Recent Image Search Techniques: Introduction

Sung-Eui Yoon

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

KAIST

Web-Scale Visual Data and Novel Applications

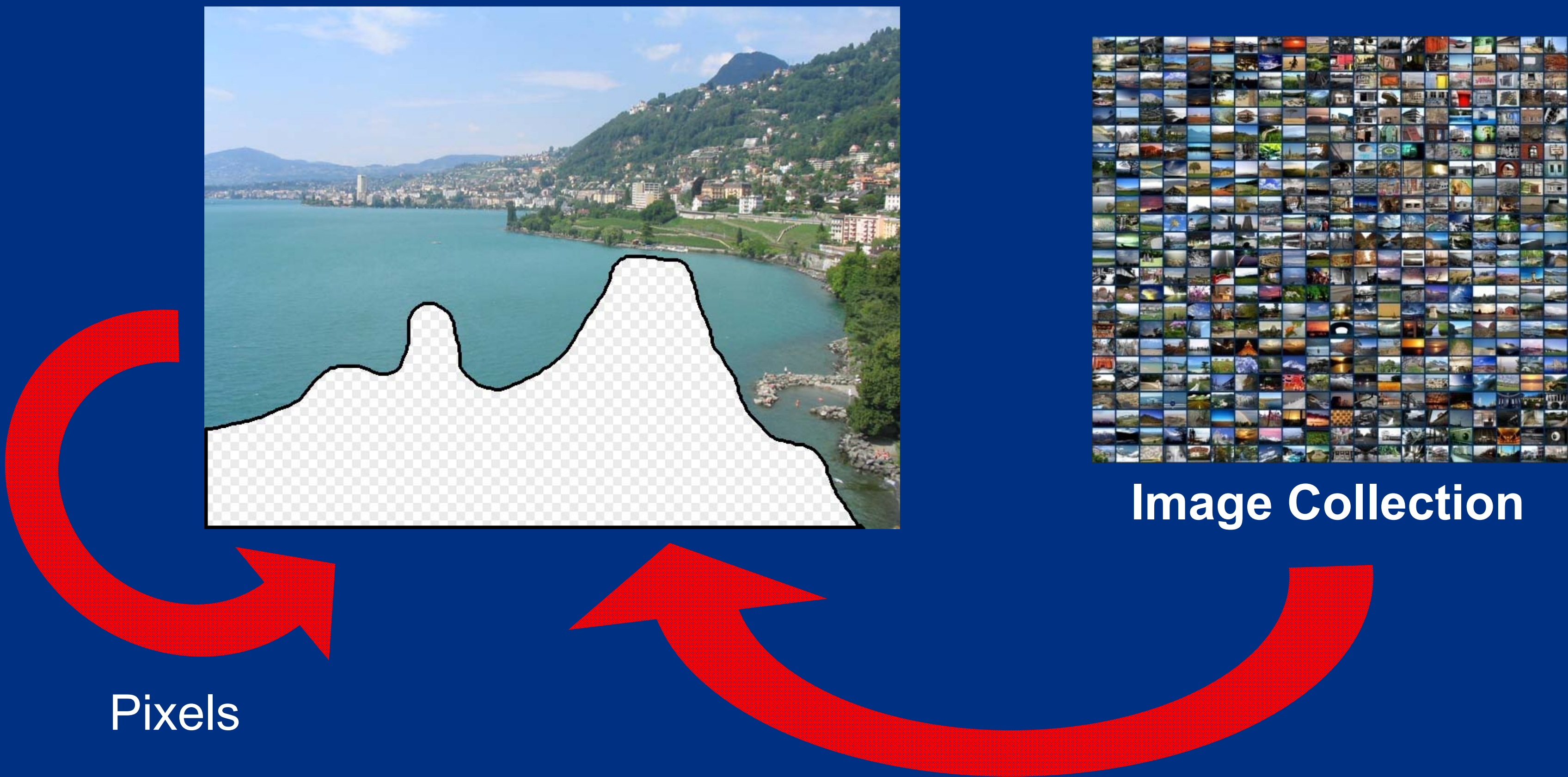
- **Visual data are widely used for various communication and, and are more widely consumed at Web and mobile devices**
  - **YouTube, Facebook, Flickr, etc.**

- **Processing them requires scalable algorithms**
- **Web-scale visual data can enable new applications (e.g., photo tourism and scene completion)**

Ack.: Hays

Photo Tourism

Scene Completion



Hays and Efros, SIGGRAPH 2007

Results

Hays and Efros, SIGGRAPH

Image Search or Content-Based Image Retrieval (CBIR)

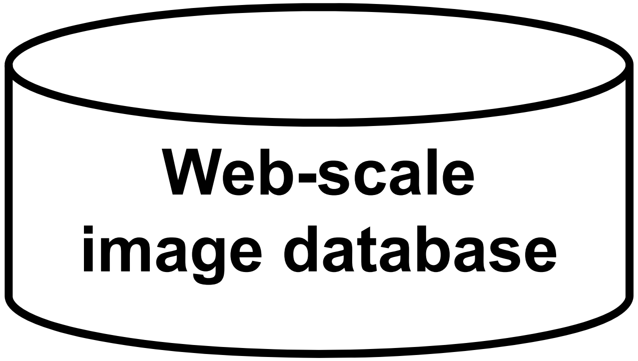
- **Identify similar images given a user-specified image or other types of inputs**

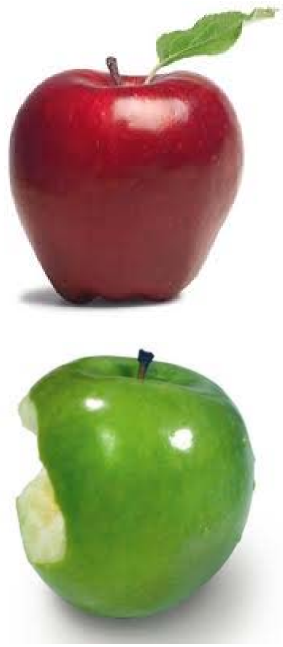
Image Search or Content-Based Image Retrieval (CBIR)

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

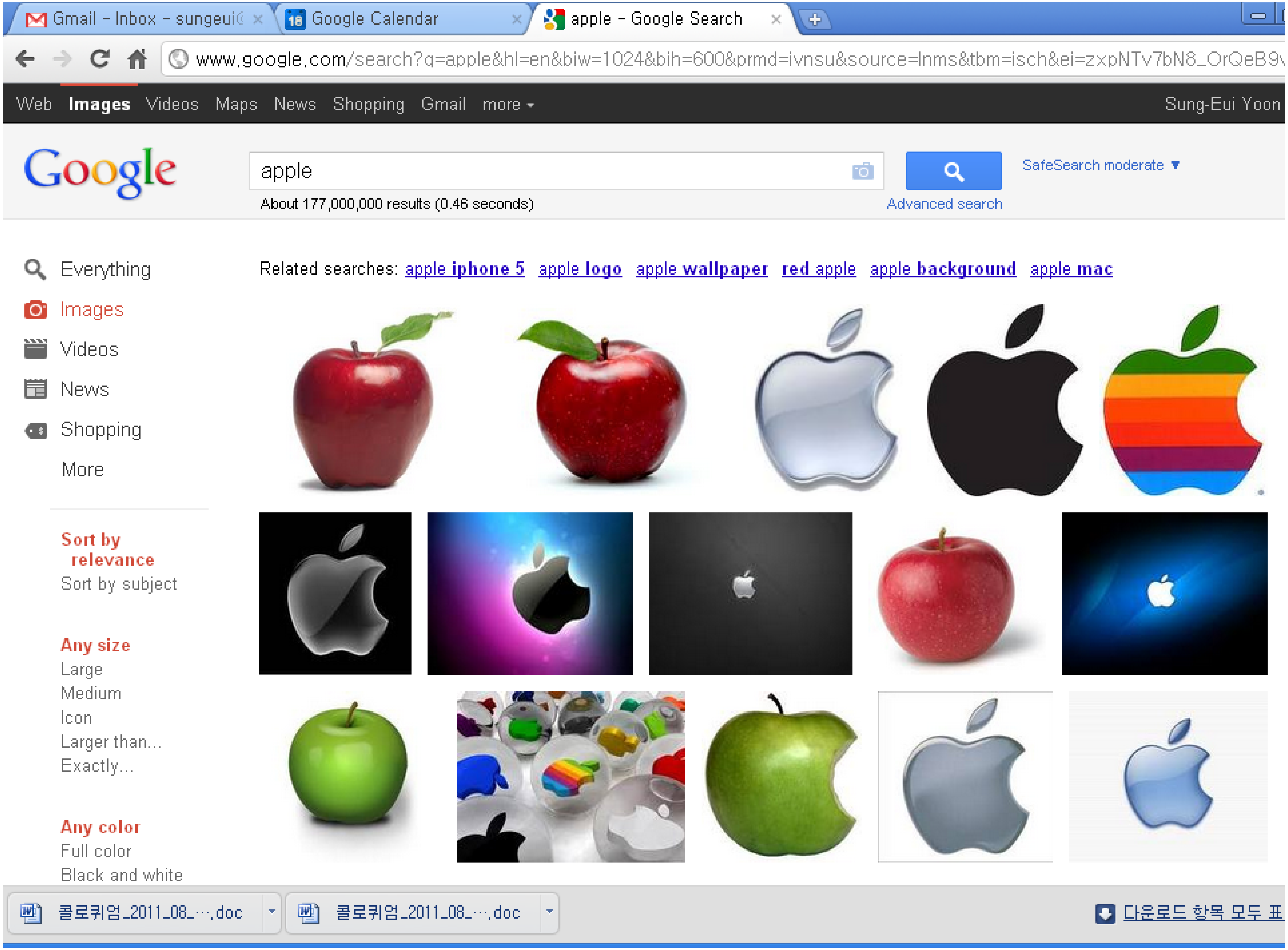
Extract image descriptors (e.g., SIFT and CNN)

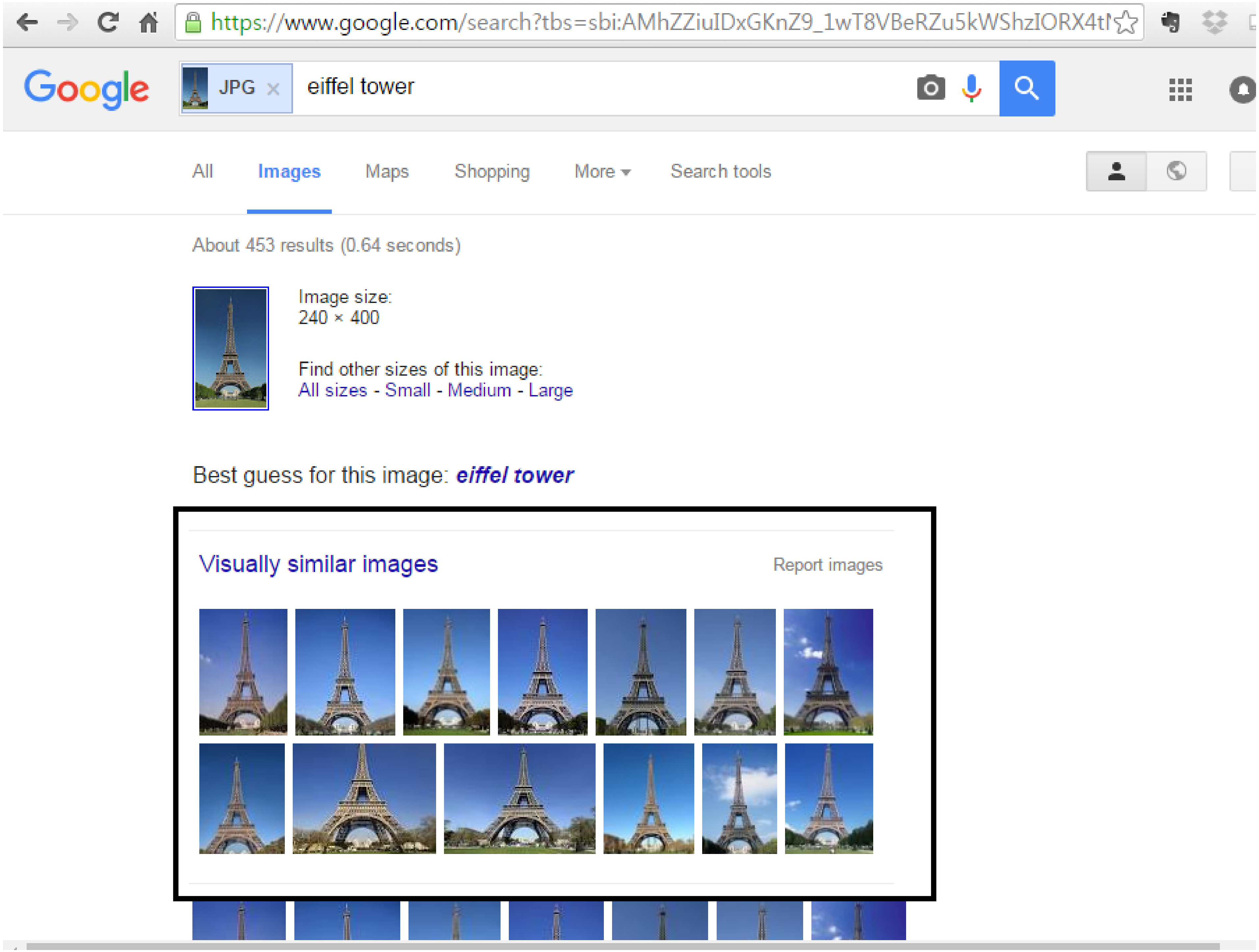
Input

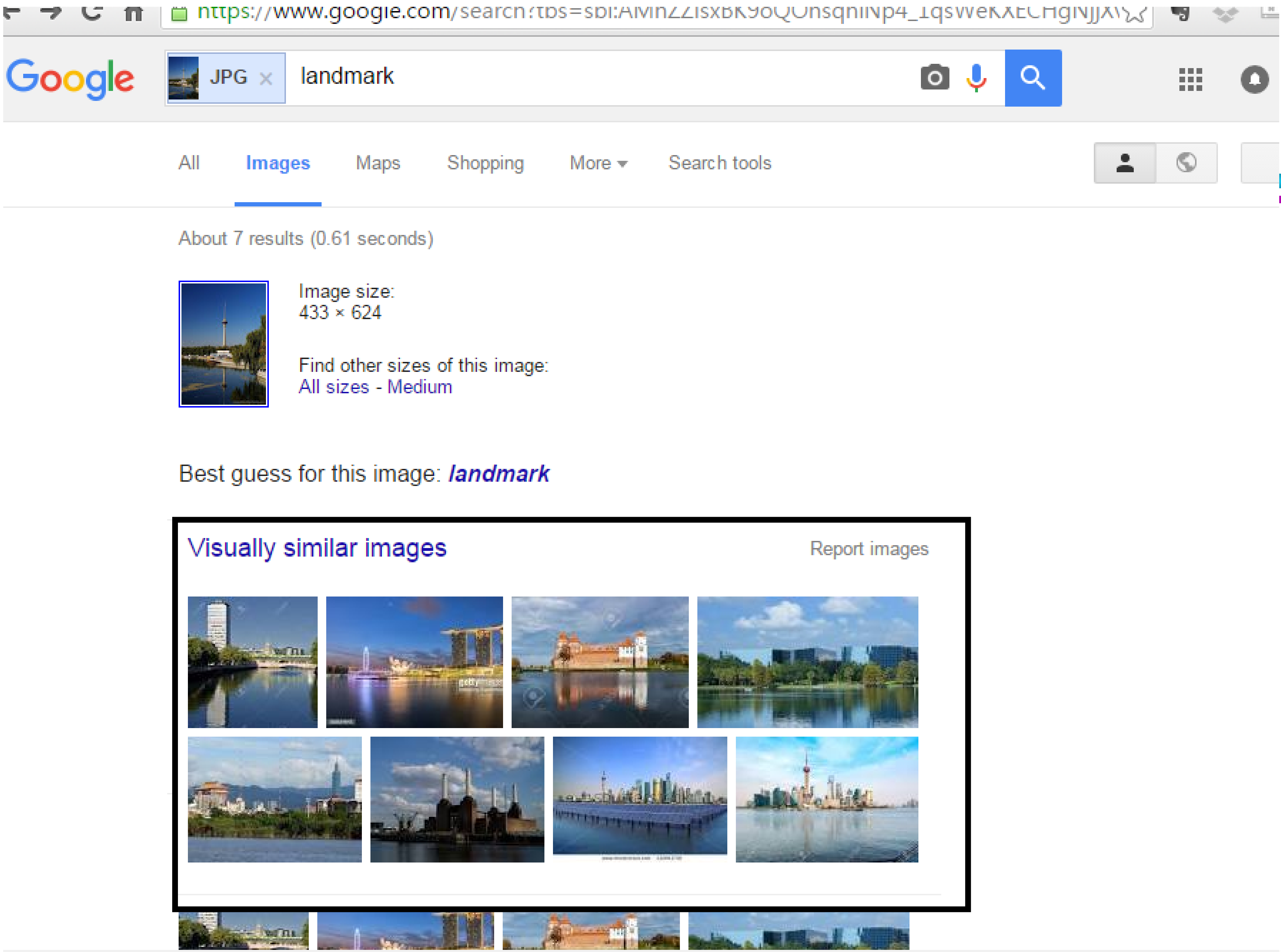




Output


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
SafeSearch moderate
About 177,000,000 results (0.46 seconds)
Advanced search
Everything
Images
Videos
News
Shopping
More
Related searches: apple iphone 5 apple logo apple wallpaper red apple apple background apple mac
Sort by relevance
Sort by subject
Any size
Large
Medium
Icon
Larger than...
Exactly...
Any color
Full color
Black and white
콜로퀴엄_2011_08_....doc
콜로퀴엄_2011_08_....doc
다운로드 항목 모두 표

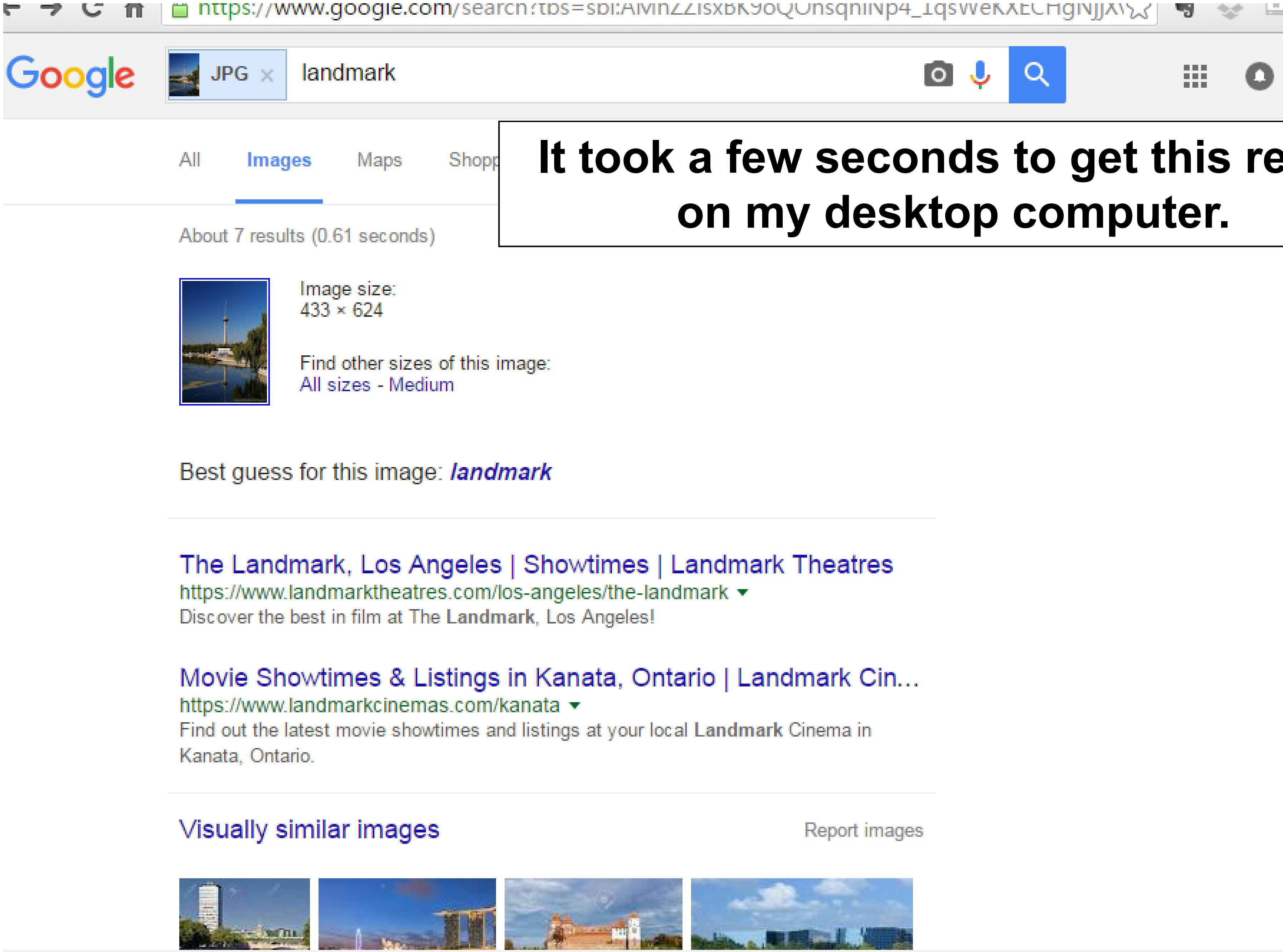

https://www.google.com/search?tbs=sbi:AMhZZiuIDxGKnZ9_1wT8VBeRZu5kWShzIORX4t
Google
JPG
eiffel tower
All
Images
Maps
Shopping
More
Search tools
About 453 results (0.64 seconds)
Image size:
240 × 400
Find other sizes of this image:
All sizes - Small - Medium - Large
Best guess for this image: eiffel tower
Visually similar images
Report images



Google
JPG
landmark
All
Images
Maps
Shopping
More
Search tools
About 7 results (0.61 seconds)
Image size:
433 × 624
Find other sizes of this image:
All sizes - Medium
Best guess for this image: landmark
Visually similar images
Report images


Gmail - Inbox - sungeui
Google Calendar
Google Search
www.google.com/search?tbs=sbi:AMhZZiuyKQN8VbdVKrlWLNXXzF8NwQ3U6nZjHoGicvPiP_1LltJ1Ts9EF9aMliLZ52wdbl9cGW
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 ابن 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

Gmail - Inbox - sungeui
Google Calendar
Google Search
www.google.com/search?tbs=sbi:AMhZZiuyKQN8VbdVKrlWLNXXzF8NwQ3U6nZjHoGicvPiP_1LltJ1Ts9EF9aMliLZ52wdbl9cGW
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
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
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.

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

Applications

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

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

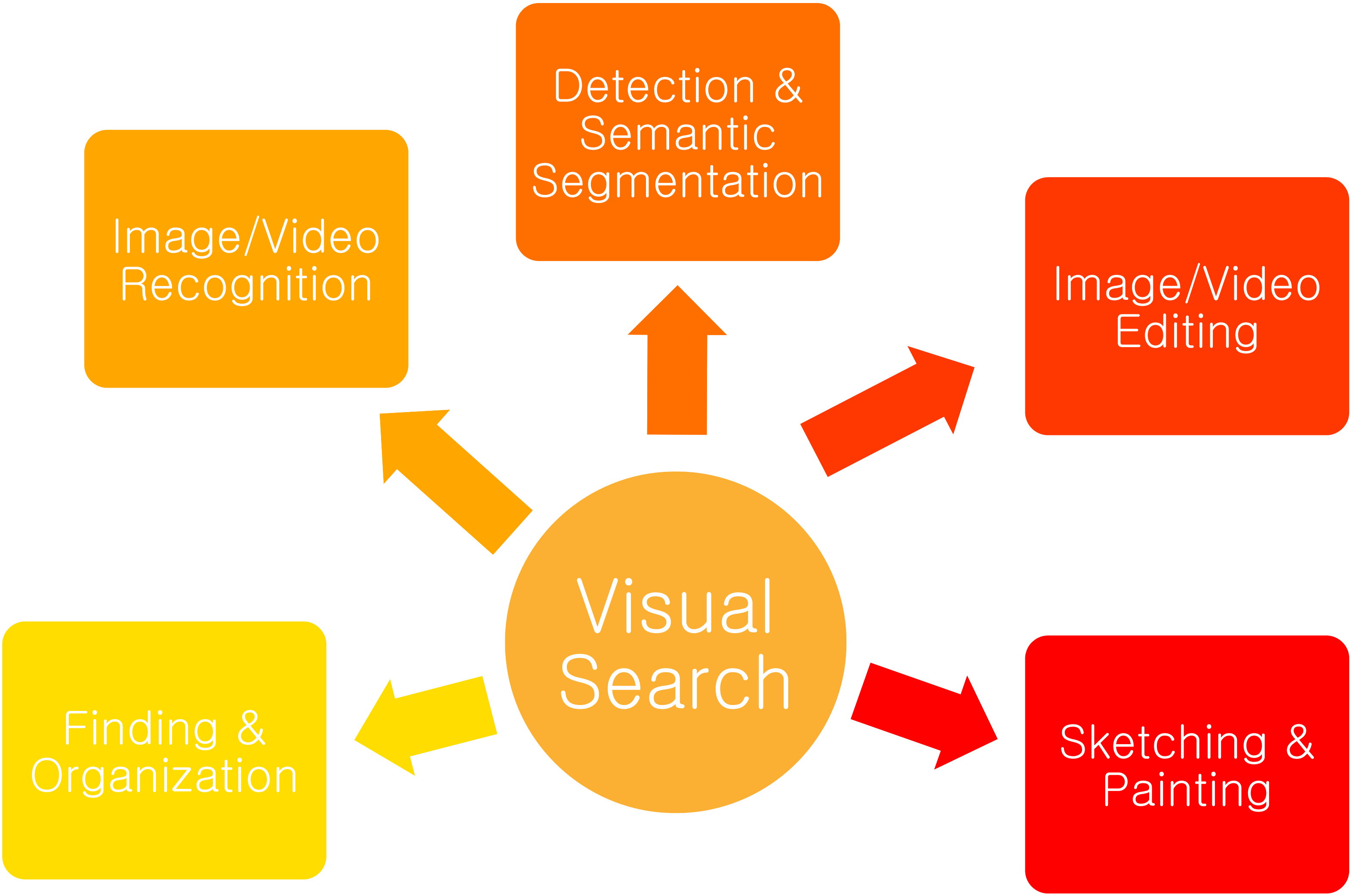
Object Detection

Landmark or Location Detection

query **City-scale image DB**

Chen et al.

Possible Application Domains



Issues of Web-Scale Image Search

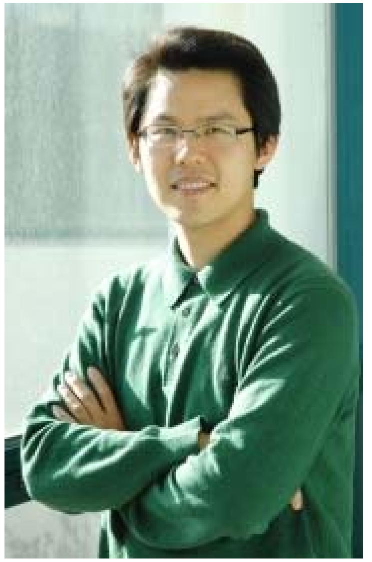
- Accuracy issues
- Memory issues
- Handling dynamic databases of images
- Novel applications?

About the Tutorial

- **We focus on the following things:**
  - **Broad understanding on image search**
  - **In-depth study on recent large-scale image search**

Tutorial Schedule

- 2:00pm, basic materials, S. Yoon
  - Introduction on image search
  - Compact representations of images



- 3:30pm, 15min break

- 3:45pm, recent techniques, Z. Lin
  - Indexing scheme for large-scale image search
  - Applications

- 5:15pm, the end

Other Related Tutorials

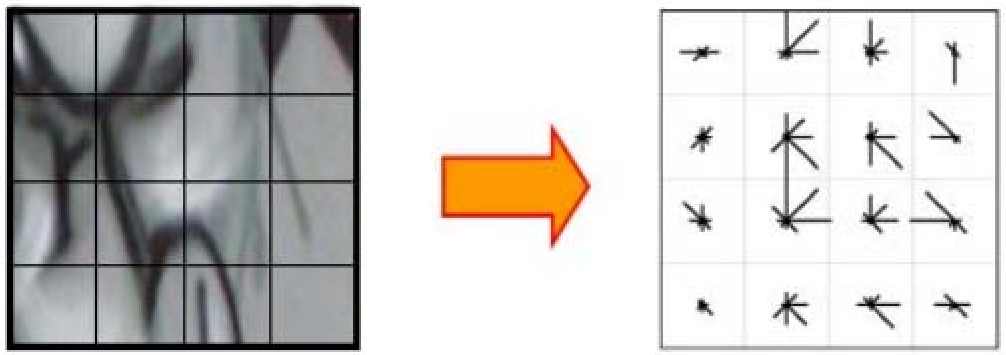
- Compact Features for Visual Search
  - Today, morning
  - Rongrong Ji, Wei Liu, Yue Gao

Key Components of Image Search

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

Image Representations

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



David G. Lowe. "Distinctive image features from scale-invariant keypoints." *IJCV* 60 (2), pp. 91-110, 2004.

Image Retrieval

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

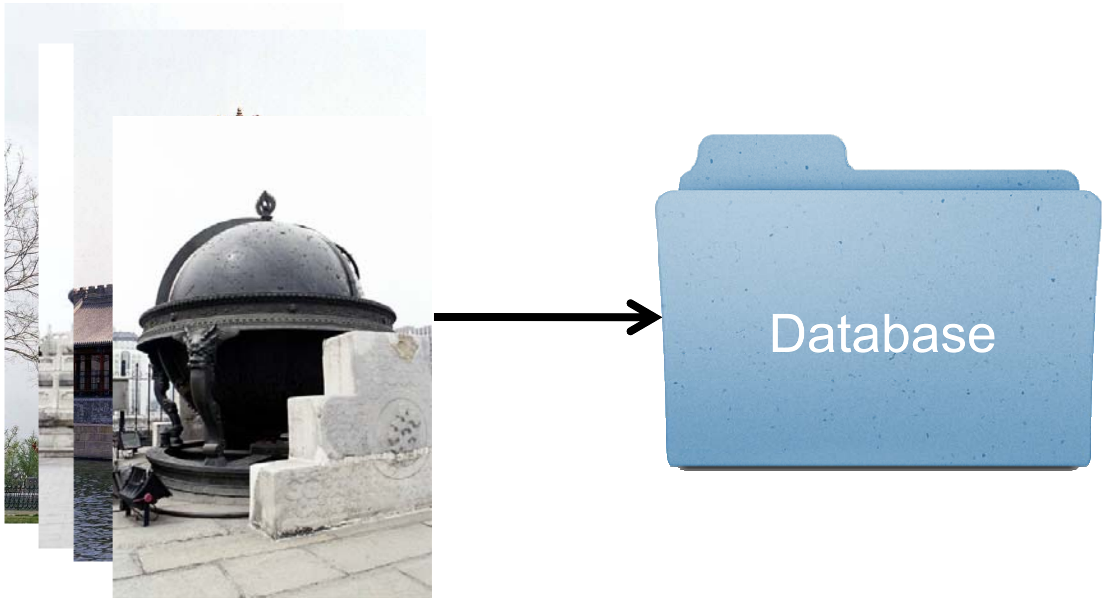
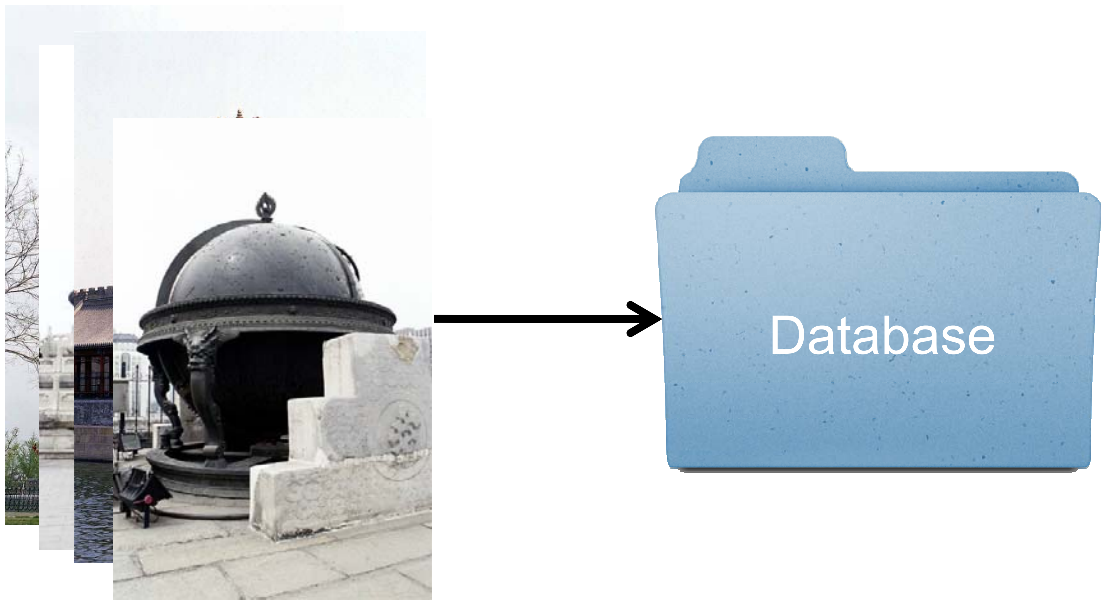


Image Retrieval

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



Index schemes:
vocabulary trees,
hashing, and
inverted files

Image Retrieval: Runtime Procedure

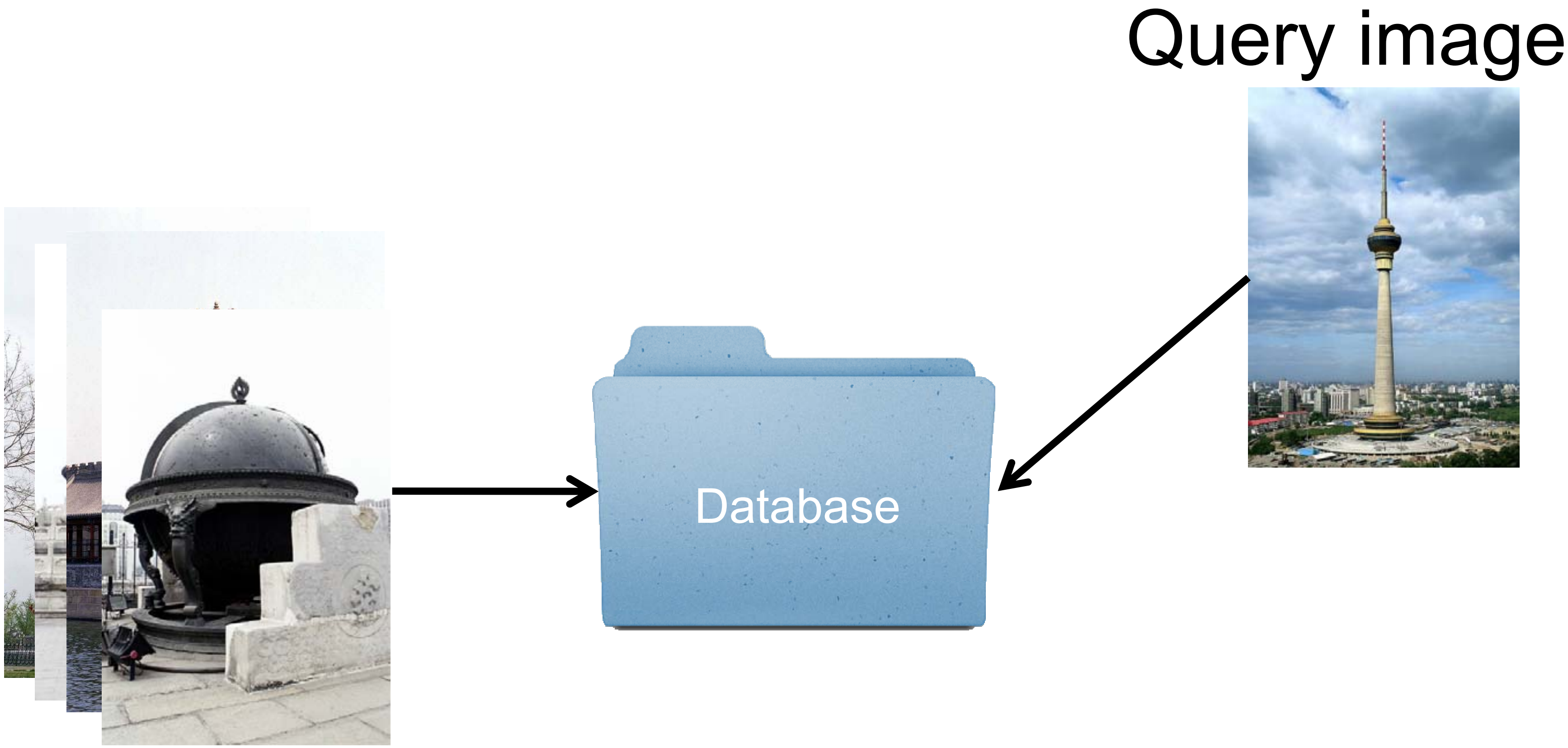
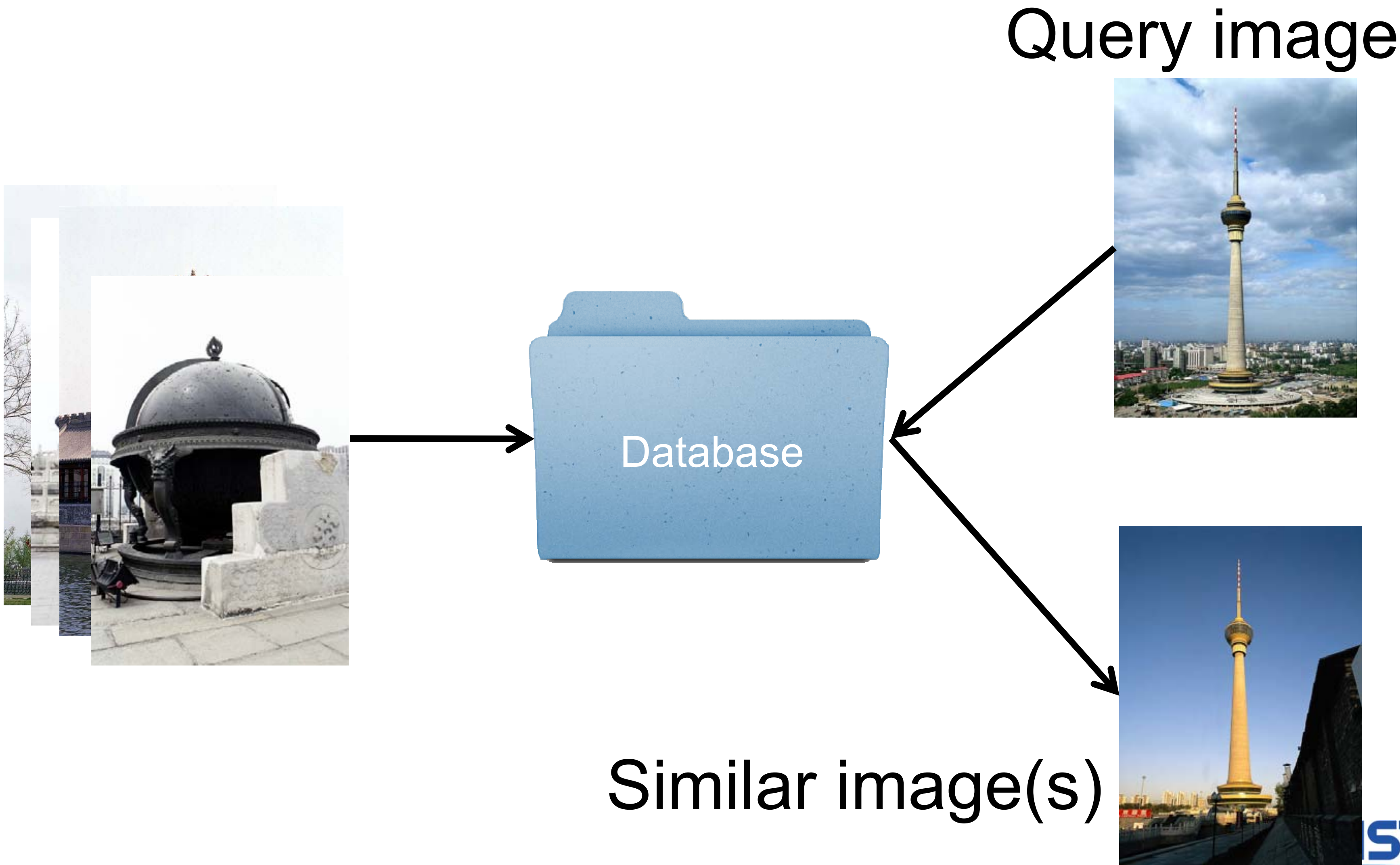


Image Retrieval: Runtime Procedure



Post-Processing

Image Retrieval with Spatially Constrained Similarity Measure

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

Classification through Image Search

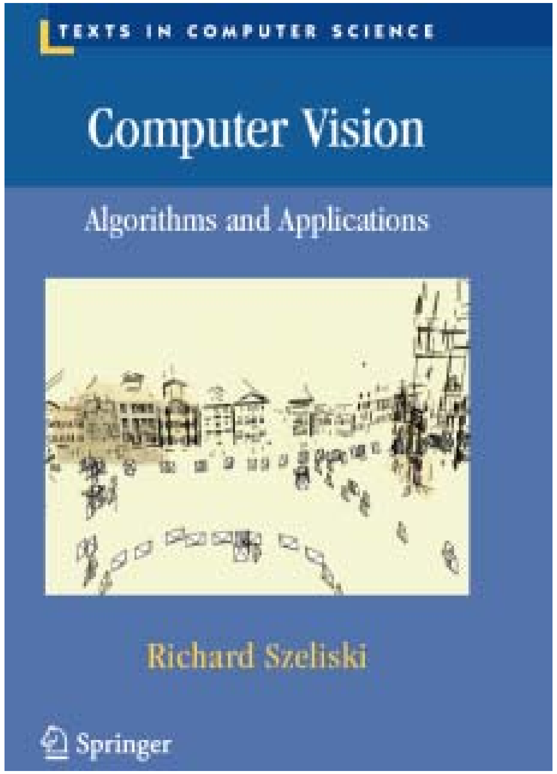
- **Image search**
  - **Find images that have smaller distances to the query**

- **Classification**
  - **Find classes that have smaller distances to the query**
  - **Utilize labels**

- **Classification using image search**
  - **Naïve Bayes Nearest Neighbor (NBNN) [Boiman et al., 08]**
  - **Image classification and Retrieval are ONE [Xie et al., 15]**

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**
  - **http://sglab.kaist.ac.kr/~sungeui/IR/**
  - **My own ongoing write-up**

KAIST

Recent Image Search Techniques: Bag-of-Words (BoW) and Inverted File

Sung-Eui Yoon

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

KAIST

Review and Outline

- **Introduction on image search**
  - **Bag-of-visual-Word (BoW) model**
  - **Inverted index**
- **Compact representations of images**

Bag-of-visual-Words (BoW)

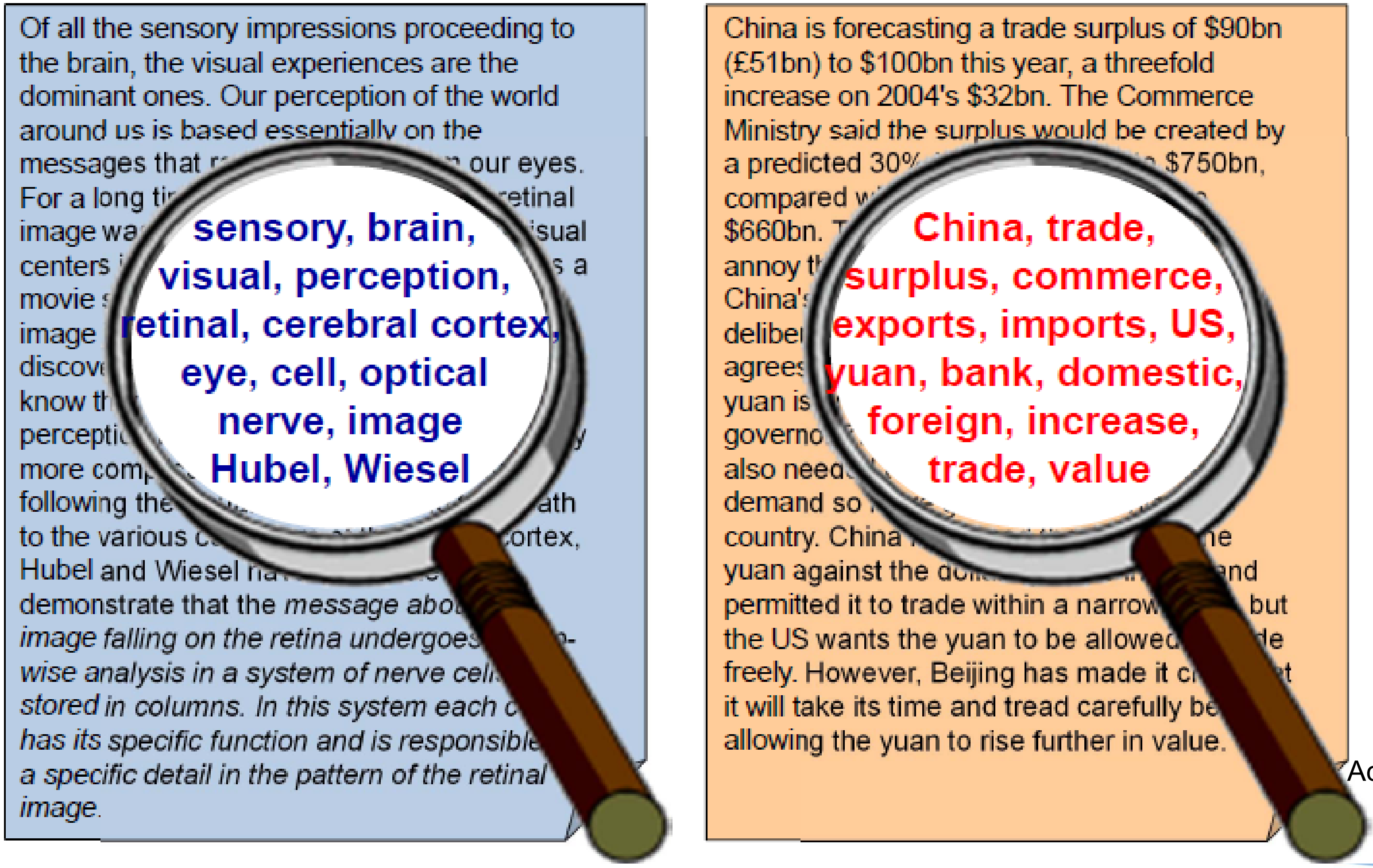
- **Inspired by text based search**

Object → Bag of 'words'

Ack: Fei-Fie Li

Bag-of-visual-Words (BoW)

- **Inspired by text based search**

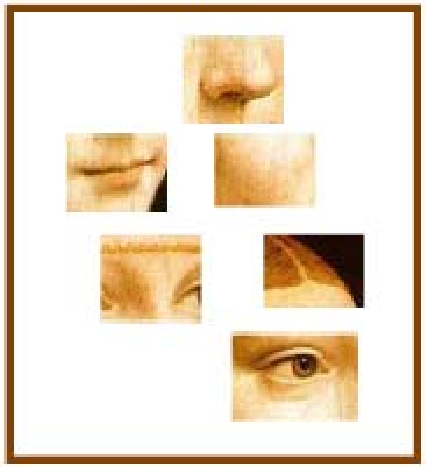


Ack: Fei-Fie Li

definition of “BoW”

– Independent features

face



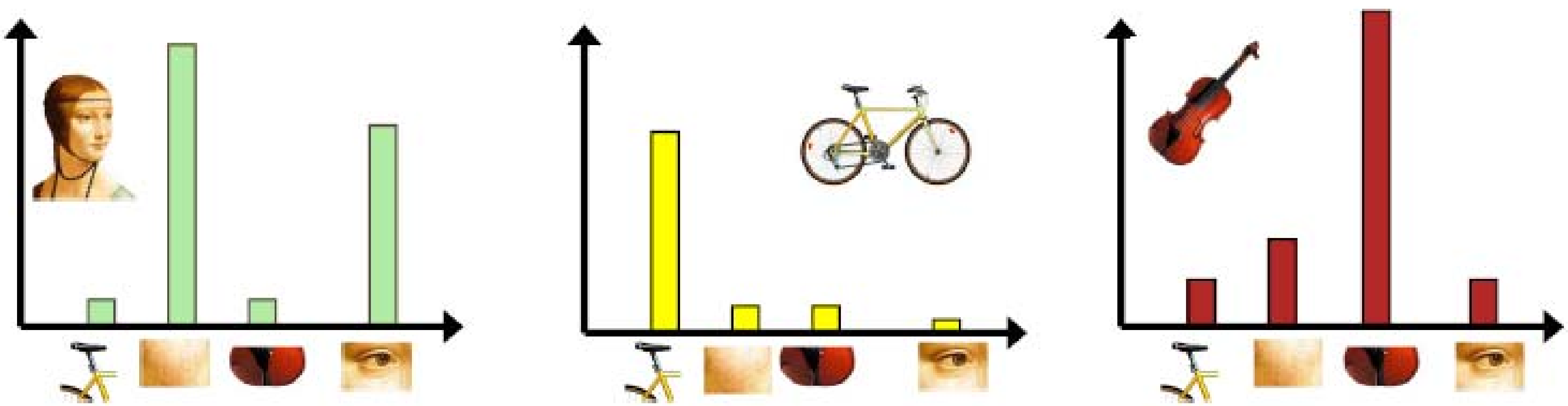
bike

violin

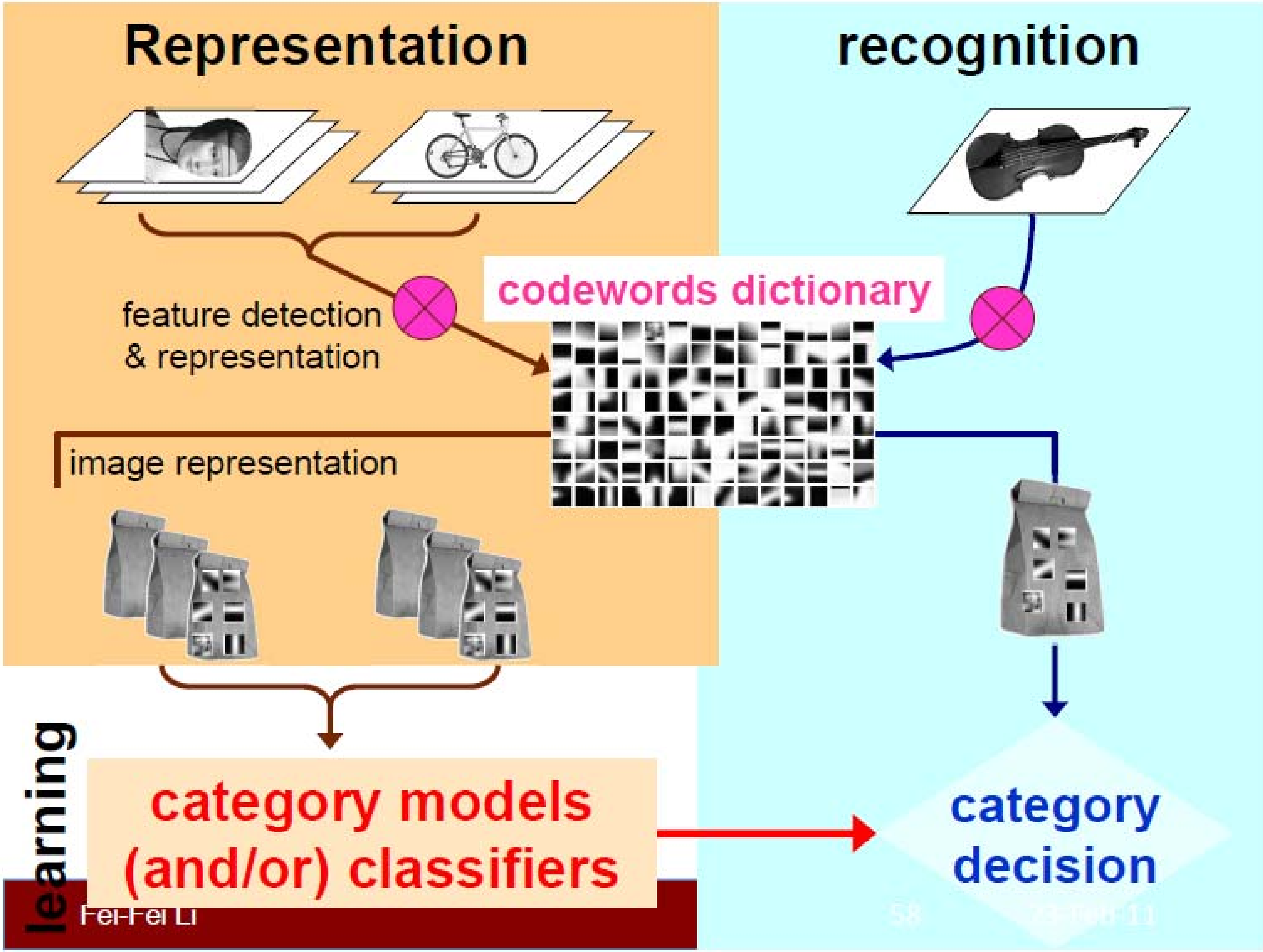
Ack: Fei-Fie Li

definition of "BoW"

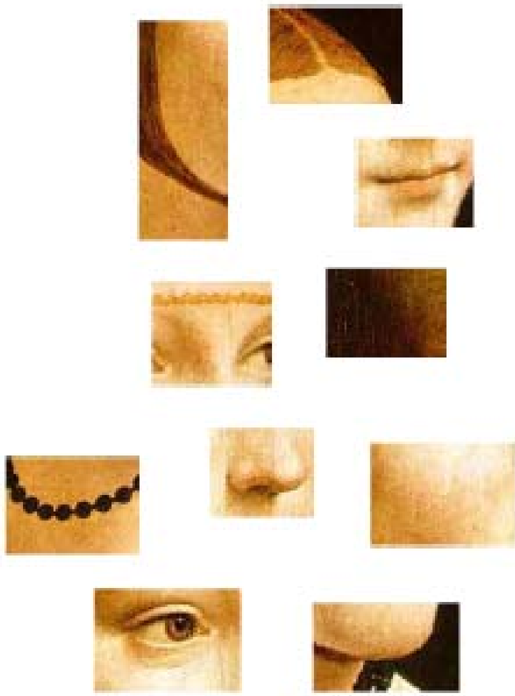
– Independent features
– histogram representation



codewords dictionary

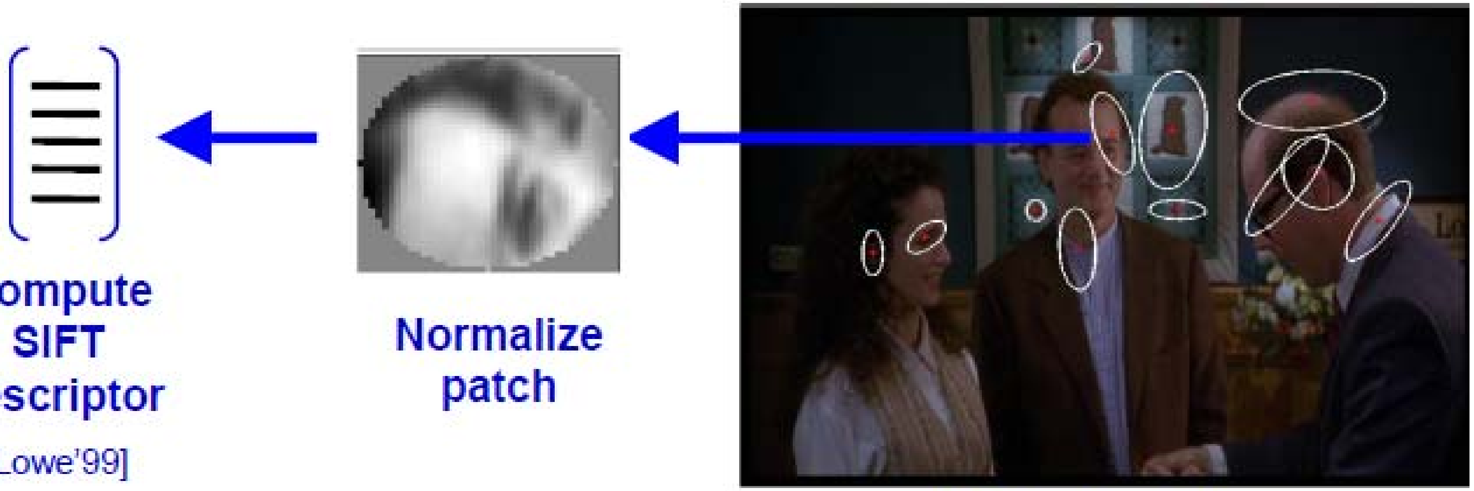

Representation
recognition
feature detection
& representation
codewords dictionary
image representation
learning
category models
(and/or) classifiers
category
decision
Fei-Fei Li
58

1.Feature detection and representation



Ack: Fei-Fie Li

1.Feature detection and representation



Compute SIFT descriptor

[Lowe'99]

Normalize patch

Detect patches

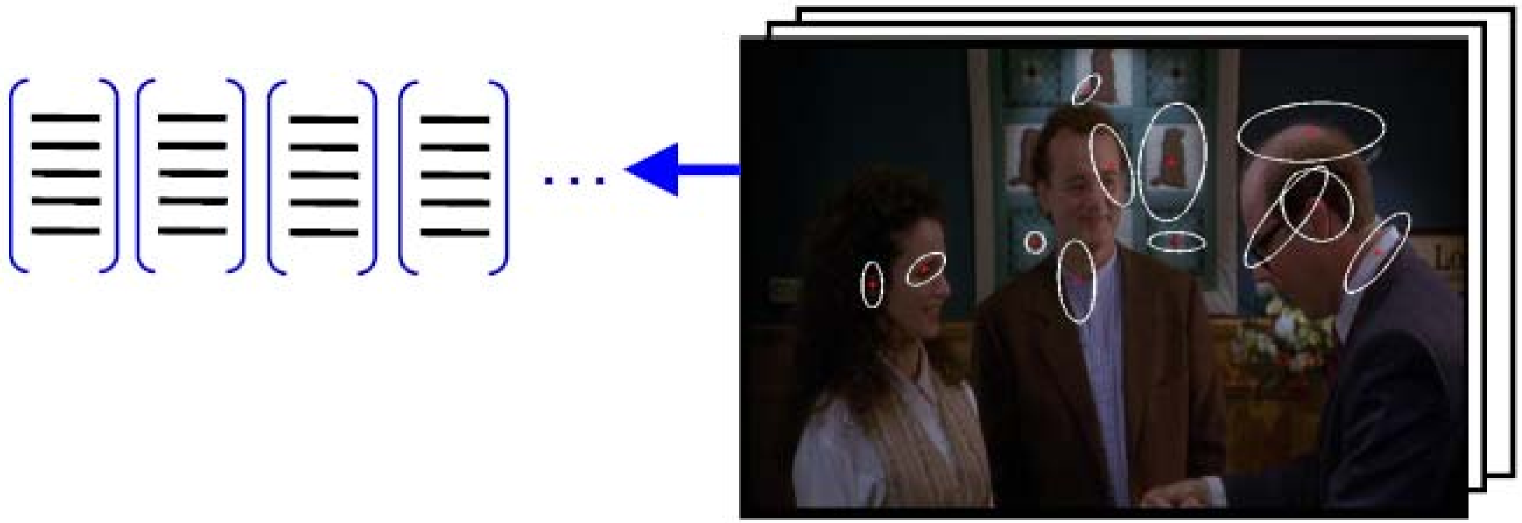
[Mikojaczyk and Schmid '02]

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

[Sivic & Zisserman, '03]

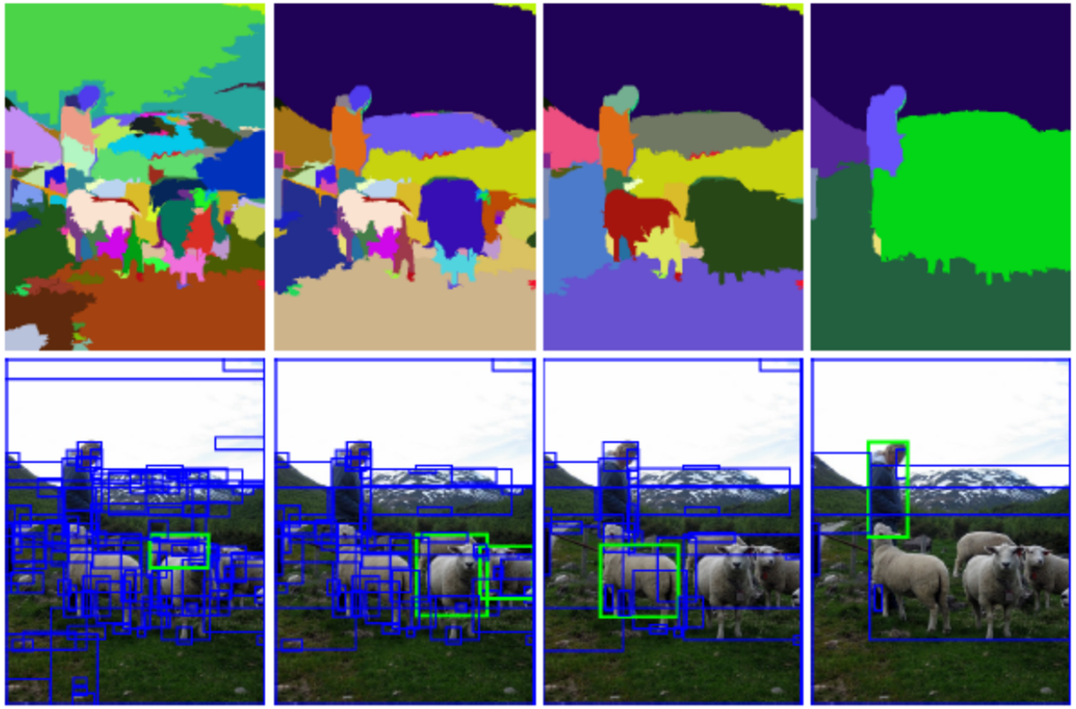
Slide credit: Josef Sivic

1.Feature detection and representation



Region Proposals

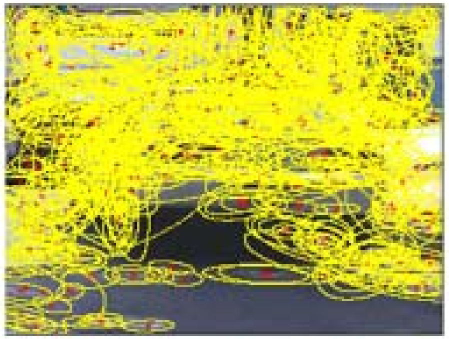
- **Adopted commonly by many recognition approaches**



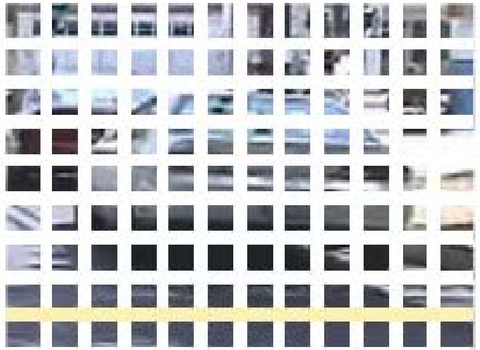
Identify different regions as candidates of objects
Selective Search, Uijlings et al.

Representation

- Building blocks: Sampling strategies

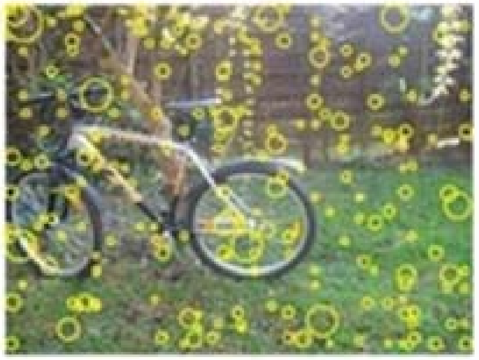


Interest operators



Dense, uniformly

Multiple interest operators

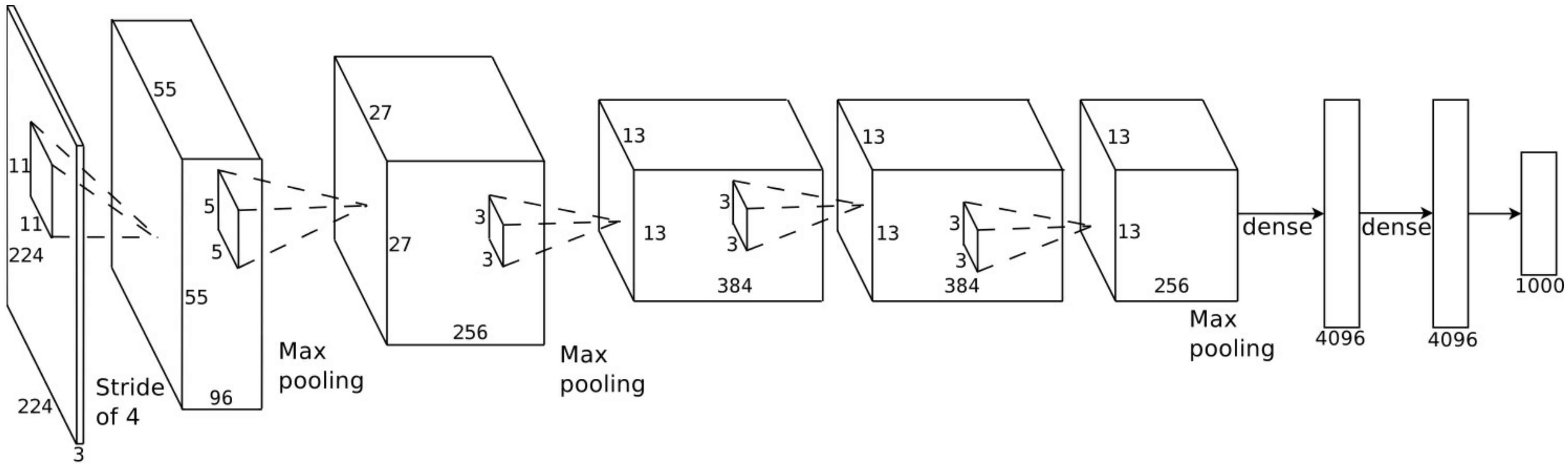


Randomly

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

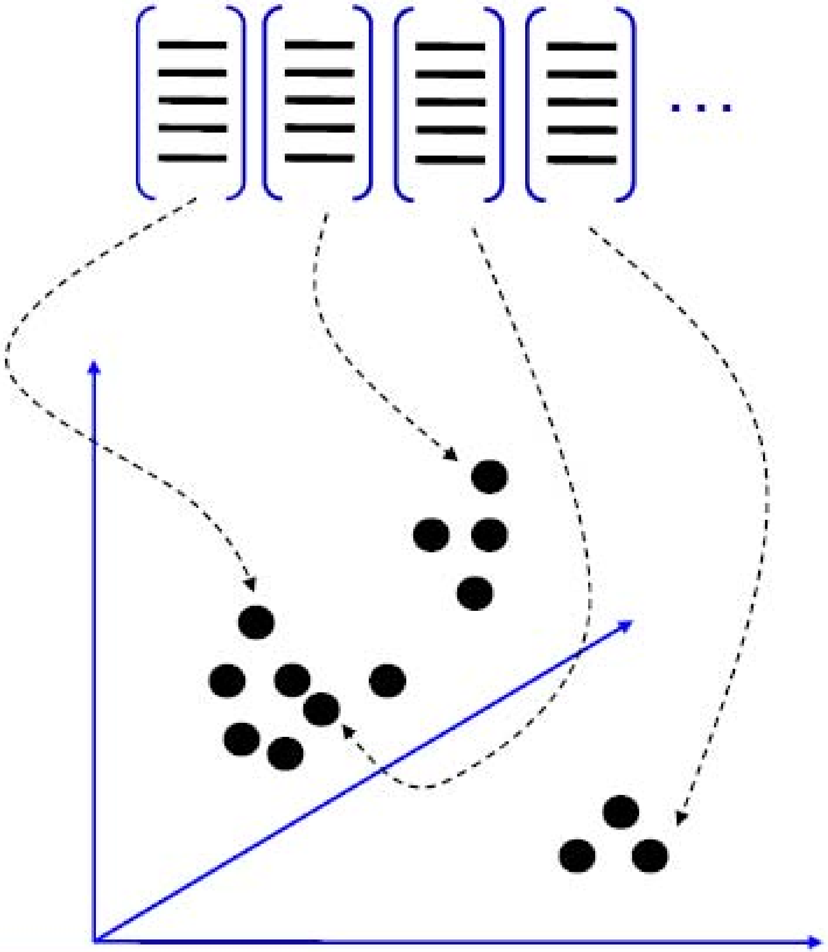
Convolutional Neural Network (CNN)

- **Features from some layers of CNNs**



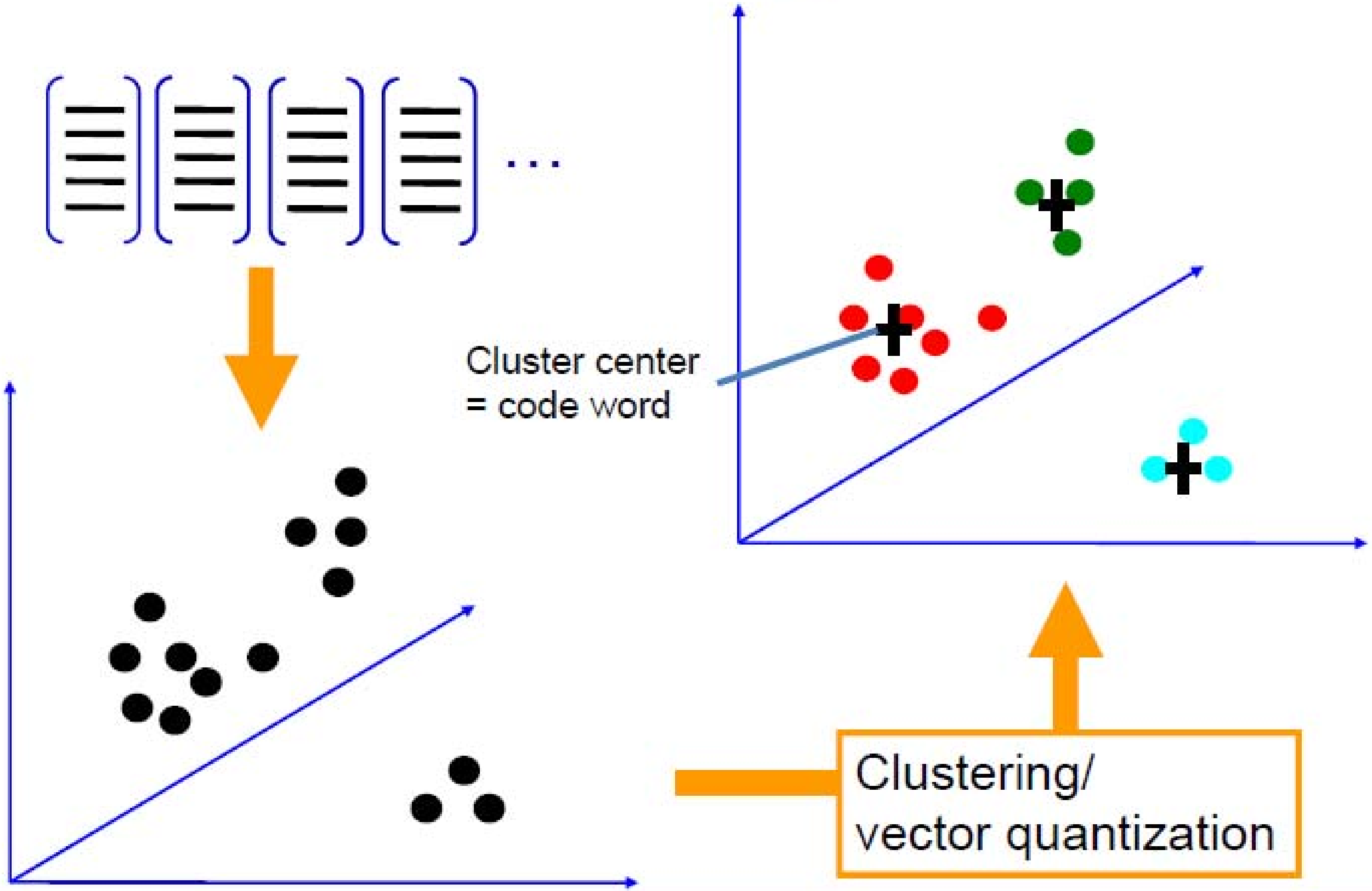
System from Krizhevsky et al., NIPS 2012

2. Codewords dictionary formation



Ack: Fei-Fie Li

2. Codewords dictionary formation

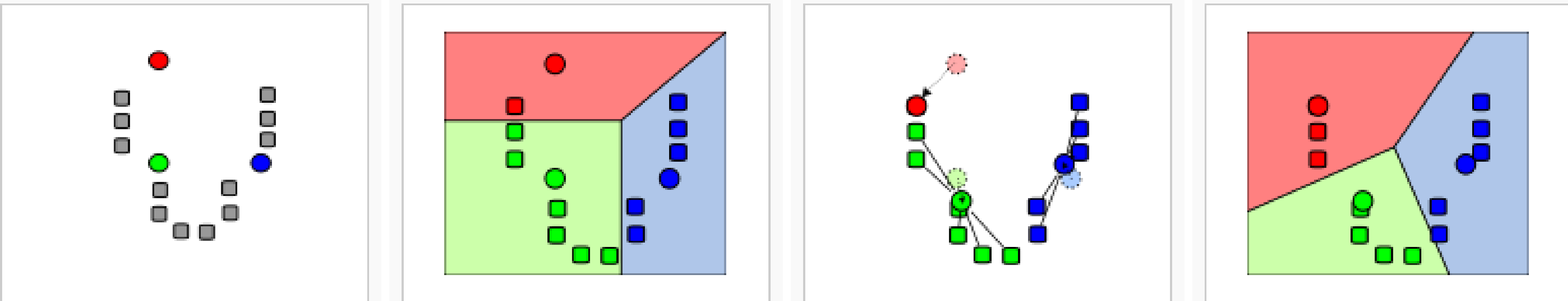


K-Means Clustering

- **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



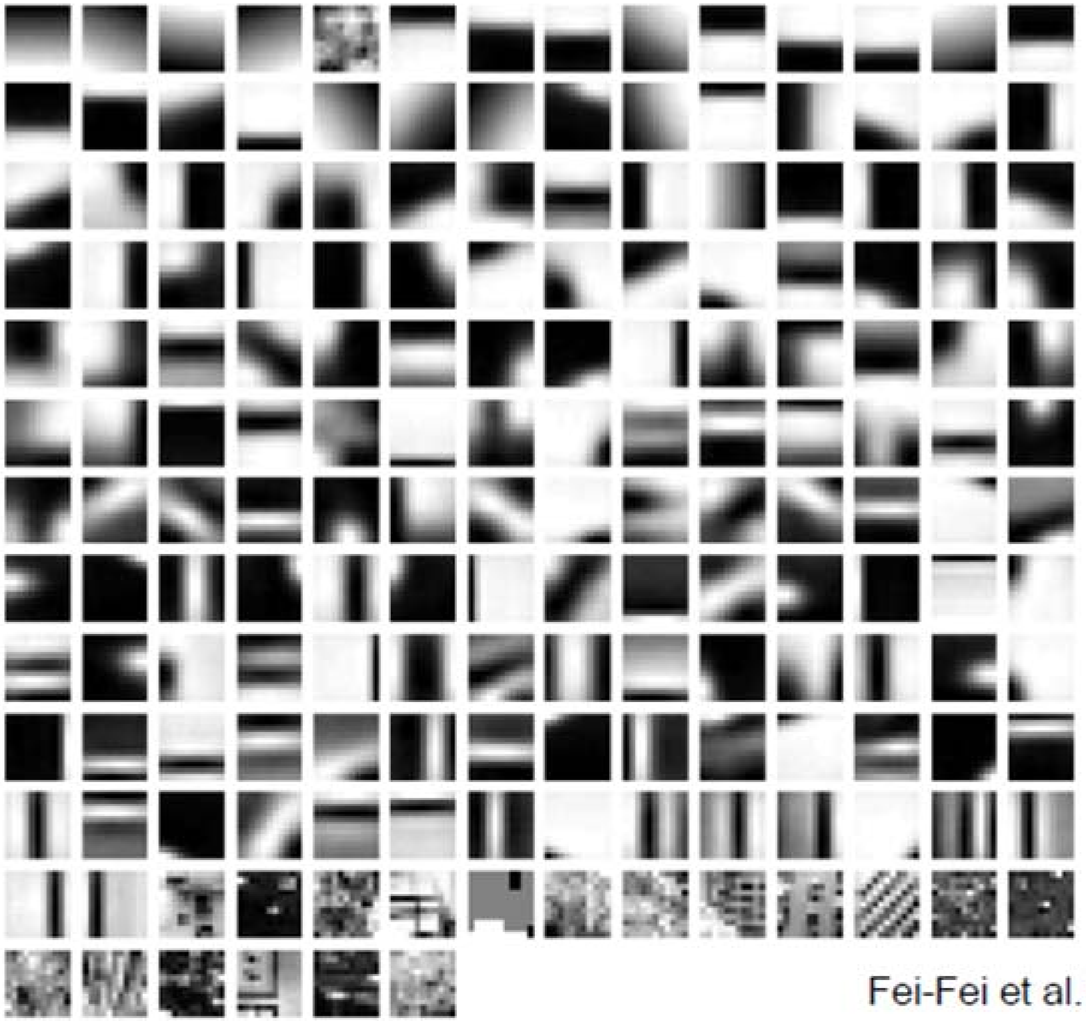
1) *k* initial "means" (in this case *k*=3) are randomly selected from the data set (shown in color).

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.

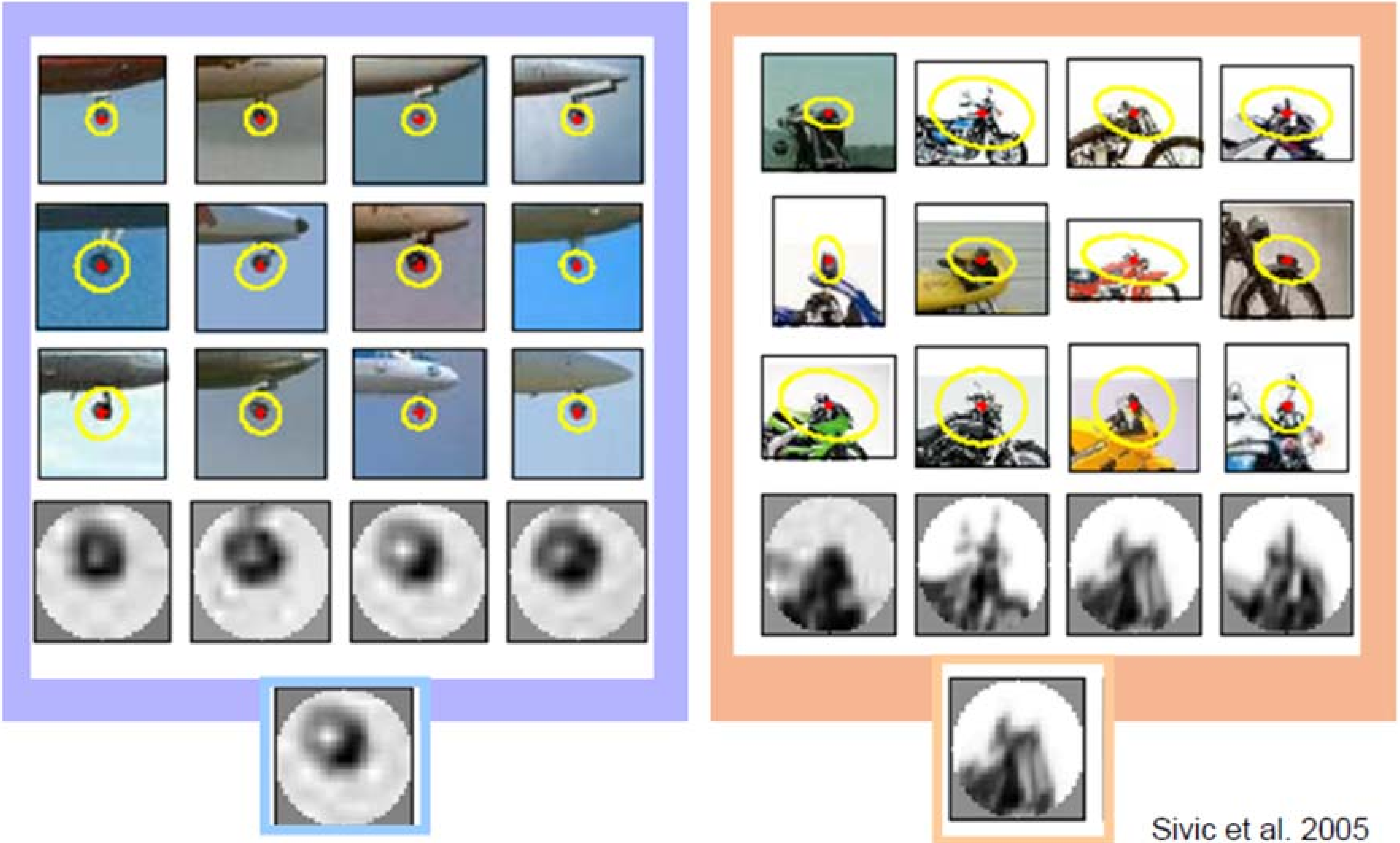
4) Steps 2 and 3 are repeated until convergence has been reached.

Codewords Dictionary Formation



Fei-Fei et al. 2005

Image Patch Examples of Codewords

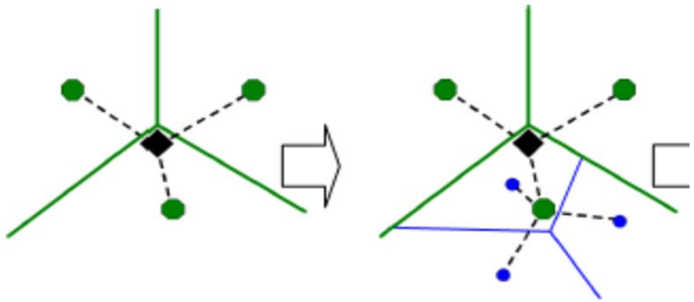


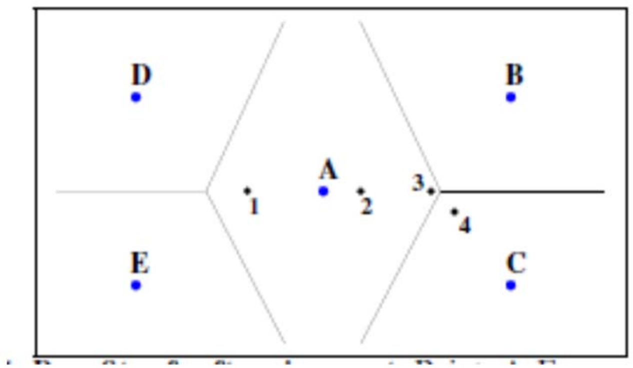
Sivic et al. 2005

Issues of Visual Vocabulary

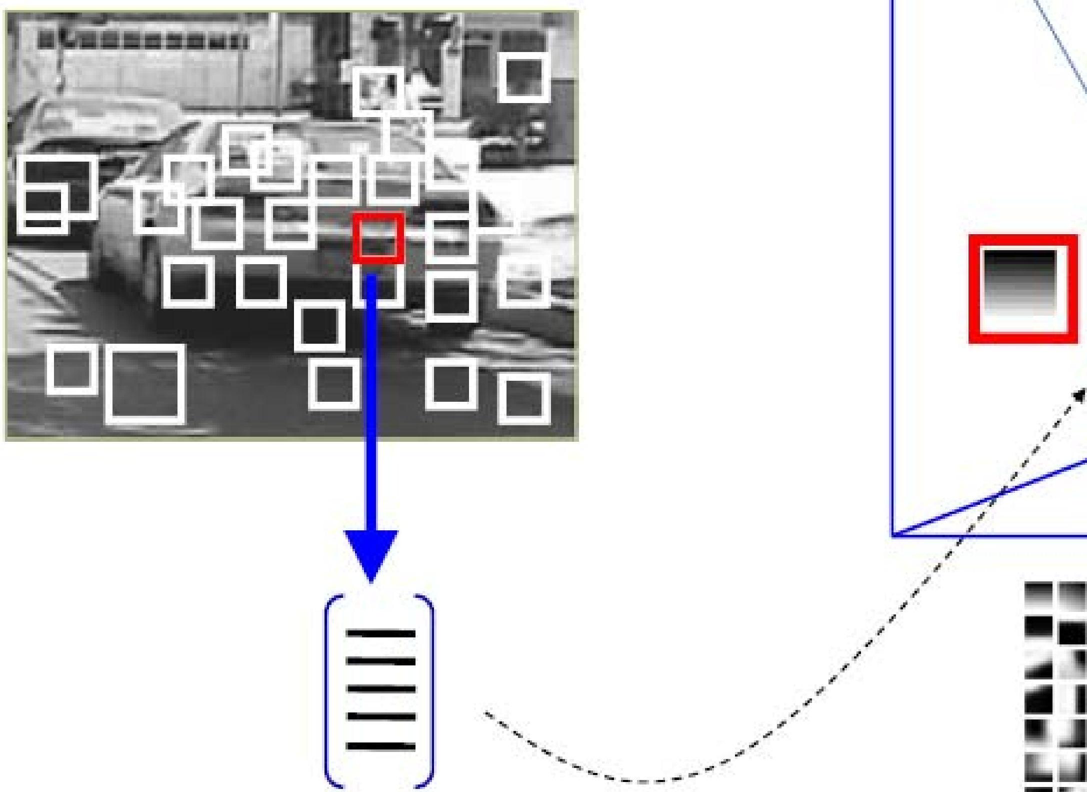
- Related to quantization
  - Too many words: quantization artifacts
  - Too small words: not representative
- K-means also takes long computation times

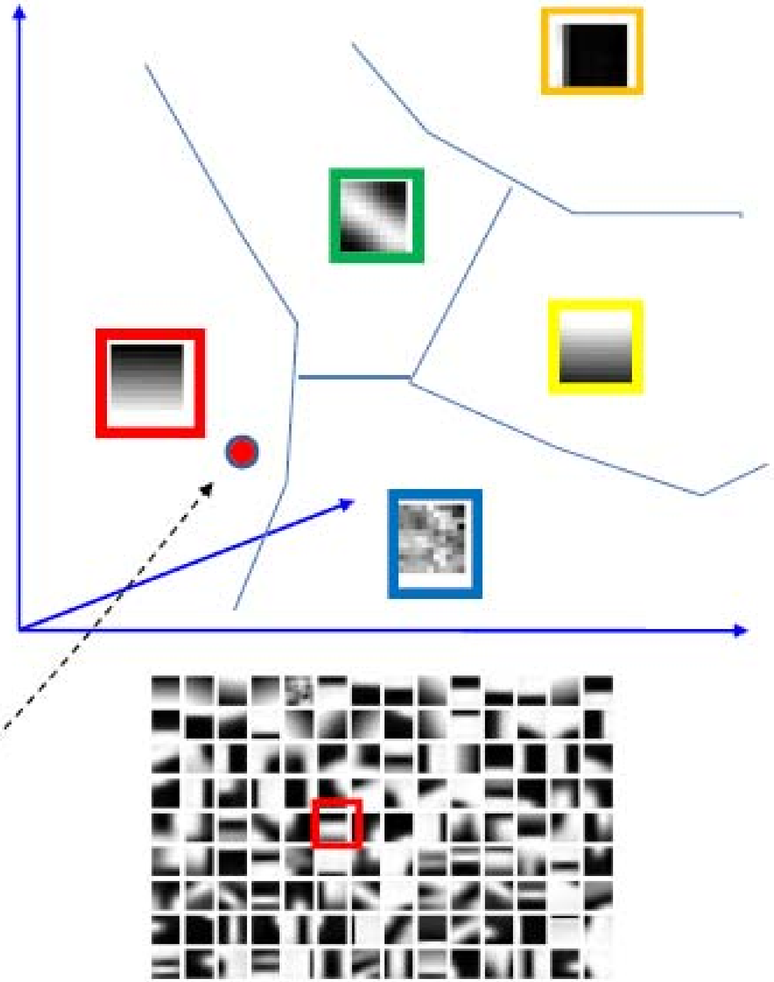
- Alternatives
  - Faster performance: vocabulary tree, Nister et al.
  - Low quantization artifacts: soft quantization, Philbin et al.





3. Bag of word representation



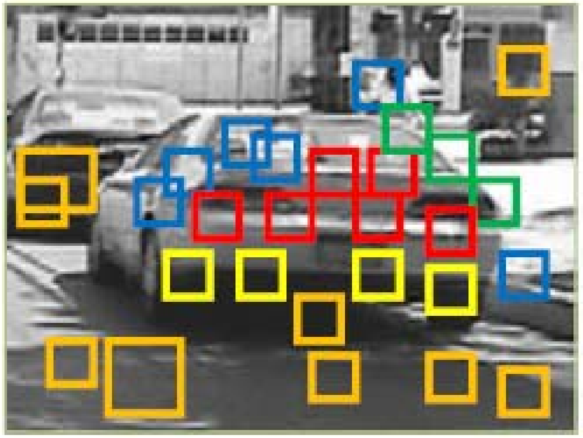


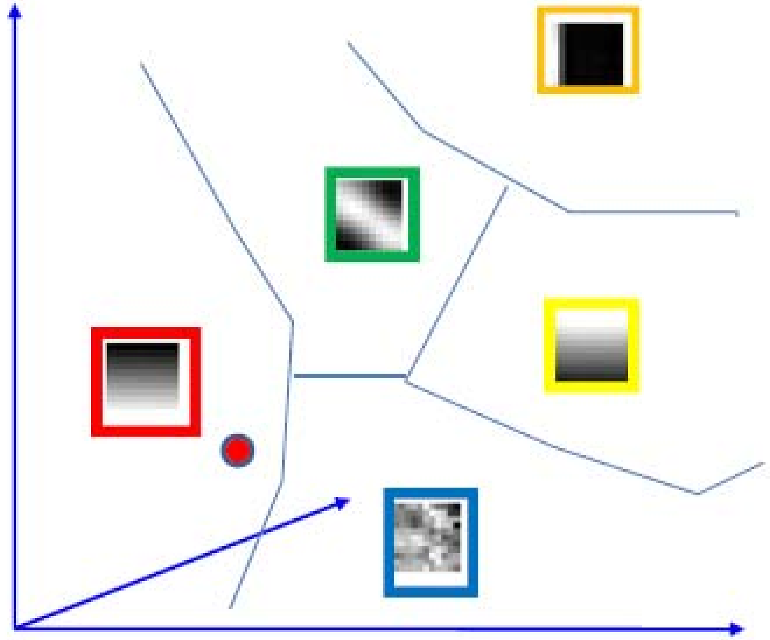
Codewords dictionary

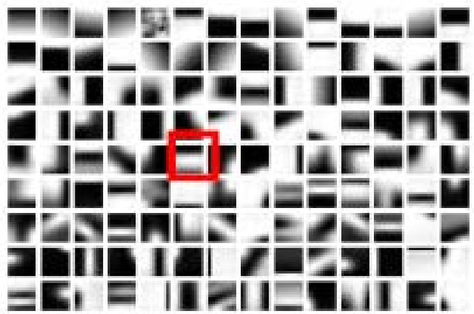
- Nearest neighbors assignment
- K-D tree search strategy

Ack: Fei-Fie Li

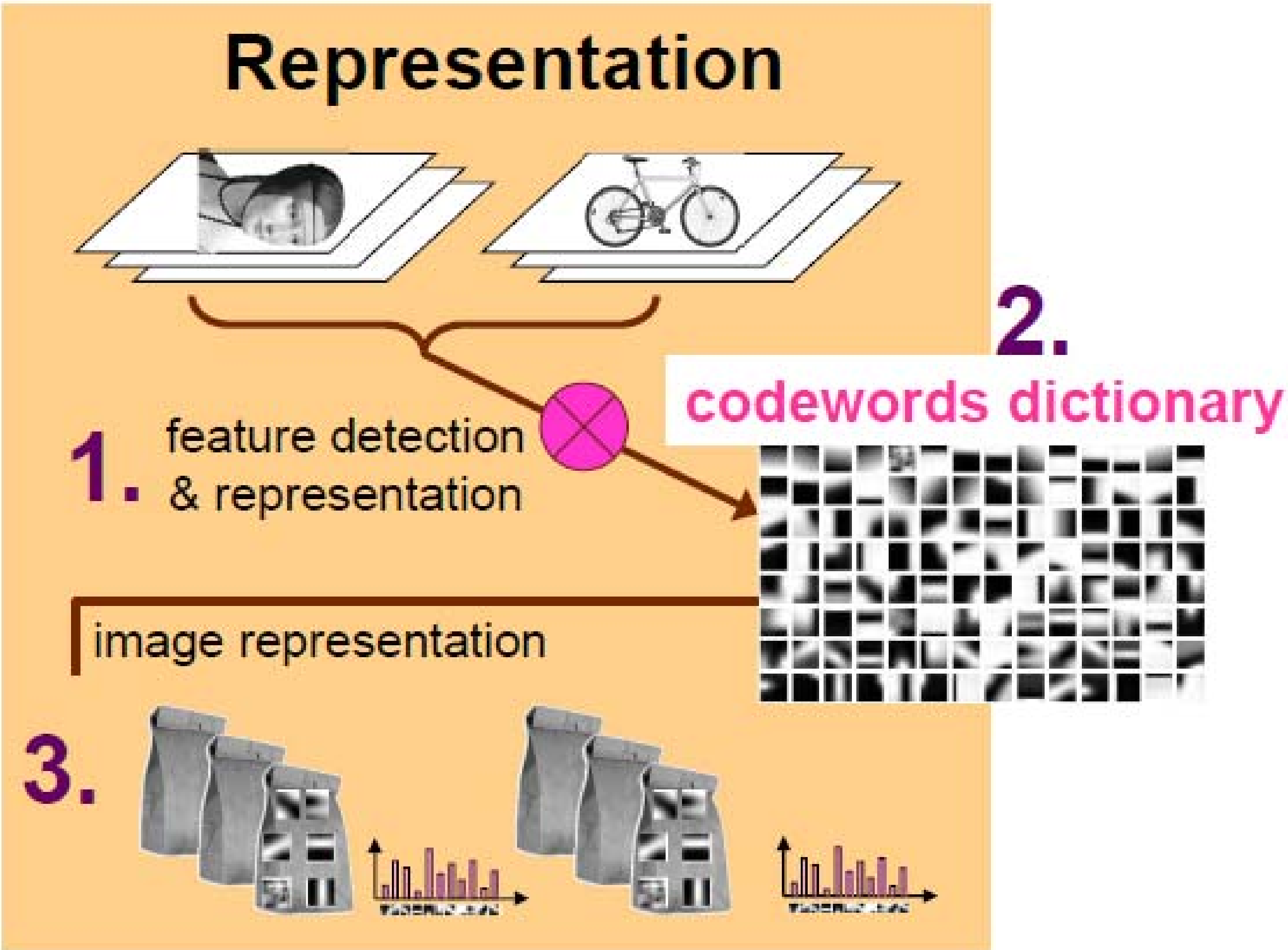
3. Bag of word representation



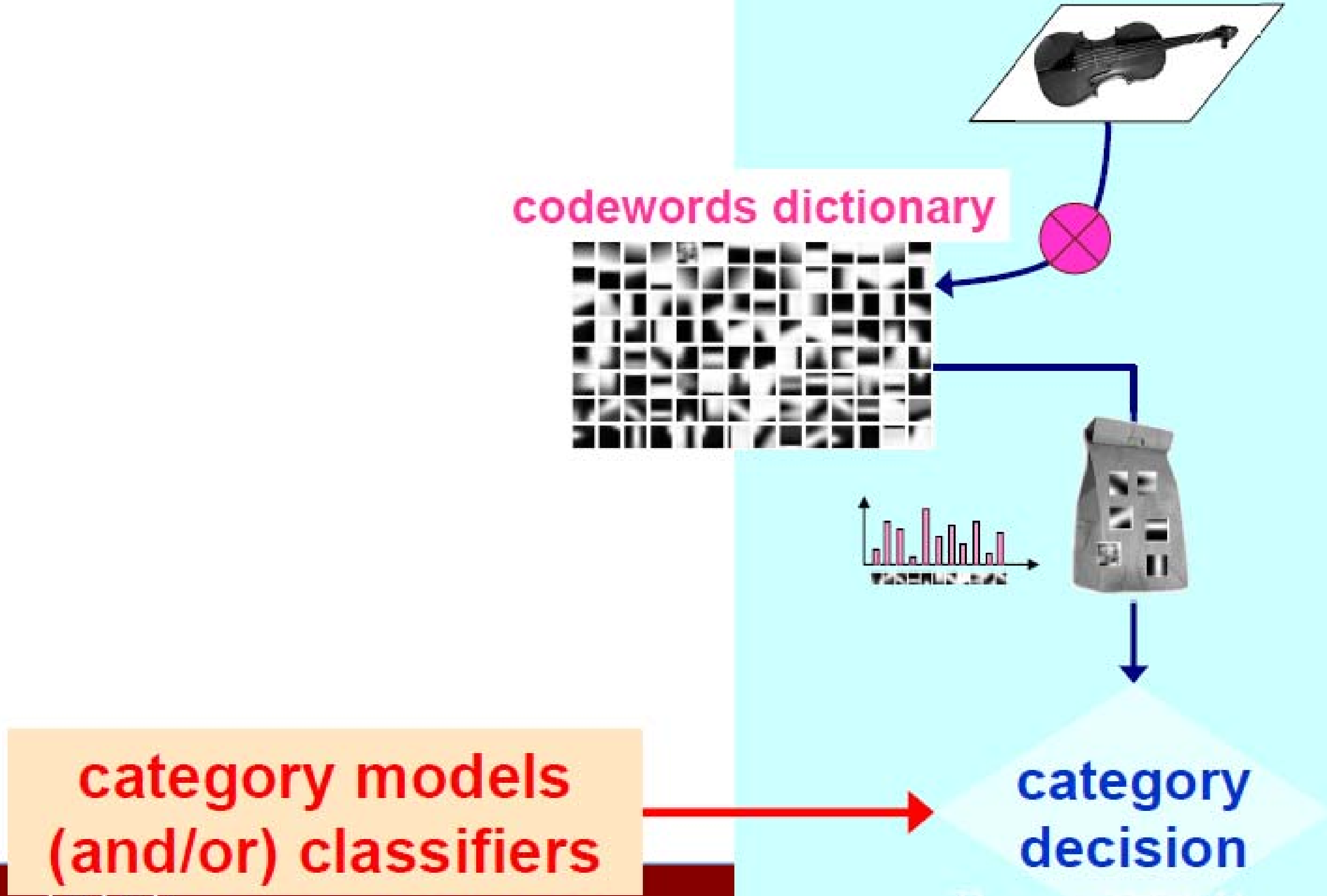




Codewords dictionary


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

Learning and Recognition



Similarity and Distance Functions

- **L1 or Euclidean distance**

$$L1\,(h_1, h_2) = \sum_i \left| h_1^i - h_2^i \right|$$

- χ^2 distance

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

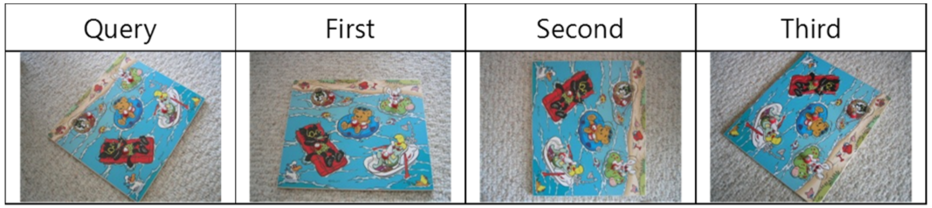
- Quadratic distance (*cross-bin*)

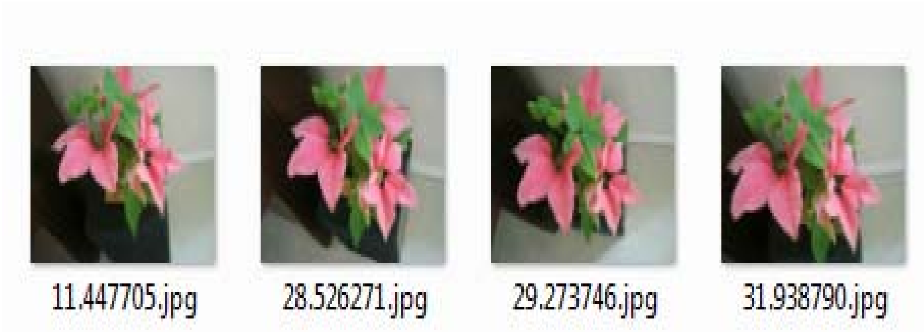
$$D(h_1, h_2) = \sum_{i,j} A_{ij}\,(h_1(i) - h_2(j))^2$$

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

Programming Assignment (PA2)

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

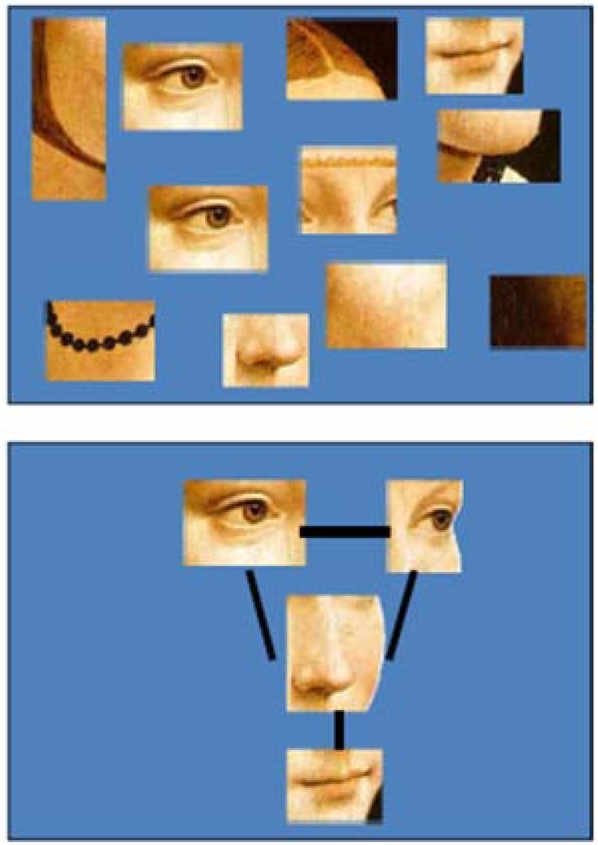




Problems of BoW Model

- No spatial relationship between words

- How can we perform segmentation and localization?

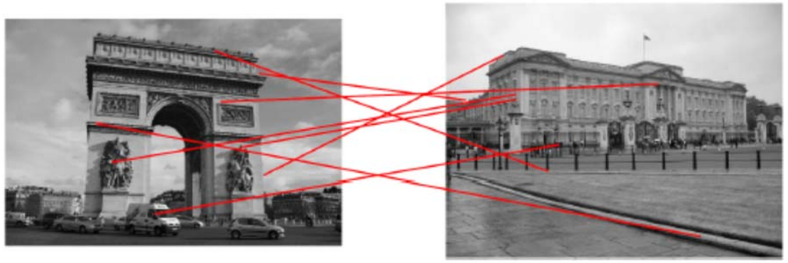


Ack.: Fei-Fei Li

Post-Processing or Reranking

Post-Processing

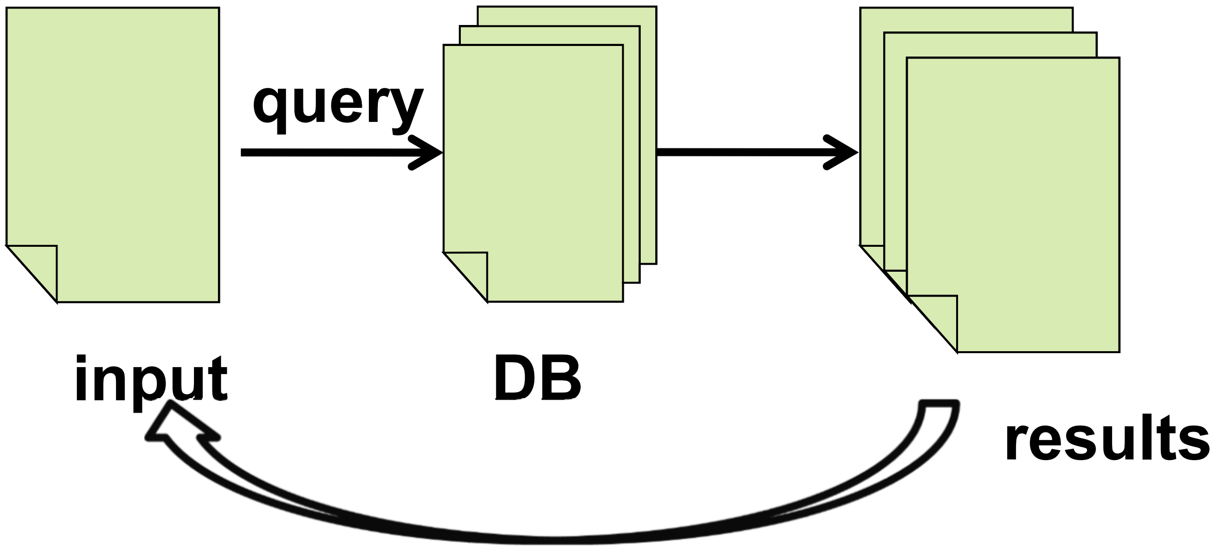
- **Geometric verification**
  - **RANSAC**



Matching w/o spatial matching

(Ack: Edward Johns et al.)

- **Query expansion**



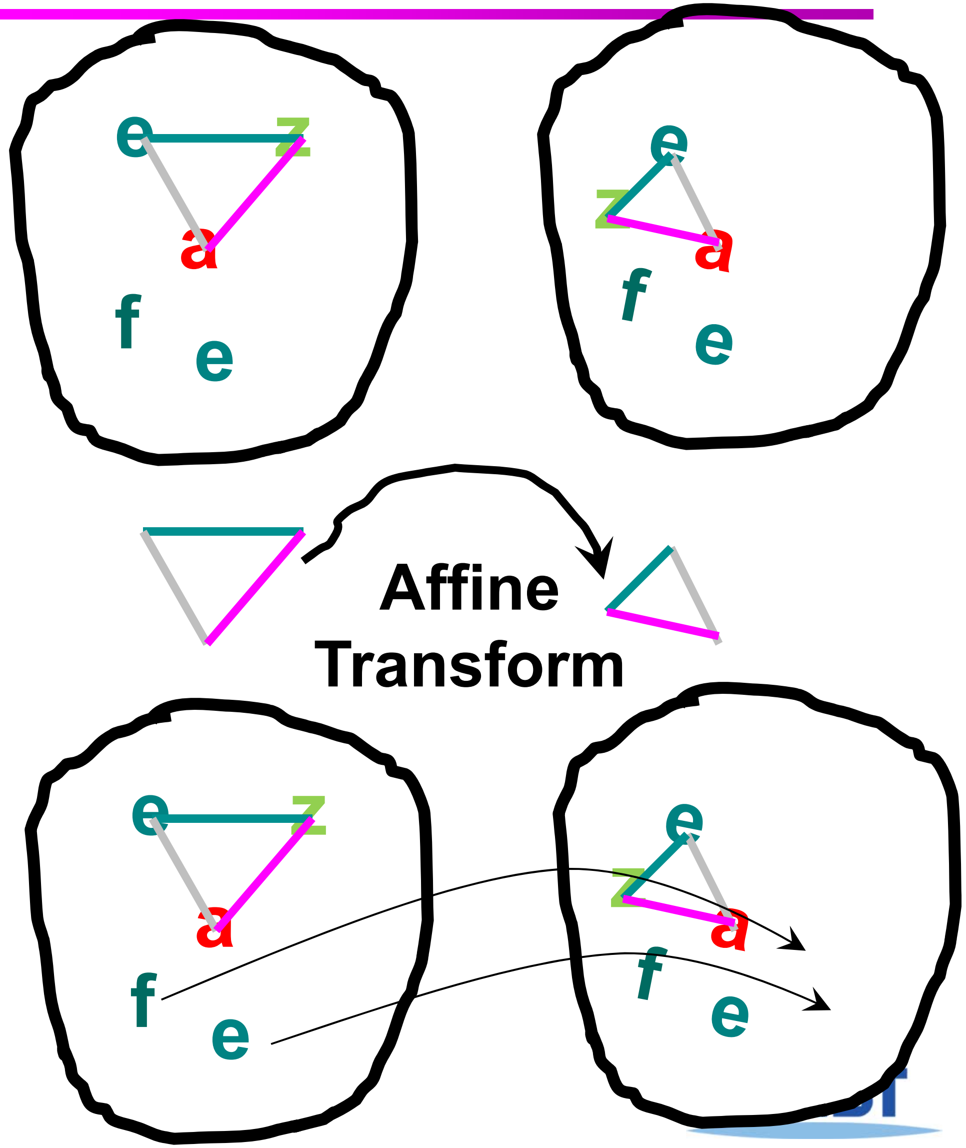
Geometric Verification using RANSAC for Affine Transform

Repeat N times:

- Randomly choose 3 matching pairs

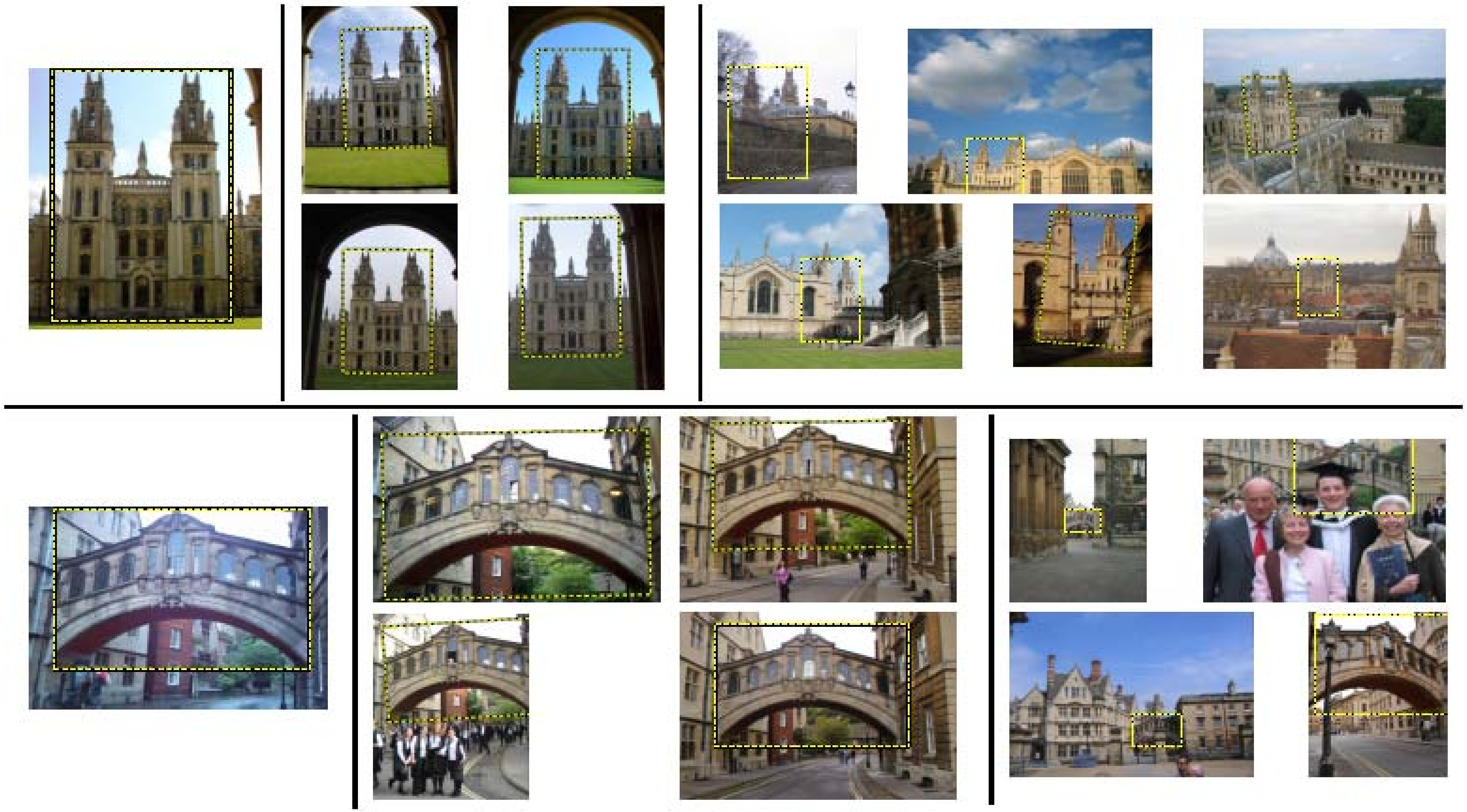
- Estimate transformation

- Predict remaining points and count "inliers"



Ack.: Derek Hoiem (UIUC)

Query Expansion [Chum et al. 07]



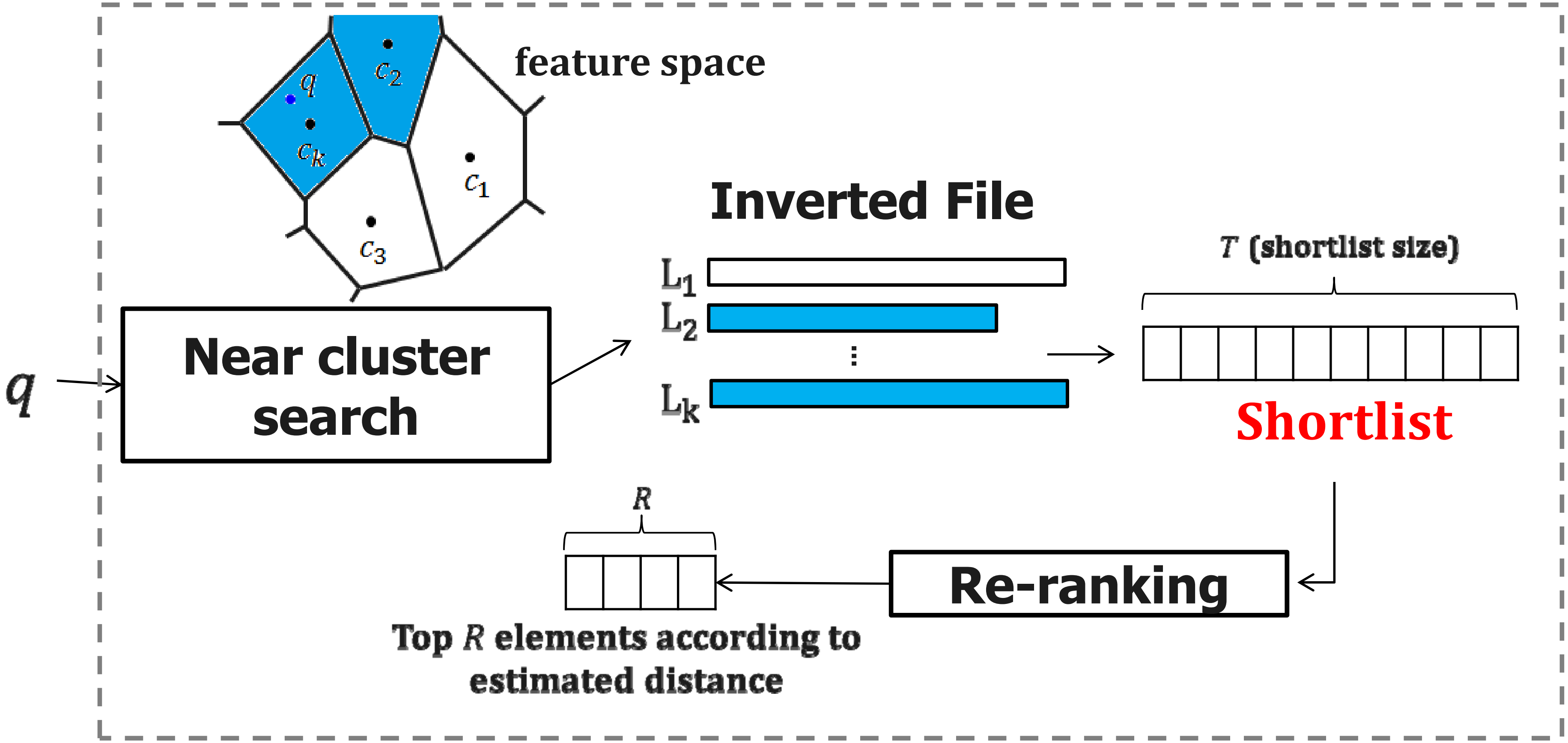
Original query

Top 4 images

Expanded results that were not identified by the original query

Efficient Search: Inverted File

- **For each word, list images containing the word**



Ack.: Dr. Heo KAIST

Scalability

- **Issues with billions of images?**
  - **Too much memory**
  - **Still low accuracy**

Recent Image Search Techniques:
Hashing Techniques

Sung-Eui Yoon
Associate Professor
KAIST

http://sglab.kaist.ac.kr

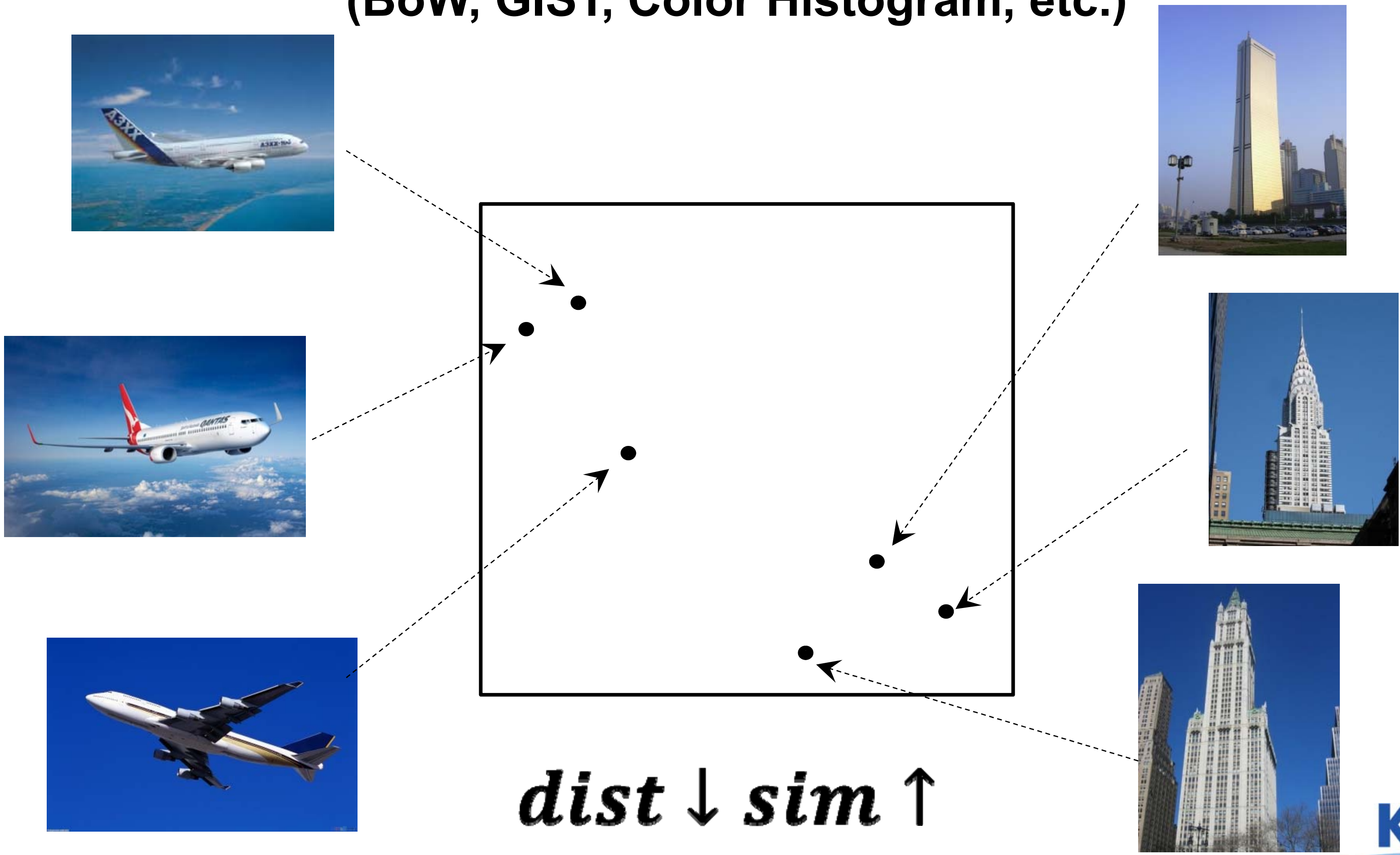
Image Retrieval

Finding visually similar images

Image Descriptor

High dimensional point

(BoW, GIST, Color Histogram, etc.)

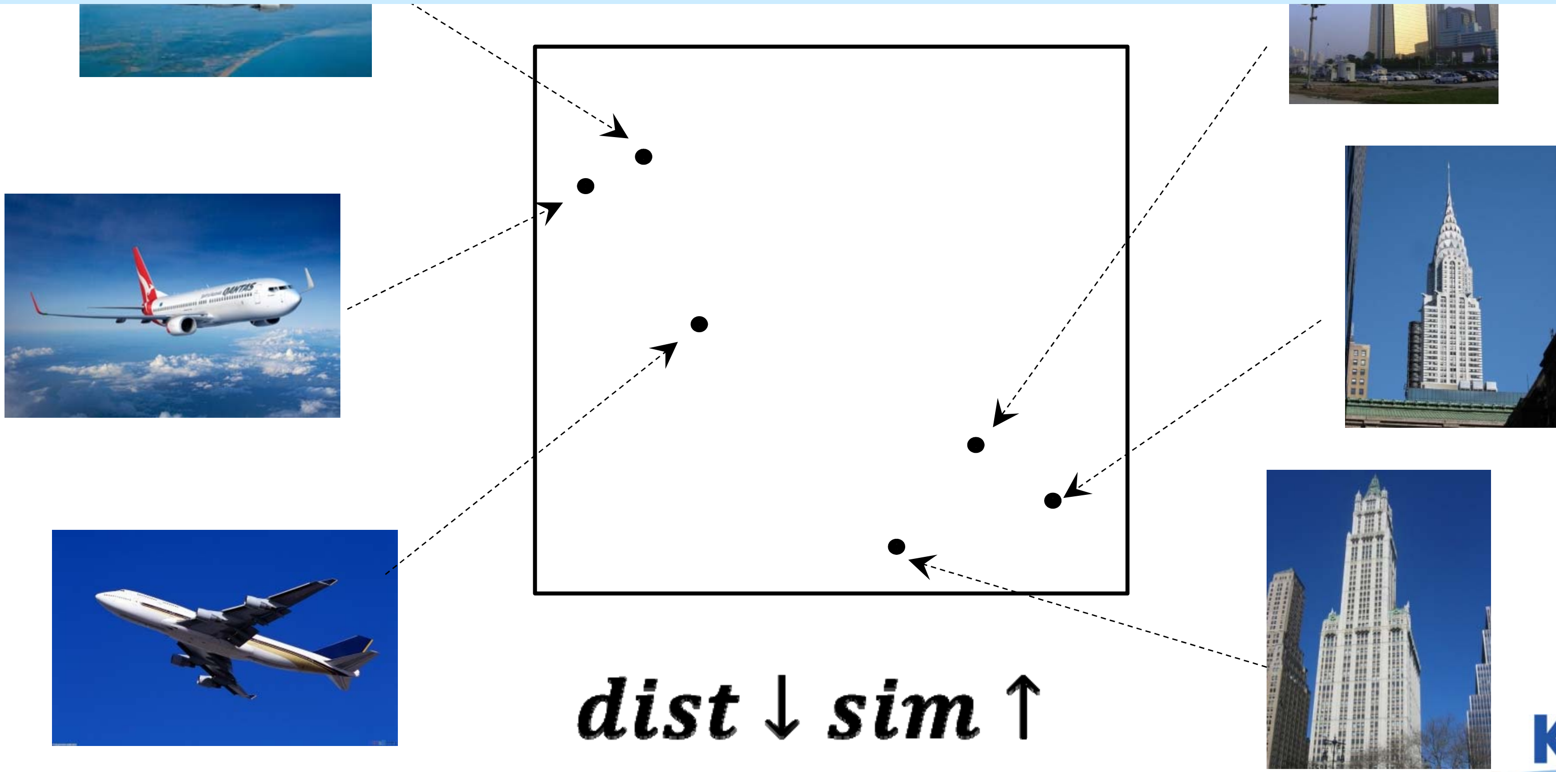


$dist \downarrow \; sim \uparrow$

Image Descriptor

High dimensional point

Nearest neighbor search (NNS) in high dimensional space

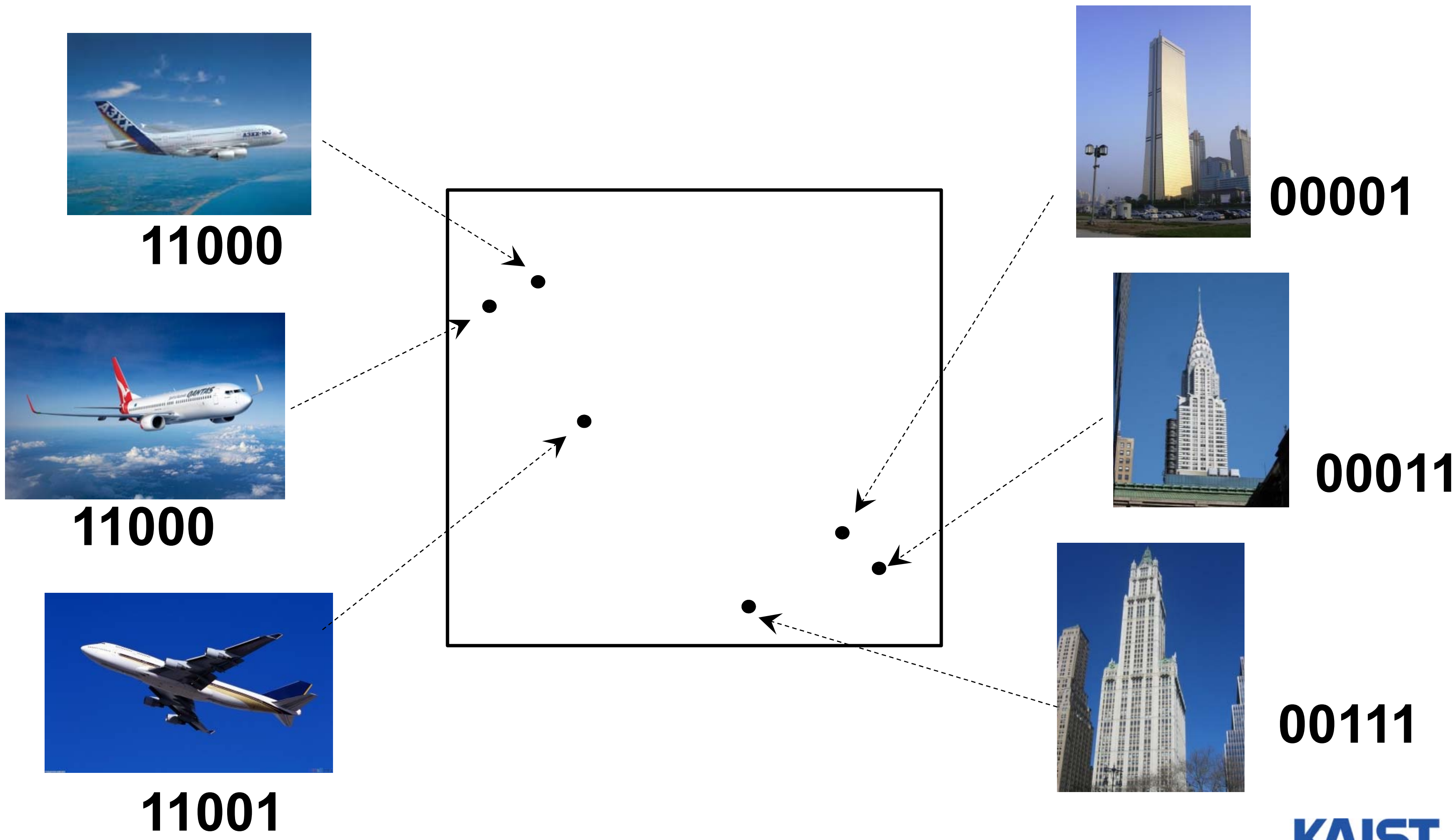


Challenge

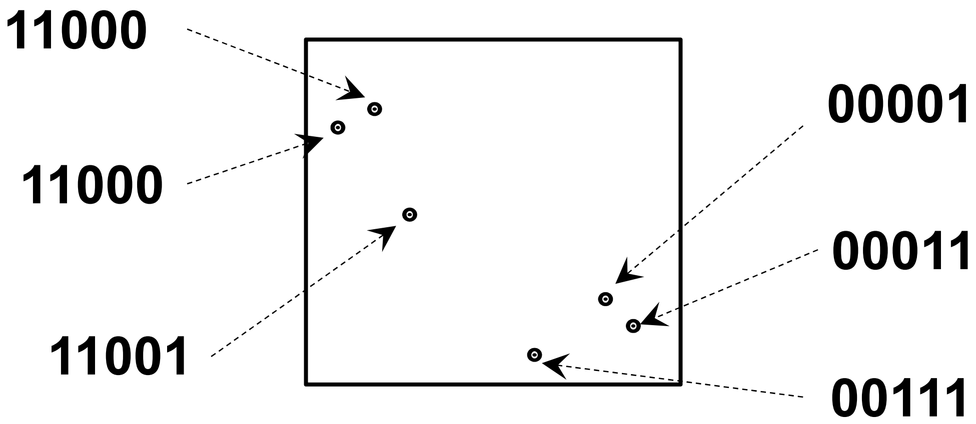
	BoW	GIST
Dimensions	1000+	300+
1 image	4 KB+	1.2 KB+
1B images	3 TB+	1 TB+

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

Binary Code



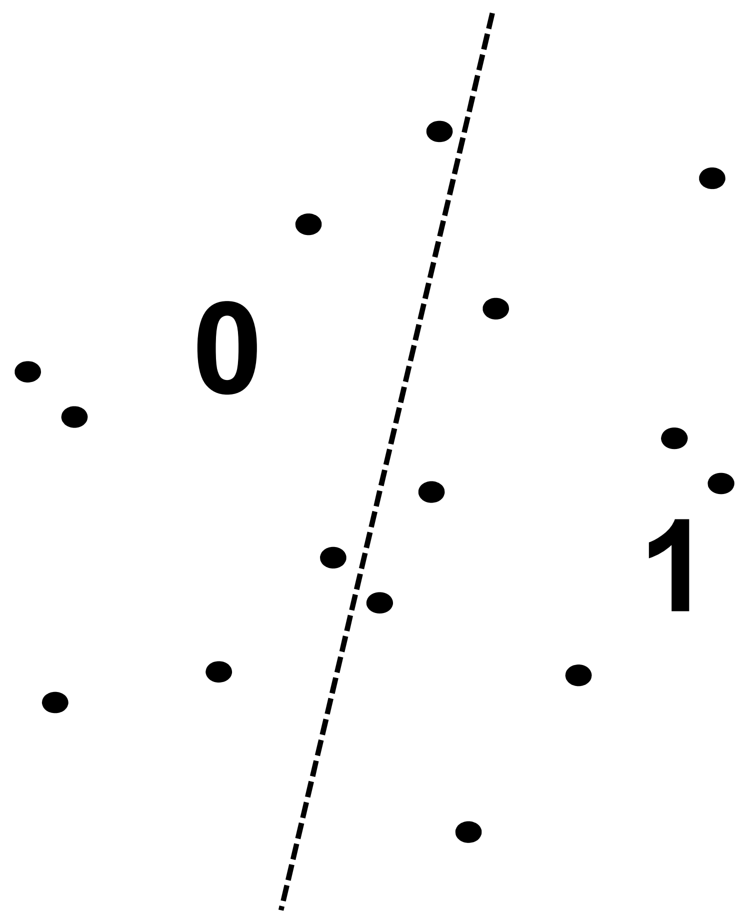
Binary Code



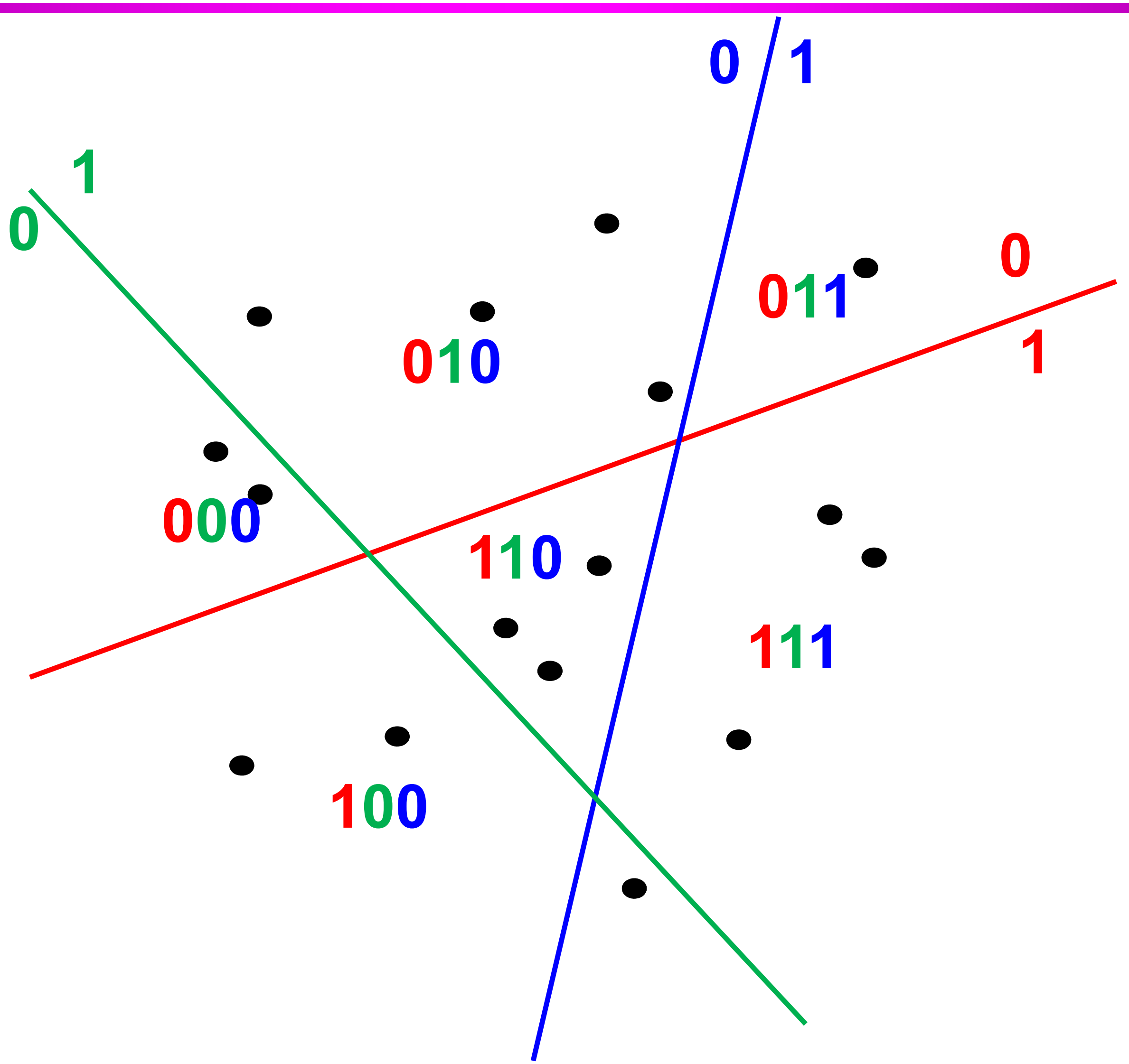
* Benefits

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

Hyper-Plane based Binary Coding



Hyper-Plane based Binary Coding

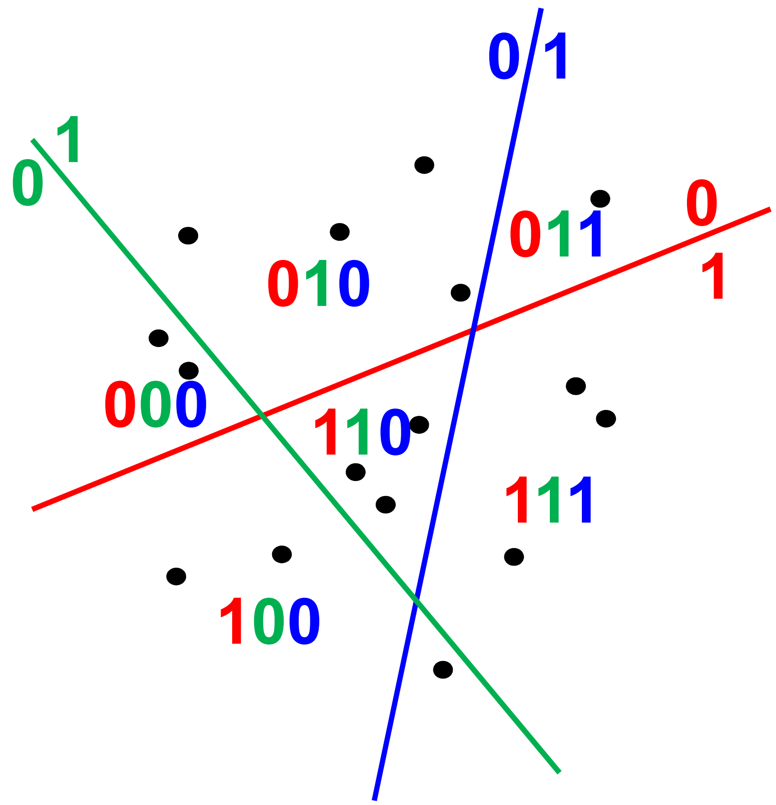


KAIST

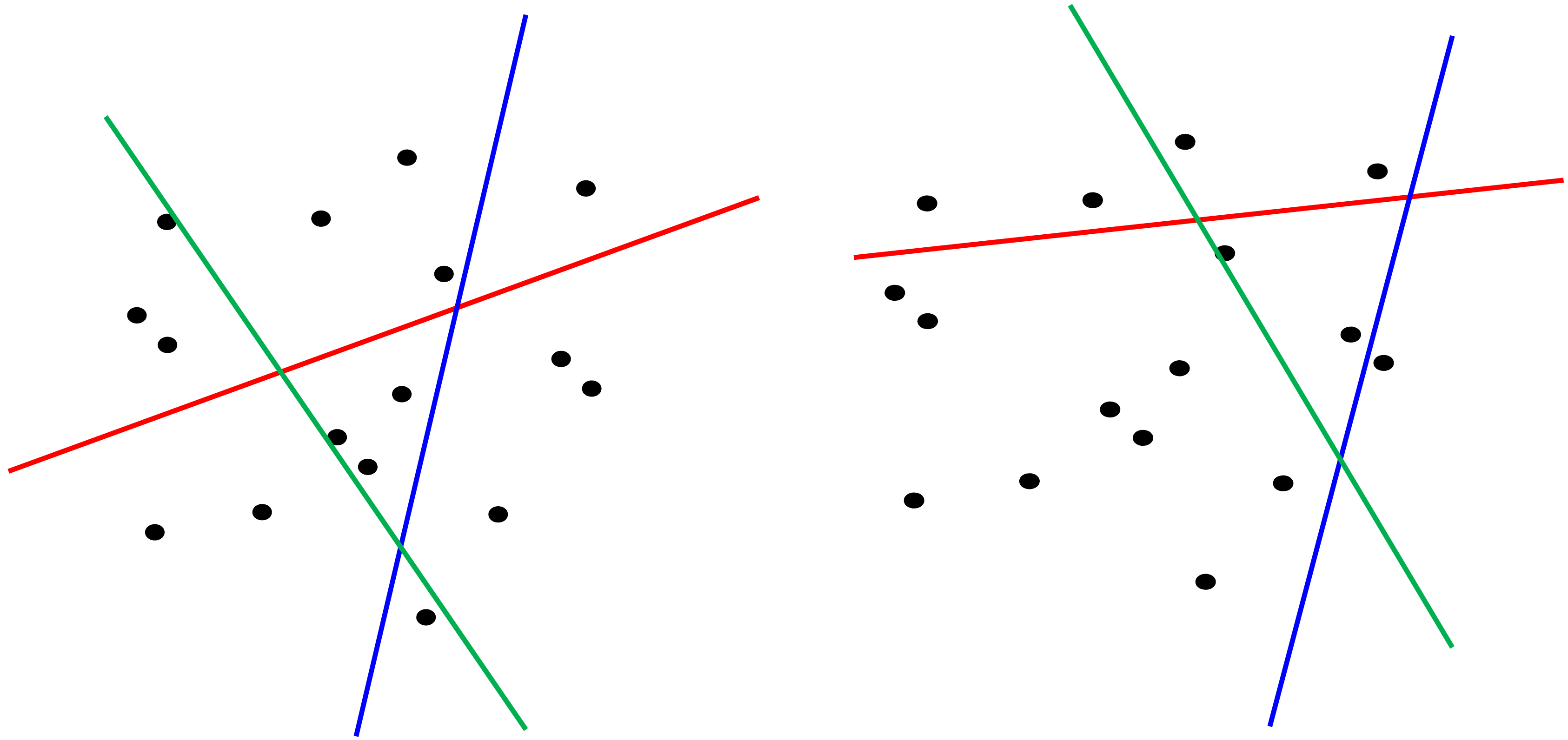
Distance between Two Points

- **Measured by bit differences, known as Hamming distance**
- **Efficiently computed by XOR bit operations**

$$d_{hd}(b_i, b_j) = |b_i \oplus b_j|$$



Good and Bad Hyper-Planes



Previous work focused on how to determine good hyper-planes

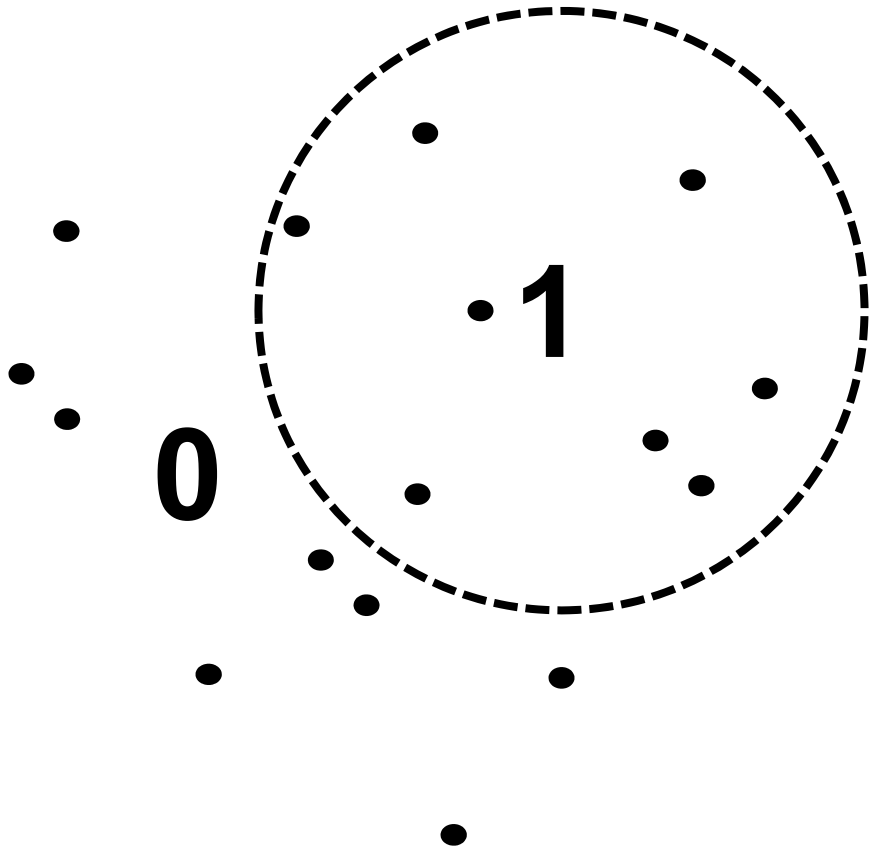
Components of Spherical Hashing

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

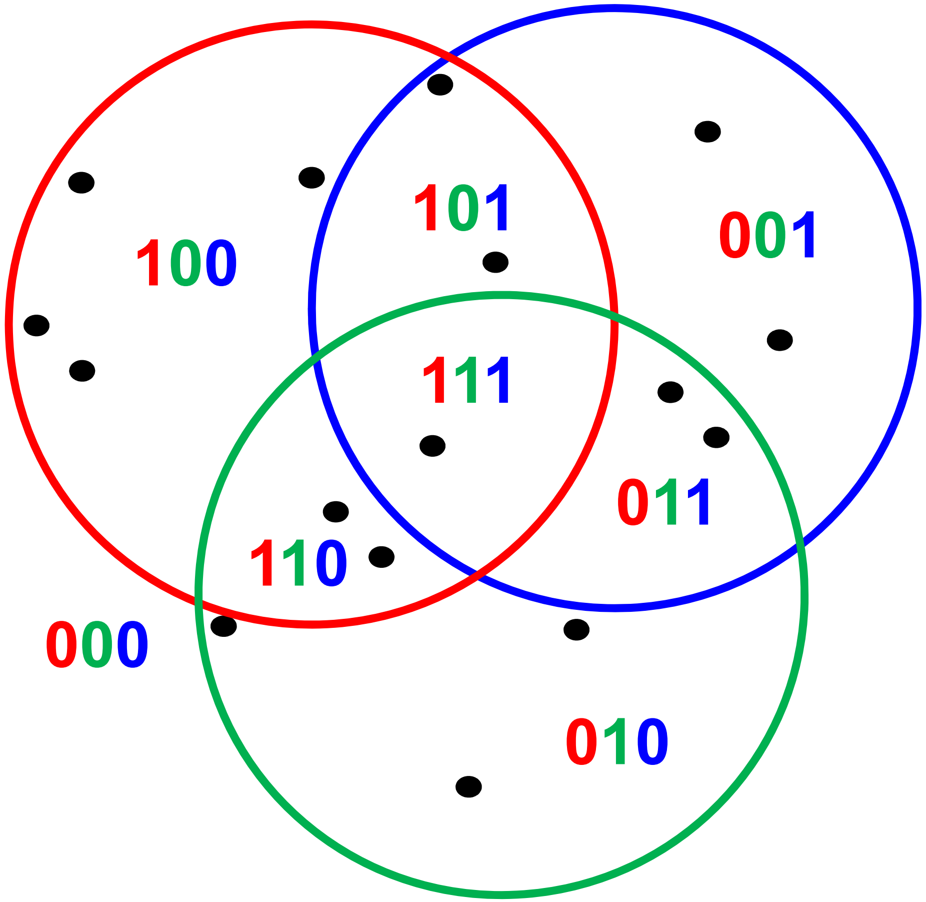
Components of Spherical Hashing

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

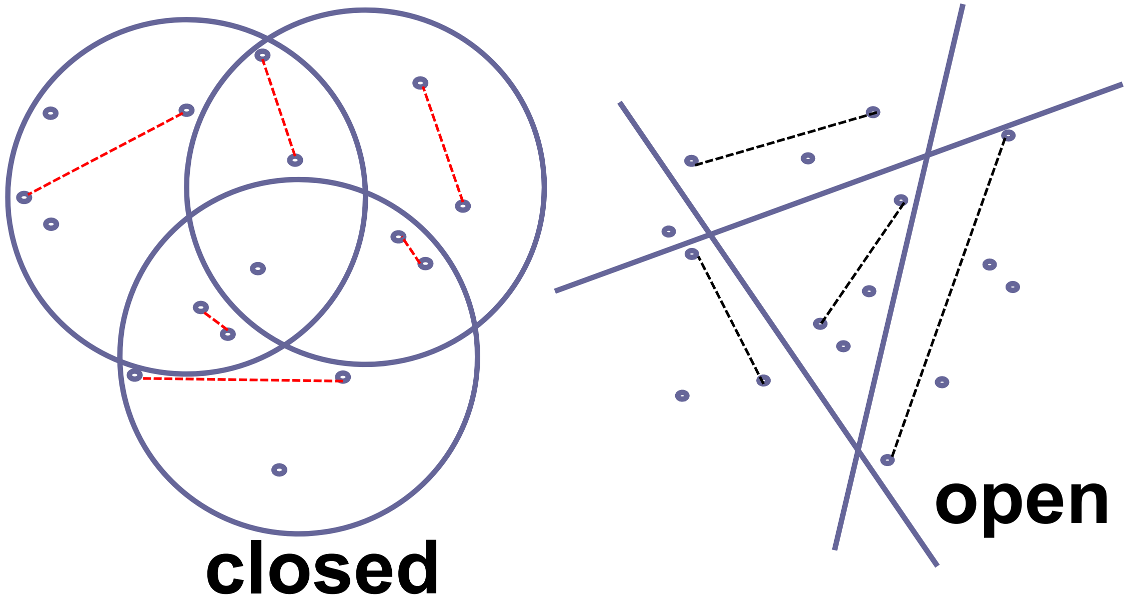
Spherical Hashing [Heo et al., CVPR 12]

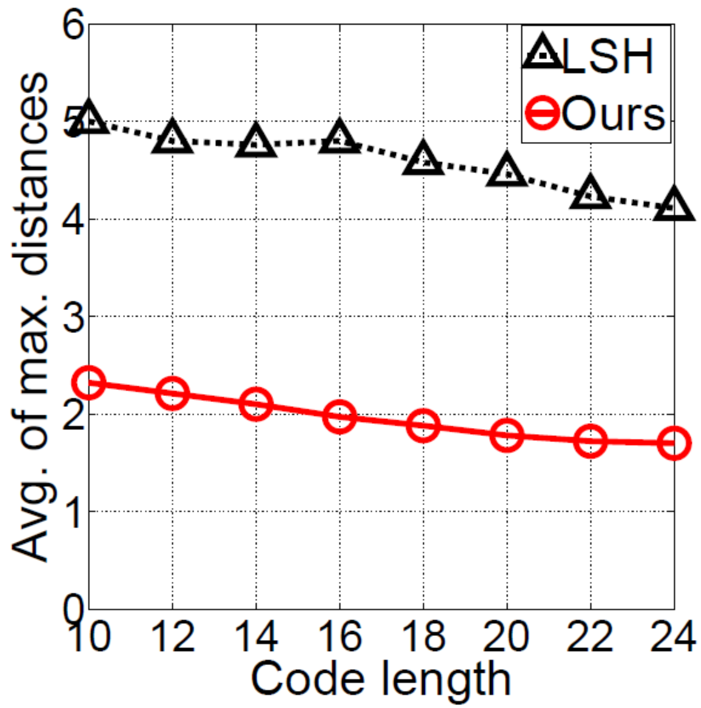


Spherical Hashing [Heo et al., CVPR 12]



Hyper-Sphere vs Hyper-Plane





Average of maximum distances within a partition:

- Hyper-spheres gives tighter bound!

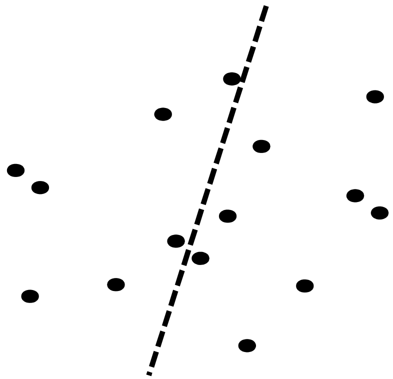
Components of Spherical Hashing

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

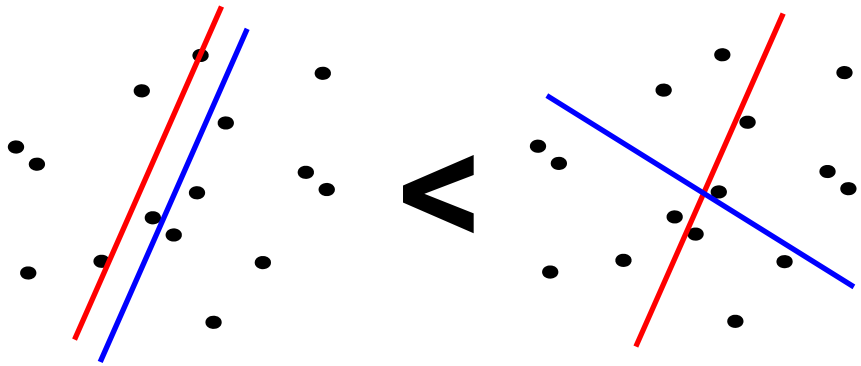
KAIST

Good Binary Coding [Yeiss 2008, He 2011]

1. **Balanced partitioning**

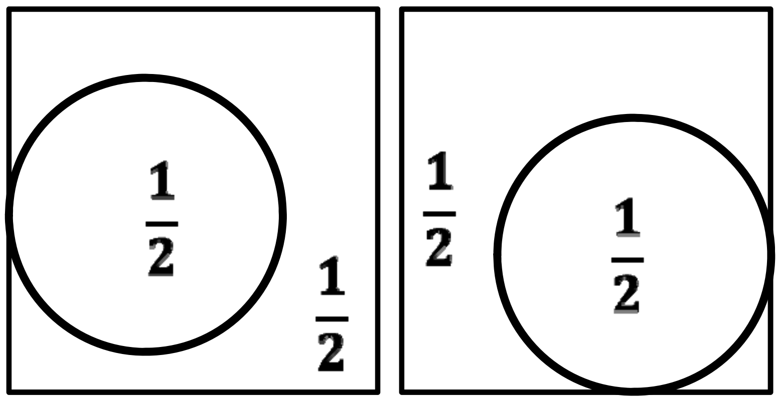


2. Independence

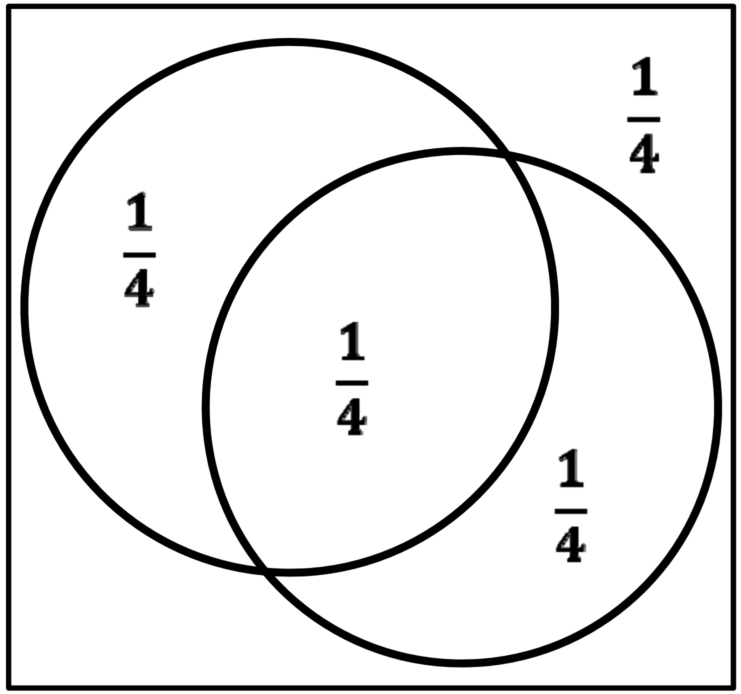


Intuition of Hyper-Sphere Setting

1. Balance

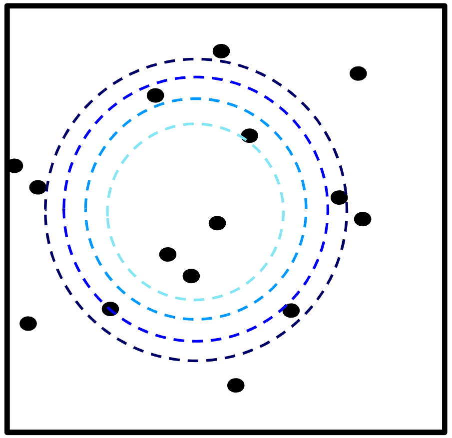


2. Independence

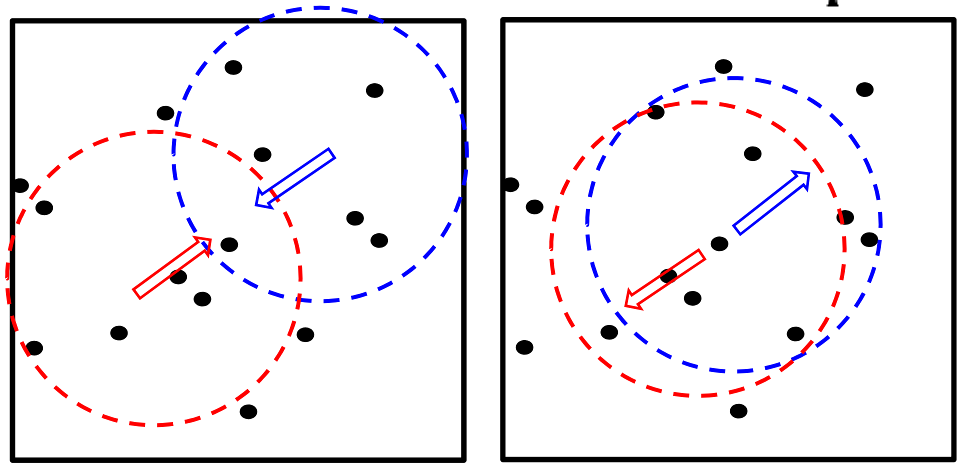


Hyper-Sphere Setting Process

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



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

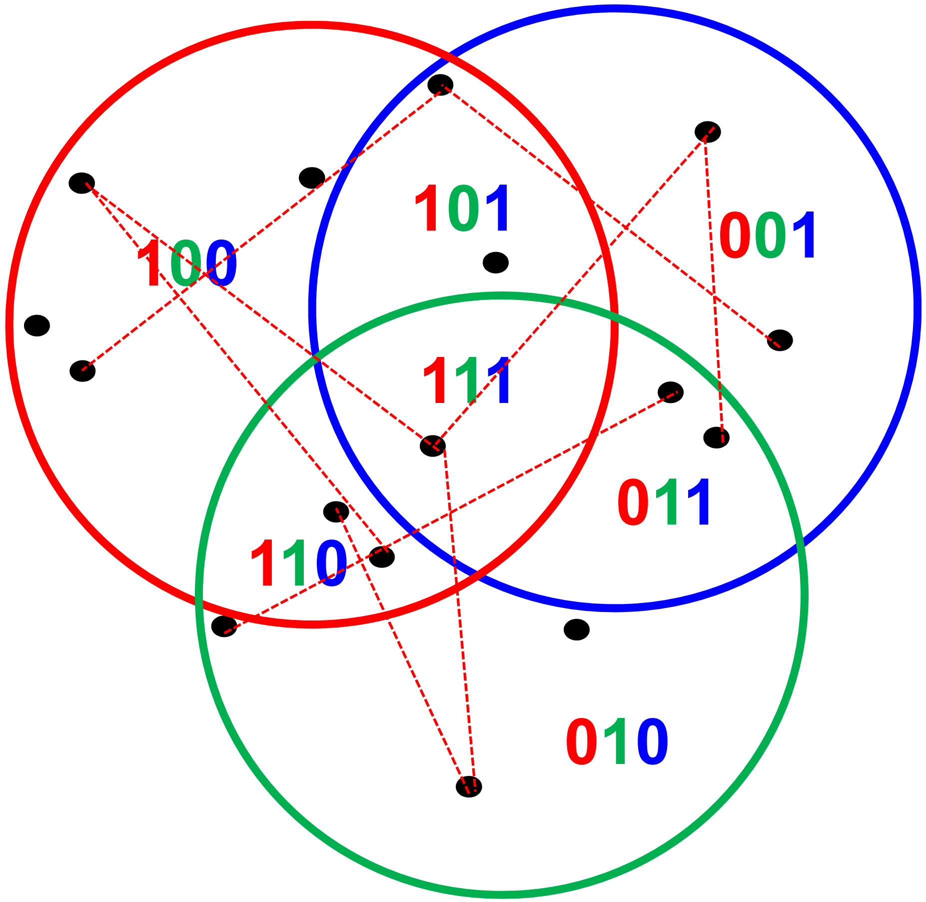


Iteratively repeat step 1, 2 until convergence.

Components of Spherical Hashing

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

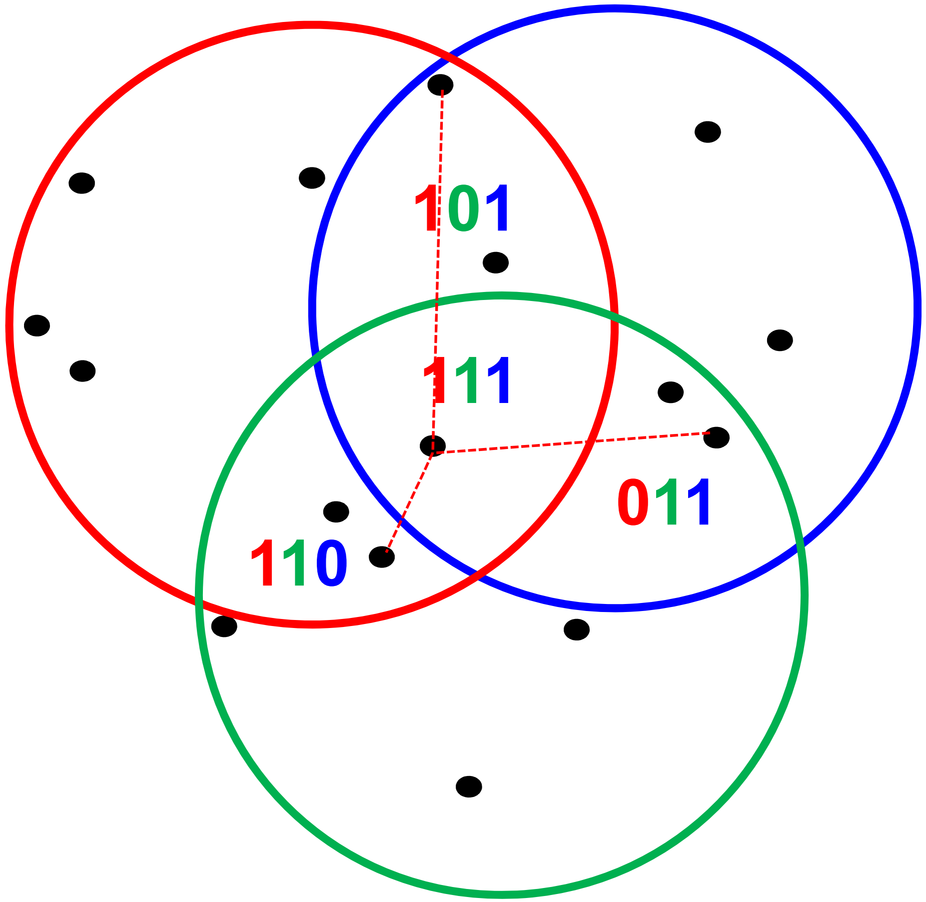
Max Distance and Common ‘1’



Common ‘1’s

: 1

Max Distance and Common '1'

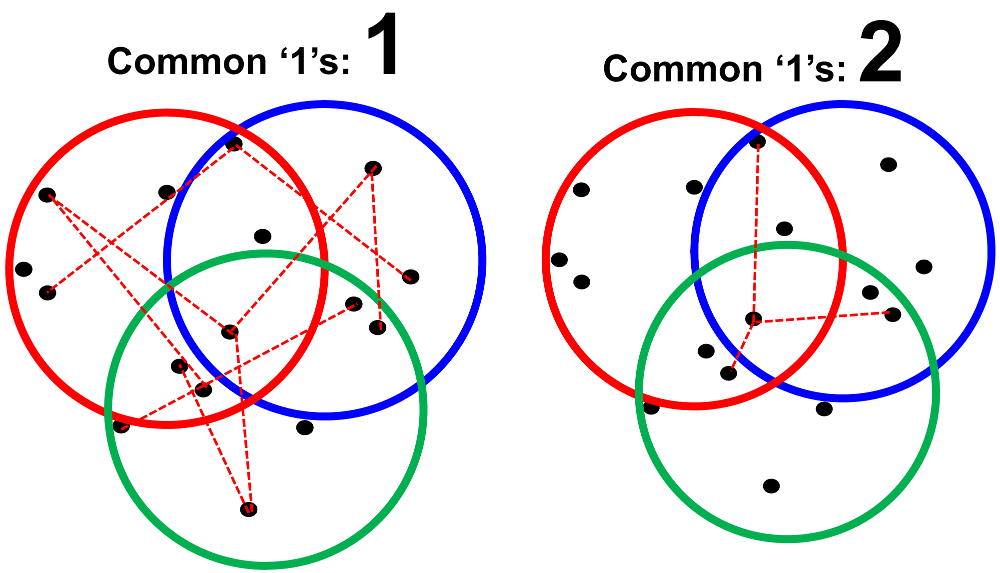


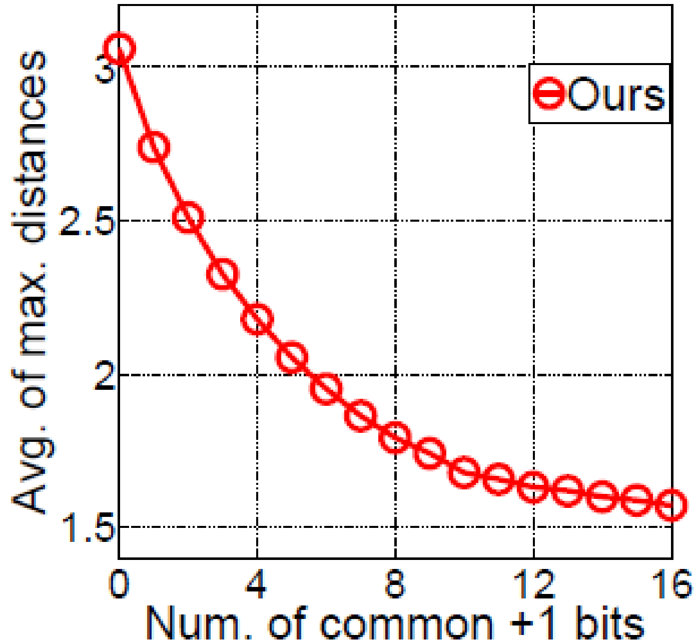
Common '1's

: **2**

KAIST

Max Distance and Common ‘1’





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

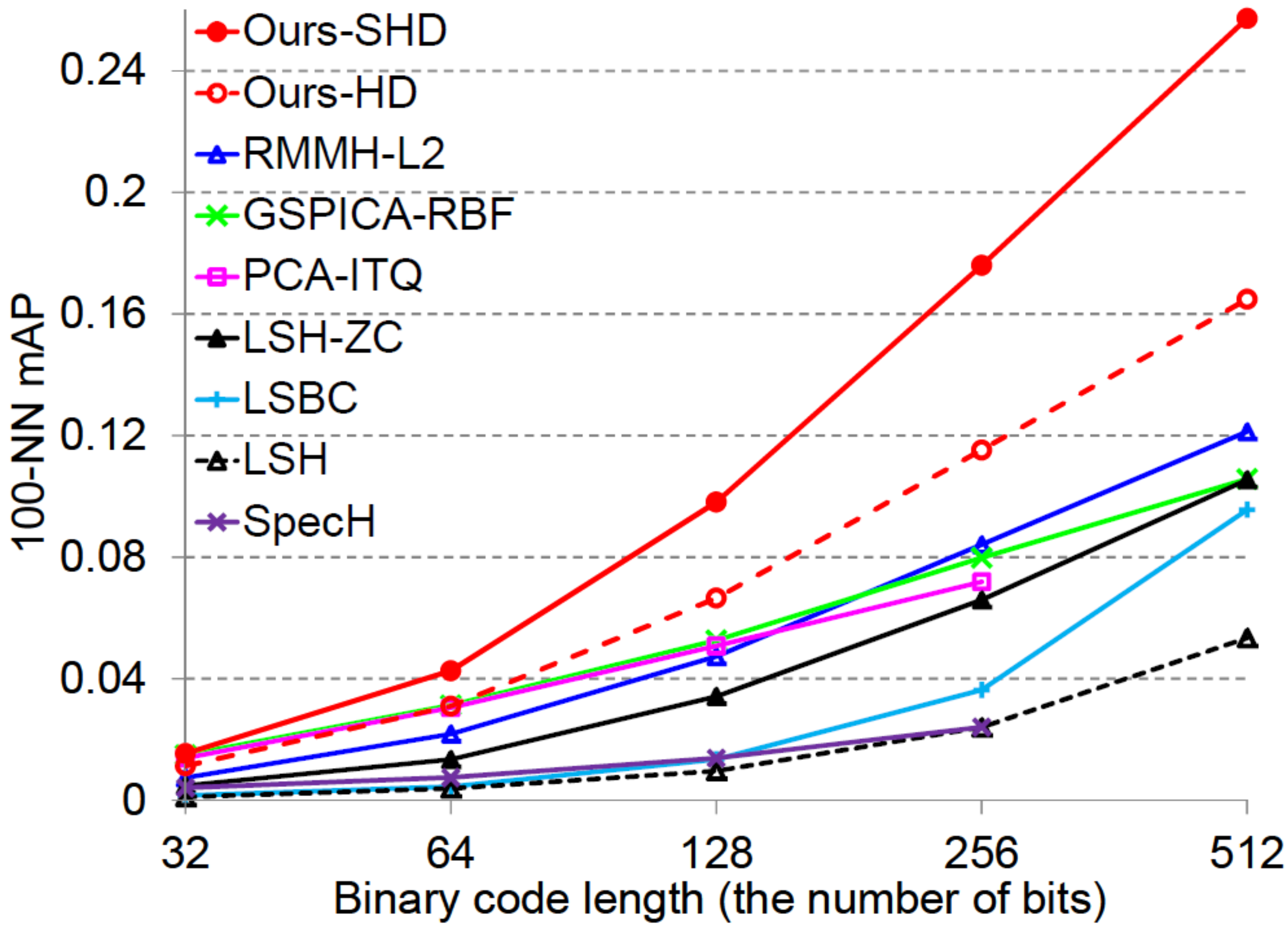
Spherical Hamming Distance (SHD)

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

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

KAIST

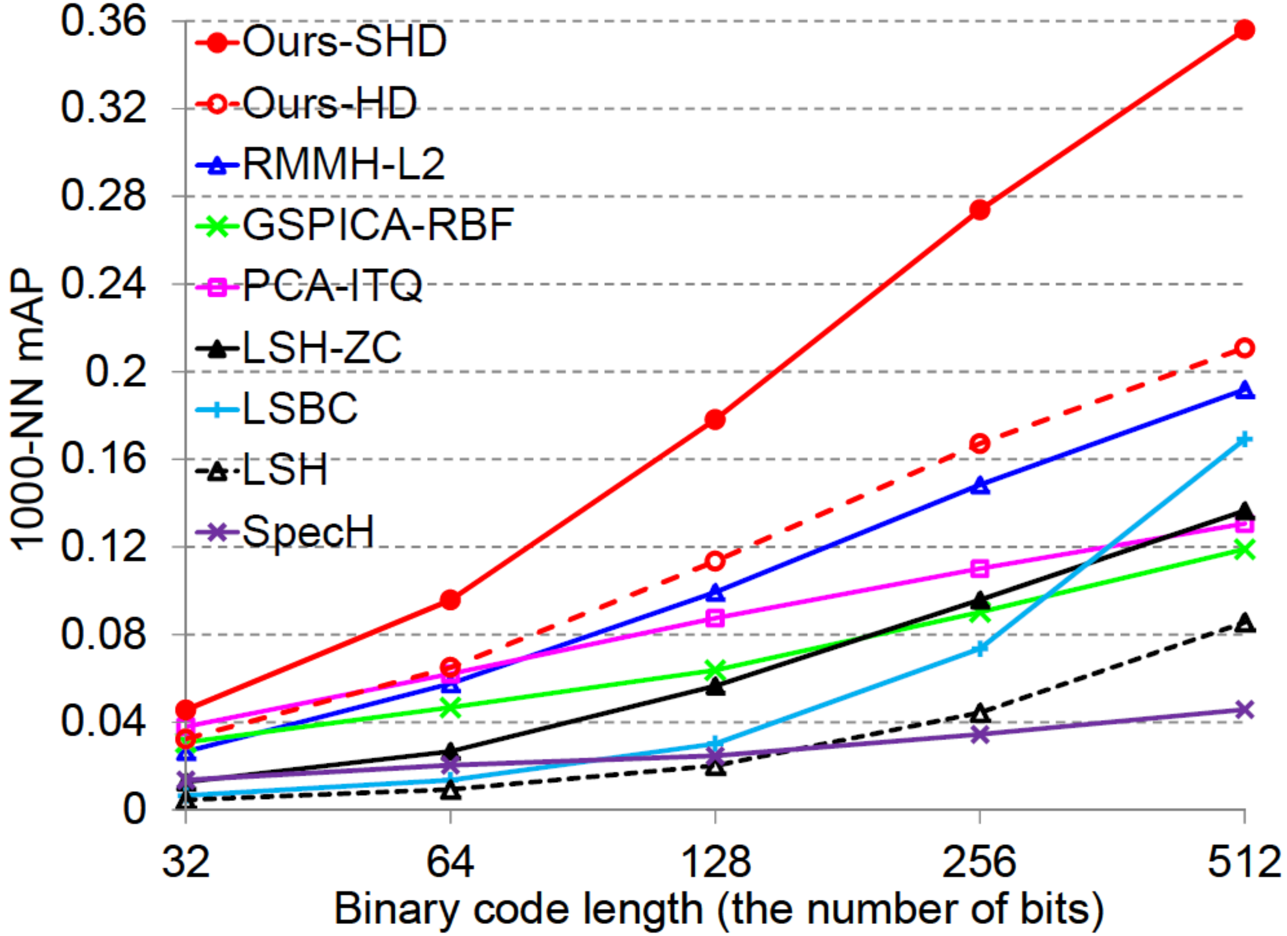
Results



384 dimensional 1 million GIST descriptors
Source codes are available

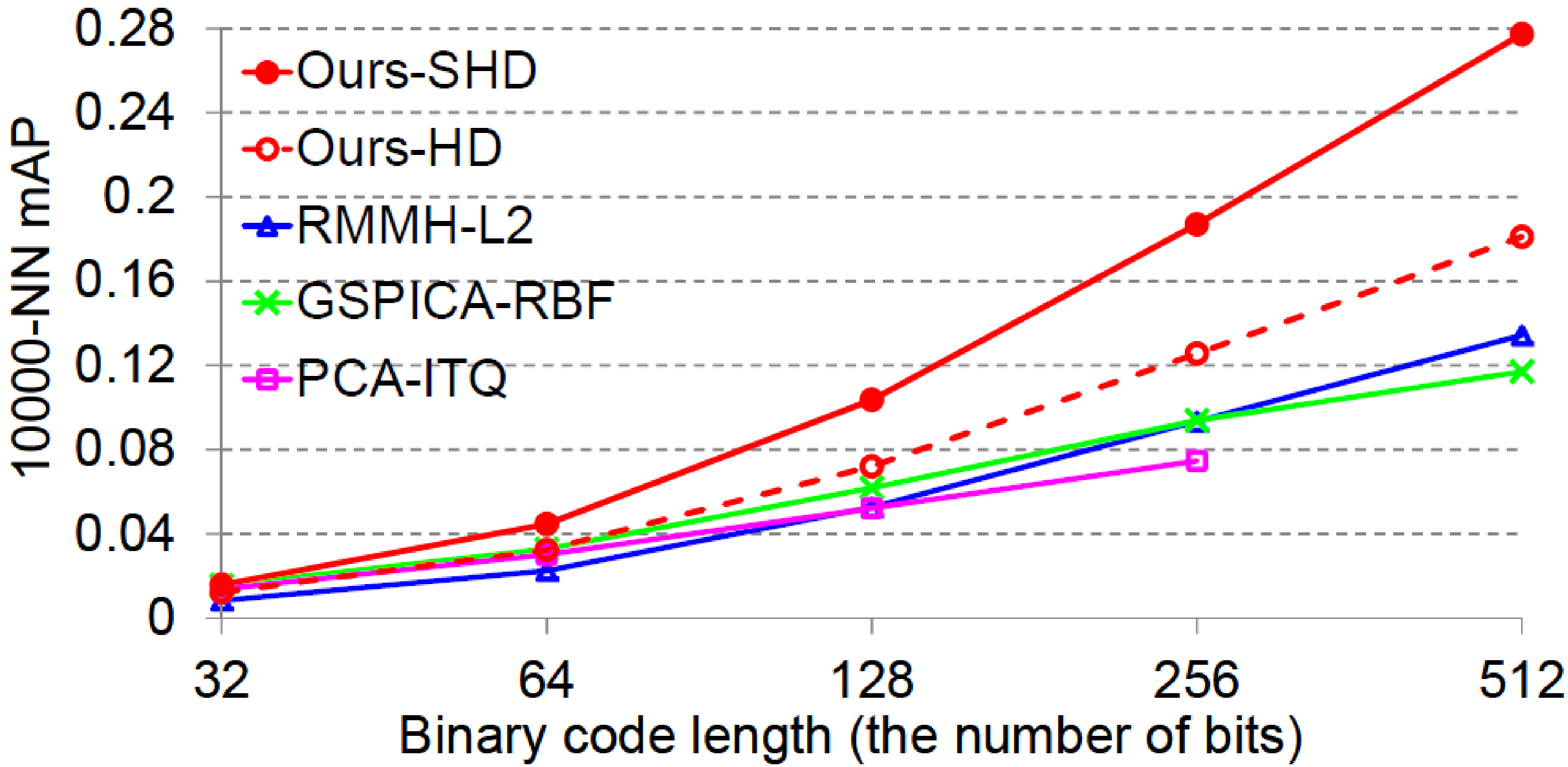
KAIST

Results



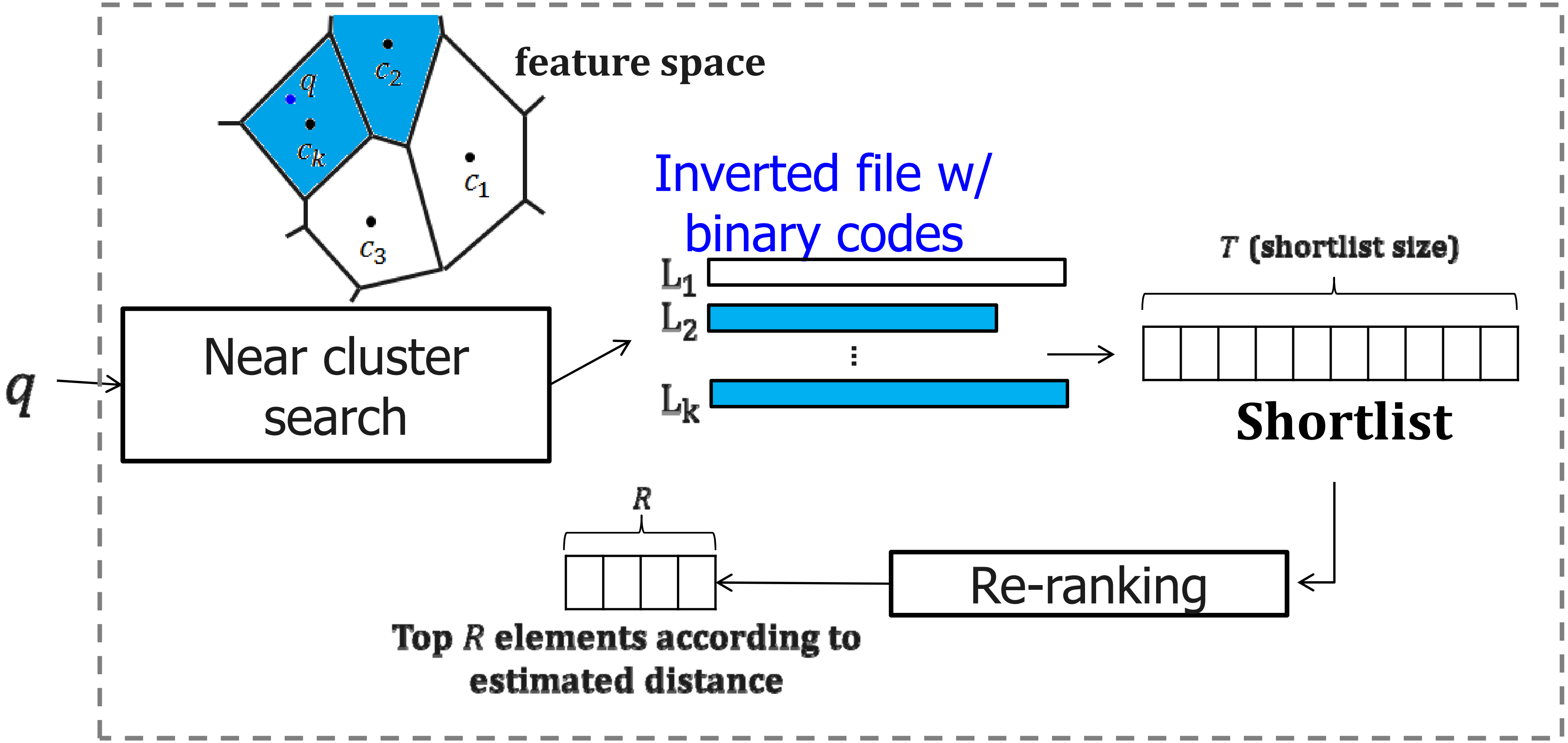
960 dimensional 1 million GIST descriptors

Results



384 dimensional 75 million GIST descriptors

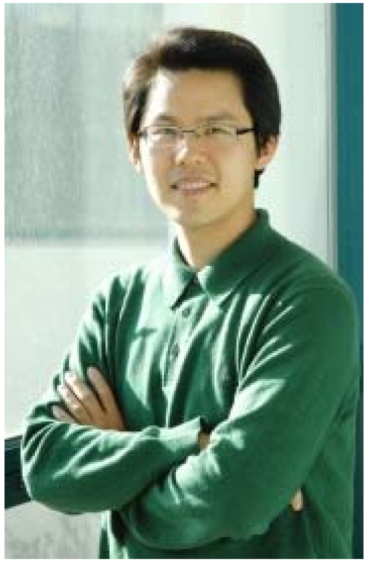
Summary



Ack.: Dr. Heo

Tutorial Schedule

- **2:00pm, basic materials, S. Yoon**
  - **Introduction on image search**
  - **Compact representations of images**



- **3:30pm, 15min break**

- **3:45pm, recent techniques, Z. Lin**
  - **Indexing scheme for large-scale image search**
  - **Applications**

- **5:15pm, the end**