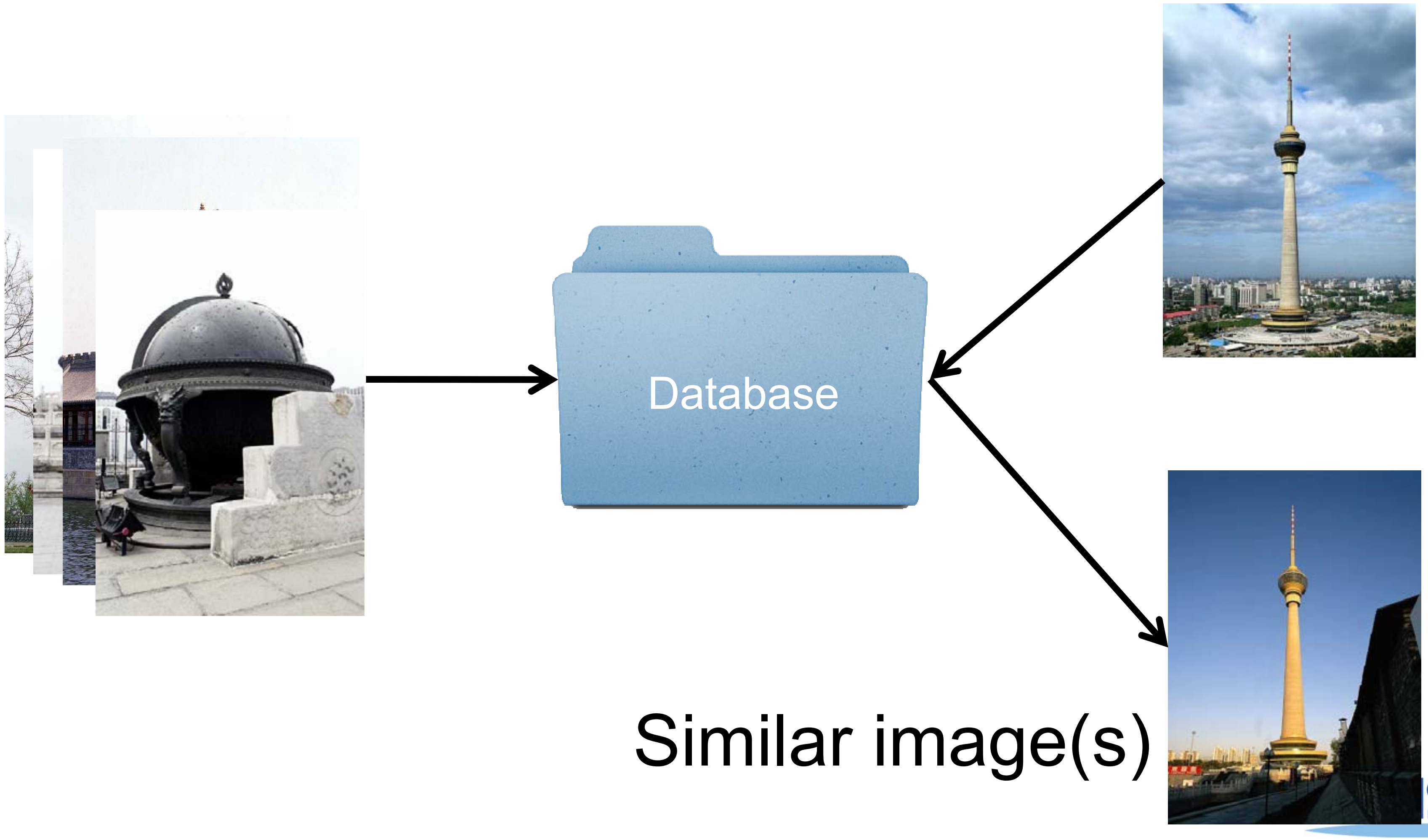
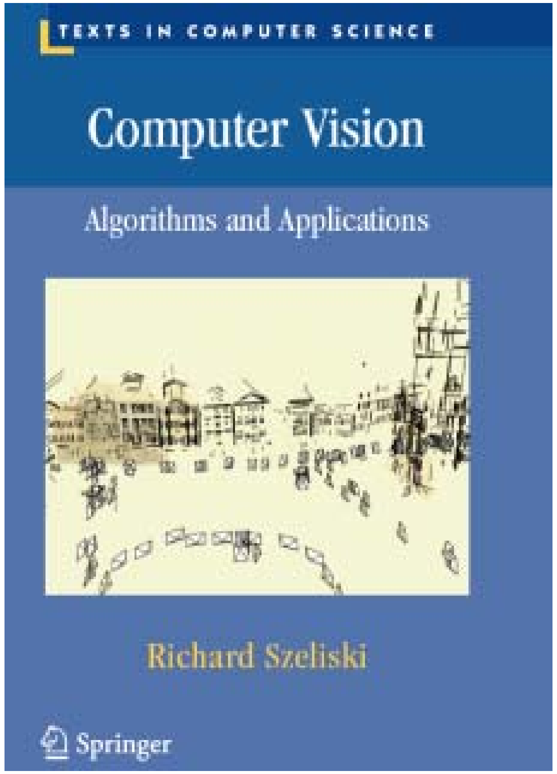


Image Retrieval with Spatially Constrained Similarity Measure

[Xiaohui Shen, Zhe Lin, Jon Brandt, Shai Avidan and Ying Wu, CVPR 2012]

Resource

- **Reference**
  - **Computer vision: algorithms and applications**
    - **Its file is available (http://szeliski.org/Book/)**



Other Resources

- **Technical papers**
  - **CVPR, ICCV, ECCV, ACM MM, SIGGRAPH, etc.**
  - **Computer vision resource (http://www.cvpapers.com/)**
- **Course homepages**
- **Google or Google scholar**

Bag-of-Words (BoW) Models

Sung-Eui Yoon
(윤성의)

Course URL:
http://sglab.kaist.ac.kr/~sungeui

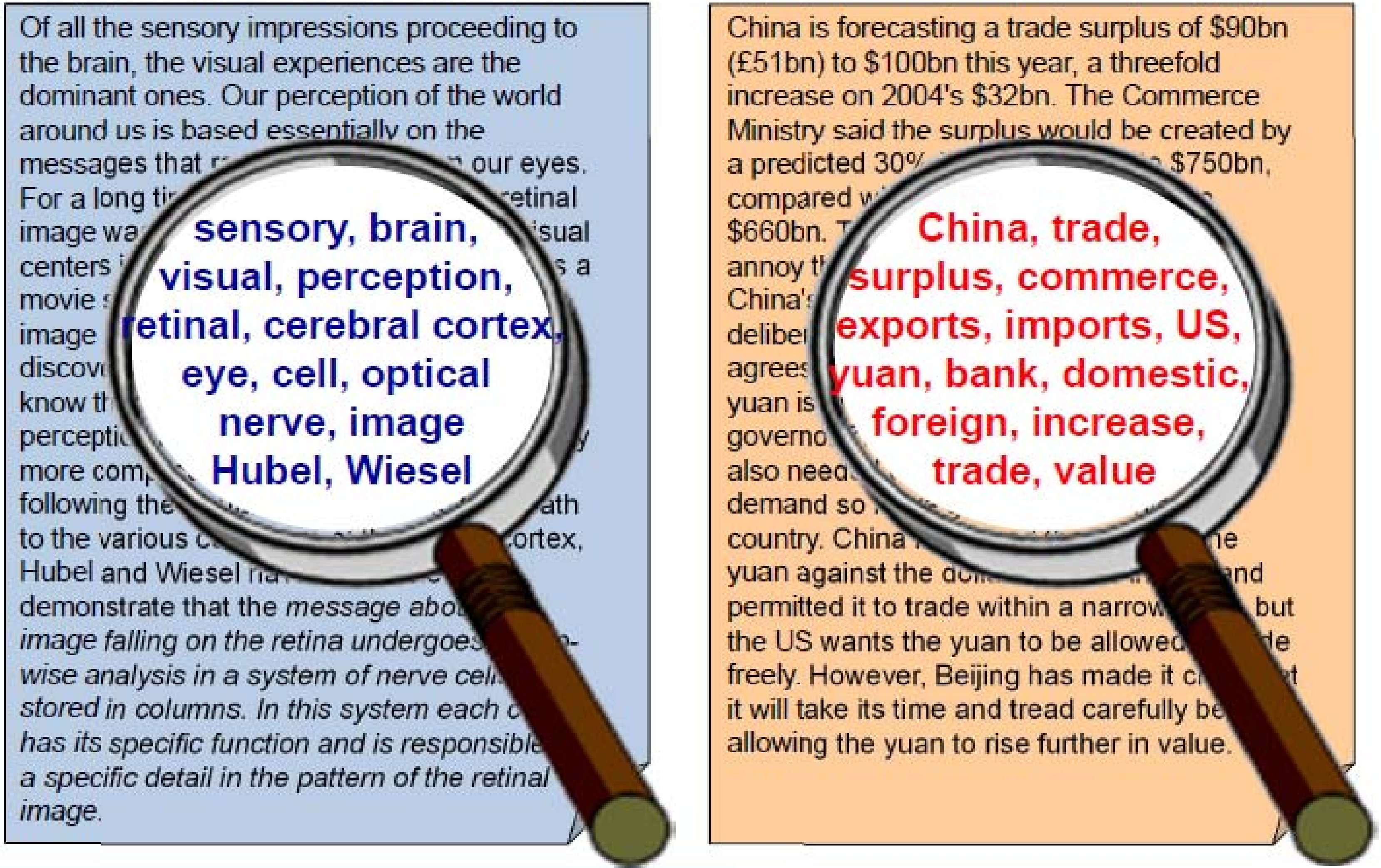
What we will learn today?

- Bag of Words models
  - Basic representation
  - Different learning and recognition algorithms

Object → Bag of 'words'

Fei-Fei Li

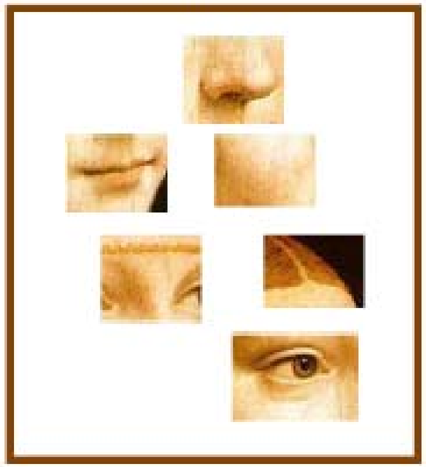
Analogy to documents



definition of "BoW"

– Independent features

face

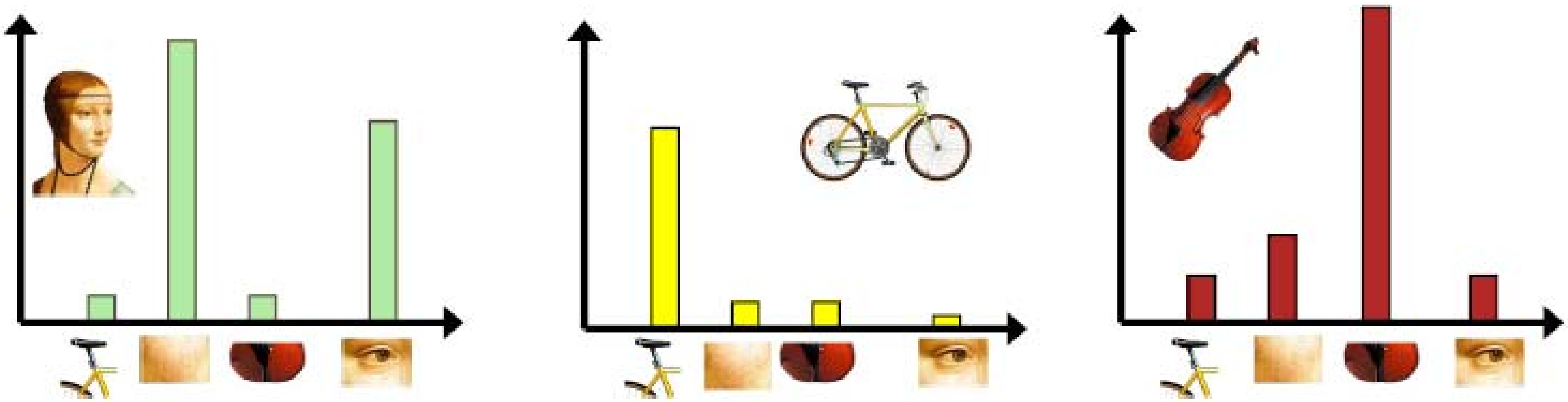


bike

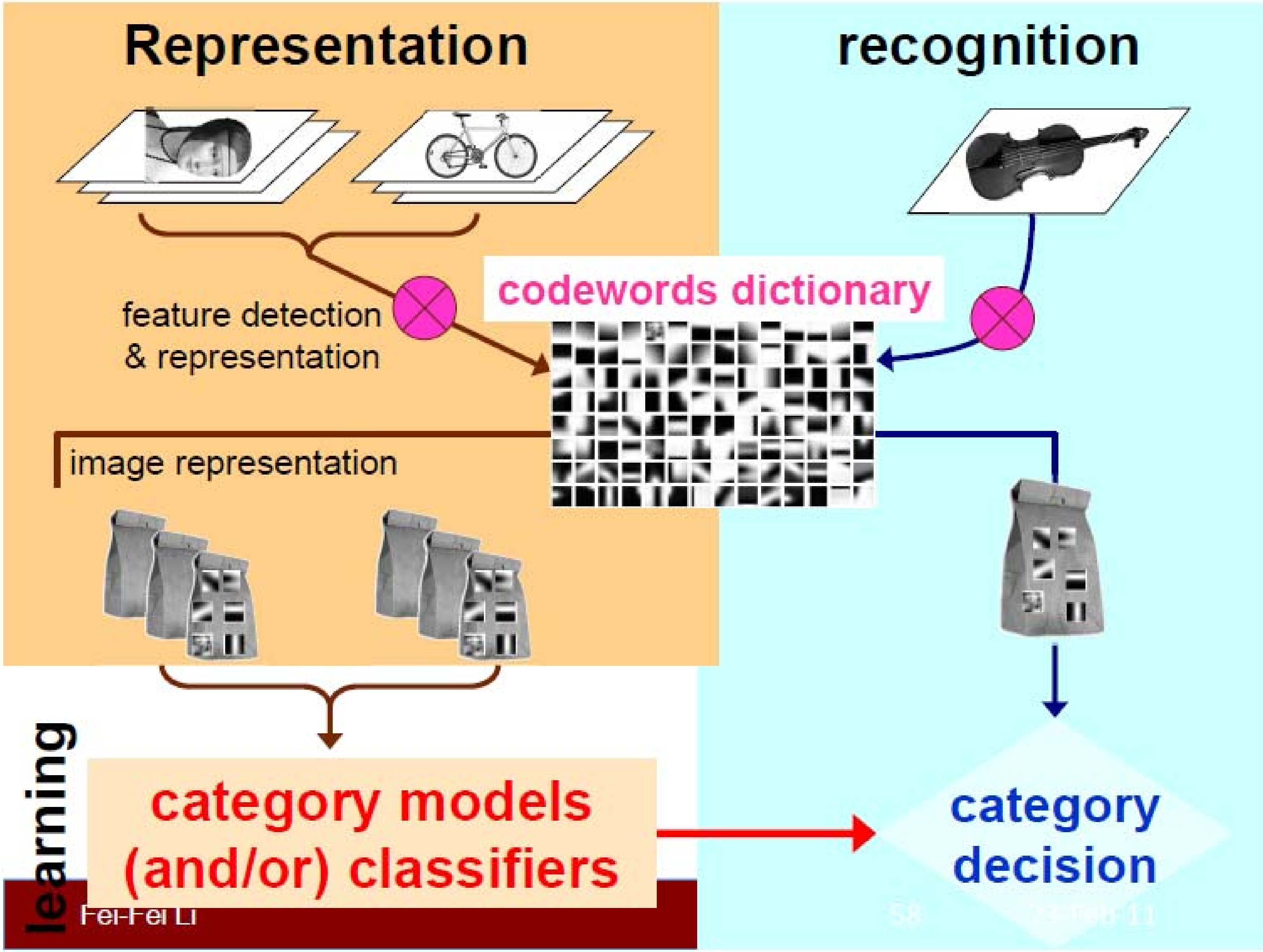
violin

definition of "BoW"

– Independent features
– histogram representation

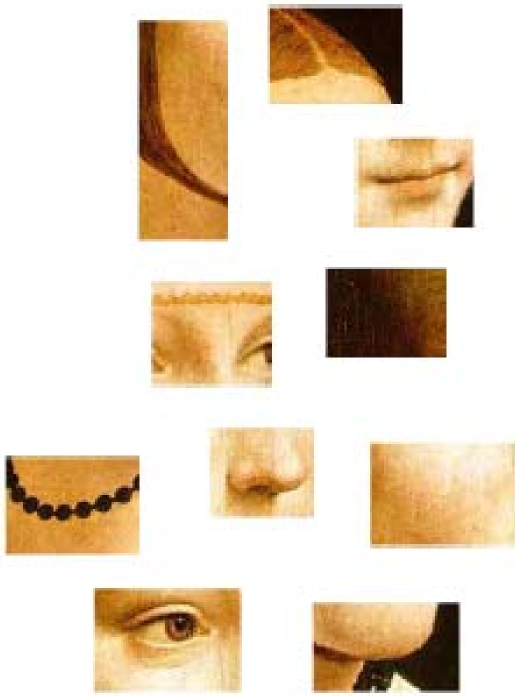


codewords dictionary



Representation
recognition
feature detection
& representation
codewords dictionary
image representation
learning
category models
(and/or) classifiers
category
decision
Fei-Fei Li
58

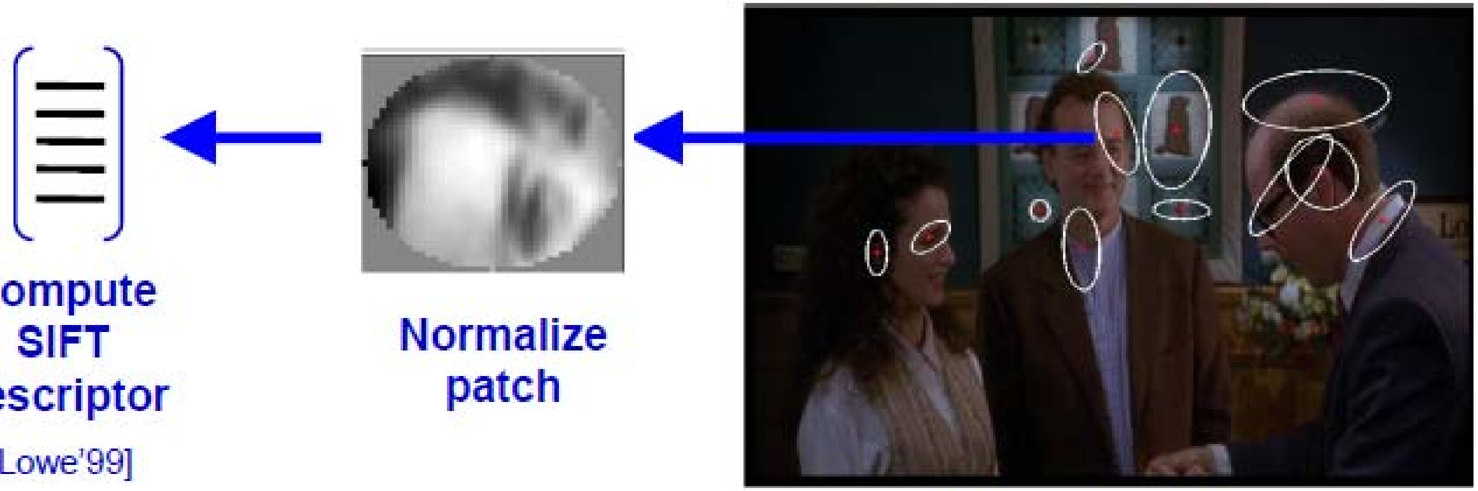
1.Feature detection and representation



1.Feature detection and representation

- Regular grid
  - Vogel & Schiele, 2003
  - Fei-Fei & Perona, 2005
- Interest point detector
  - Csurka, Bray, Dance & Fan, 2004
  - Fei-Fei & Perona, 2005
  - Sivic, Russell, Efros, Freeman & Zisserman, 2005
- Other methods
  - Random sampling (Vidal-Naquet & Ullman, 2002)
  - Segmentation based patches (Barnard, Duygulu, Forsyth, de Freitas, Blei, Jordan, 2003)

1.Feature detection and representation



Compute SIFT descriptor

[Lowe'99]

Normalize patch

Detect patches

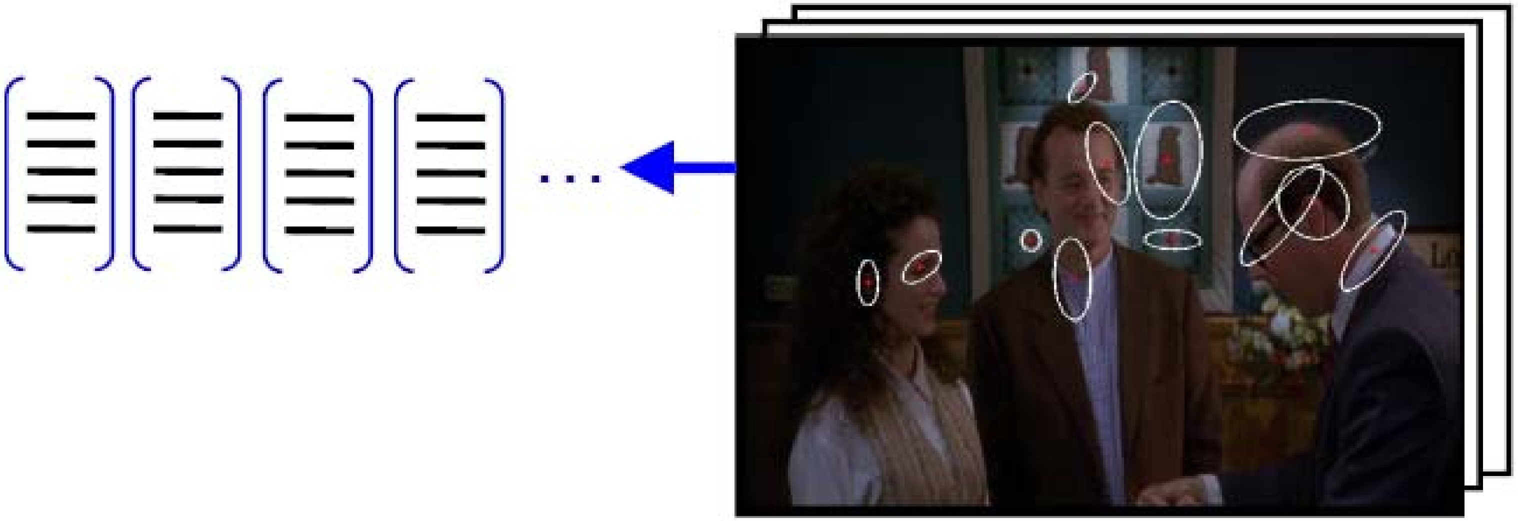
[Mikojaczyk and Schmid '02]

[Mata, Chum, Urban & Pajdla, '02]

[Sivic & Zisserman, '03]

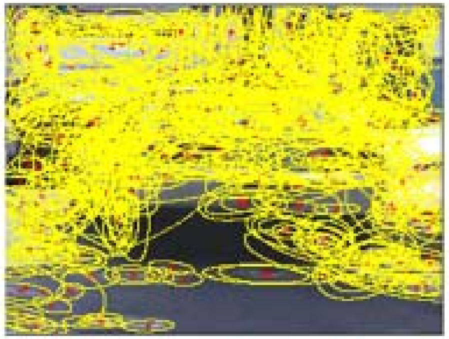
Slide credit: Josef Sivic

1.Feature detection and representation

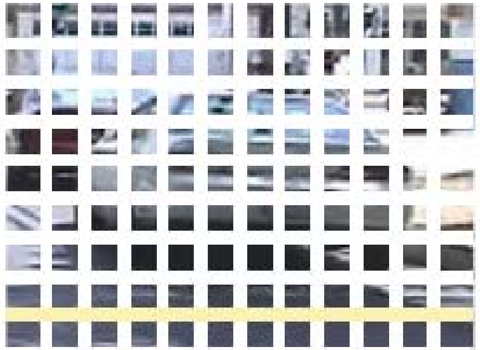


Representation

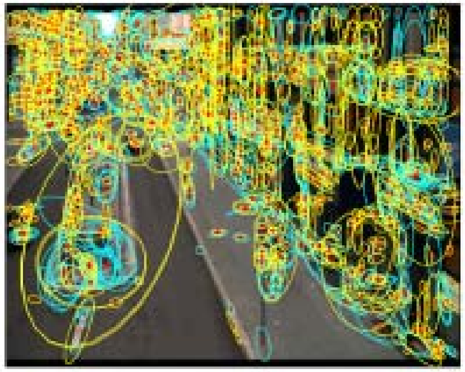
- Building blocks: Sampling strategies



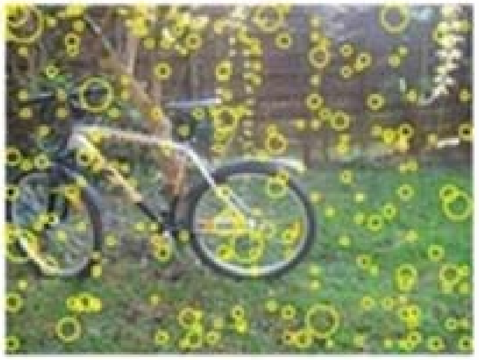
Interest operators



Dense, uniformly



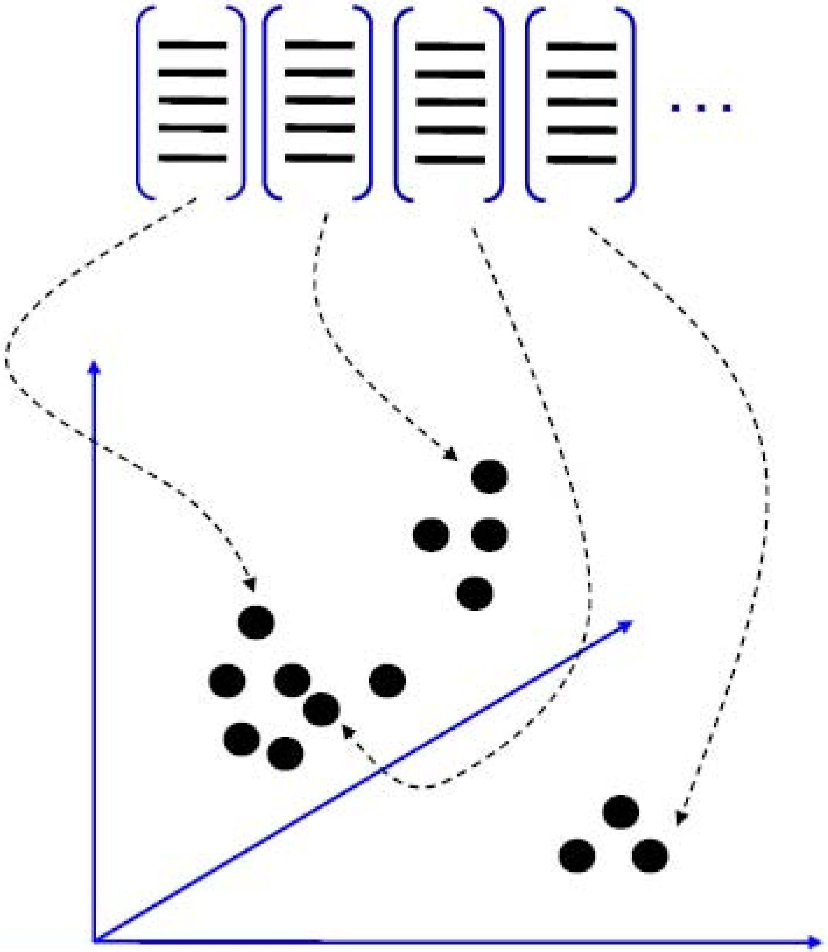
Multiple interest operators



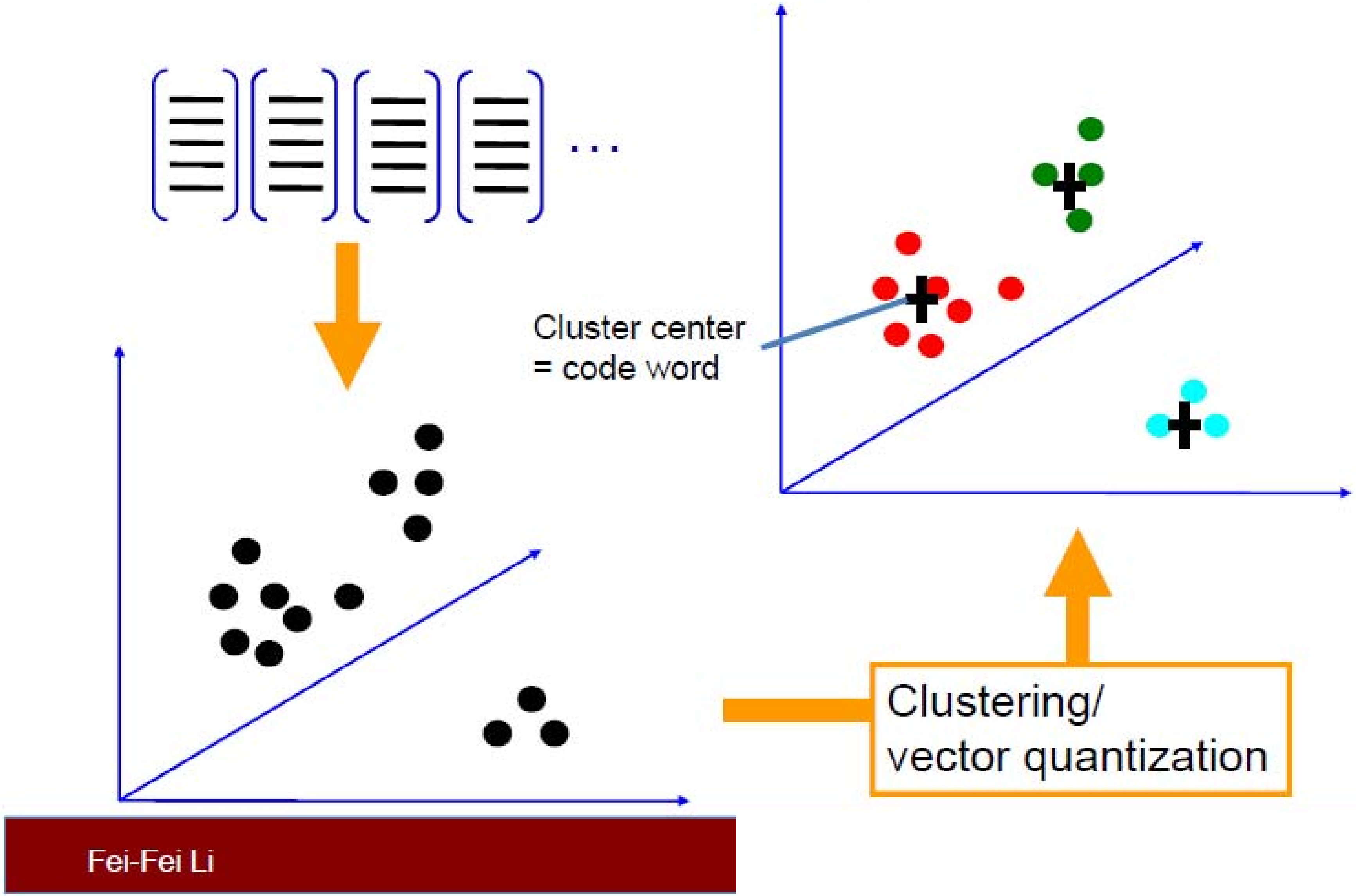
Randomly

Image credits: L. Fei-Fei, E. Nowak, J. Sivic

2. Codewords dictionary formation



2. Codewords dictionary formation

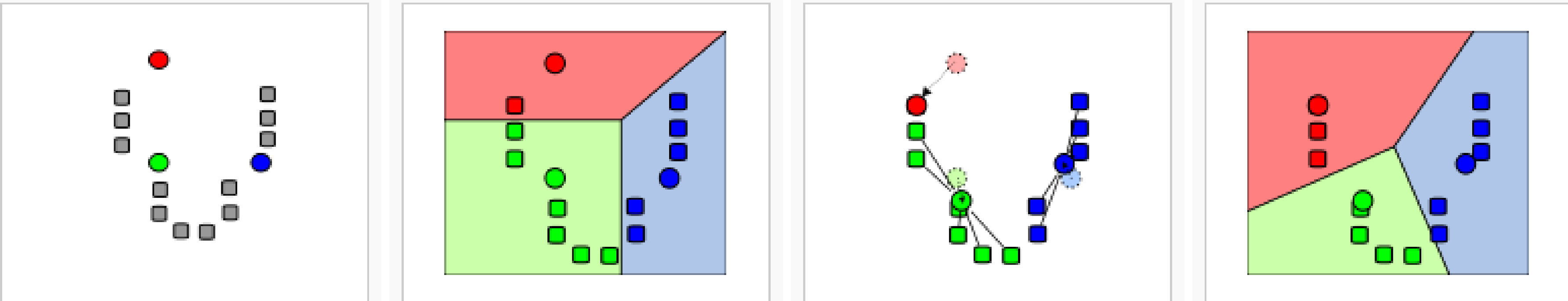


K-Means Clustering

- **Minimizing the within-cluster sum of squares (WCSS)**

$$\underset{\mathbf{S}}{\arg\min} \sum_{i=1}^{k} \sum_{\mathbf{x}_j \in S_i} \left\| \mathbf{x}_j - \boldsymbol{\mu}_i \right\|^2$$

Demonstration of the standard algorithm



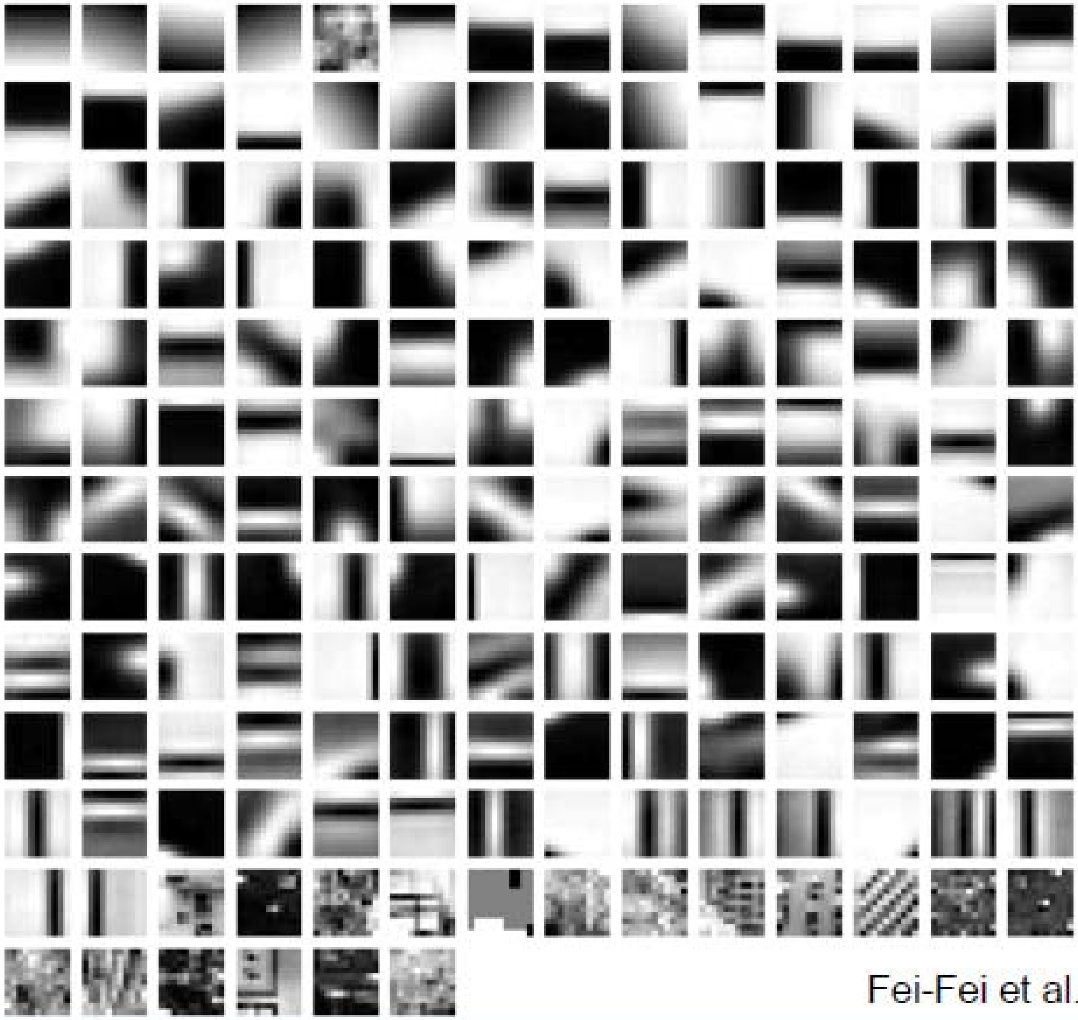
1) *k* initial "means" (in this case *k*=3) are randomly selected from the data set (shown in color).

2) *k* clusters are created by associating every observation with the nearest mean. The partitions here represent the Voronoi diagram generated by the means.

3) The centroid of each of the *k* clusters becomes the new means.

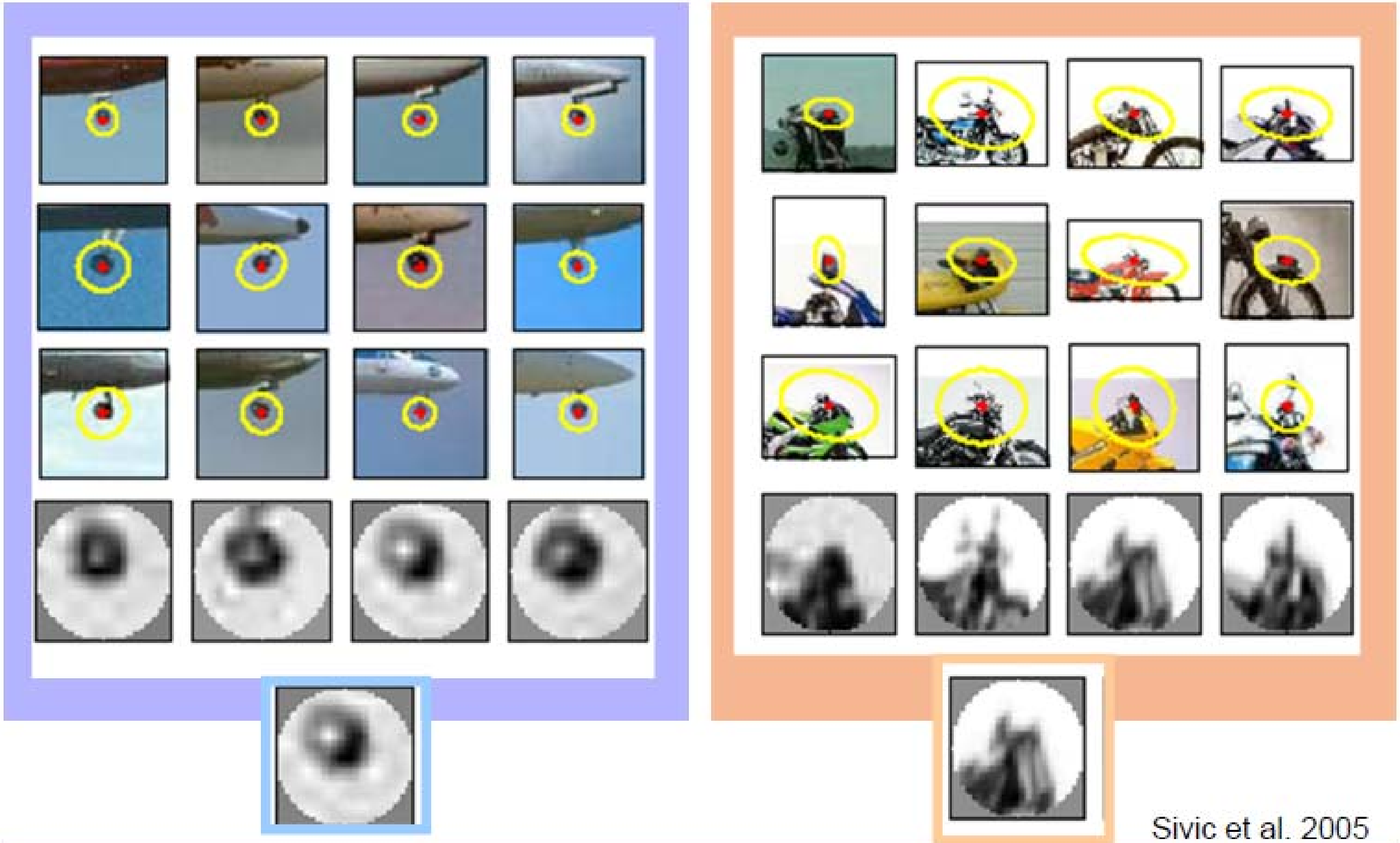
4) Steps 2 and 3 are repeated until convergence has been reached.

2. Codewords dictionary formation



Fei-Fei et al. 2005

Image patch examples of codewords

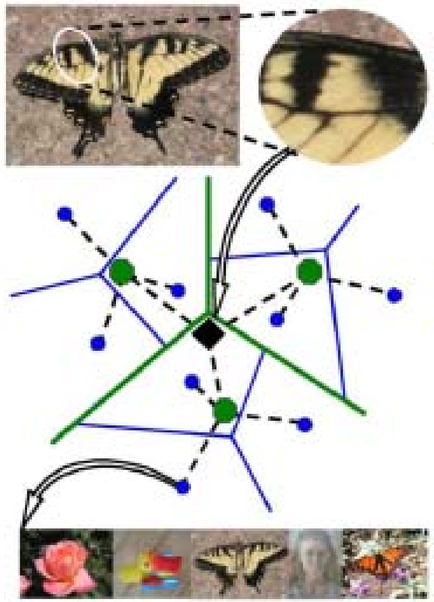


Sivic et al. 2005

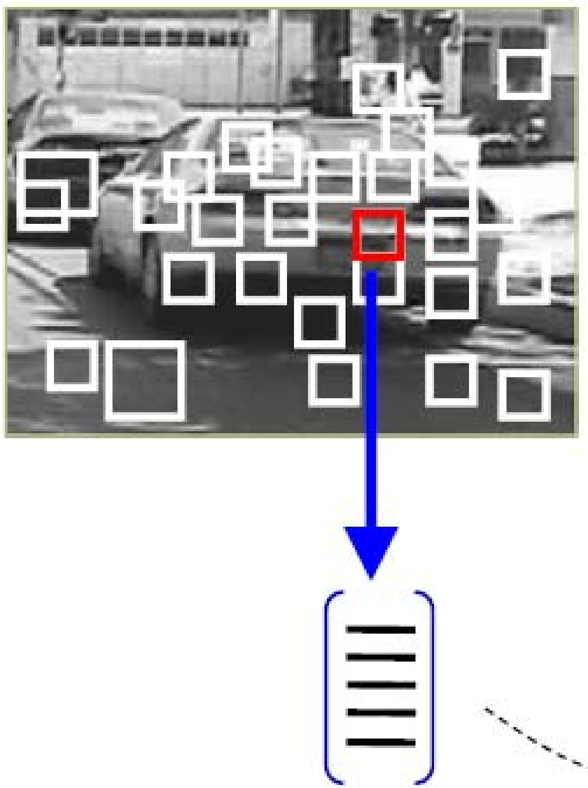
Visual vocabularies: Issues

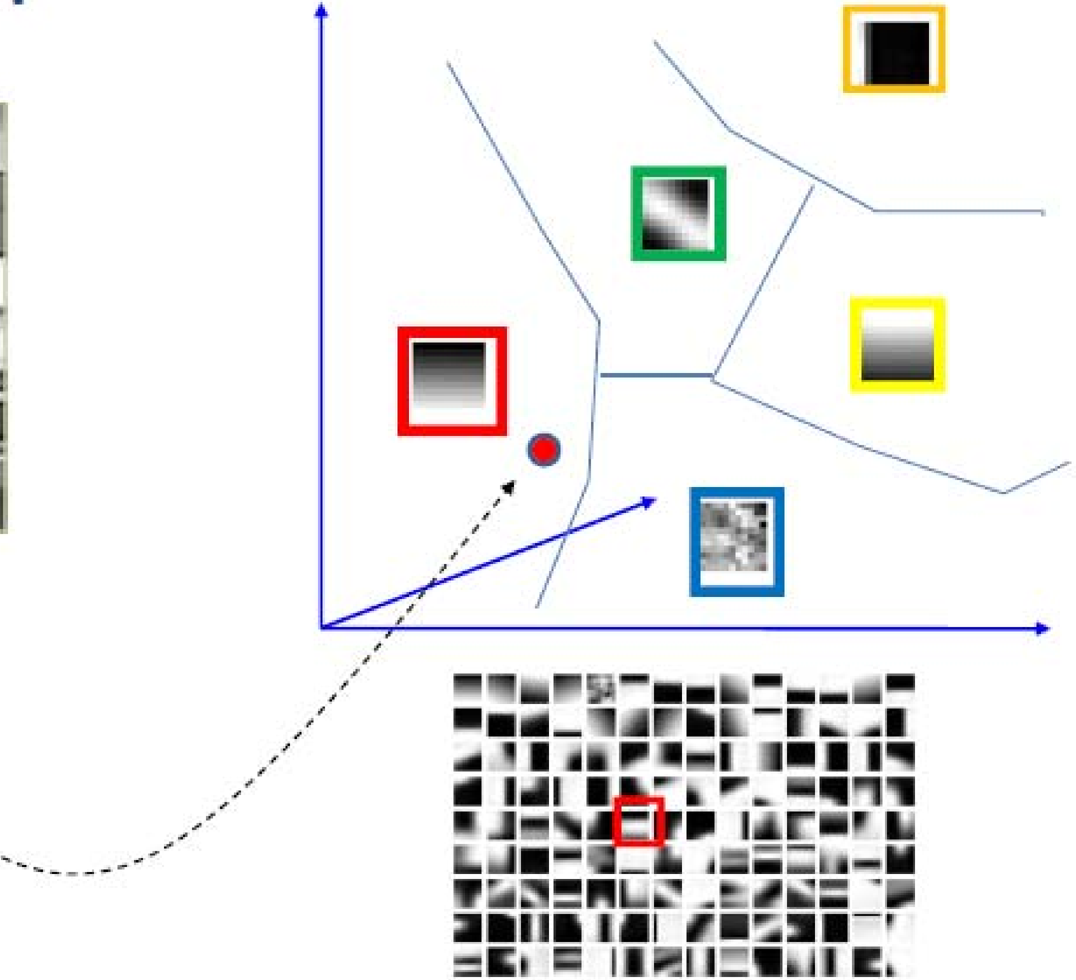
- How to choose vocabulary size?
  - Too small: visual words not representative of all patches
  - Too large: quantization artifacts, overfitting

- Computational efficiency
  - Vocabulary trees
    (Nister & Stewenius, 2006)



3. Bag of word representation

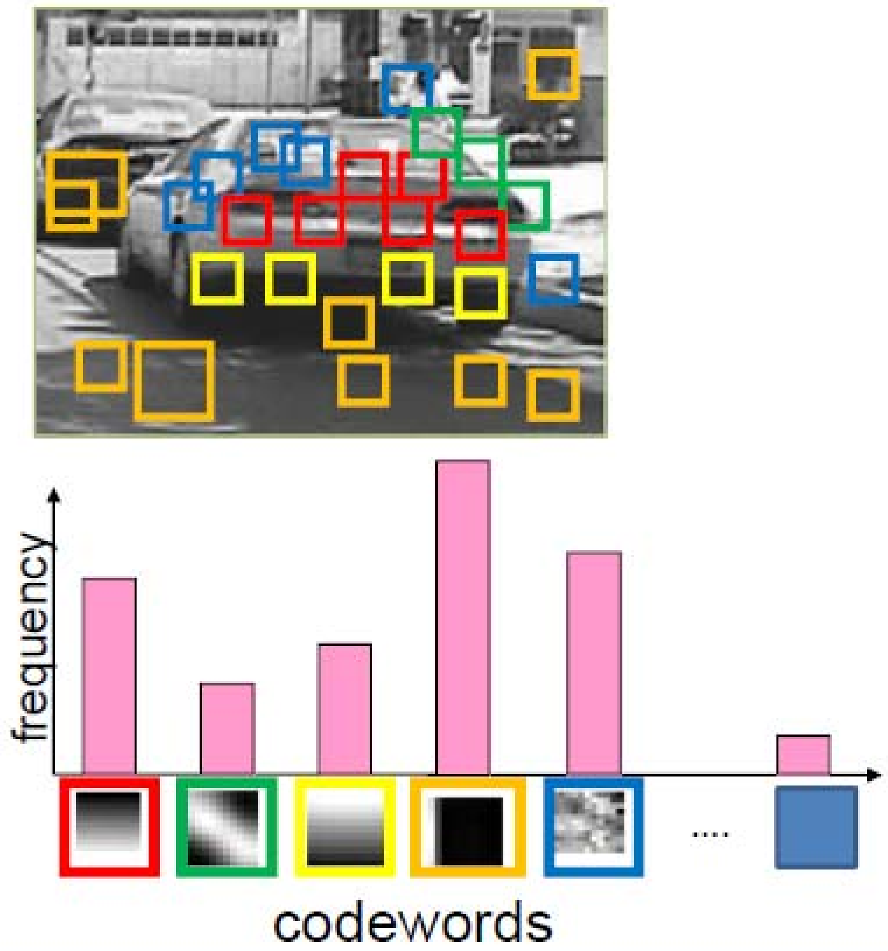


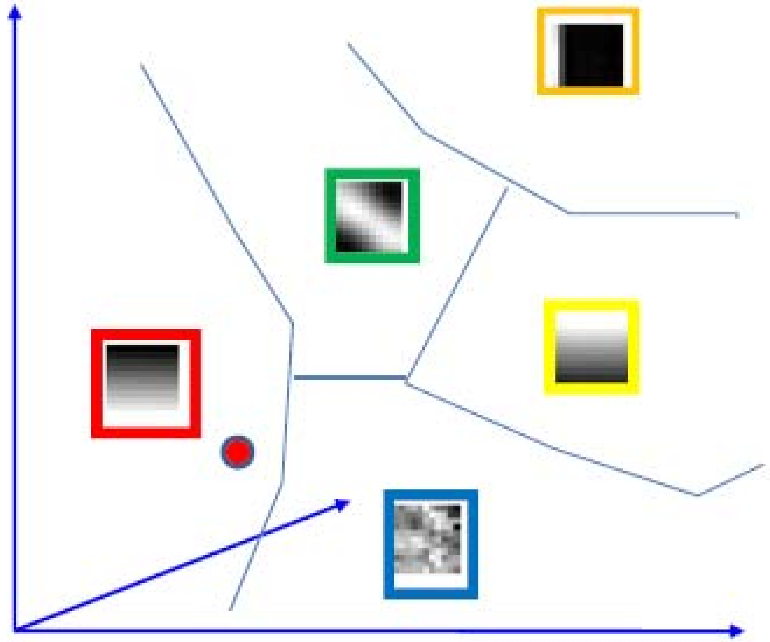


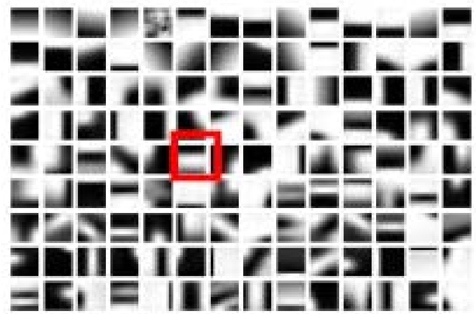
- Nearest neighbors assignment
- K-D tree search strategy

Codewords dictionary

3. Bag of word representation

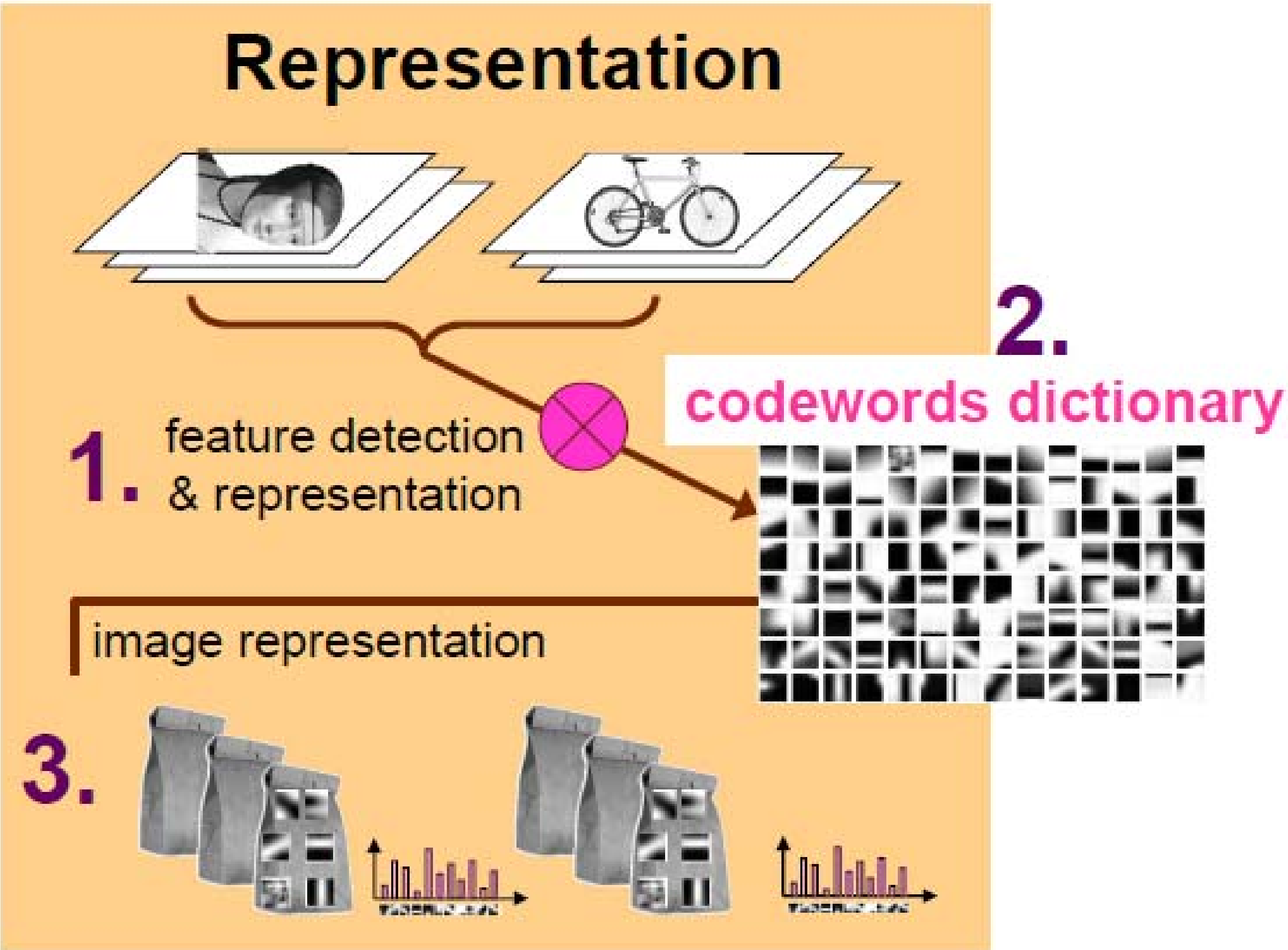




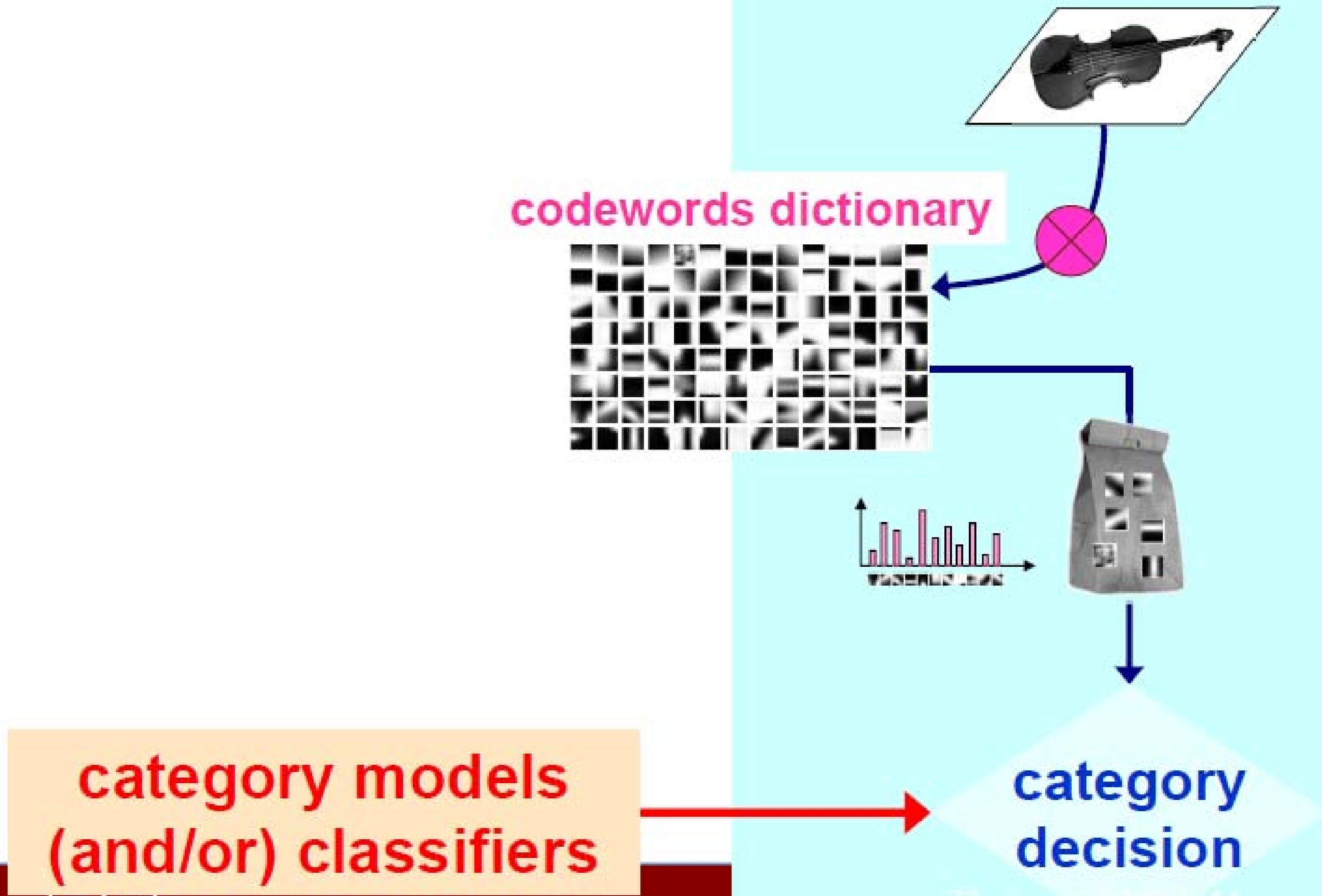


Codewords dictionary

Fei-Fei Li

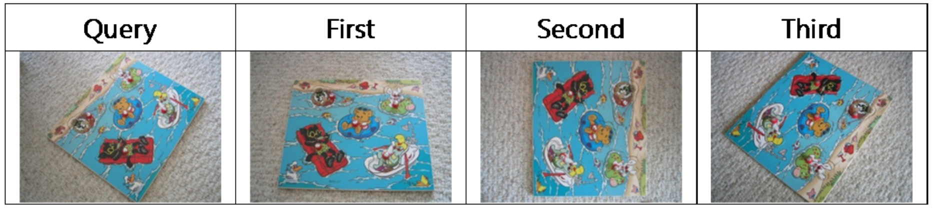

Representation
2.
codewords dictionary
1.
feature detection
& representation
image representation
3.

Learning and Recognition



PA2

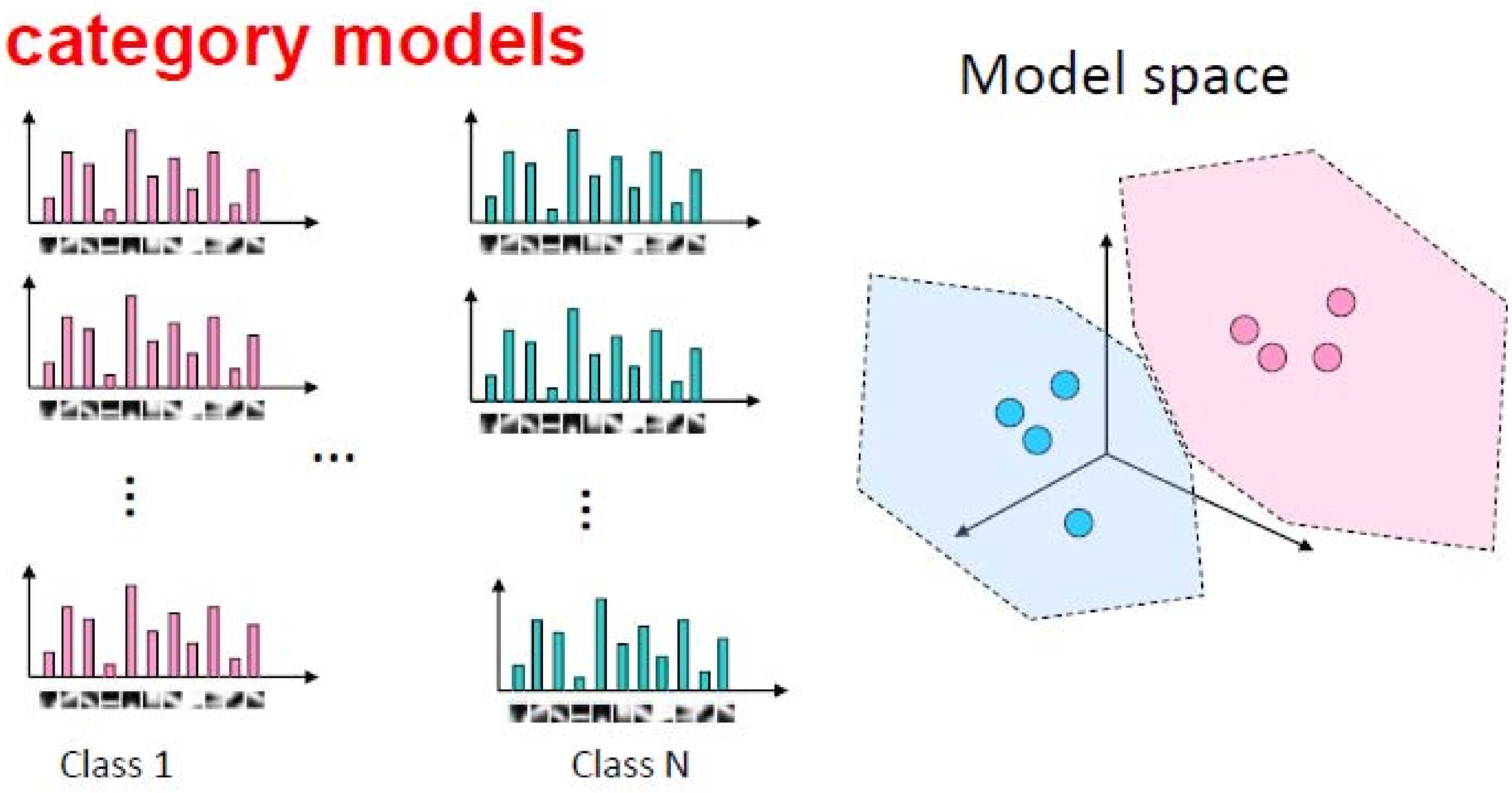
- Understand and implement a basic image retrieval system
- Use the original UKBenchmark



Learning and Recognition

- **Nearest neighbor**
- **SVM**

Discriminative classifiers

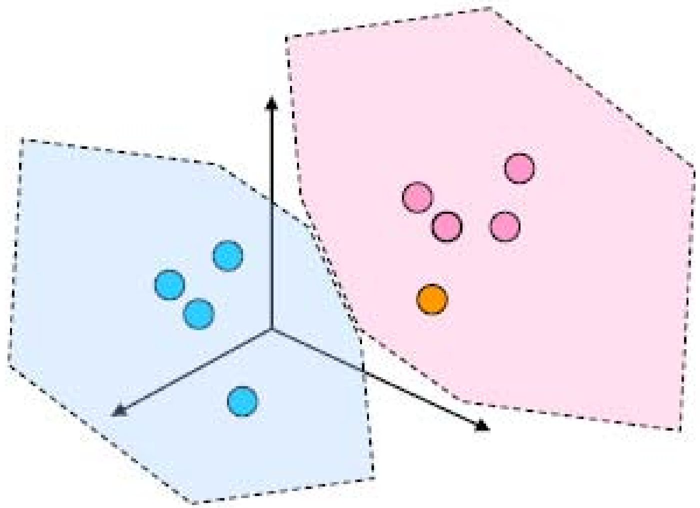


Discriminative classifiers

Query image

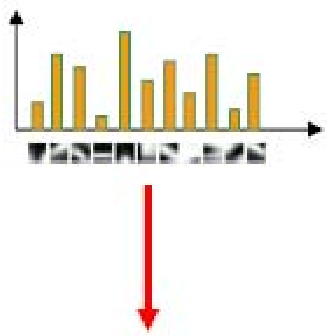
Winning class: pink

Model space



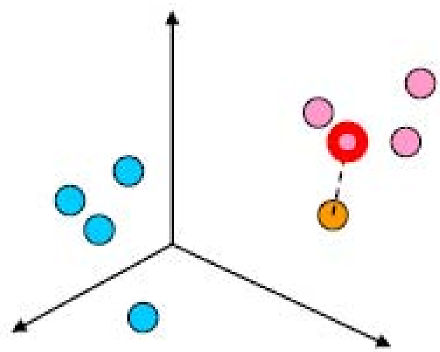
Nearest Neighbors classifier

Query image



Winning class: pink

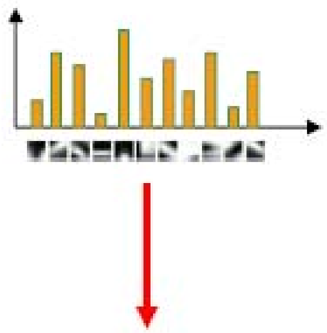
Model space



- Assign label of nearest training data point to each test data point

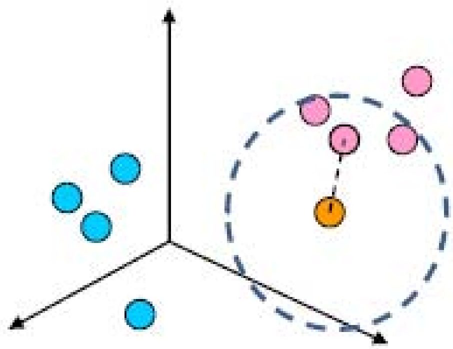
K- Nearest Neighbors classifier

Query image



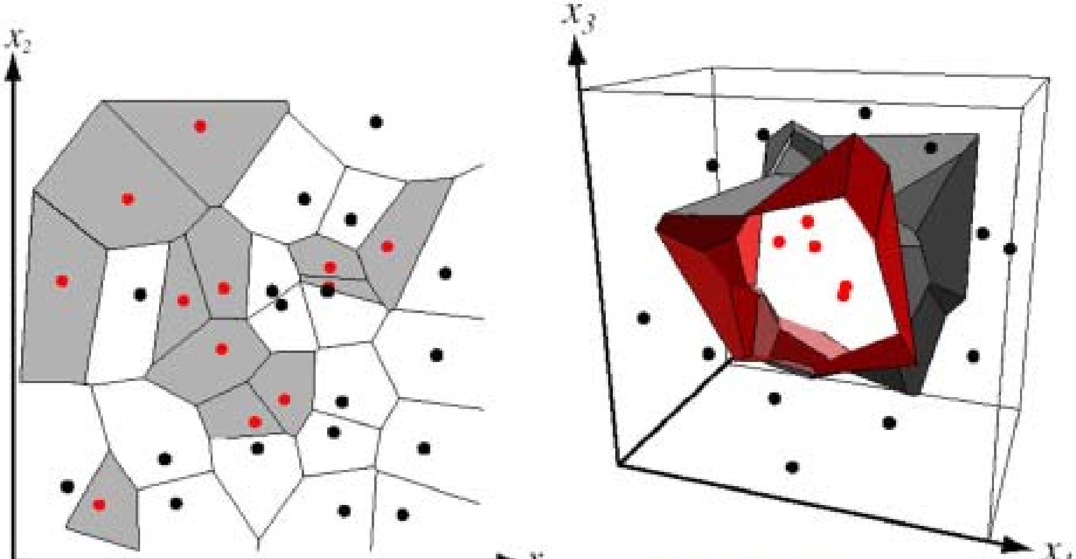
Winning class: pink

Model space



- For a new point, find the k closest points from training data
- Labels of the k points "vote" to classify
- Works well provided there is lots of data and the distance function is good

K- Nearest Neighbors classifier



from Duda *et al.*

- Voronoi partitioning of feature space for 2-category 2-D and 3-D data

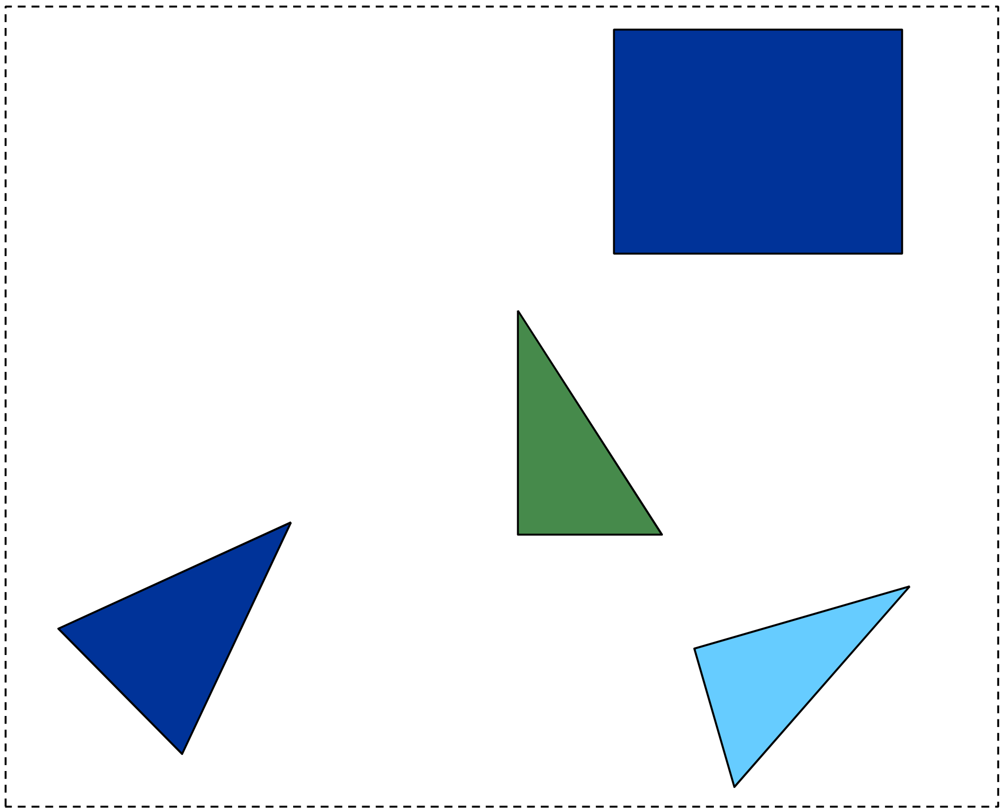
- For k dimensions: k-D tree = space-partitioning data structure for organizing points in a *k*-dimensional space
- Enable efficient search

- Nice tutorial: http://www.cs.umd.edu/class/spring2002/cmsc420-0401/pbasic.pdf

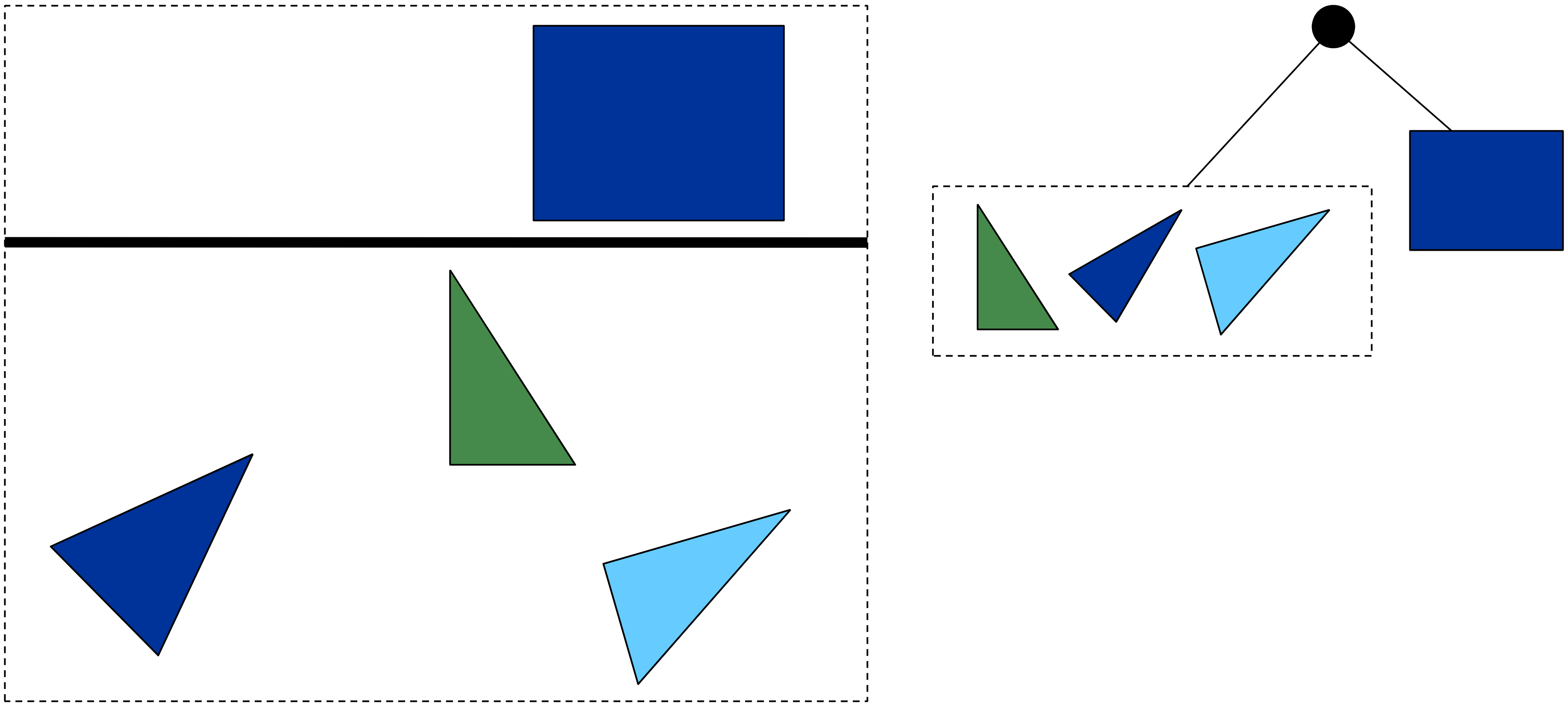
Overview of kd-Trees

- **Binary spatial subdivision (special case of BSP tree)**
- **Split planes aligned on main axis**
- **Inner nodes: subdivision planes**
- **Leaf nodes: points**

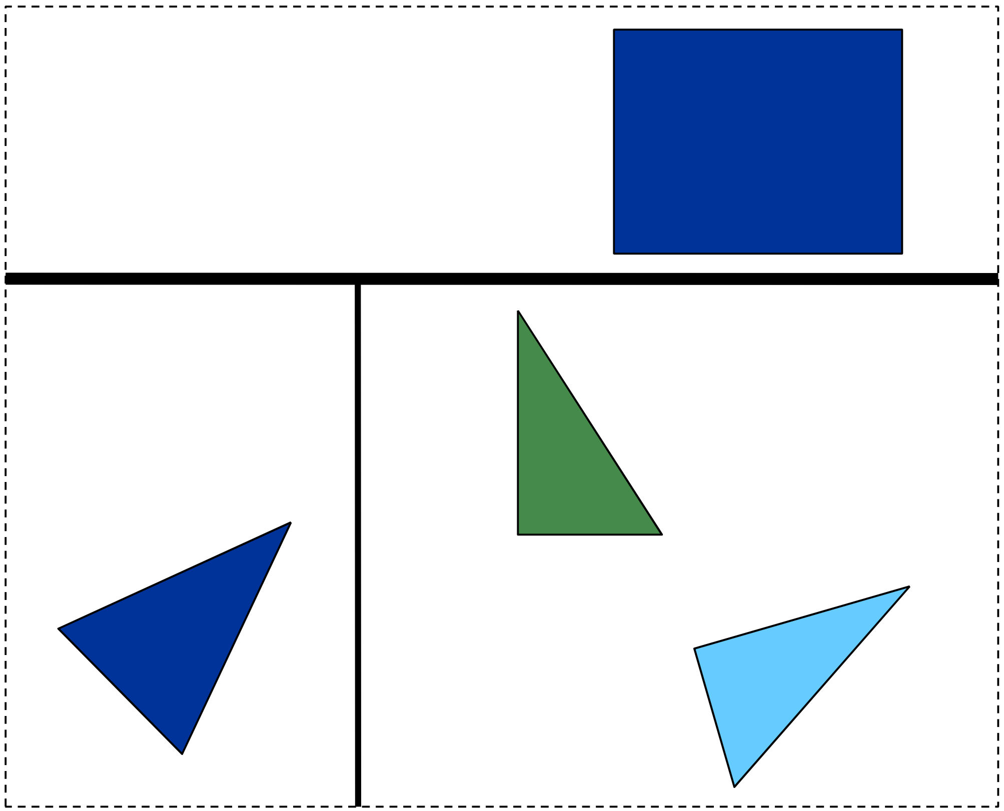
2D Example with Triangles

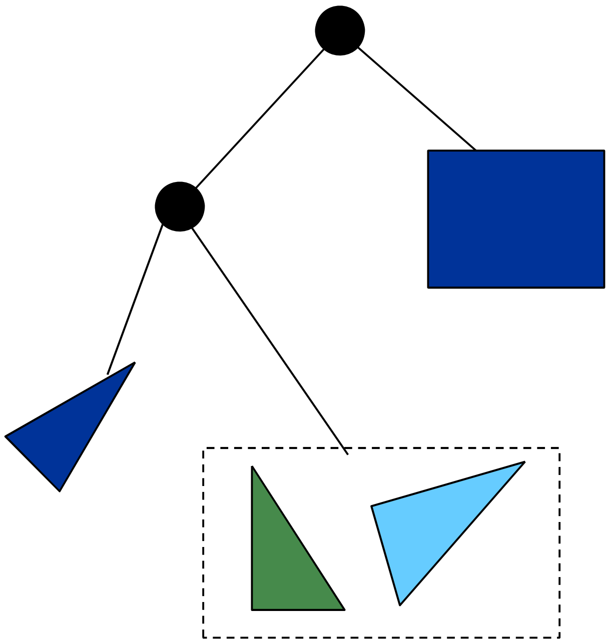


2D Example with Triangles

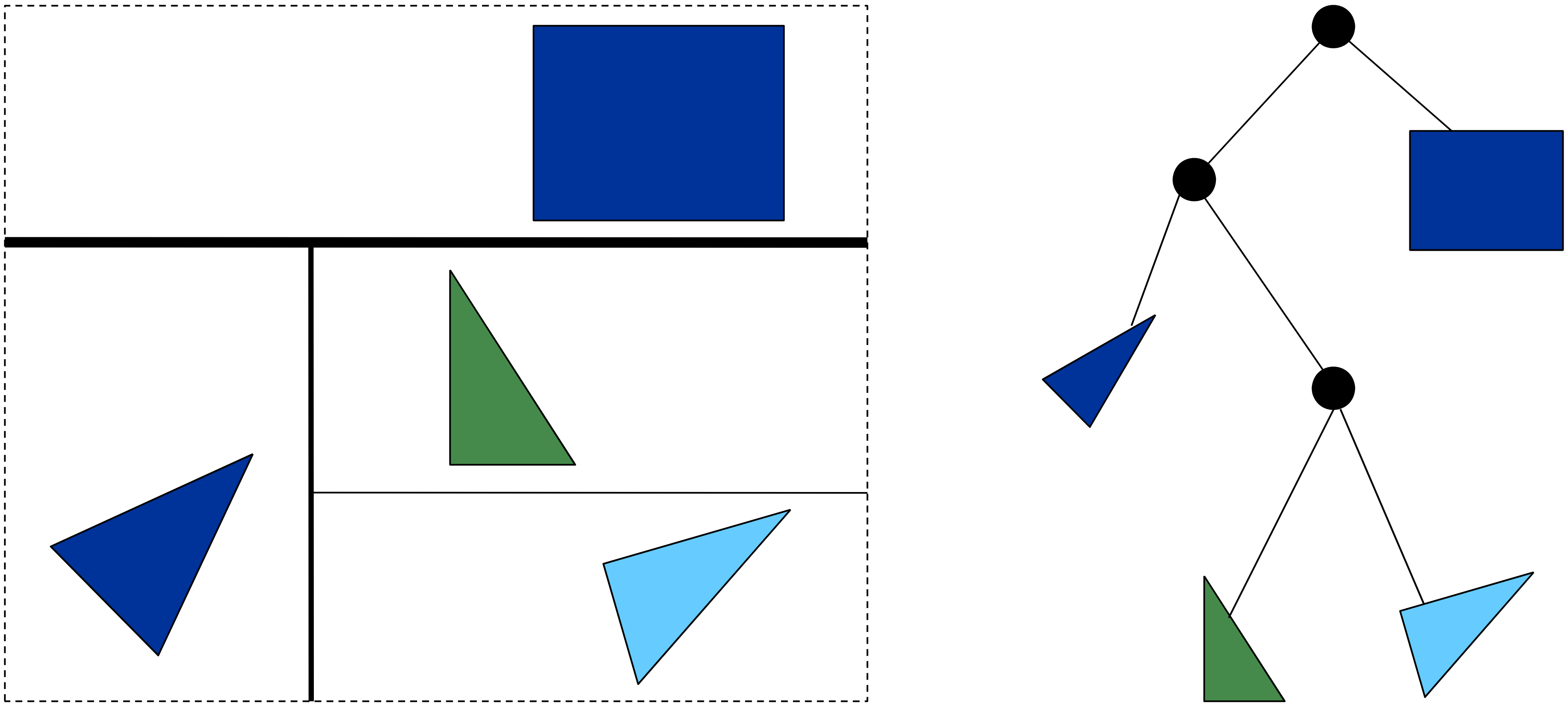


2D Example with Triangles





2D Example with Triangles



Nearest Neighbor Search with kd-tree

- Goal: find k nearest neighbors given a point
  - Commonly identify approximate, not exact nearest neighbors
- Apply a depth-first search
  - Traverse the tree with a stack
- Or, we can apply a best-bin first search
  - Traverse more promising nodes first
- Traverse until we visit a certain number of nodes

Hashing techniques

- Kd-trees are not scalable
- Hashing arise as better technology

Functions for comparing histograms

- L1 distance

$$D(h_1, h_2) = \sum_{i=1}^{N} | h_1(i) - h_2(i) |$$

- χ^2 distance

$$D(h_1, h_2) = \sum_{i=1}^{N} \frac{\left(h_1(i) - h_2(i)\right)^2}{h_1(i) + h_2(i)}$$

- Quadratic distance (*cross-bin*)

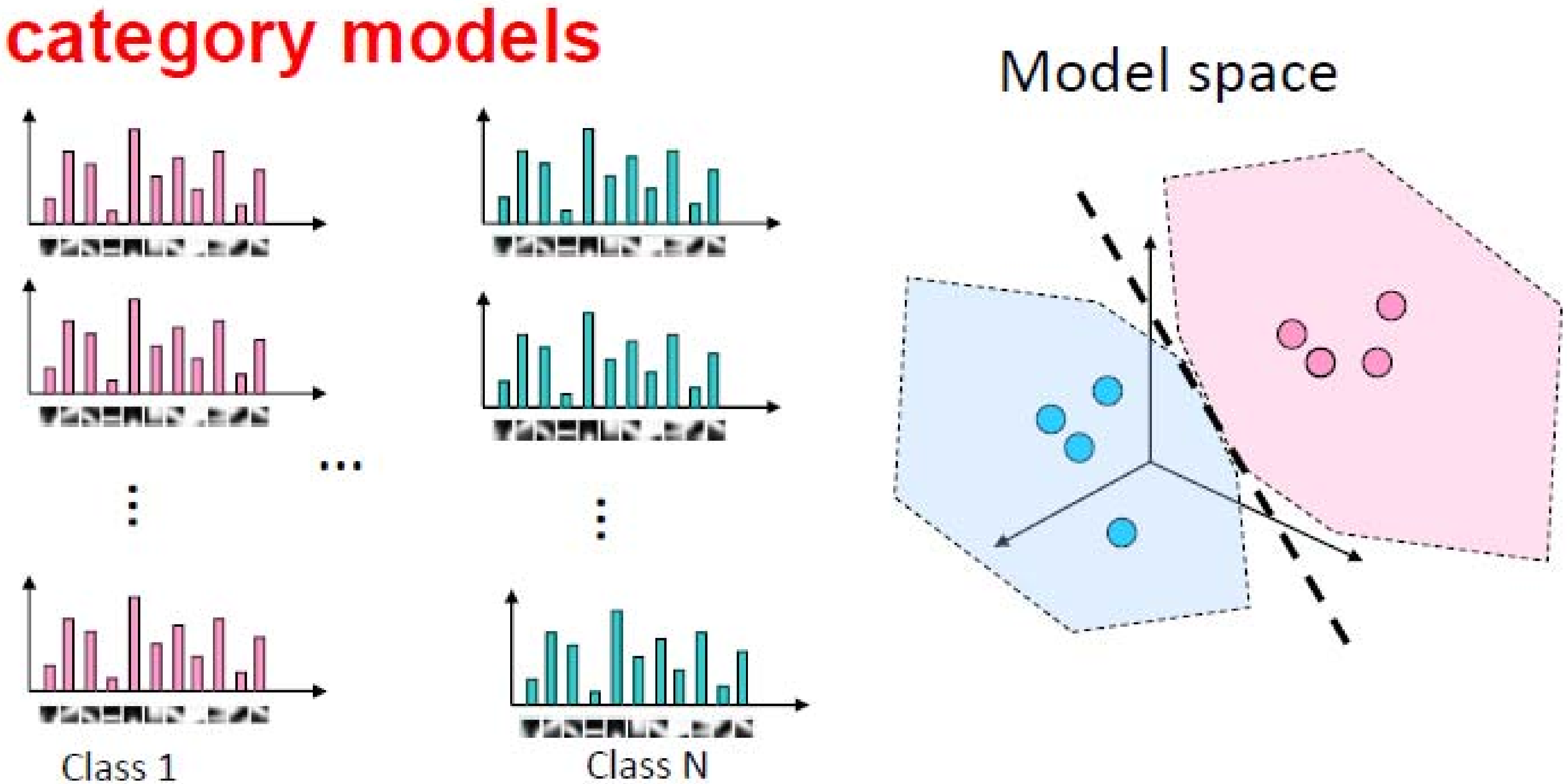
$$D(h_1, h_2) = \sum_{i,j} A_{ij} (h_1(i) - h_2(j))^2$$

Jan Puzicha, Yossi Rubner, Carlo Tomasi, Joachim M. Buhmann: Empirical Evaluation of Dissimilarity Measures for Color and Texture. ICCV 1999

Learning and Recognition

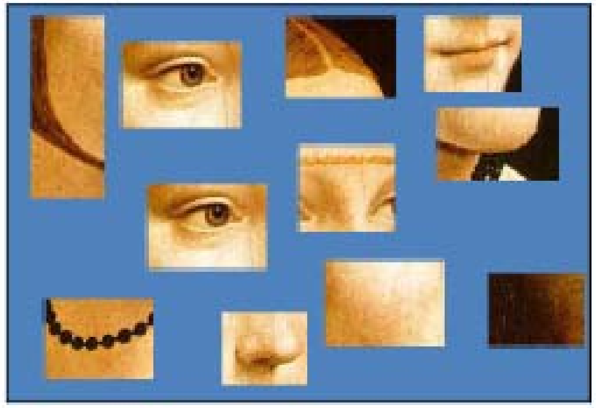
- **Nearest neighbor**
- **SVM**

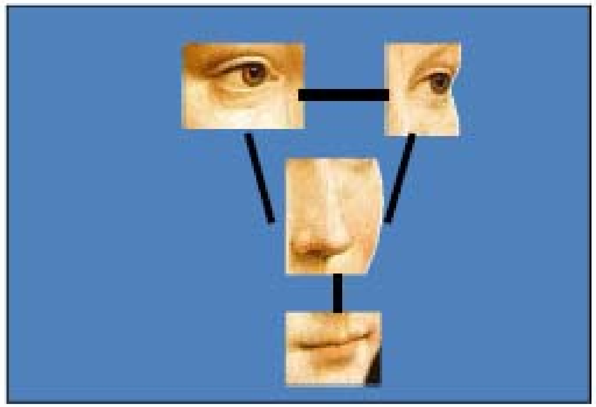
Discriminative classifiers (linear classifier)



Weakness of BoW the models

- No rigorous geometric information of the object components
- It's intuitive to most of us that objects are made of parts – no such information
- Not extensively tested yet for
  - View point invariance
  - Scale invariance
- Segmentation and localization unclear





What have we learned today?

- Bag of Words models
  - Basic representation
  - Different learning and recognition algorithms

Next Time…

- **Nearest neighbor search using hashing**

Hashing Techniques

윤성의 (Sung-Eui Yoon)
Associate Professor
KAIST

http://sglab.kaist.ac.kr

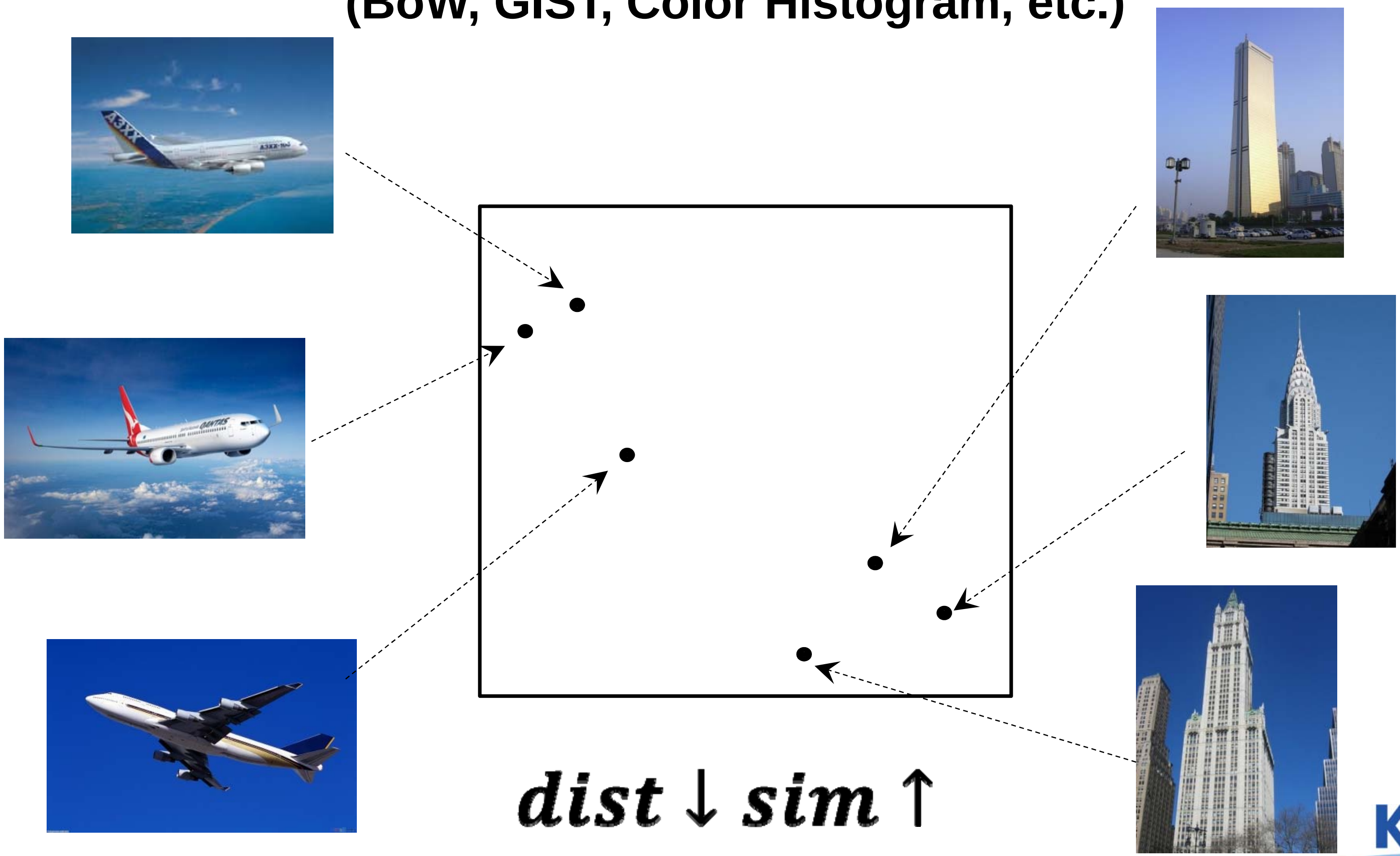
Image Retrieval

Finding visually similar images

Image Descriptor

High dimensional point

(BoW, GIST, Color Histogram, etc.)

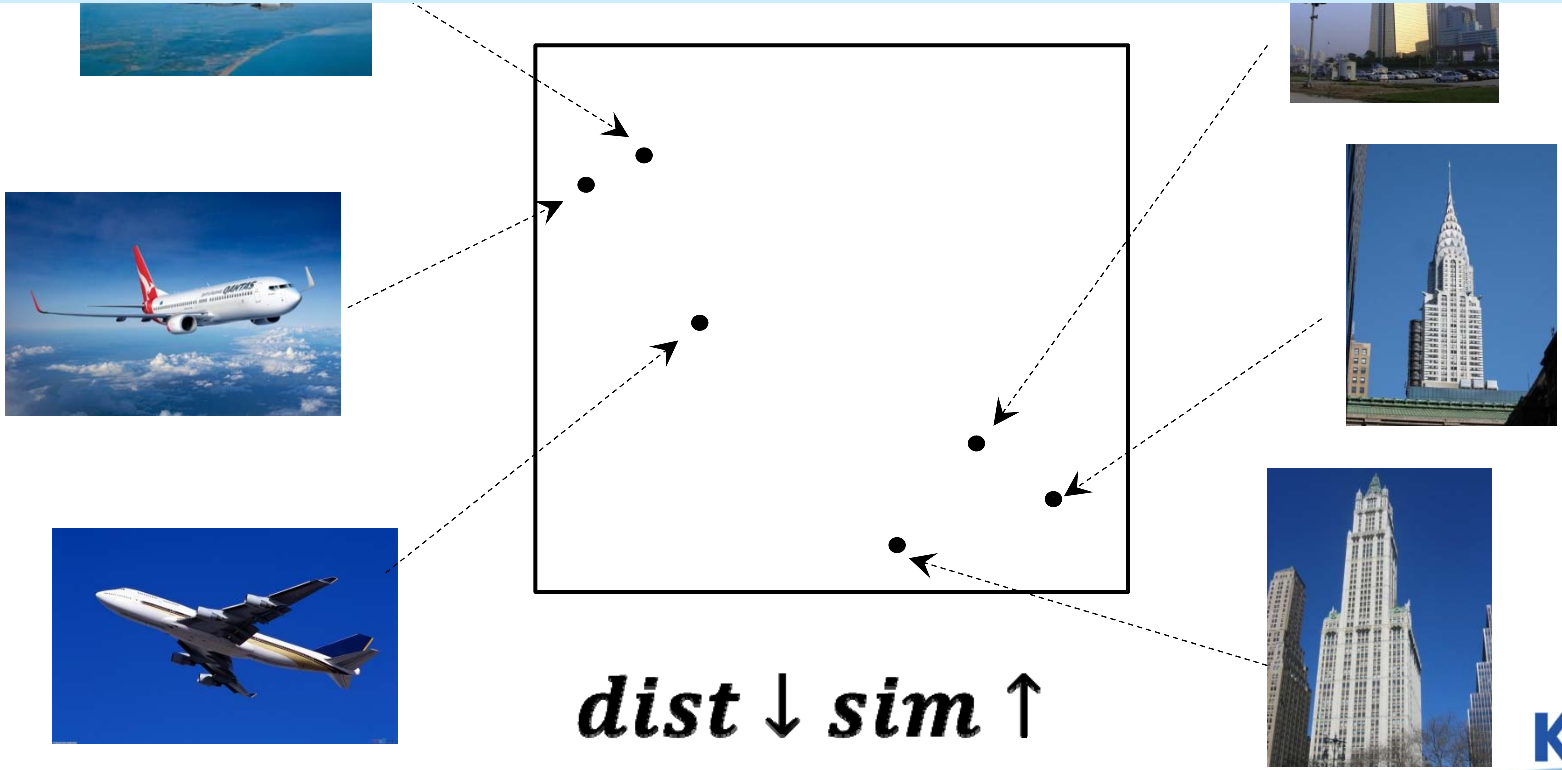


$dist \downarrow sim \uparrow$

Image Descriptor

High dimensional point

Nearest neighbor search (NNS)
in high dimensional space



$dist \downarrow \; sim \uparrow$

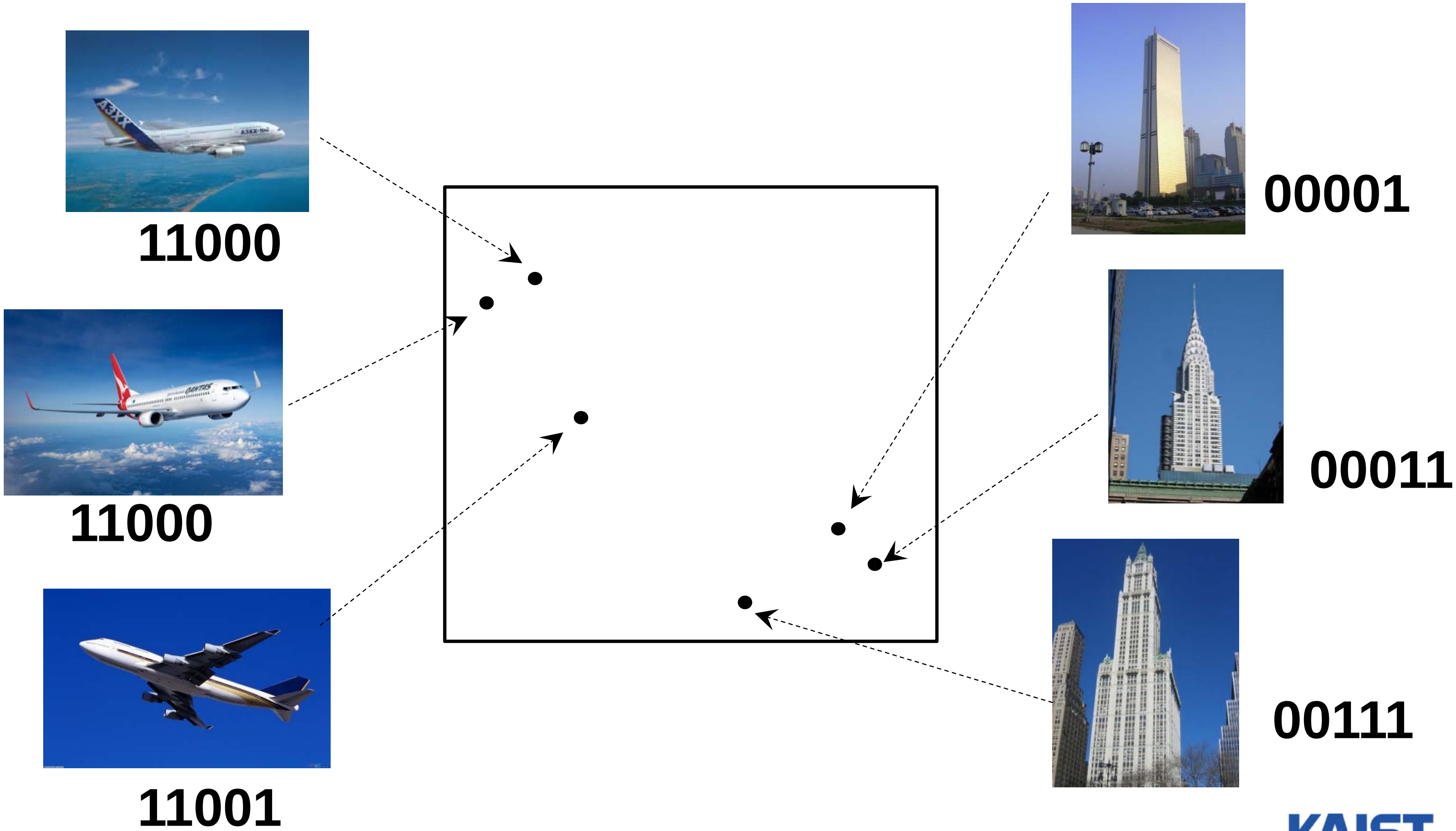
Challenge

	BoW	GIST
Dimensions	1000+	300+
1 image	4 KB+	1.2 KB+
1B images	3 TB+	1 TB+

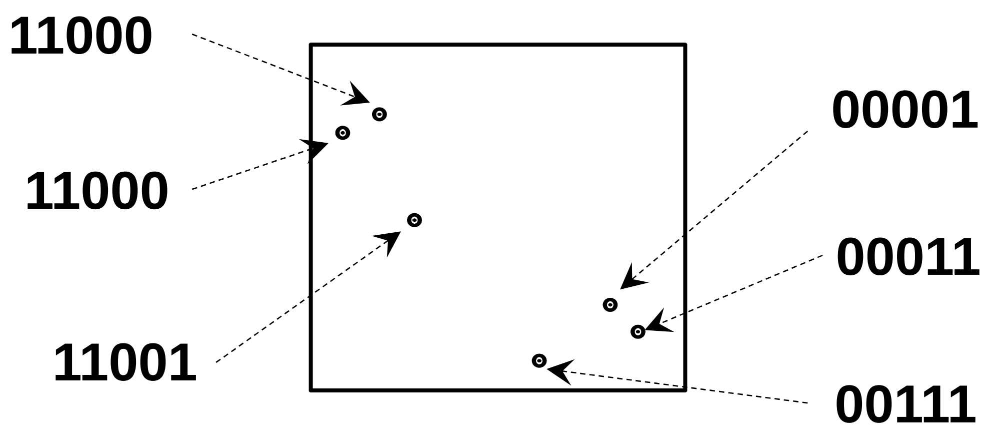
$$\frac{144\ GB\ memory}{1\ billion\ images} \approx \frac{128\ bits}{1\ image}$$

KAIST

Binary Code



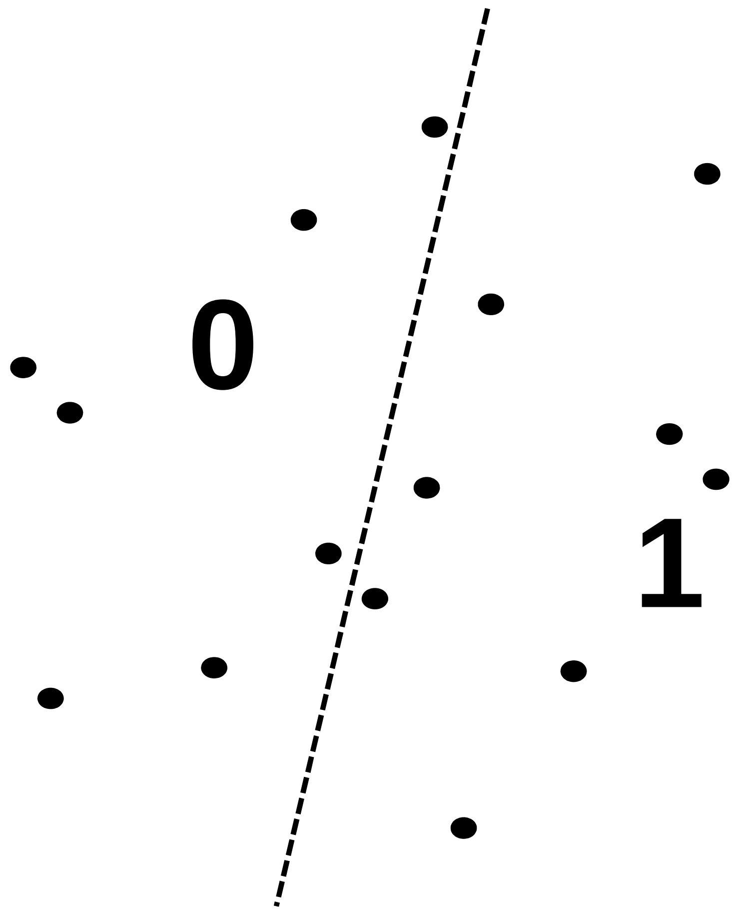
Binary Code



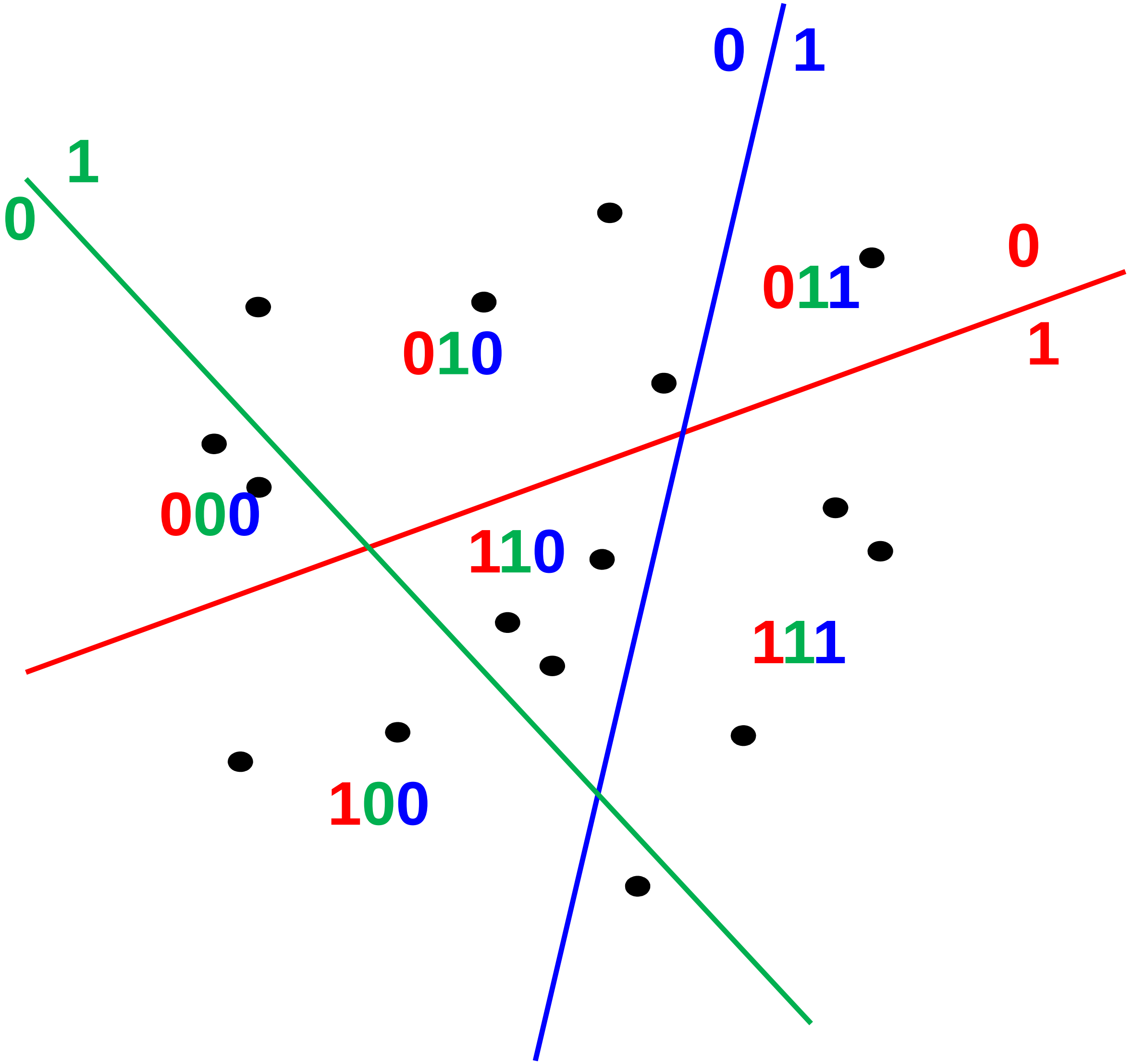
* Benefits

- Compression
- Very fast distance computation (Hamming Distance, XOR)

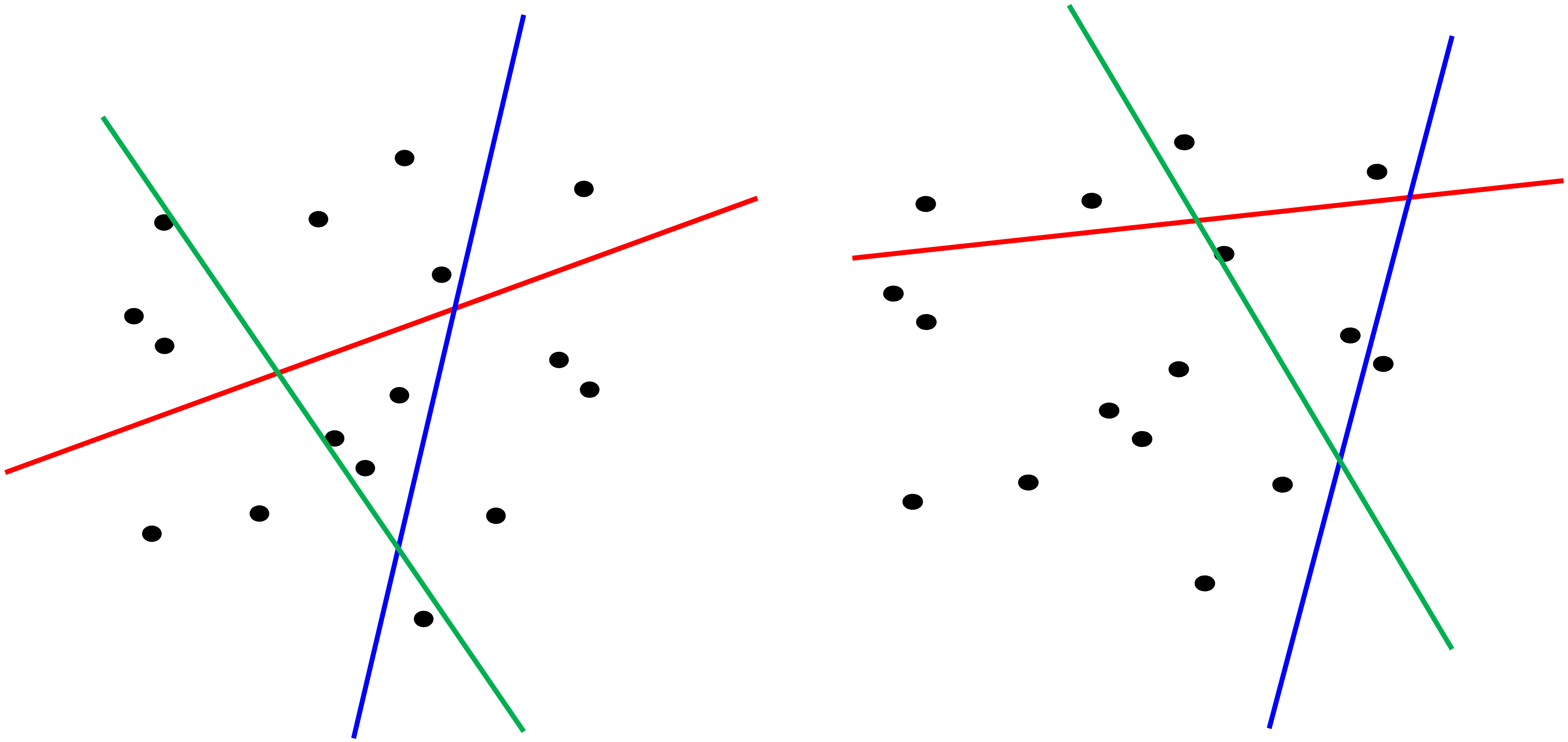
Hyper-Plane based Binary Coding



Hyper-Plane based Binary Coding



Good and Bad Hyper-Planes



Previous work focused on how to determine good hyper-planes

Previous Work

- **Random hyper-planes from a specific distribution**
  [Indyk STOC 1998, Raginsky NIPS 2009]
- **Spectral graph partitioning**
  [Yeiss, NIPS 2008]
- **Minimize quantization error**
  [Gong, CVPR 2011 oral session]
- **Independent component analysis**
  [He, CVPR 2011 oral session]
- **Support Vector Machine**
  [Joly, CVPR 2011]

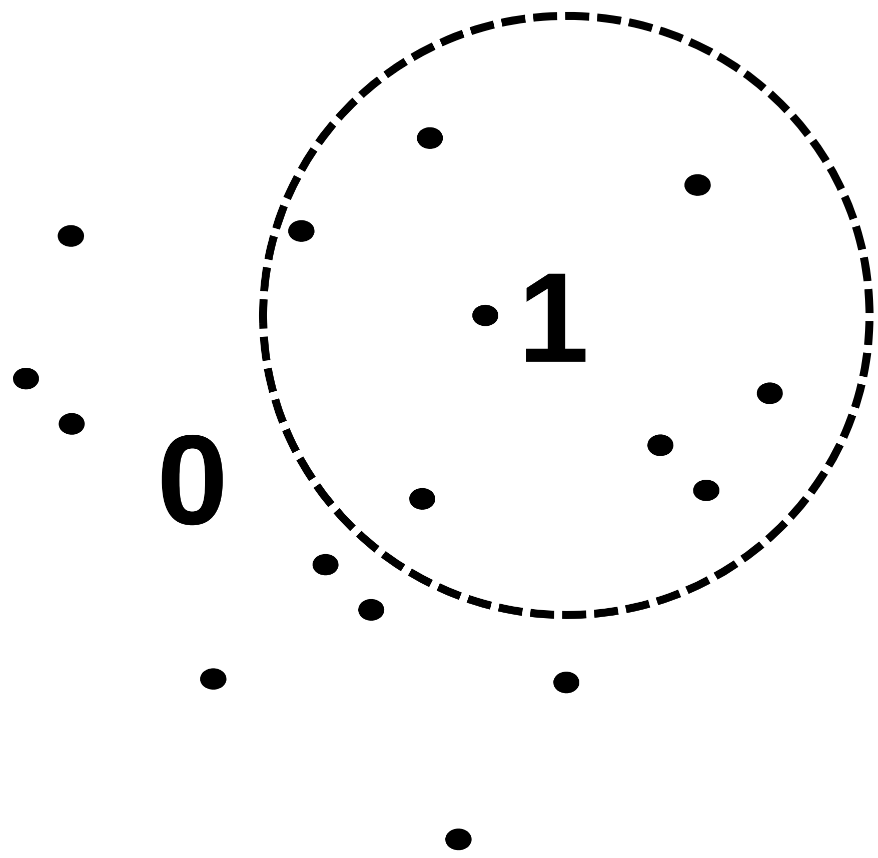
Components of Spherical Hashing

- **Spherical hashing**
- **Hyper-sphere setting strategy**
- **Spherical Hamming distance**

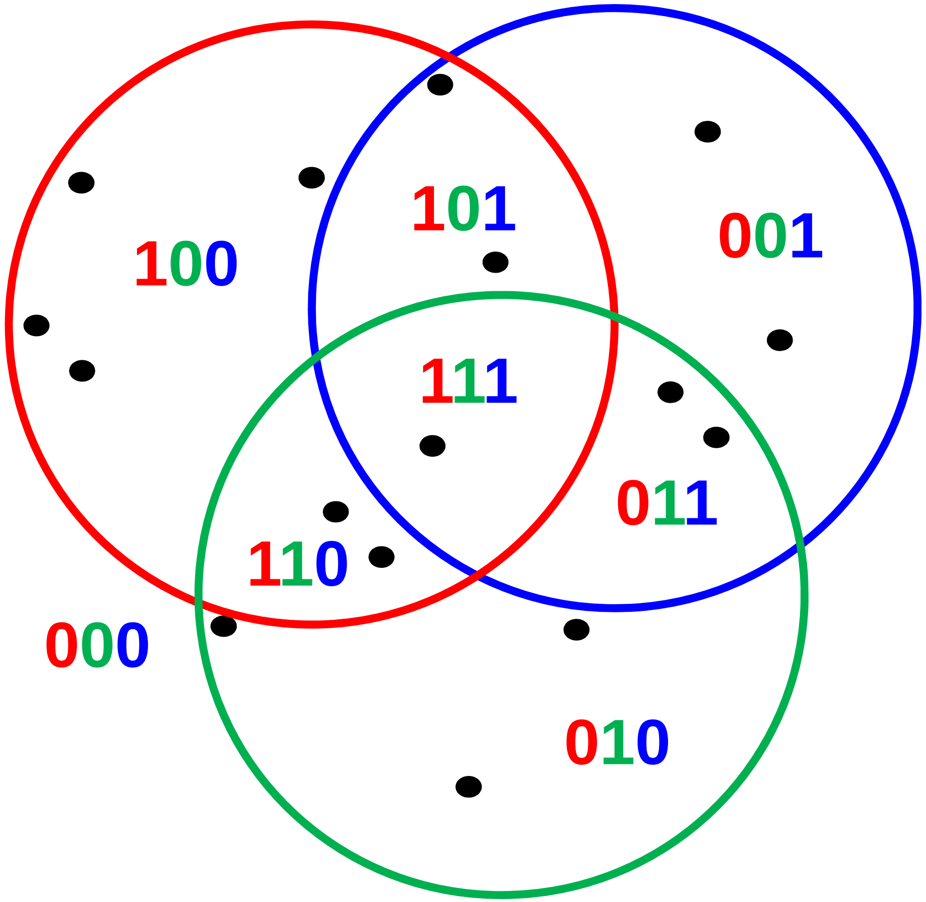
Components of Spherical Hashing

- **Spherical hashing**
- Hyper-sphere setting strategy
- Spherical Hamming distance

Spherical Hashing [Heo et al., CVPR 12]

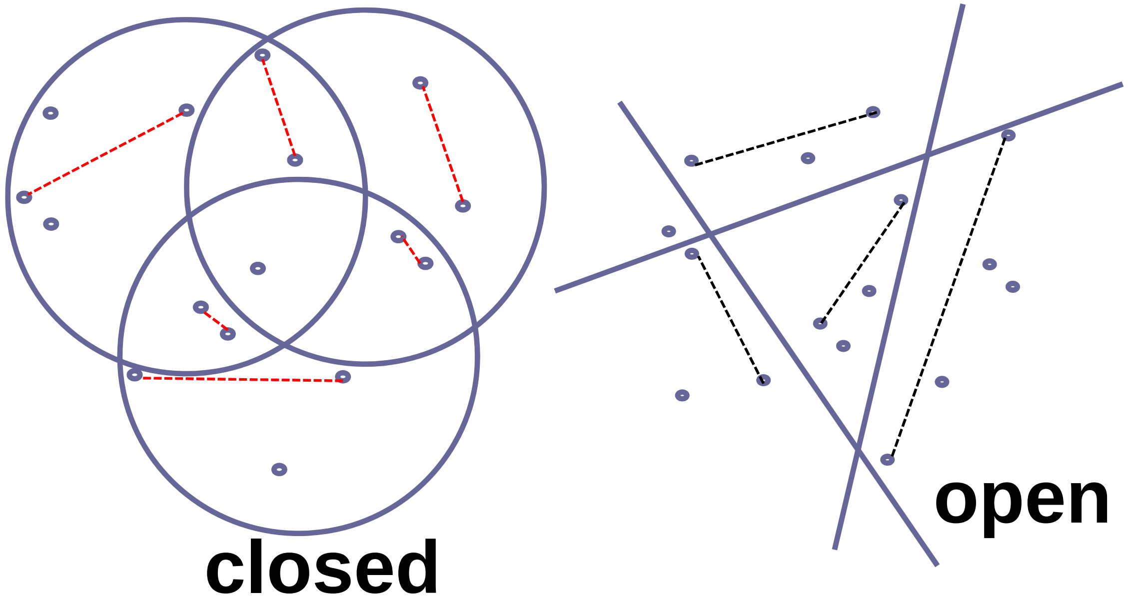


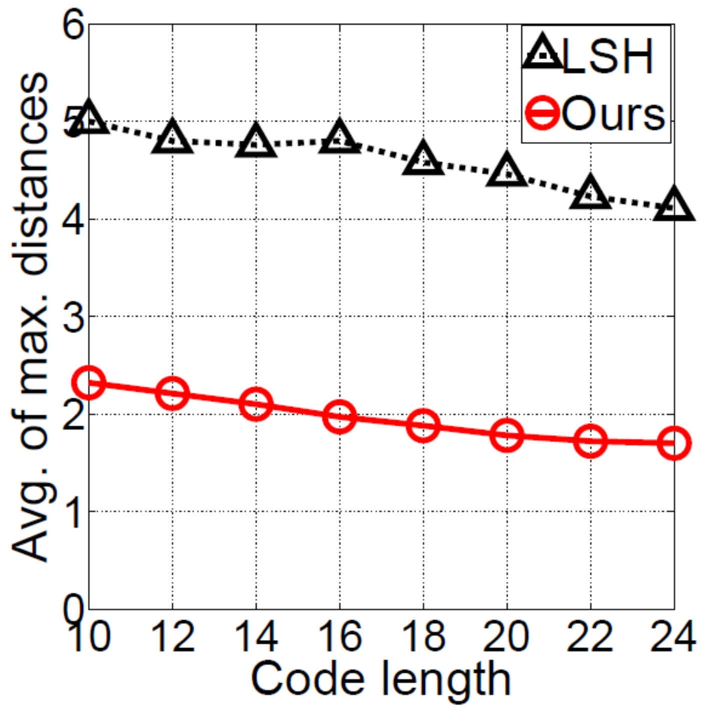
Spherical Hashing [Heo et al., CVPR 12]



KAIST

Hyper-Sphere vs Hyper-Plane





Average of maximum distances within a partition:

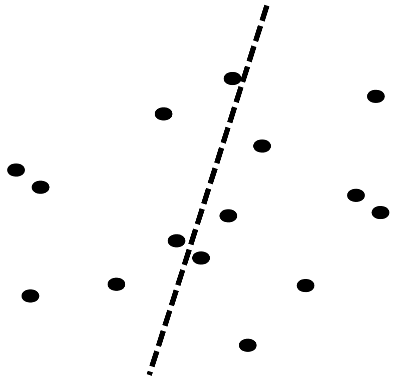
- Hyper-spheres gives tighter bound!

Components of Spherical Hashing

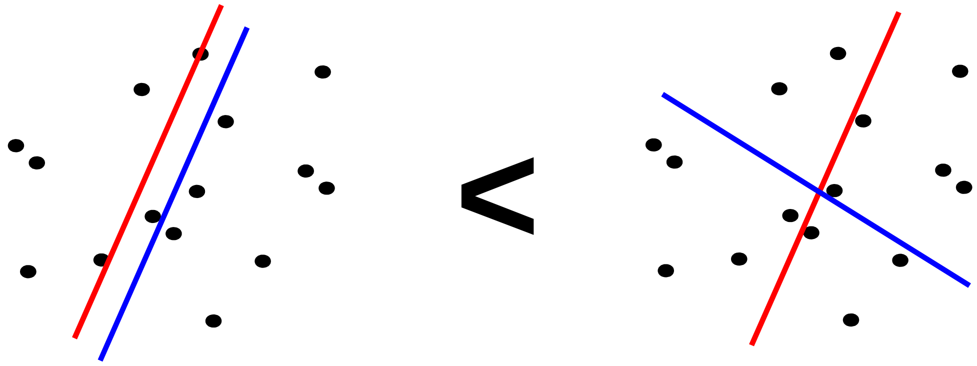
- Spherical hashing
- **Hyper-sphere setting strategy**
- Spherical Hamming distance

Good Binary Coding [Yeiss 2008, He 2011]

1. Balanced partitioning

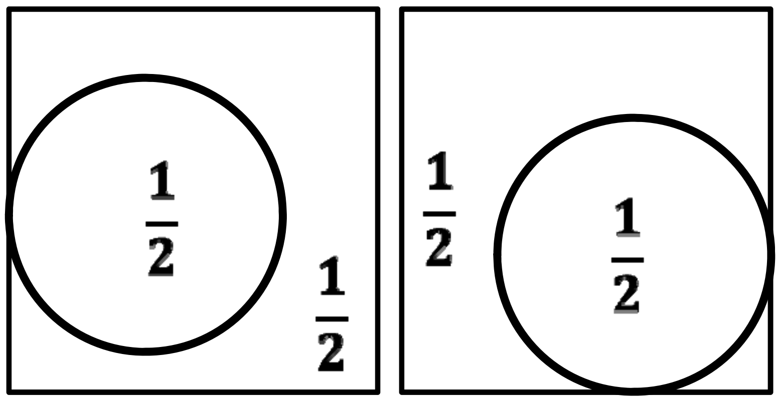


2. Independence

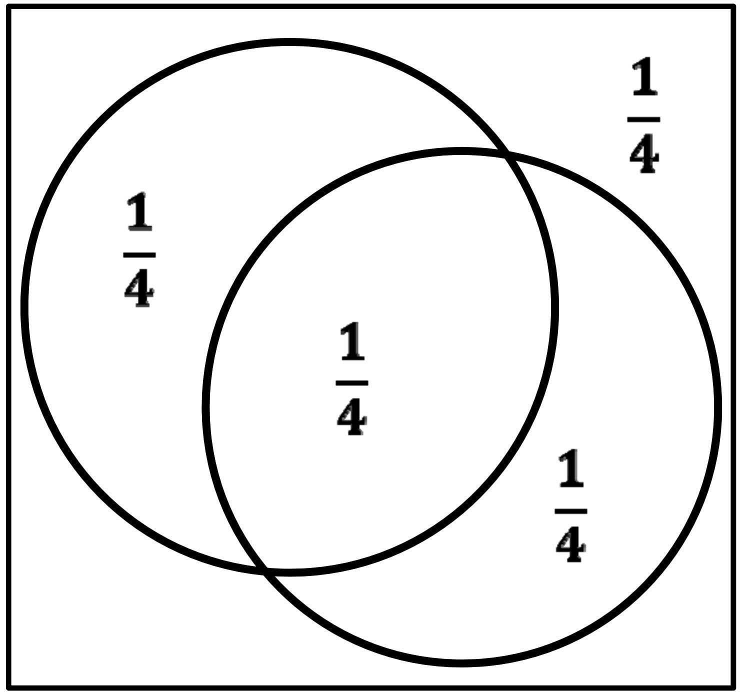


Intuition of Hyper-Sphere Setting

1. Balance

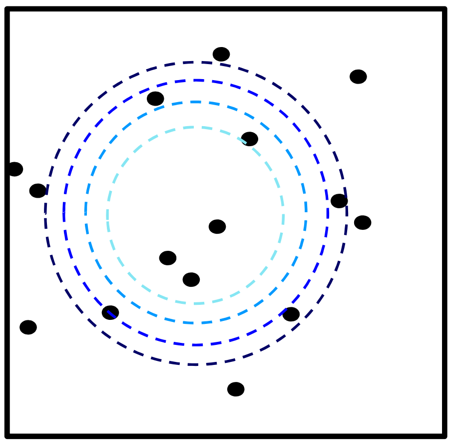


2. Independence

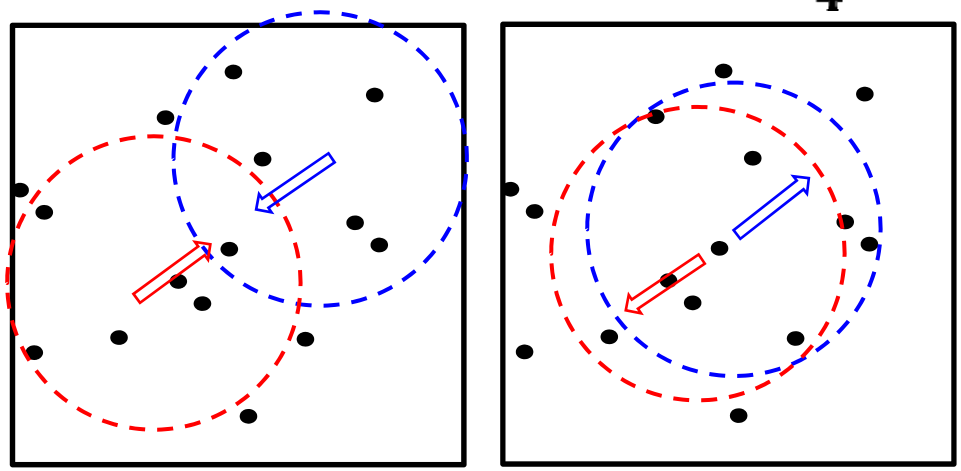


Hyper-Sphere Setting Process

1. Balance
- by controlling radius for $n(S) = \frac{N}{2}$



2. Independence
- by moving two hyper-spheres for $n(S_1 \cap S_2) = \frac{N}{4}$

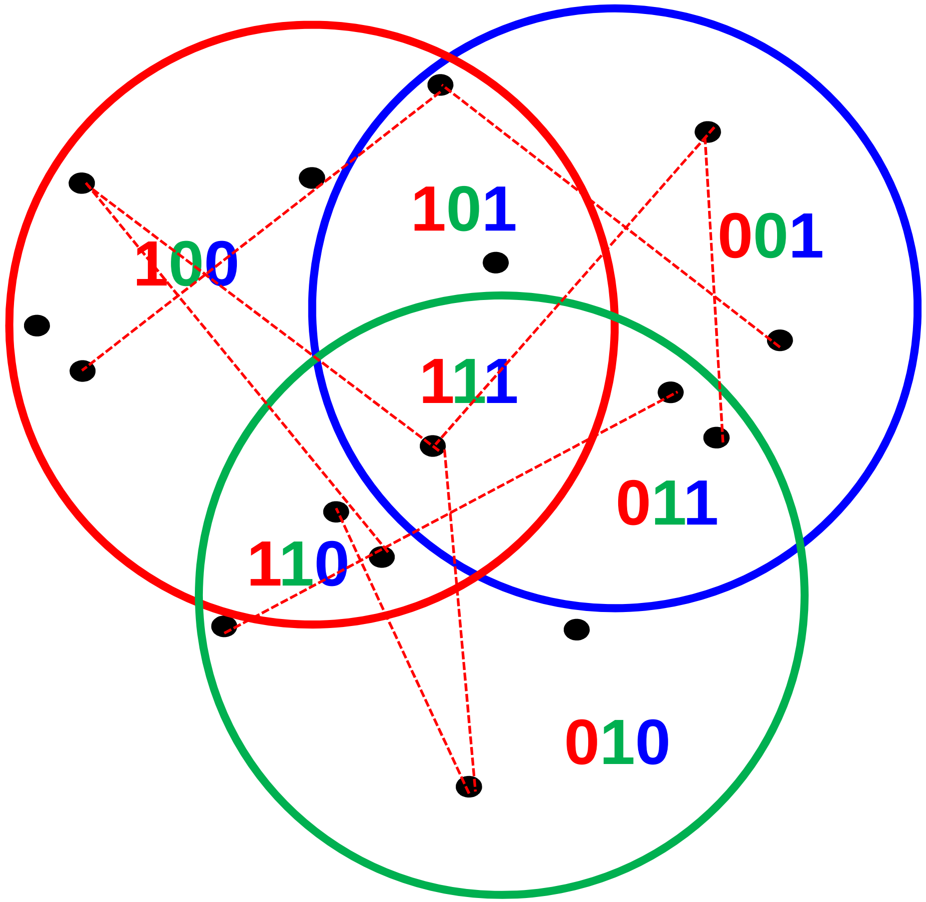


Iteratively repeat step 1, 2 until convergence.

Components of Spherical Hashing

- Spherical hashing
- Hyper-sphere setting strategy
- **Spherical Hamming distance**

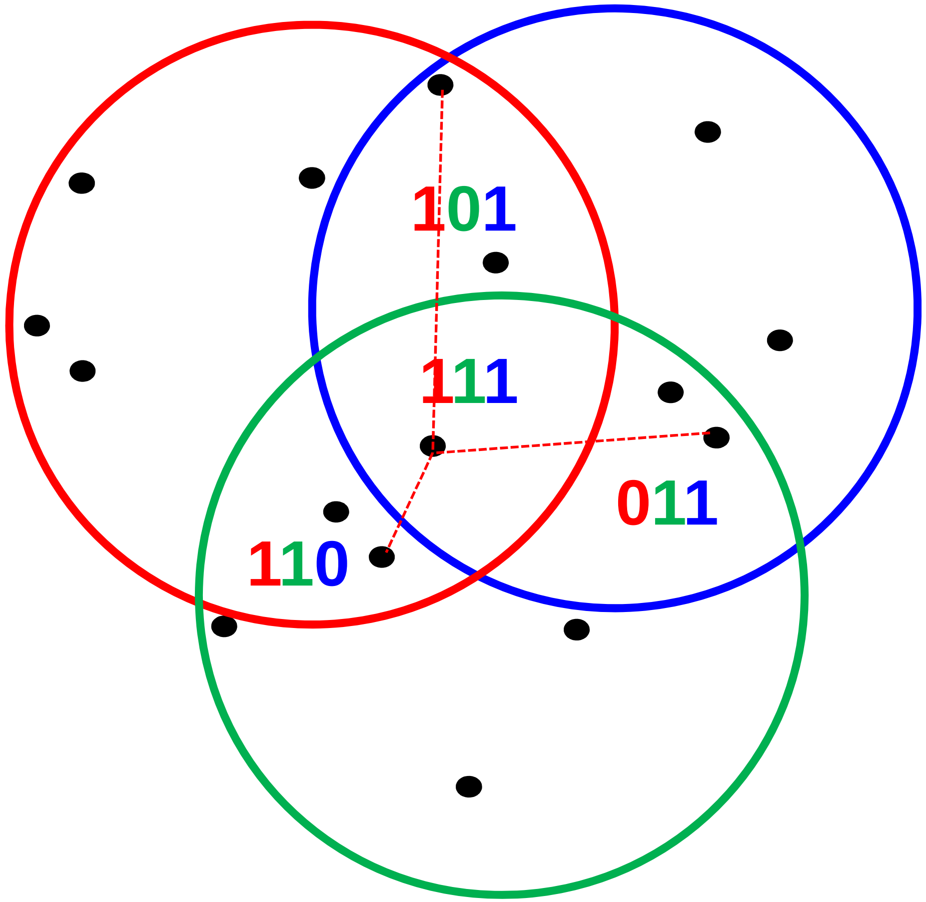
Max Distance and Common ‘1’



Common ‘1’s

: 1

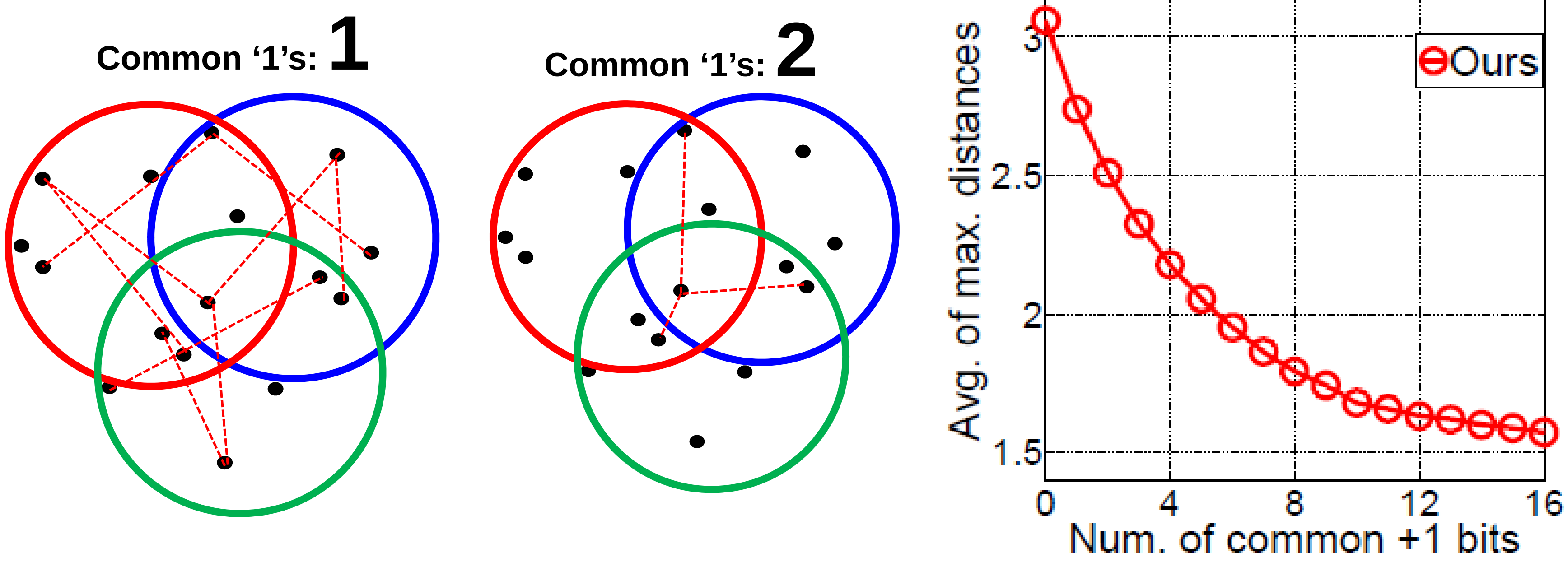
Max Distance and Common ‘1’



Common ‘1’s

: 2

Max Distance and Common '1'



Average of maximum distances between two partitions: decreases as number of common '1'

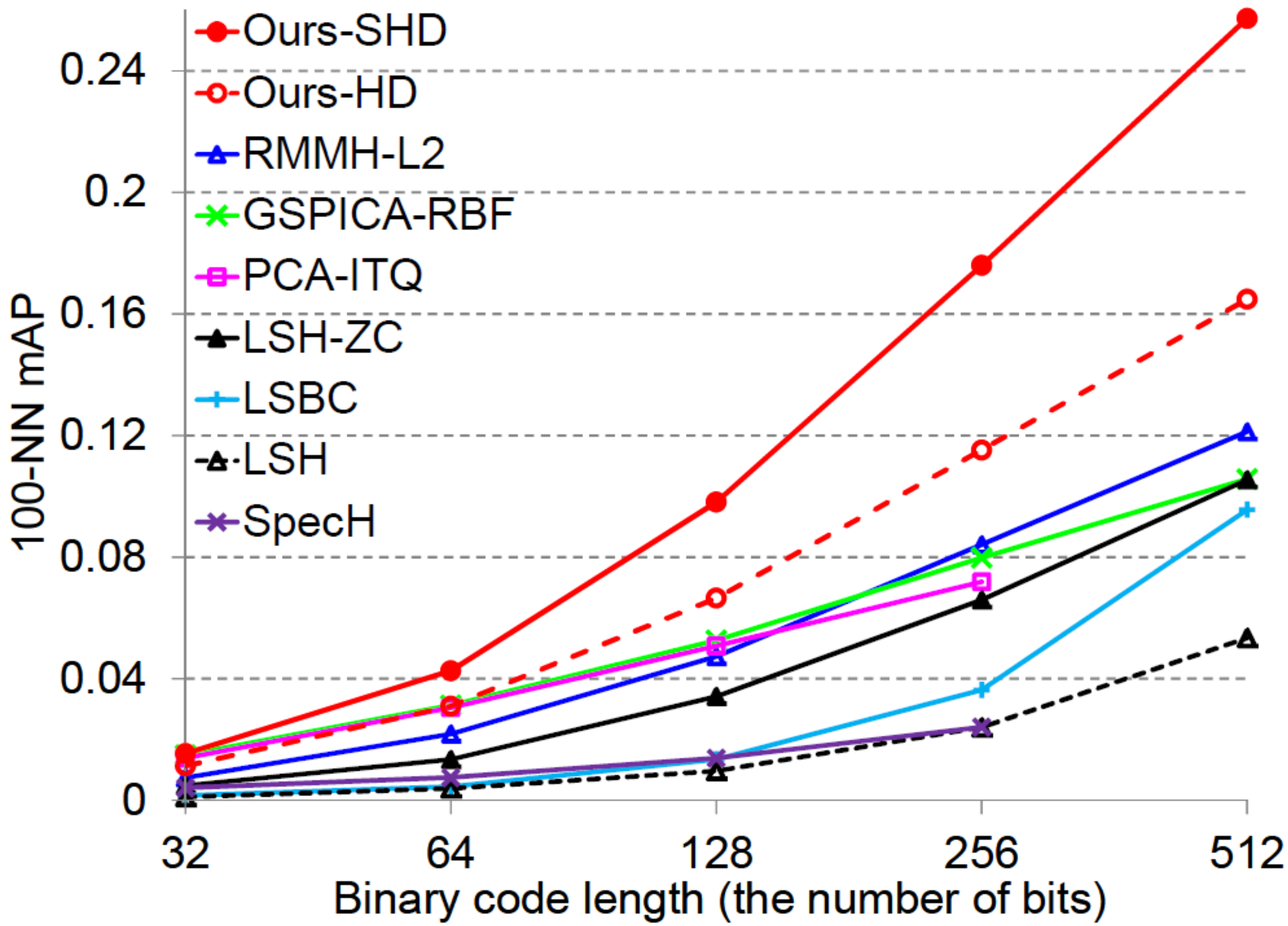
KAIST

Spherical Hamming Distance (SHD)

$$d_{shd}(b_i, b_j) = \frac{|b_i \oplus b_j|}{|b_i \wedge b_j|}$$

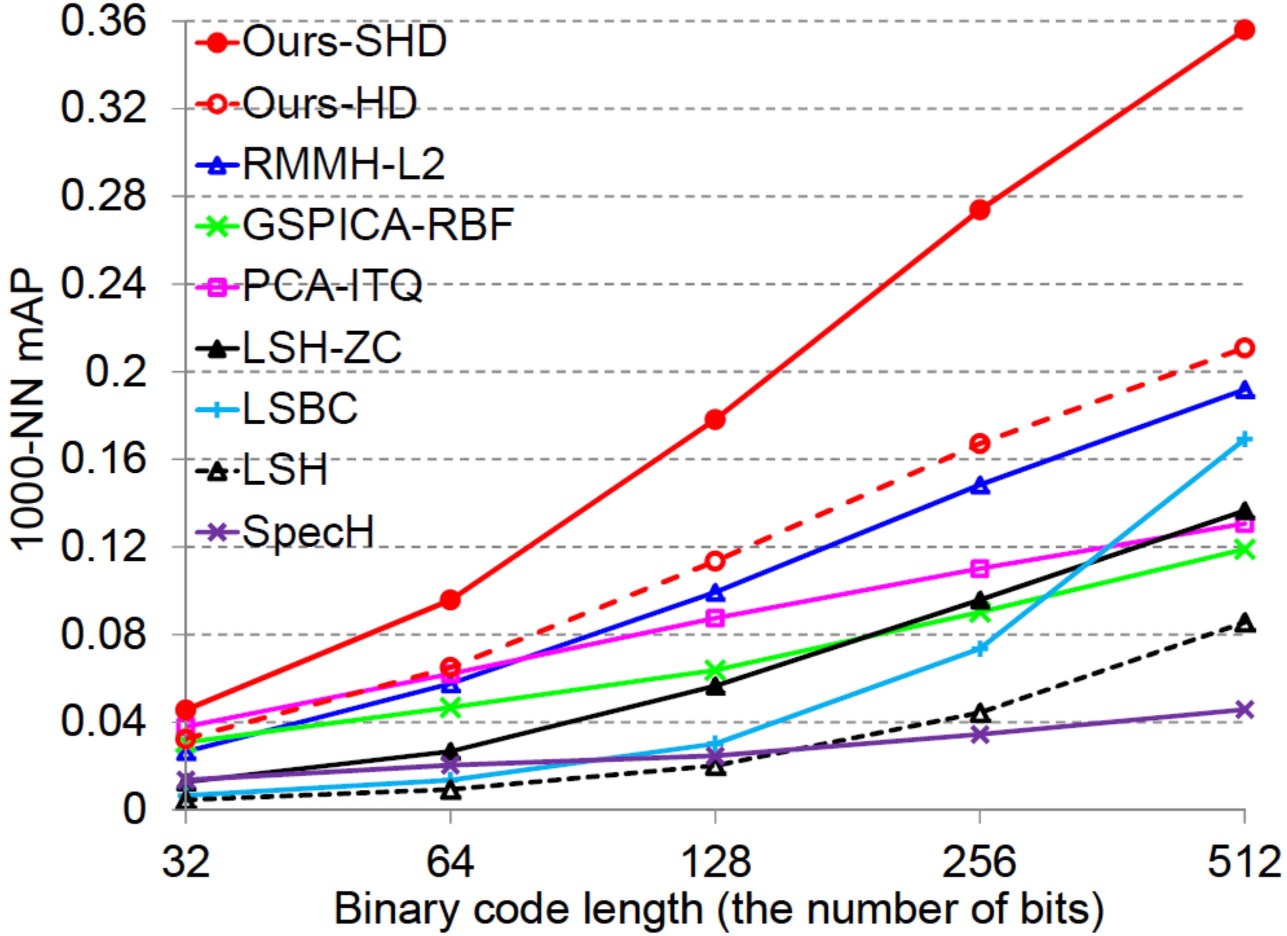
SHD: Hamming Distance divided by the number of common '1's.

Results



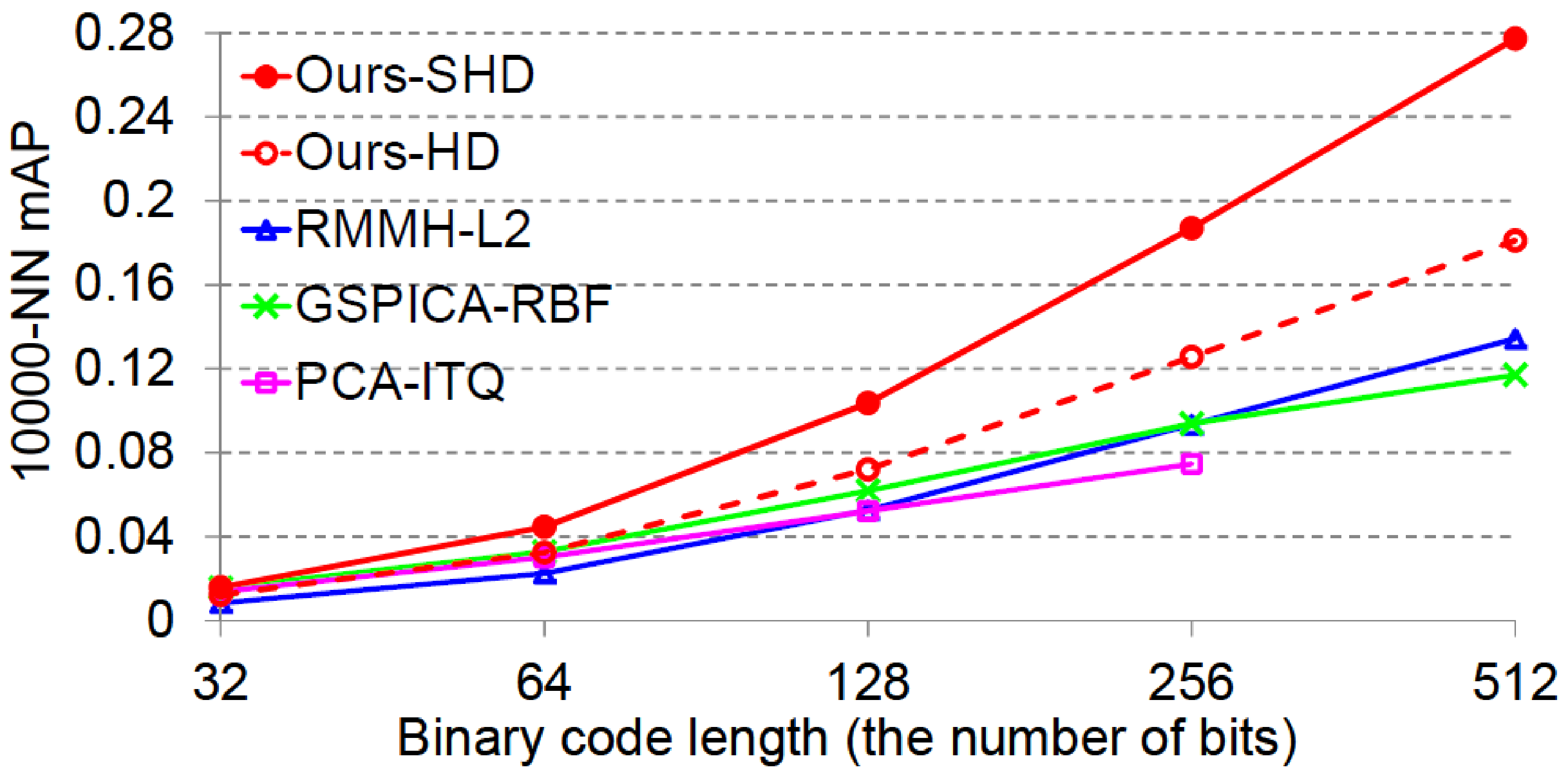
384 dimensional 1 million GIST descriptors

Results



960 dimensional 1 million GIST descriptors

Results



384 dimensional 75 million GIST descriptors

Summary

- The need of binary code embedding

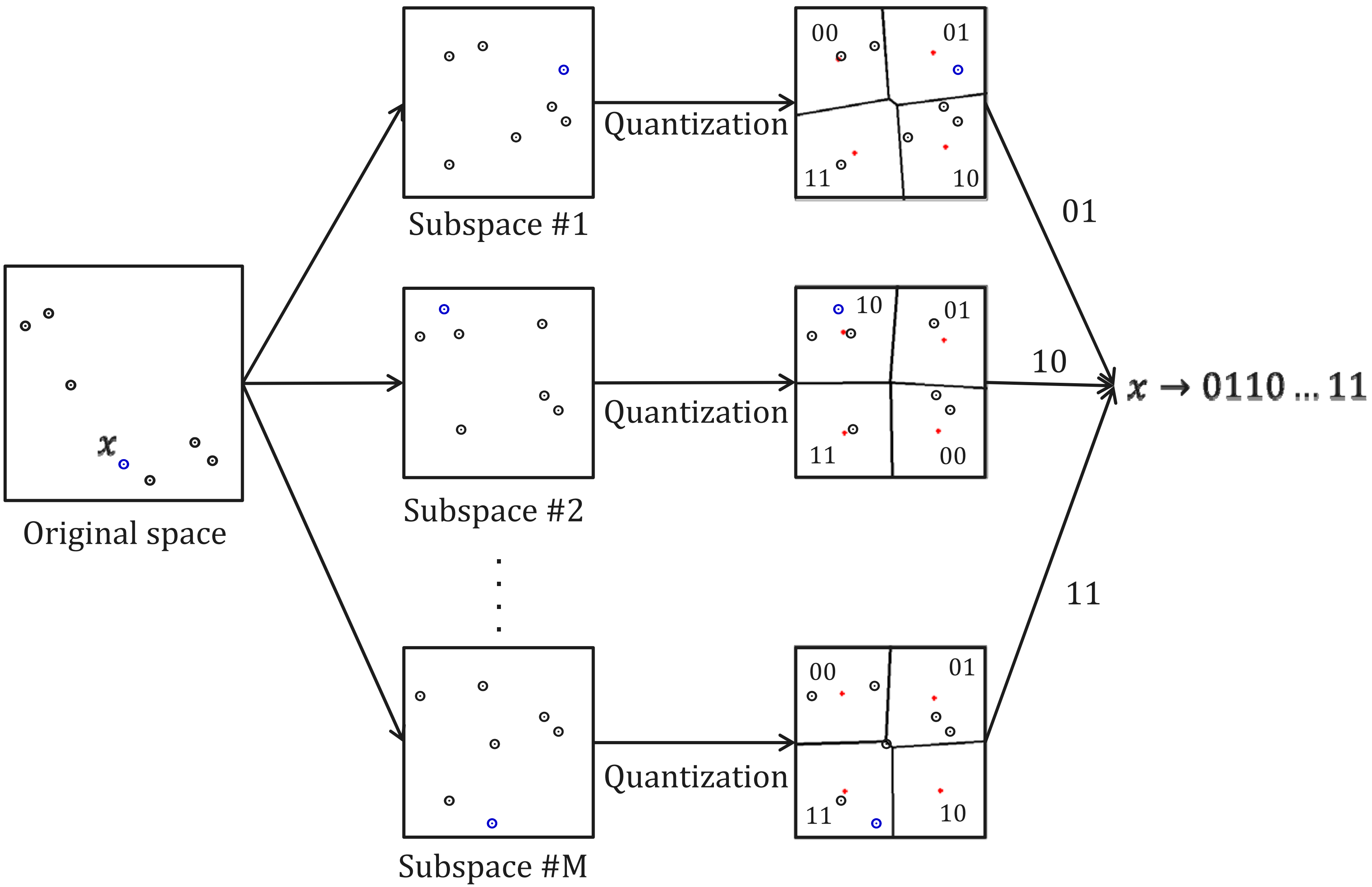
- Spherical binary code embedding
  - Uses spherical hashing for tighter bounds
  - Iterative process to achieve balance and independence
  - Spherical Hamming distance

Distance Encoded Product Quantization

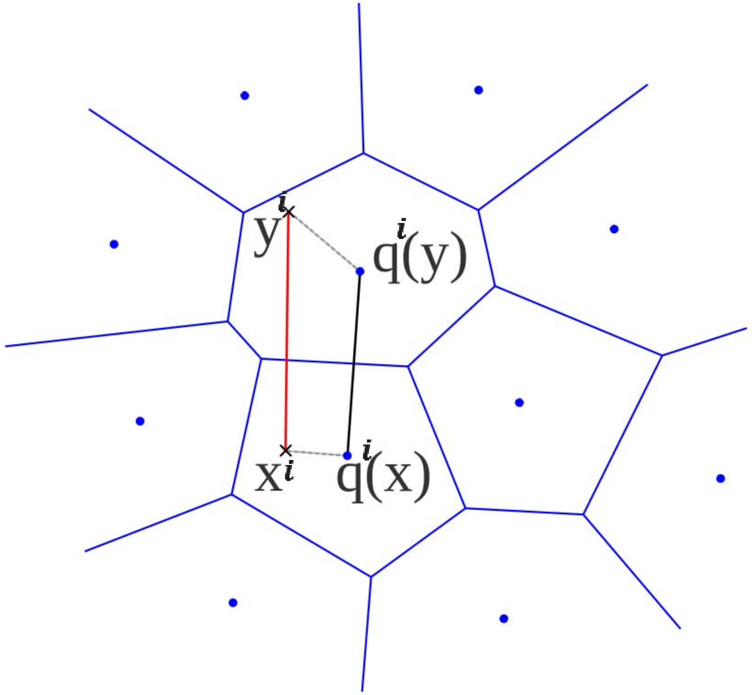
Jae-Pil Heo, Zhe Lin, and Sung-Eui Yoon

CVPR 2014

PQ: Product Quantization [Jegou et al., TPAMI 2011]

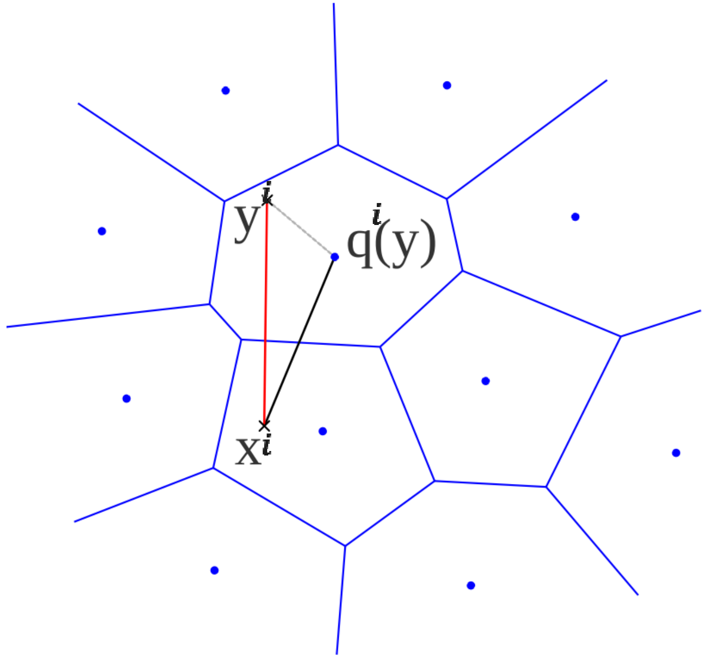


Distance Computation in PQ



Symmetric Distance

$$d_{SD}^{PQ}(x,y)^2 = \sum_{i=1}^{M} \left\| q^i(x^i) - q^i(y^i) \right\|^2$$



Asymmetric Distance

$$d_{AD}^{PQ}(x,y)^2 = \sum_{i=1}^{M} \left\| x^i - q^i(y^i) \right\|^2$$

Terms

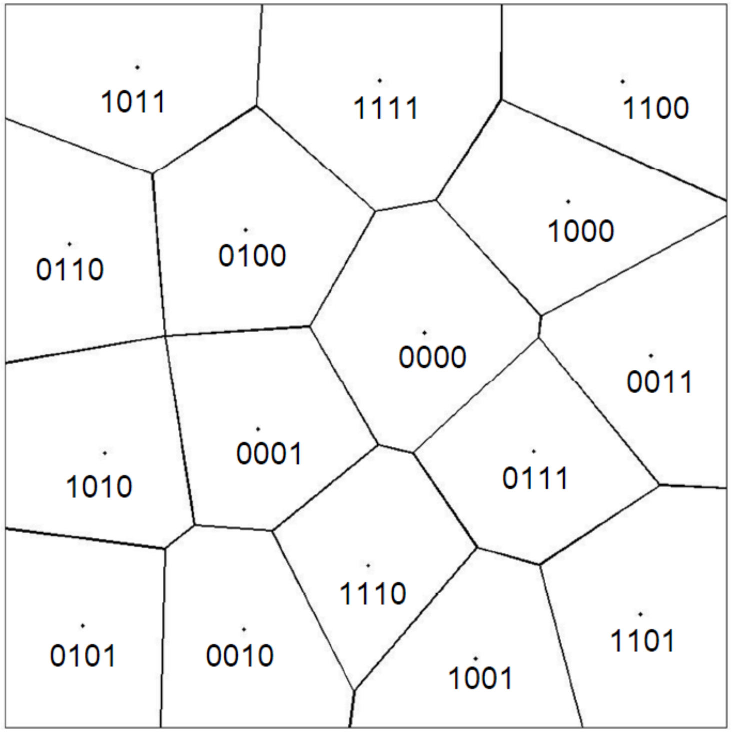
x: query, y: data, M: # of Subspaces,

q^i: quantizer in i^{th} subspace, x^i: sub-vector of x in i^{th} subspace

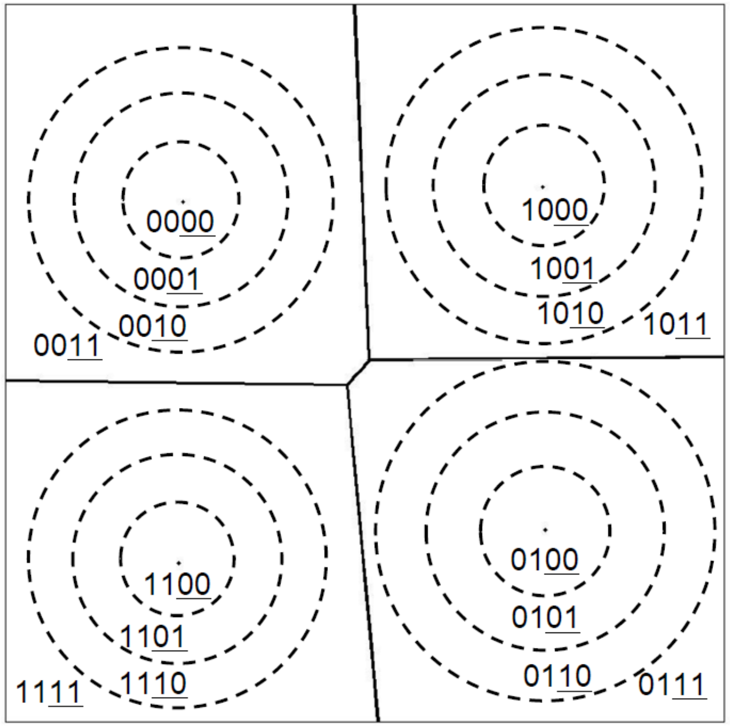
Figures are from [Jegou et al., TPAMI 2011]

DPQ: Distance Encoded PQ

- DPQ encodes quantized distance from the center as well as the cluster index in each subspace.

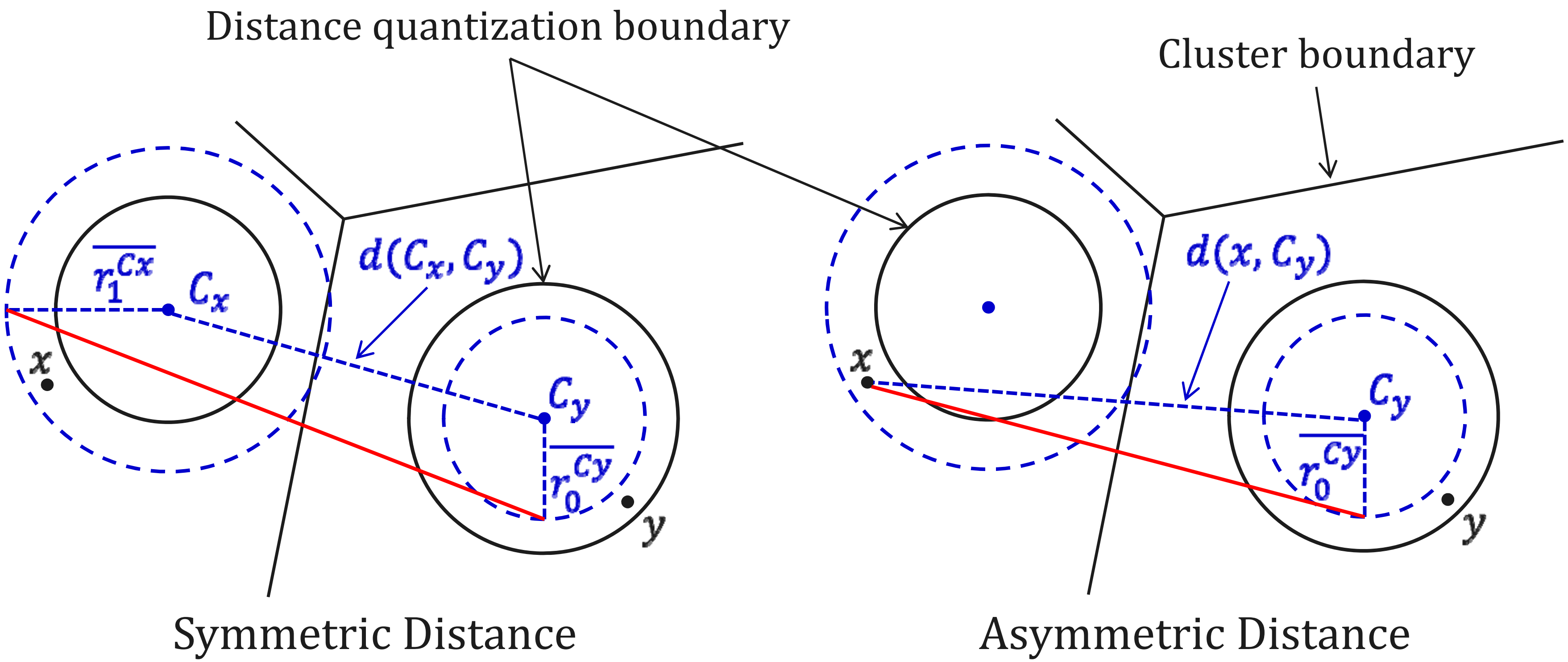


PQ example



DPQ example

Distance Computation in DPQ

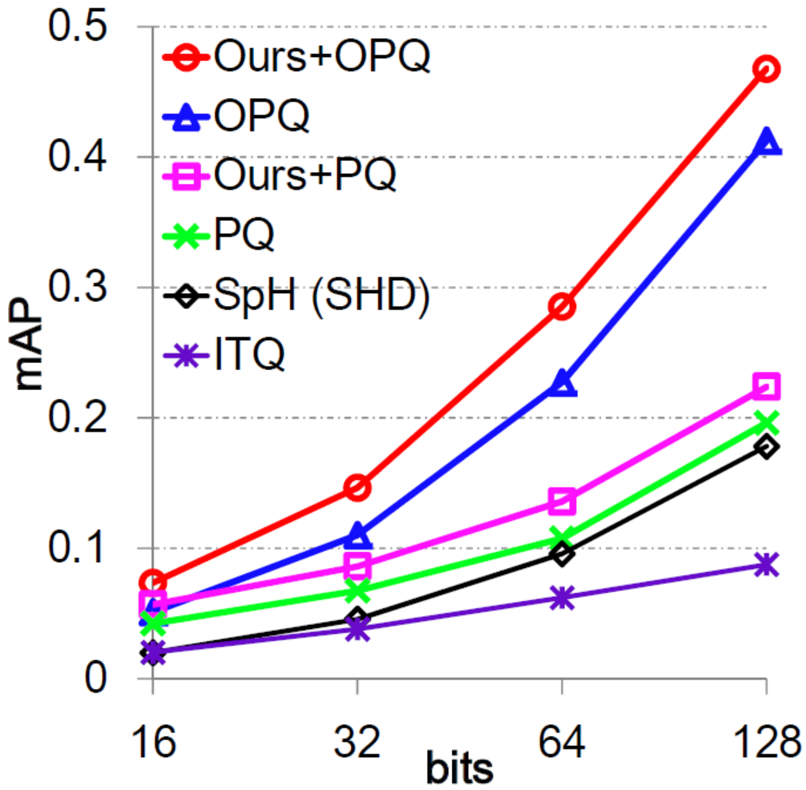


$$d_{SD}^{DPQ}(x,y)^2 = d(C_x,C_y)^2 + \overline{r_1^{Cx}}^2 + \overline{r_0^{Cy}}^2$$

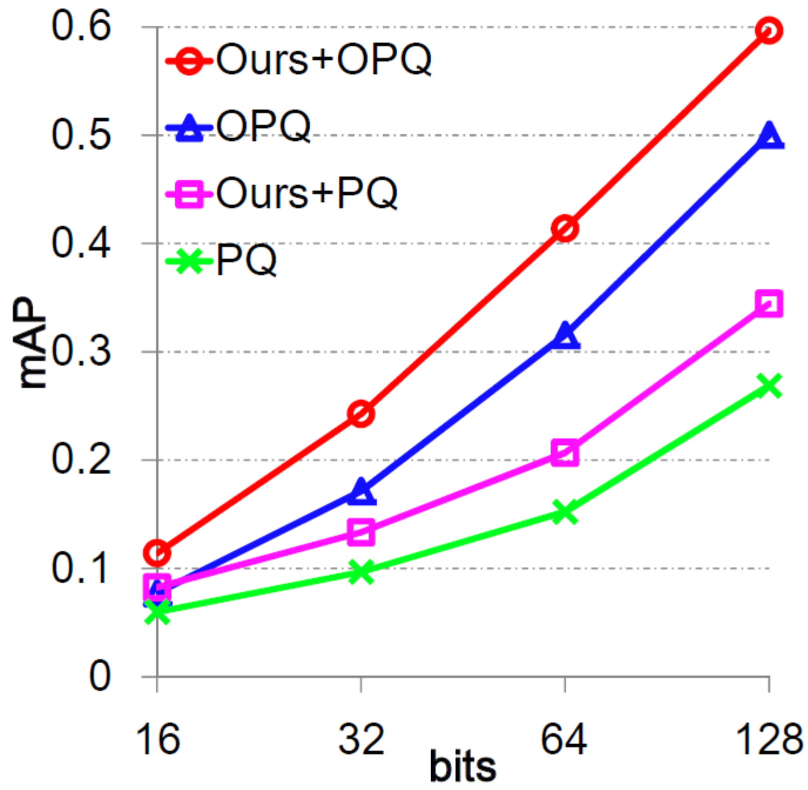
$$d_{SD}^{DPQ}(x,y)^2 = d(x,C_y)^2 + \overline{r_0^{Cy}}^2$$

$\overline{r_j^C}$: average distance from the center to points whose cluster center is C and quantized distance index is j

Results on GIST-1M-960D



Symmetric distance



Asymmetric distance

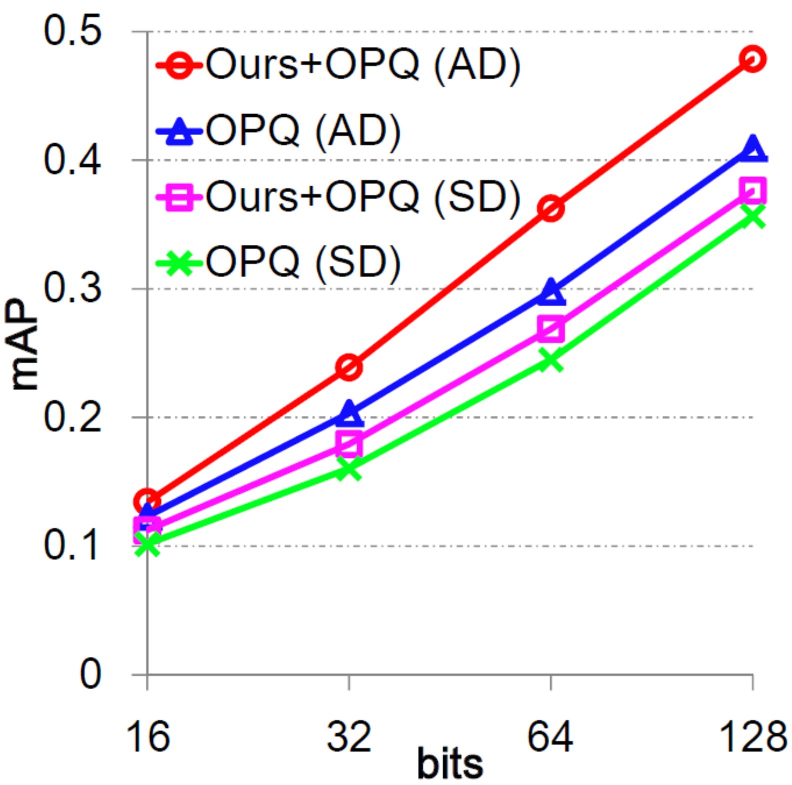
1000-nearest neighbor search mAP

OPQ: Optimized PQ [Ge et al., CVPR 2013]

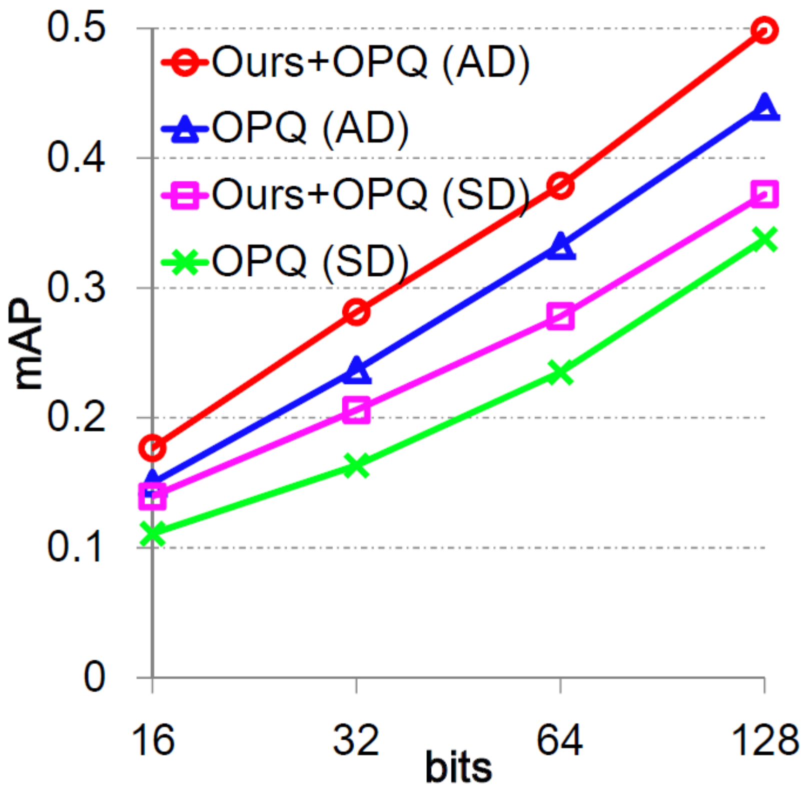
SpH: Sperical Hashing [Heo et al., CVPR 2012]

ITQ: Iterative Quantization [Gong and Lazebnik, CVPR 2011]

Results on BoW-1M-1024D



Original Data



L_2 Normalized data

1000-nearest neighbor search mAP

SD: Symmetric distance

AD: Asymmetric distance

Conclusions

- **Visual data are exploding!**
- **Image search is one of key techniques for various application including classification**
- **Processing them requires scalable algorithms**
  - **Hashing techniques for nearest neighbor search**
- **Codes are available**

http://sglab.kaist.ac.kr/software.htm