
Web-Scale Image Search and Their Applications

Sung-Eui Yoon
KAIST

<http://sglab.kaist.ac.kr>

KAIST

The KAIST logo consists of the letters 'KAIST' in a bold, blue, sans-serif font. Below the text is a light blue, horizontal oval shape that serves as a shadow or base for the letters.

Project Guidelines: Project Topics

- **Any topics related to the course theme are okay**
 - **You can find topics by browsing recent papers**

Expectations

- **Mid-term project presentation**
 - **Introduce problems and explain why it is important**
 - **Give an overall idea on the related work**
 - **Explain what problems those existing techniques have**
 - **(Optional) explain how you can address those problems**
 - **Explain roles of each member**

Expectations

- **Final-term project presentation**
 - **Cover all the materials that you talked for your mid-term project**
 - **Present your ideas that can address problems of those state-of-the-art techniques**
 - **Give your qualitatively (or intuitive) reasons how your ideas address them**
 - **Also, explain expected benefits and drawbacks of your approach**
 - **(Optional) backup your claims with quantitative results collected by some implementations**
 - **Explain roles of each members**

A few more comments

- **Start to implement a paper, if you don't have any clear ideas**
 - **While you implement it, you may get ideas about improving it**

Speaker	Novelty of the project and idea (1 ~ 5)	Practical benefits of the method (1 ~ 5)	Completeness level of the project (1 ~ 5)	Total score (3 ~ 15)	Role of each student is clear and well balanced? (Yes or No)
XXX					
YYY					

Project evaluation sheet

You name:

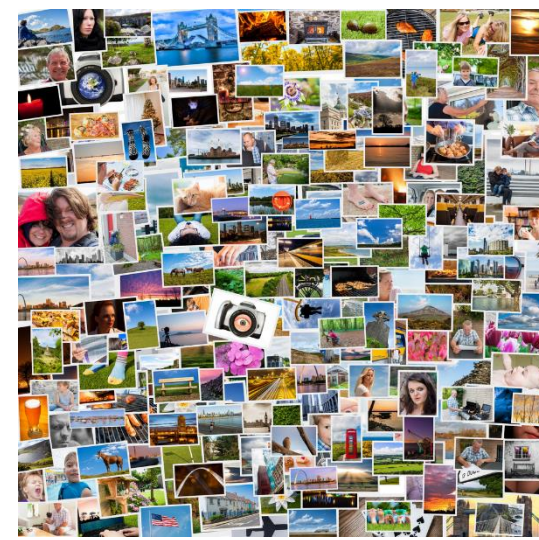
ID:

Score table: higher score is better.

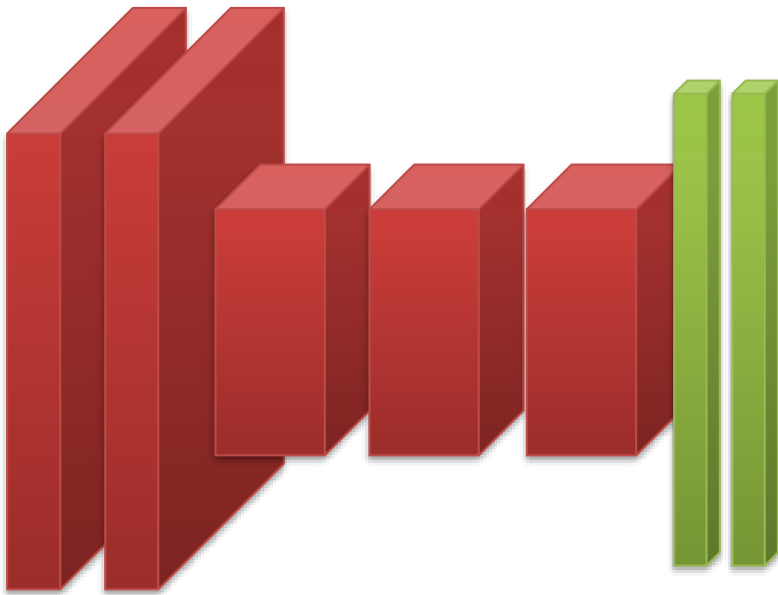
Web-Scale Visual Data and Novel Applications

- **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**

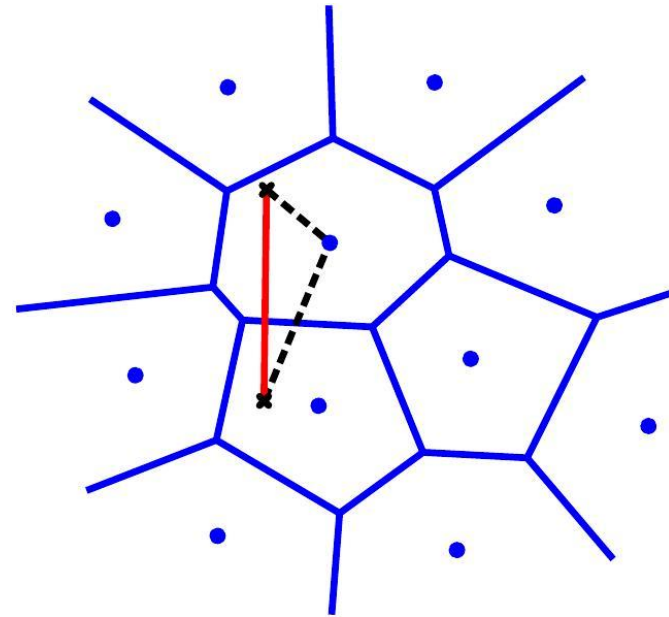
Review: Efficient Image Search



Deep Convolutional Neural Network

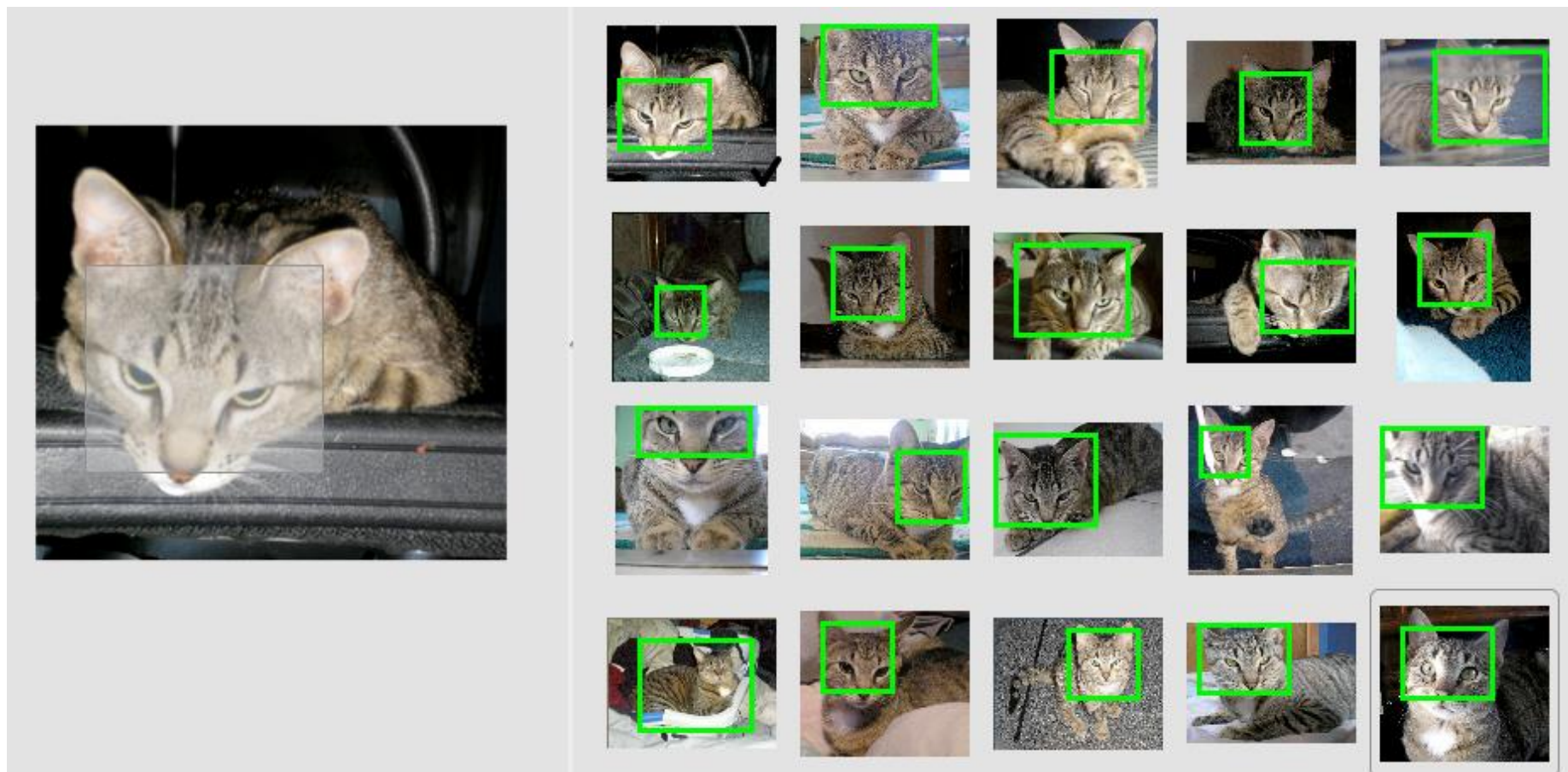


Distance Encoded Optimized PQ



Ack.: Zhe Lin

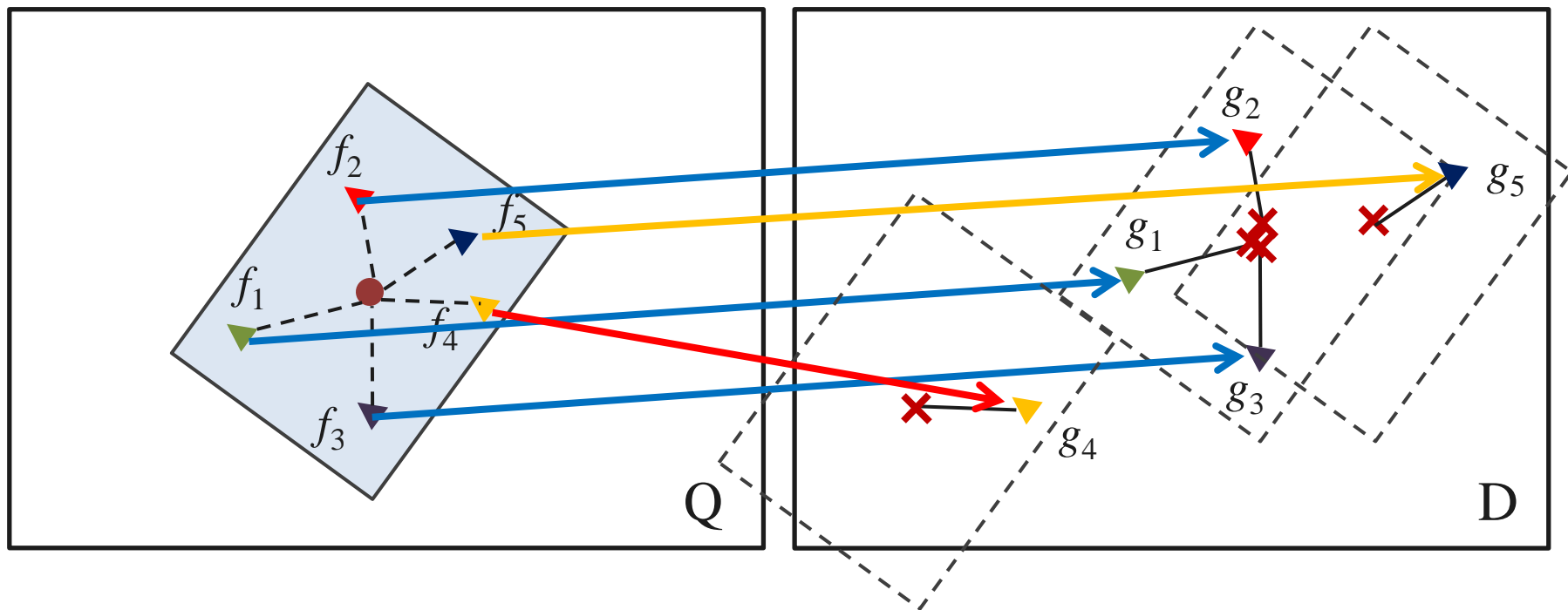
Object Retrieval and Localization



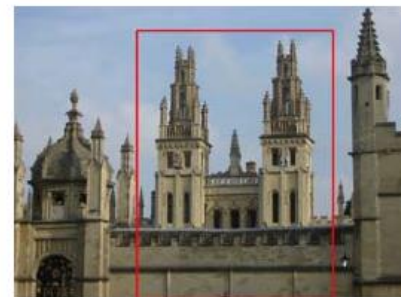
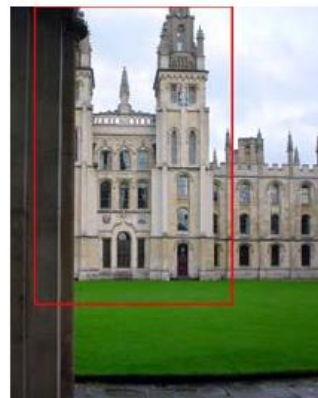
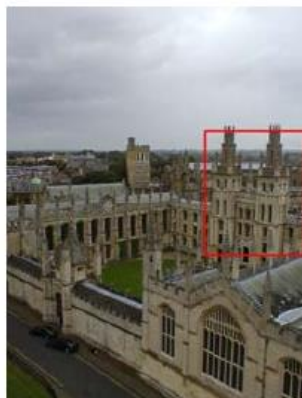
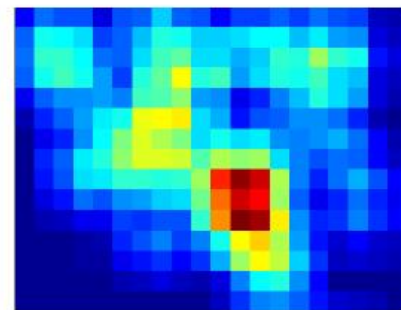
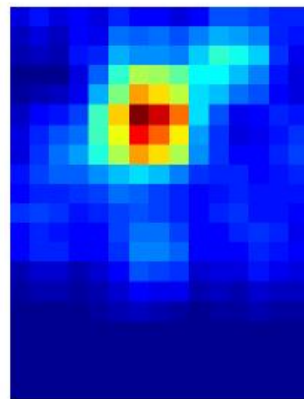
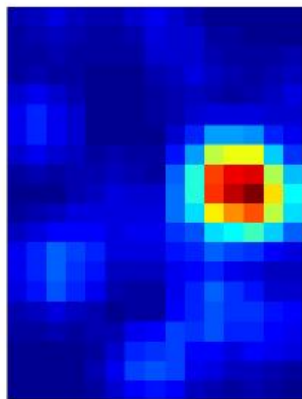
[X. Shen et al., CVPR 2012]

Object Retrieval and Localization

- Local correspondence voting for non-rigid object matching

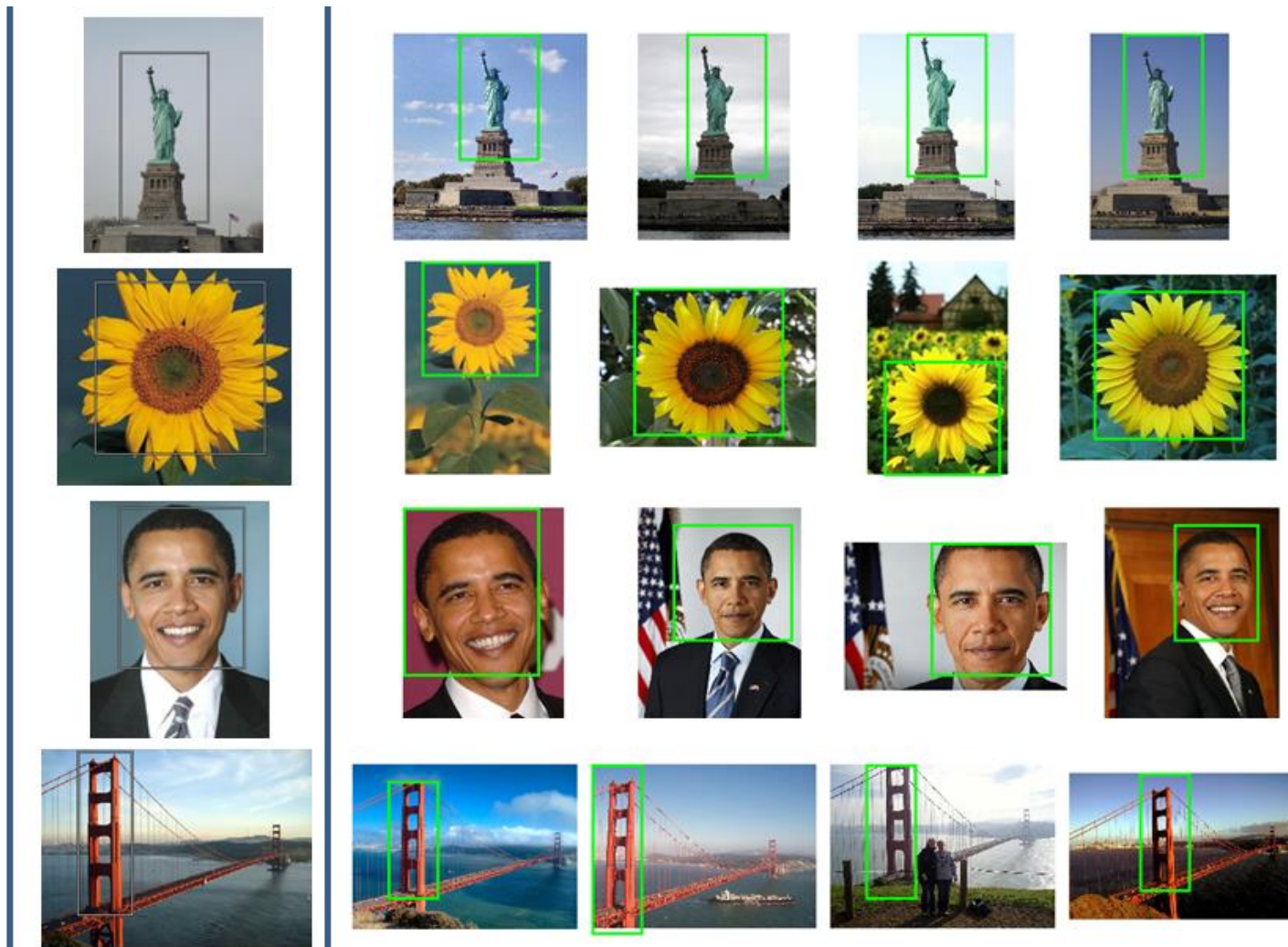


Object Retrieval and Localization



Examples of Voting Maps

Object Retrieval and Localization



Non-rigid cases

Product Image Recognition

[X. Shen et al., ECCV 2012]



Examples of product images in the database

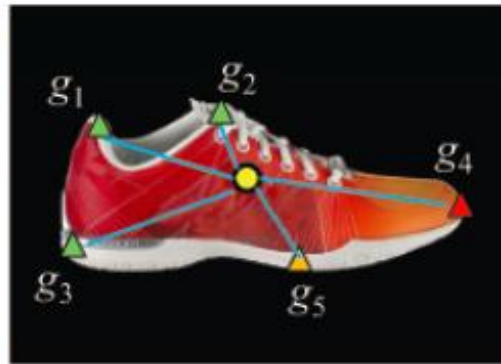


Examples of query images taken by mobile phones

Product Image Recognition



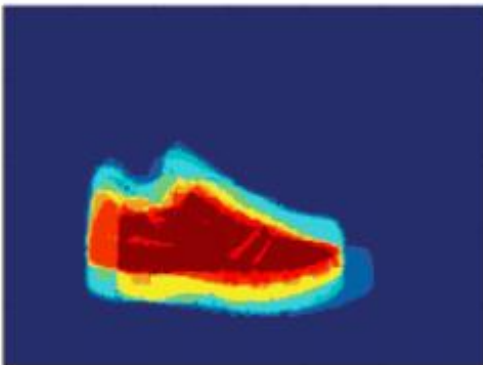
a) A query



b) DB image



c) A vote map



d) Aggregated voting maps



e) Tri map



f) Segmented result

Product Image Recognition



Images

Support map

Extraction

GrabCut w/ manual init.

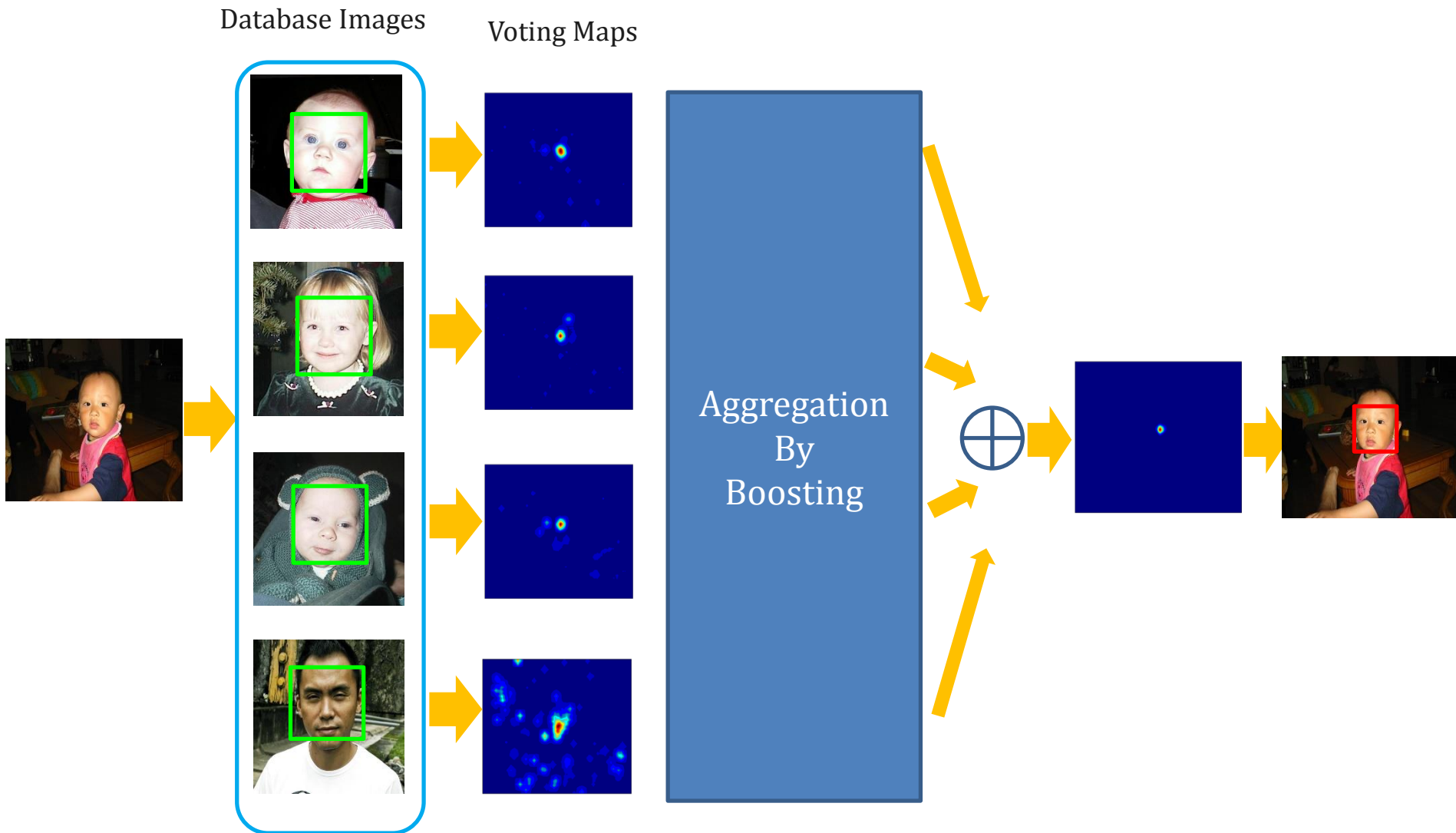
Face Detection by Image Retrieval



[X. Shen et al., CVPR 2013]

[H. Li et al., CVPR 2014]

Face Detection by Image Retrieval

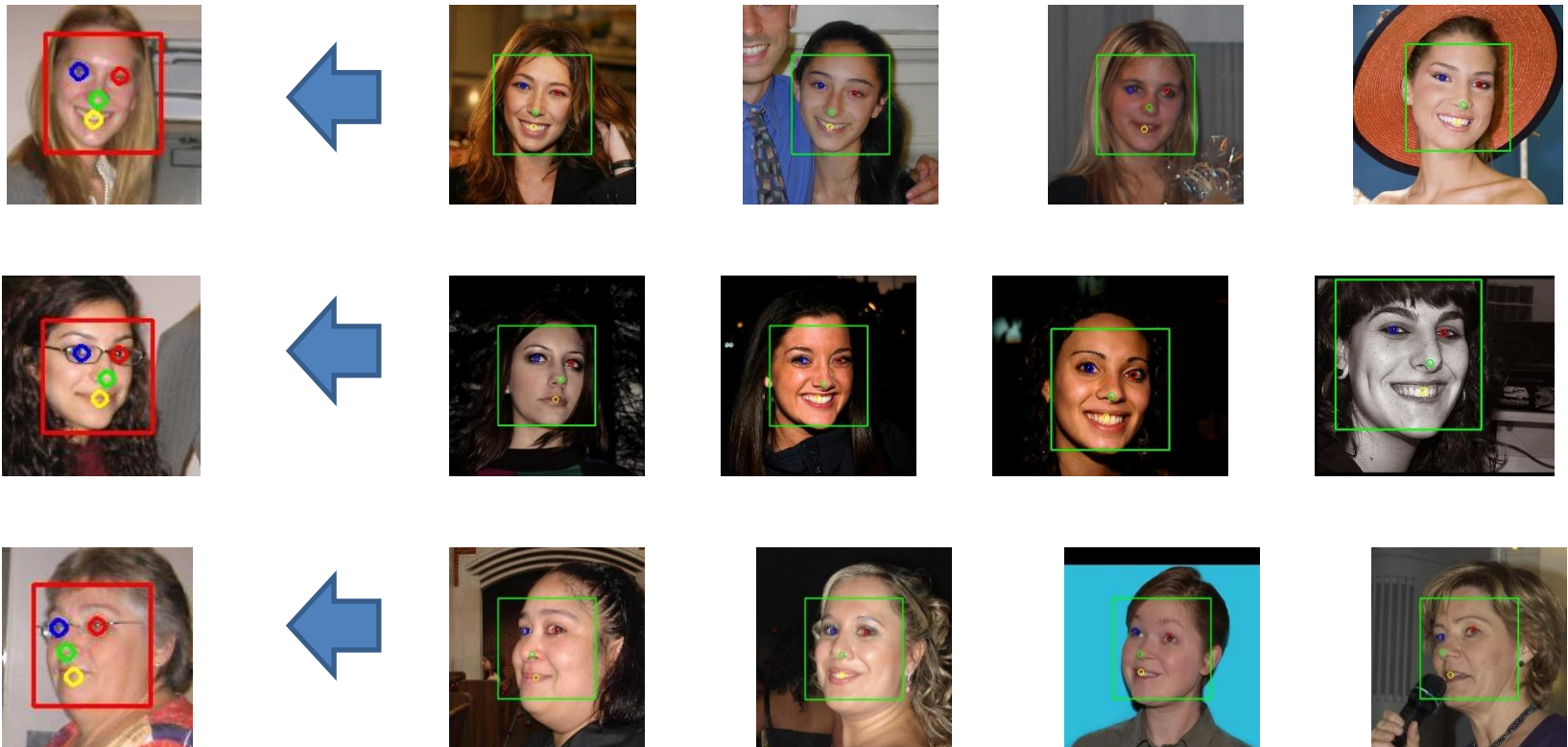


Face Detection by Image Retrieval



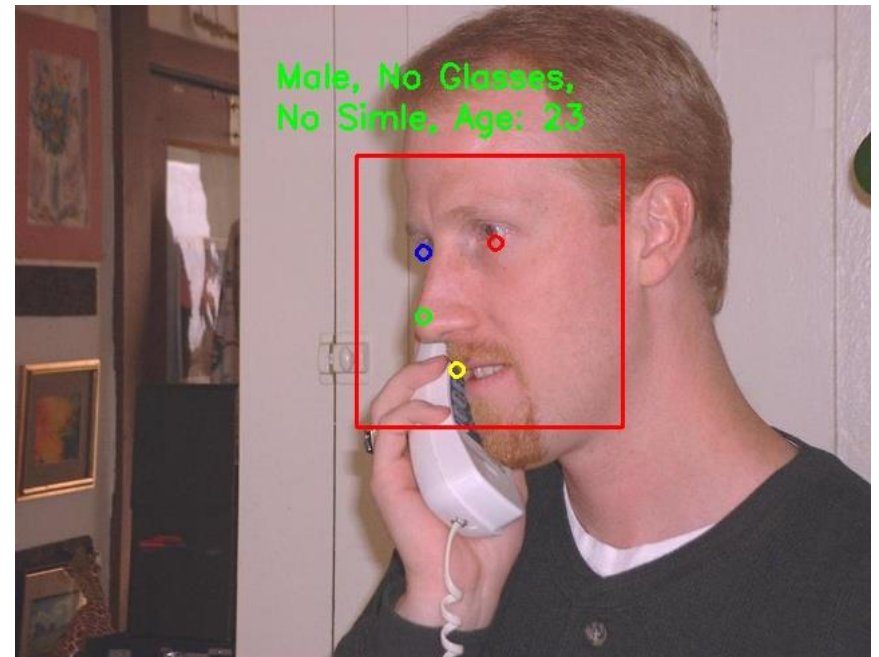
Example detection results

Facial Attribute Recognition



transfer landmark, pose, age, gender, expression...

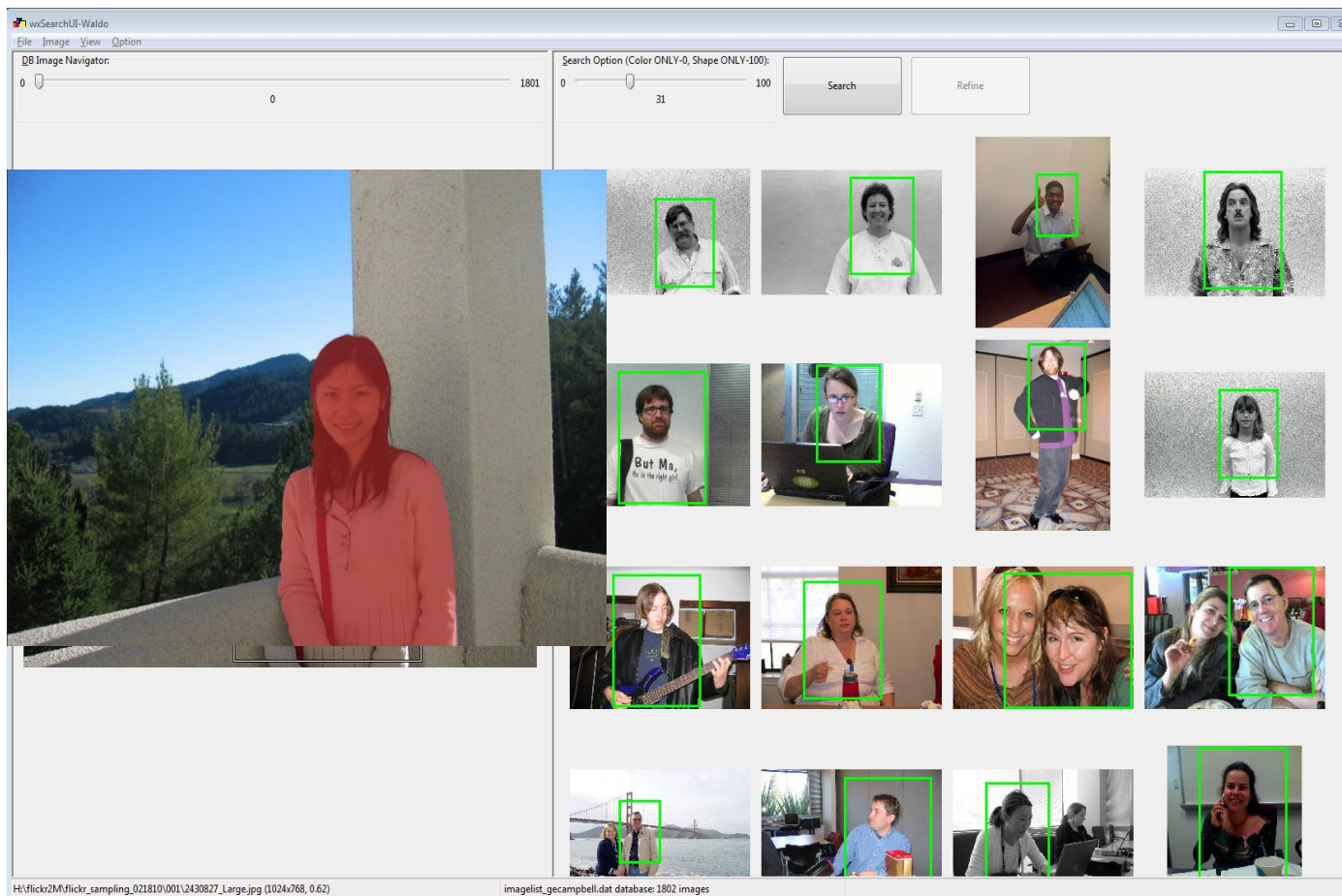
Facial Attribute Recognition



Data-Driven Object Segmentation

[J. Yang et al. CVPR 2014]

Find seg. examples and transfer

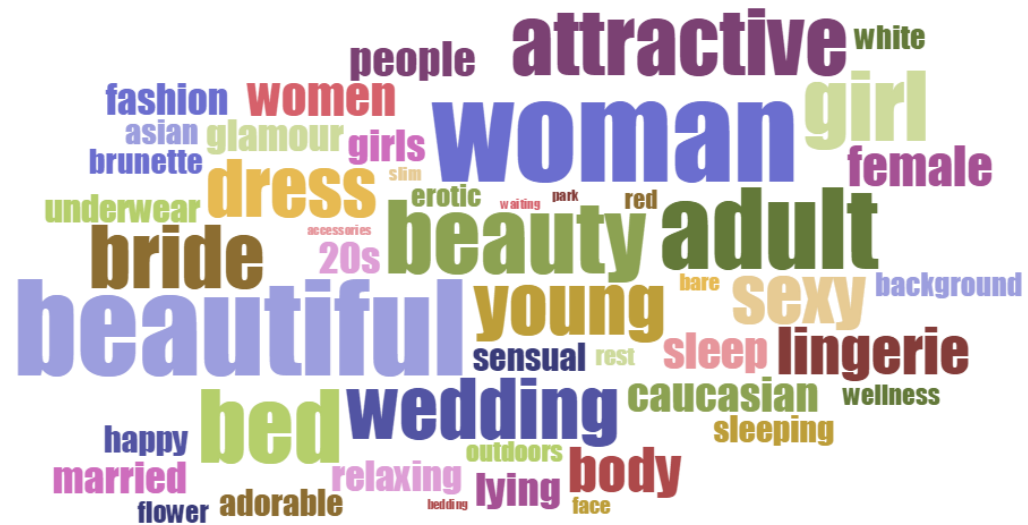


Data-Driven Automatic Cropping

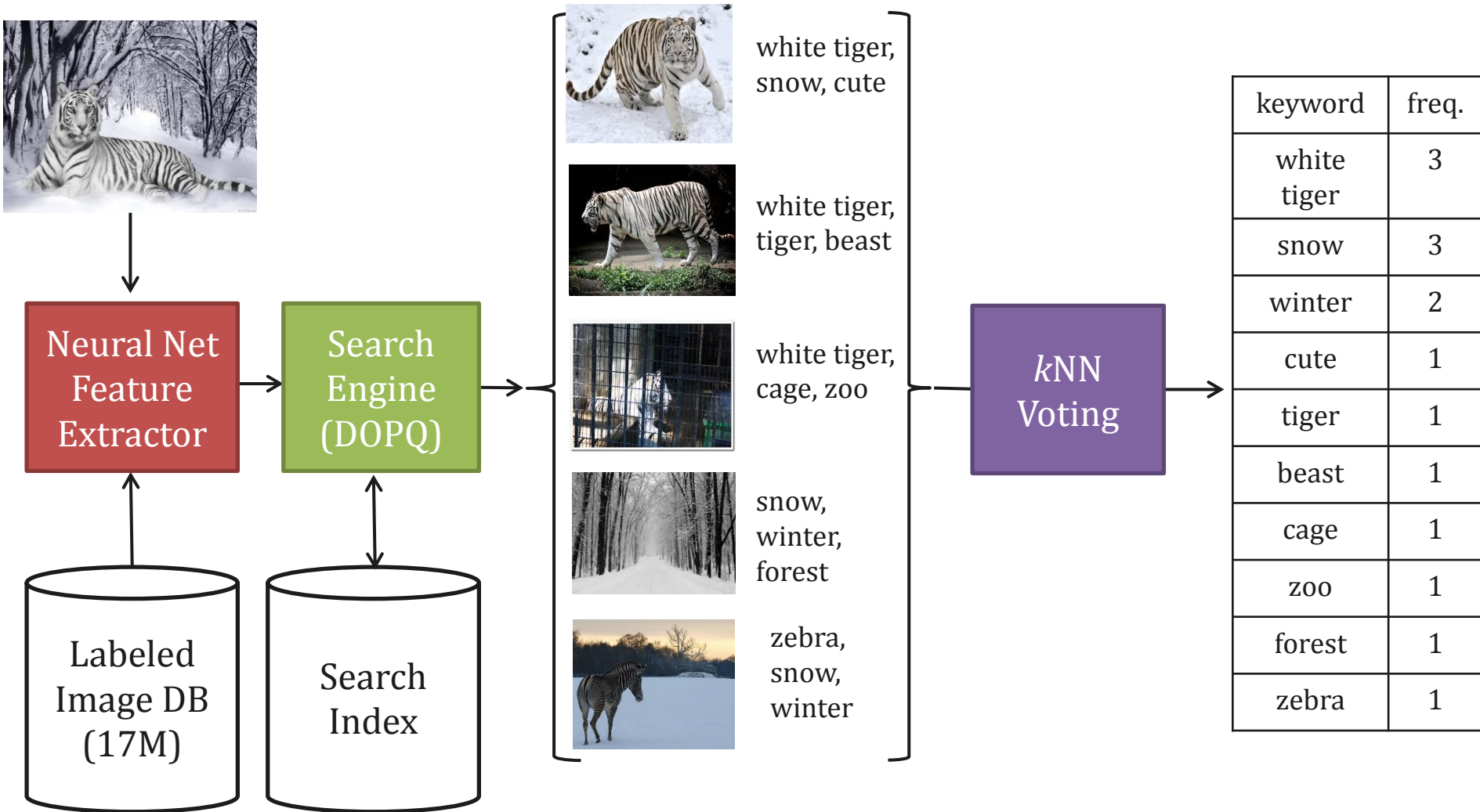
[A. Samii et al. CGF 2015]



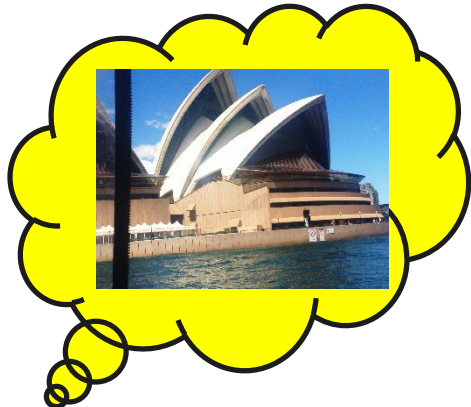
Automatic Image Tagging




Deep-kNN Tagging System



Free-Text Image Search



 sydney opera house

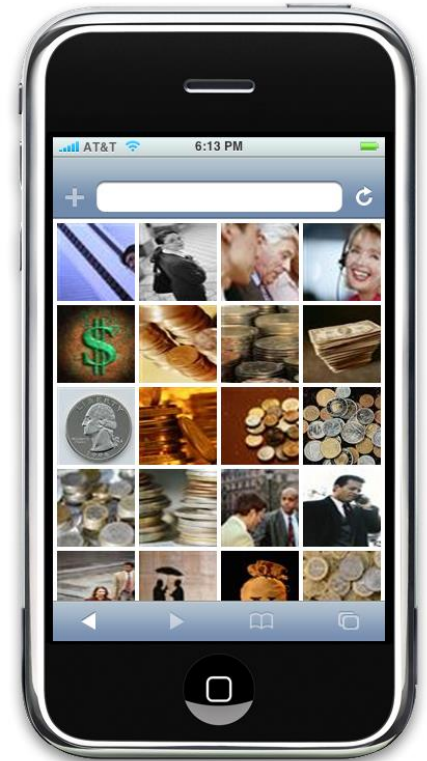
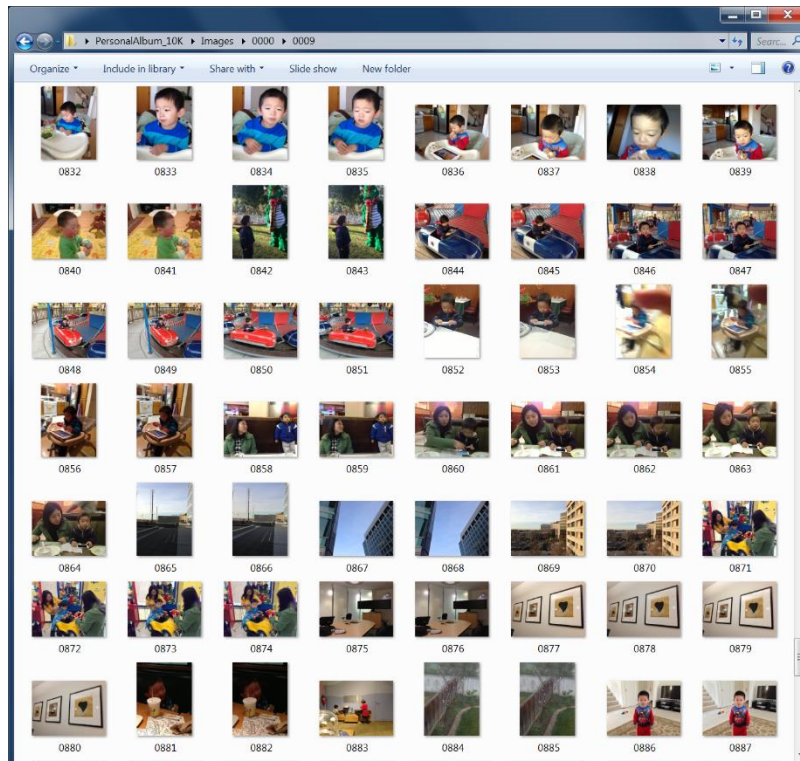


Image Recommendation: Collaborative Feature Learning from Social Media

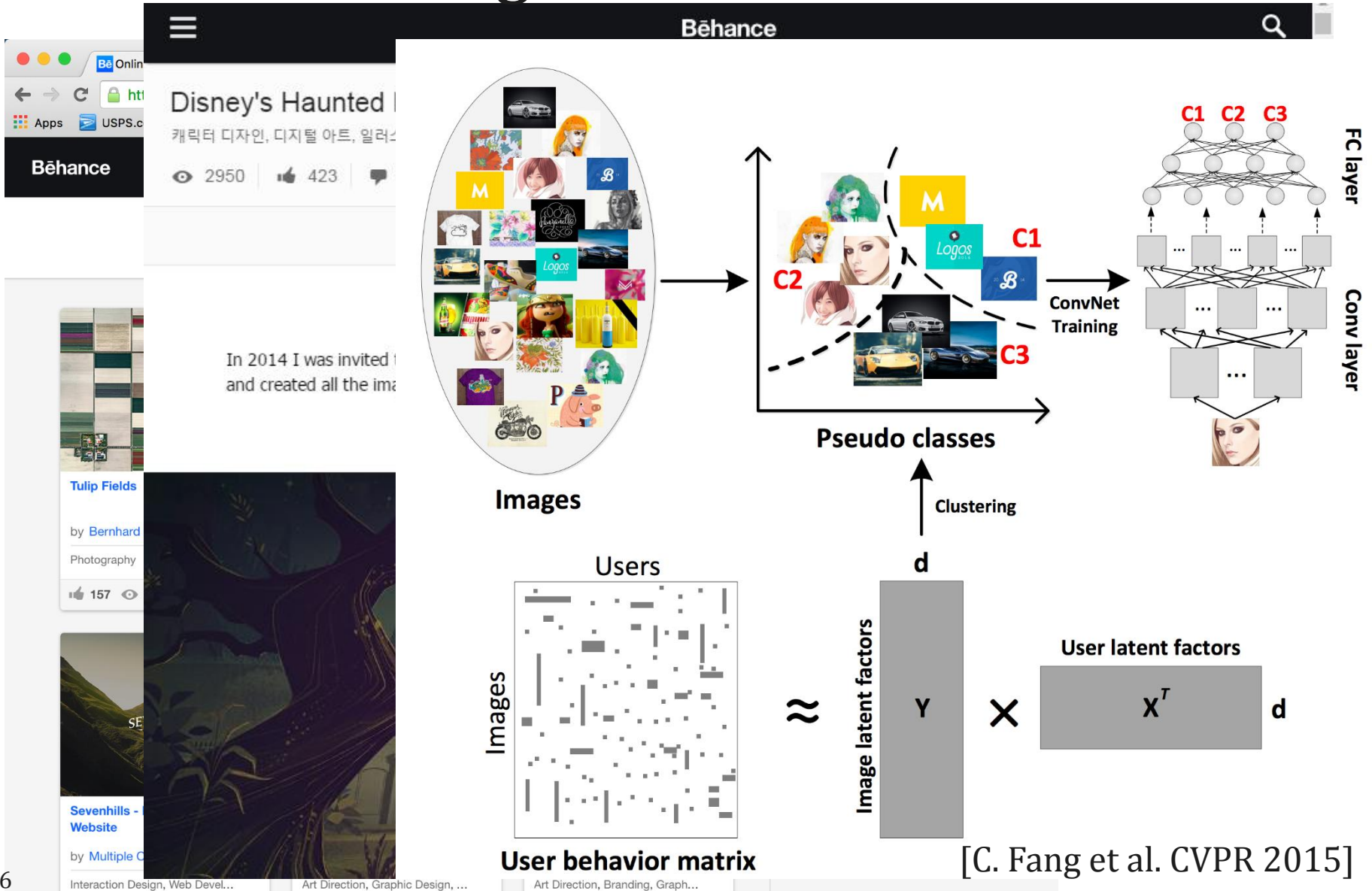
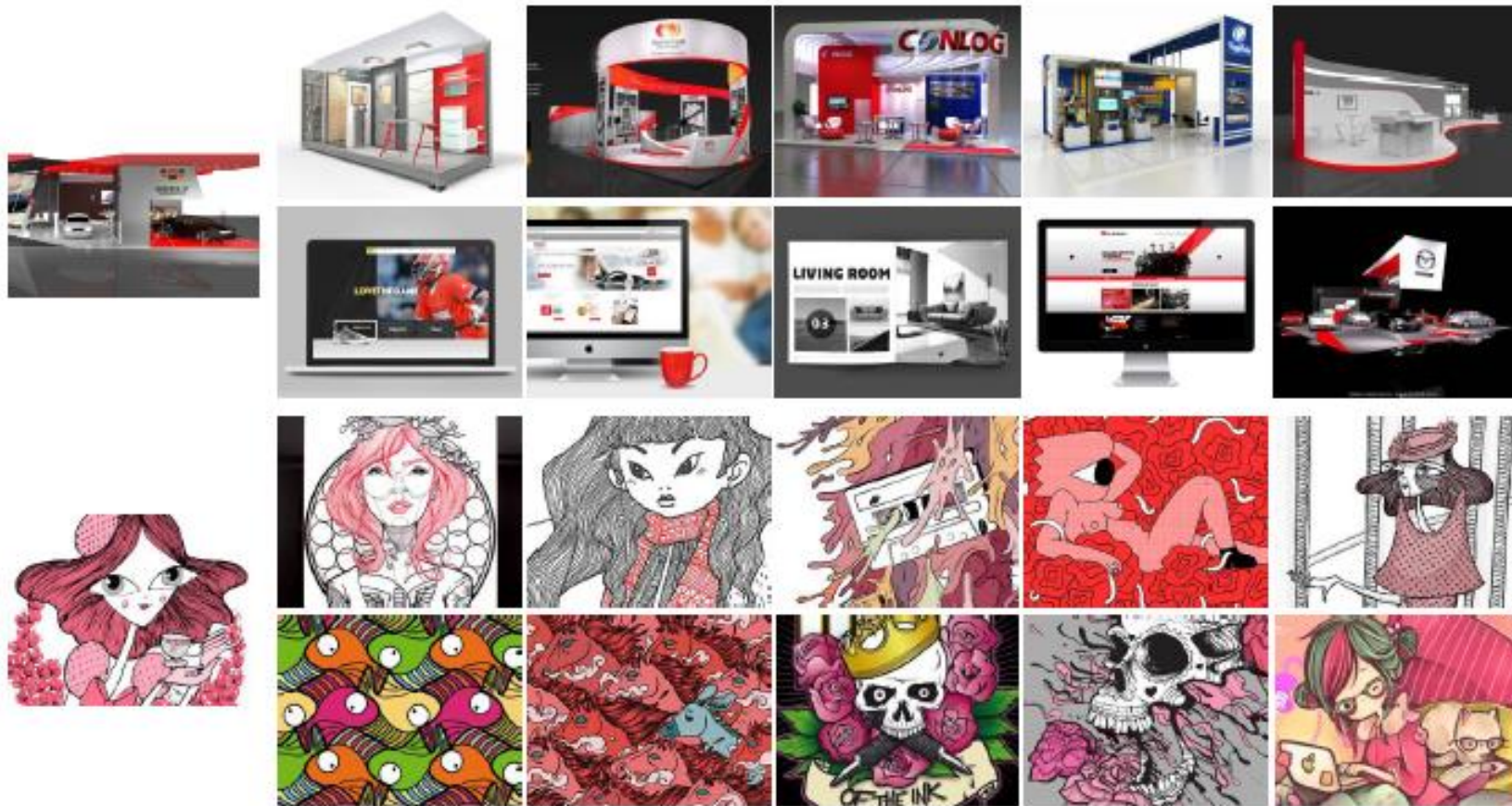


Image Recommendation: Collaborative Feature Learning from Social Media



ours

Features
from
ImageNet

Image Retrieval based Image Watermarking [IWDW11]

- Exhaustive watermark matching
 - Sequential one-to-one comparison
 - Time-consuming job



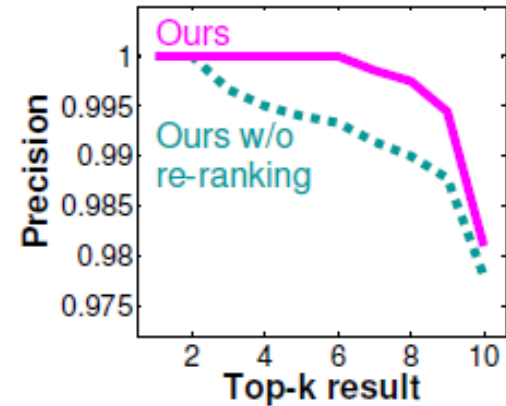
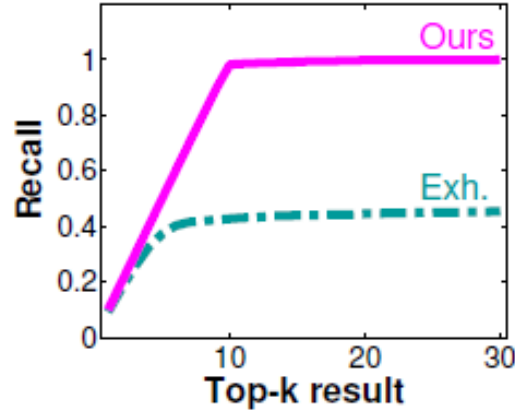
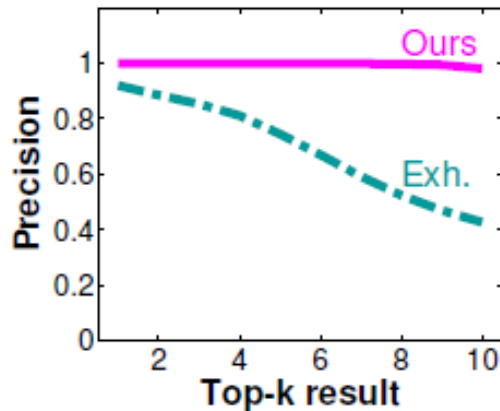
- **Image Retrieval based Image watermarking (IRIW)**

- Reduce search domain by image search
- Achieve performance enhancement



Result

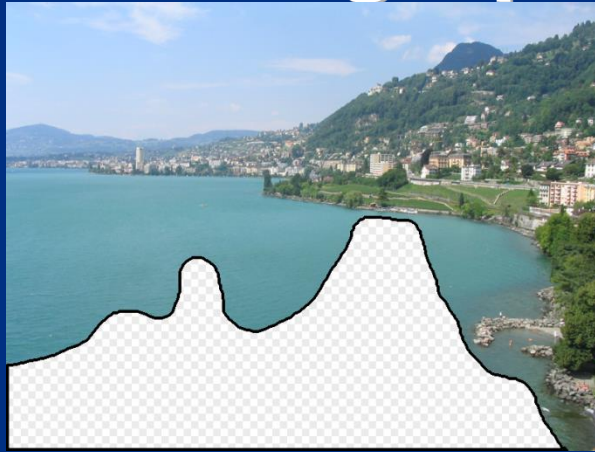
- Accuracy (100 tests)



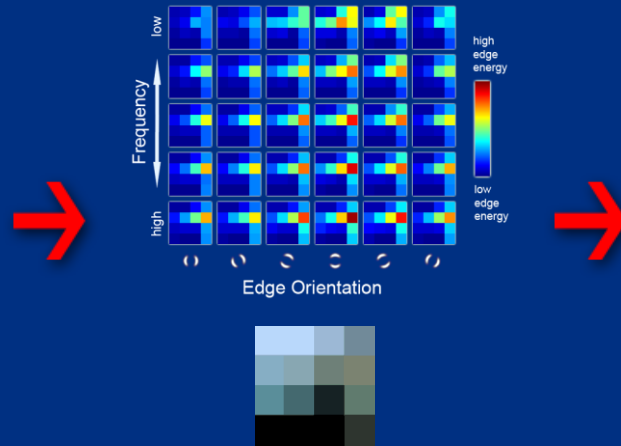
$$\text{Precision} = \frac{\# \text{of}(I \cap R)}{\# \text{of}(R)} \quad I : \text{ground truth set}$$

$$\text{Recall} = \frac{\# \text{of}(I \cap R)}{\# \text{of}(I)} \quad R : \text{result set}$$

Scene Completion using Millions of Photographs [SIG. 07]



Input image



Scene Descriptor



Image Collection



20 completions



Context matching
+ blending



200 matches

Results

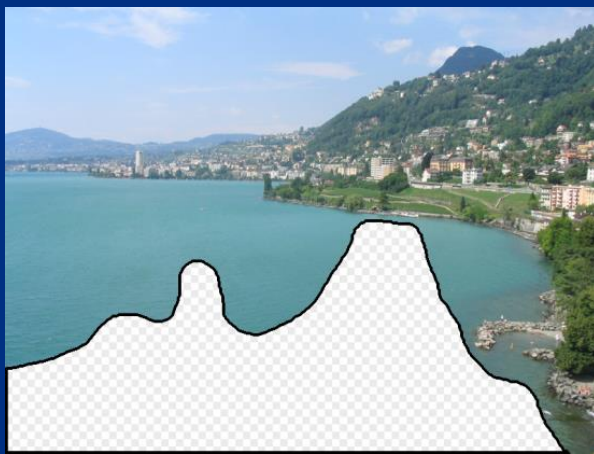






Photo Tourism [SIG. 11]





15,464



37,383

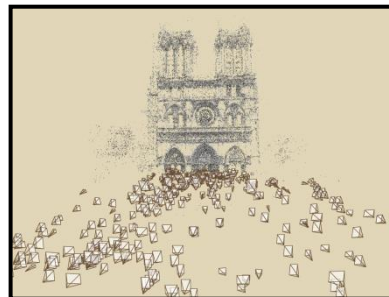
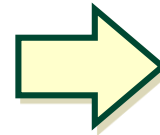
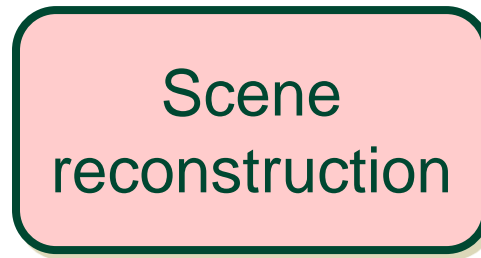
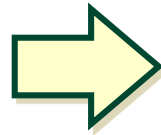


76,389

Photo Tourism overview



Input photographs



Relative camera positions and orientations
Point cloud
Sparse correspondence



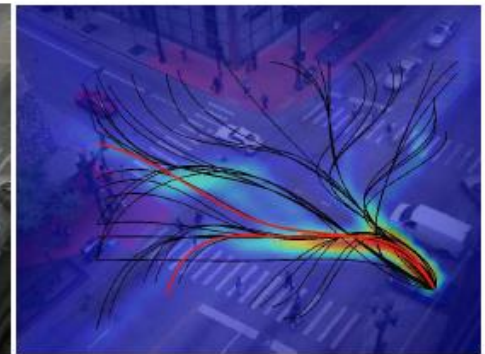
Photo Explorer

Visual Prediction

- **Predict possible actions by:**
 - **Identify similar patches in the training videos based on NNS**
 - **Propagating them in the query image**



(a) Original Image



(b) Prediction Heatmap

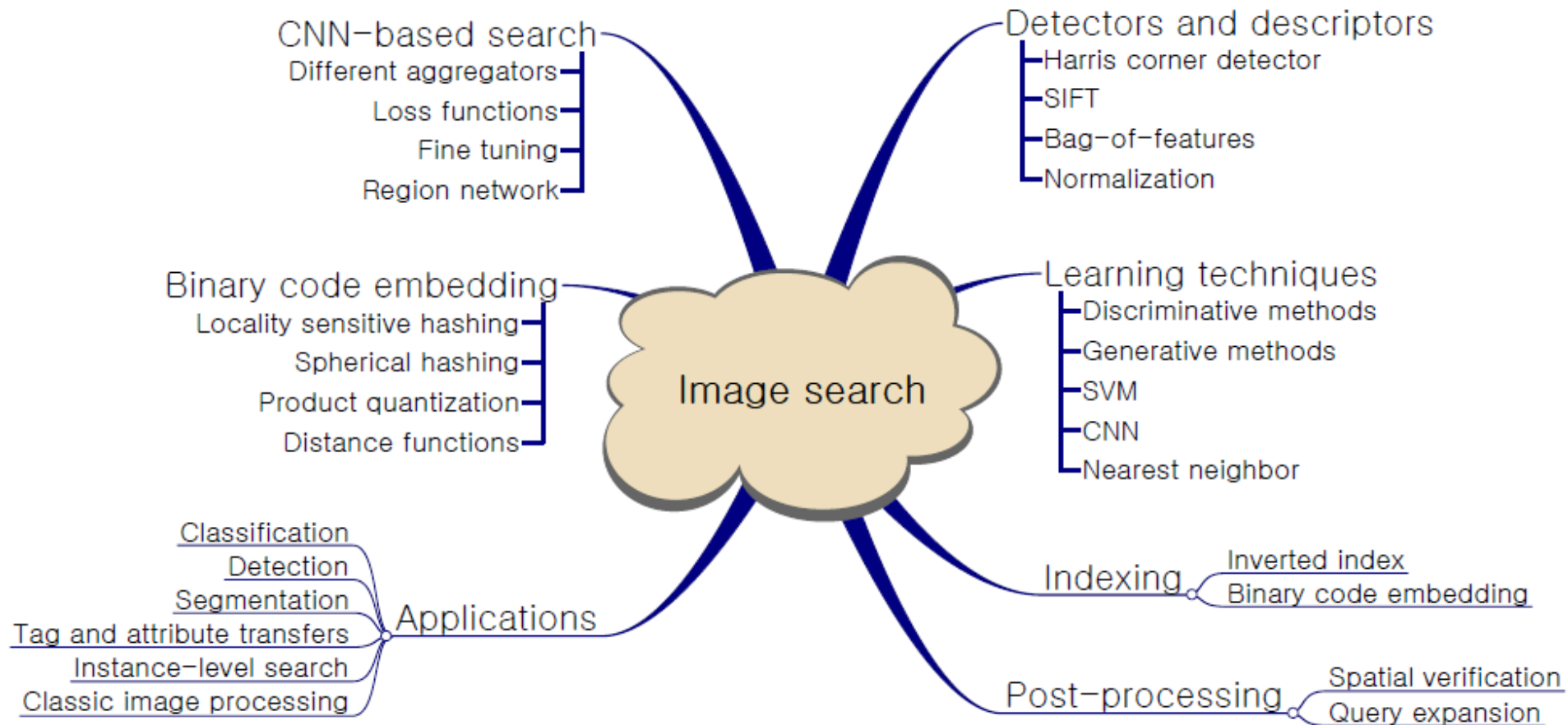


(c) Predicted Path -1



(d) Predicted Path -2

Summary



Conclusions

- **Visual data are more widely used for various communication and are thus associated at Web**
- **Processing them requires scalable algorithms**
- **Web-scale visual data can enable new applications**
- **Examples**
 - **Photo tourism**
 - **Scene completion**
 - **Image-retrieval based image watermarking**