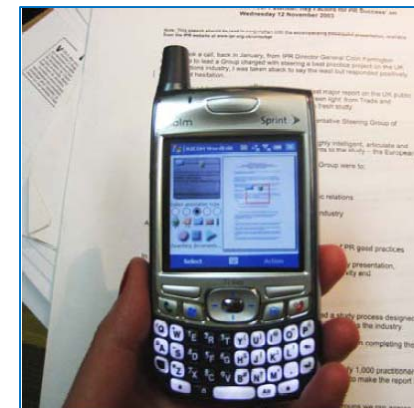


Visual Search

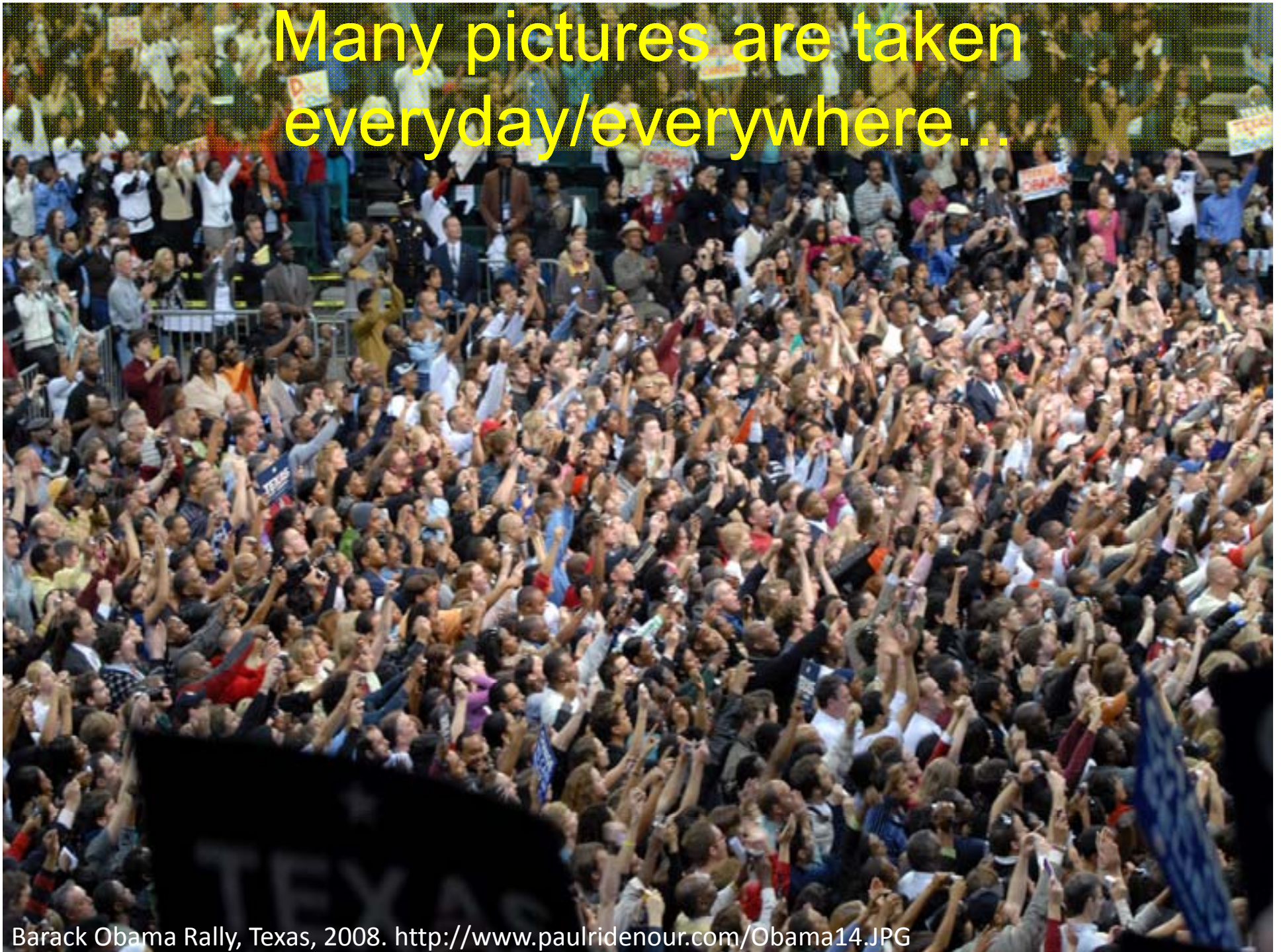
Shih-Fu Chang

Department of Electrical Engineering
Columbia University

Boston University ECE Lecture, September 2010

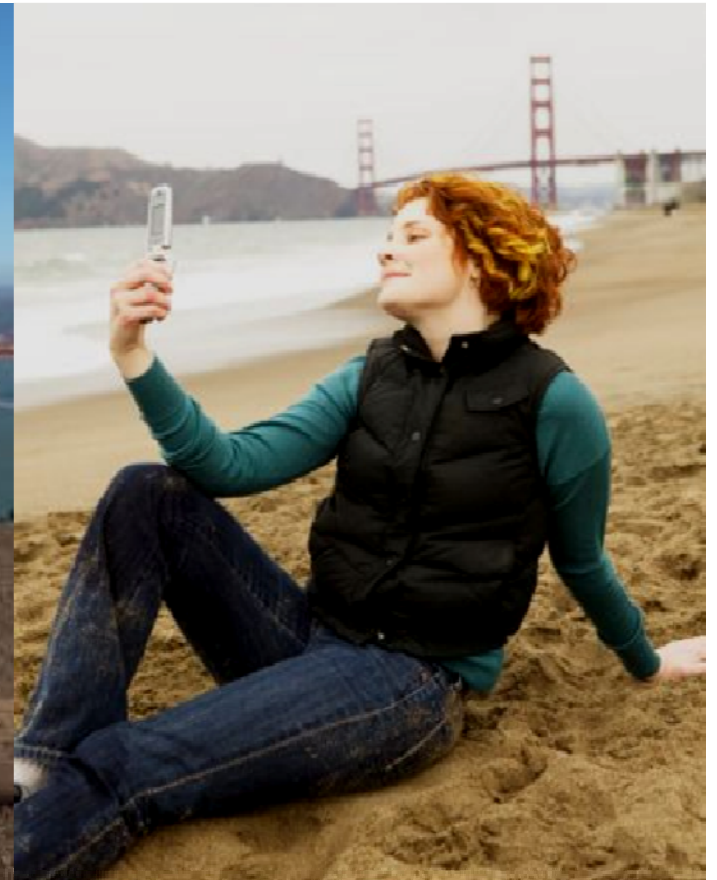


Many pictures are taken
everyday/everywhere...



Barack Obama Rally, Texas, 2008. <http://www.paulridenour.com/Obama14.JPG>

Taking pictures everyday/everywhere...



<http://cache1.asset-cache.net/xc/200545511-001.jpg?v=1&c=NewsMaker&k=2&d=CED48661B87C5DBFF5EA7A9E4218F9EE1F6F6178A68B340C>

<http://cache4.asset-cache.net/xc/200483959-001.jpg?v=1&c=NewsMaker&k=2&d=EDF6F2F4F969CEBD9A551141E9DF12C0E47ADEC7FBB32857DF0450484851C07200123AA3B5A18ED0>

Taking pictures everyday/everywhere...



miami.eater.com



2.bp.blogspot.com



digital-photography-school.com

But we are not good at organizing them ...

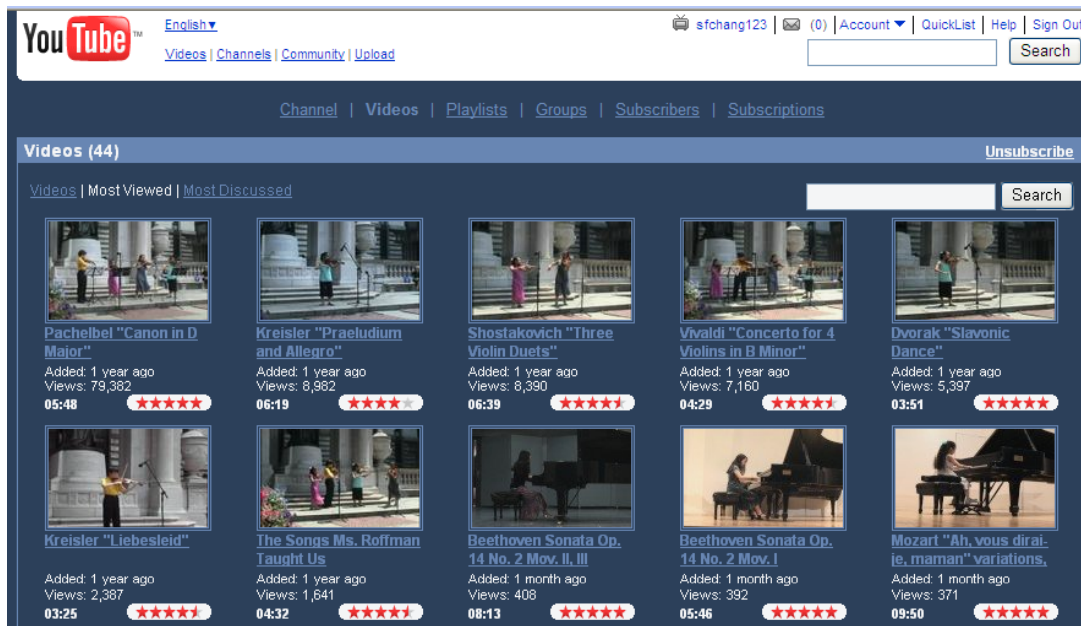


- Most photos and videos remain in shoebox or e-shoebox
- But we love to share them ...

Example:

My family's video channel

- Sharing ~ 100 videos on Youtube
- This single video has been viewed >130,000 times
 - **More popular than all of my published papers!!!**



The screenshot shows a YouTube channel page for 'sfchang123'. The channel has 44 videos. The top navigation bar includes 'English', 'Videos', 'Channels', 'Community', and 'Upload'. The main content area displays a grid of 10 video thumbnails, each with a title, duration, and star rating. The videos are:

Title	Duration	Views	Added	Rating
Pachelbel "Canon in D Major"	05:48	79,382	1 year ago	★★★★★
Kreisler "Praeludium and Allegro"	06:19	8,982	1 year ago	★★★★★
Shostakovich "Three Violin Duets"	06:39	8,390	1 year ago	★★★★★
Vivaldi "Concerto for 4 Violins in B Minor"	04:29	7,160	1 year ago	★★★★★
Dvorak "Slavonic Dance"	03:51	5,397	1 year ago	★★★★★
Kreisler "Liebesleid"	03:25	2,387	1 year ago	★★★★★
The Songs Ms. Roffman Taught Us	04:32	1,641	1 year ago	★★★★★
Beethoven Sonata Op. 14 No. 2 Mov. II, III	08:13	408	1 month ago	★★★★★
Beethoven Sonata Op. 14 No. 2 Mov. I	05:46	392	1 month ago	★★★★★
Mozart "Ah, vous dirai-je, maman" variations	09:50	371	1 month ago	★★★★★



The screenshot shows a YouTube video player for the video 'Pachelbel "Canon in D Major"'. The video is playing at 1:40 / 5:48. The video has a rating of 5 stars (★★★★★) based on 191 ratings and has been viewed 79,390 times. The video content shows a group of musicians performing the Canon in D Major on a stage.

video

Challenge after capture and upload

It will be nice to tag ...



Performance, classical music, ensemble, concert; Pachelbel Canon in D Major, romantic; young musicians, girls, boys; outdoor, Bryant park in NYC, statue, columns, stage, flowers; Instruments, audio equipments, violins; . . .

- **Tagging is boring and hard**
- **Each uploaded photo has only 0.97 tag on average**
(Naaman, Yahoo!/Rutgers)

Scarce Tags → Faulty Search Engines



















“Manhattan Cruise”

Google™ [Advanced Image Search](#)
[Moderate SafeSearch is on](#) [Preferences](#)

Images Showing: Results 1 - 18 of about 676,000 (0.13 seconds)

[Manhattan Vacation](#) [Manhattan Cruise](#) [New York Harbor Cruises](#) Sponsored Links

www.Expedia.com **Manhattan** is Calling. Book Your Vacation at Expedia Today. www.viator.com/new-york-cruises Book 4-star rated New York City sightseeing **cruises** on Viator. AllNewYorkTours.com/Cruises Harbor **Cruises** in NYC. Sightseeing & Dinner, Sunsets, Skyline Views!

 <p>... Manhattan skyline. 800 x 600 - 95k wirednewyork.com [More from wirednewyork.com]</p>	 <p>Manhattan Dinner Cruises 360 x 276 - 18k - jpg www.manhattan-dinner-cruises.com</p>	 <p>> Manhattan Cruise Ship.jpg 722 x 480 - 110k - jpg www.pbase.com</p>	 <p>Cape Liberty Cruise Port in Bayonne ... 800 x 600 - 72k wirednewyork.com</p>	 <p>Cruises thrive at Manhattan, ... 364 x 425 - 116k - jpg www.newsday.com</p>	 <p>... Cruise departed their Manhattan ... 420 x 300 - 69k - jpg www.poplife.biz</p>
 <p>Manhattan cruise 500 x 375 - 105k - jpg www.flickr.com</p>	 <p>... Aboard A Manhattan Harbor Cruise 400 x 300 - 23k www.ny1.com</p>	 <p>... father tom cruise in manhattan 500 x 318 - 31k www.theinsider.com</p>	 <p>... Manhattan Scenic Cruise, ... 300 x 240 - 62k - jpg www.xperiencedays.com</p>	 <p>Manhattan Cruise terminal, New York 450 x 335 - 53k - jpg www.skyscrapercity.com [More from www.nycruise.com]</p>	 <p>PortOfManhattan 464 x 348 - 34k - jpg www.cruisedirectorsinc.com</p>
 <p>Manhattan Cruise Terminal 320 x 363 - 4k - gif www.nycruiseinfo.com</p>	 <p>Cruise ship Zenith of Celebrity ... 800 x 600 - 73k wirednewyork.com</p>	 <p>... the Manhattan Bridge [corrected] ... 480 x 640 - 63k famousankles.com</p>	 <p>The Manhattan Cruise Terminal, ... 450 x 348 - 98k - gif www.panynj.gov</p>	 <p>Suri Cruise is Made For Manhattan 300 x 300 - 31k - jpg justjared.buzznet.com</p>	 <p>View of Manhattan from Cruise Ship, ... 800 x 594 - 46k - jpg www.panoramio.com</p>



















Scarce Tags → Faulty Search Engines

“Cruise ship in Manhattan”

Google [Advanced Image Search](#)
[Moderate SafeSearch is on](#) [Preferences](#)

Images Showing: Results 1 - 18 of about 236,000 (0.17 seconds)

Manhattan Vacation [www.Expedia.com](#) Manhattan is Calling. Book Your Vacation at Expedia Today. **Cruise In Manhattan** [www.viator.com/new-york-cruises](#) Book 4-star rated New York City sightseeing cruises on Viator. **Manhattan Cruises** [www.CitySightsNY.com/Cruises](#) Hop on Board - New York City Harbor **Manhattan Cruises**. Save \$5 Online. Sponsored Links

 <p>Cape Liberty Cruise Port in Bayonne ... 800 x 600 - 72k wirednewyork.com [More from www.wirednewyork.com]</p>	 <p>Cruise ship Zenith of Celebrity ... 800 x 600 - 73k wirednewyork.com</p>	 <p>Cruise ship strikes Manhattan pier 450 x 258 - 37k - jpg www.nydailynews.com</p>	 <p>cruise ship manhattan 500 x 375 - 98k - jpg www.flickr.com [More from flickr.com]</p>	 <p>"Carnival Miracle" Cruise Ship ... 500 x 375 - 137k - jpg www.flickr.com</p>	 <p>S.S. Manhattan Vintage Passenger ... 400 x 400 - 32k www.zazzle.com</p>
 <p>... be on a cruise ship, in Hell, ... 339 x 510 - 88k - jpg ilovemonsterbash.blogspot.com</p>	 <p>Pedicabs from Manhattan ... 320 x 320 - 26k - jpg www.bizbash.com</p>	 <p>... City Cruise - cruise ships in ... 450 x 450 - 40k - jpg www.dezeen.com</p>	 <p>... Smooth Cruise around Manhattan 485 x 322 - 60k www.smoothjazznewyork.com</p>	 <p>Cruise Ship Crashes Into Manhattan ... 400 x 300 - 15k - jpg www.ny1.com</p>	 <p>Cruise ship leaves Manhattan ... 3446 x 414 - 307k - jpg www.panoramio.com</p>
 <p>Cruise Ship Crashes Into Manhattan ... 640 x 429 - 29k - jpg www.cruise crazies.com</p>	 <p>Cruise Ship Crashes Into Manhattan ... 636 x 429 - 20k - jpg www.cruise crazies.com</p>	 <p>... Authorities say a cruise ship ... 440 x 503 - 44k - jpg www.1010wins.com</p>	 <p>Cruise ship Carnival Triumph sails ... 800 x 600 - 71k - jpg wirednewyork.com</p>	 <p>Cruise Ship History: UNITED STATES ... 898 x 1129 - 564k - jpg cruiselinehistory.com</p>	 <p>Crystal Symphony a Cruise Ship ... 3008 x 2000 - 295k - jpg radonic.wordpress.com</p>

The other side of coin

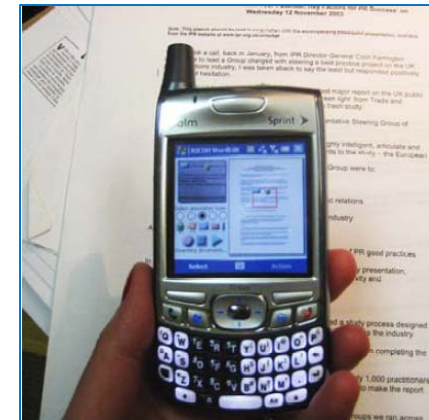
Visual Search: use pixels to find information



product search



landmark search



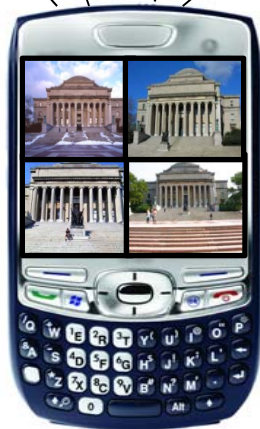
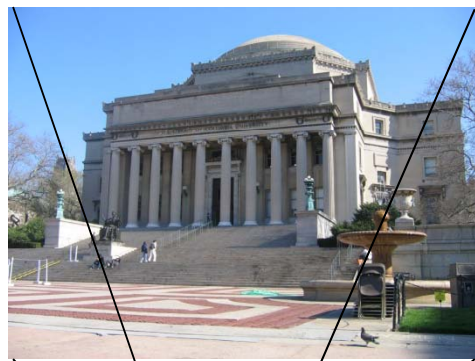
document search

Recent commercial applications:

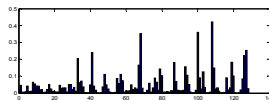
[iPhone App SnapTel](#)

[Google Goggles](#)

How Does It Work?



1. Take a picture



2. Image feature extraction



3. Send to server via SMS
5. Send most similar images back

4. Feature matching with database images

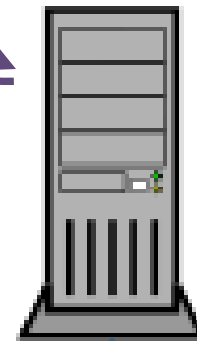


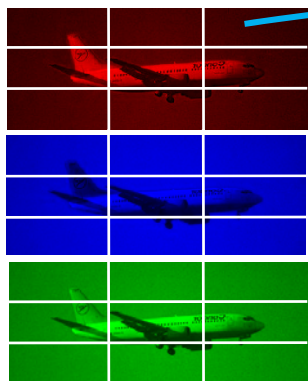
Image Database



Basic Image Features



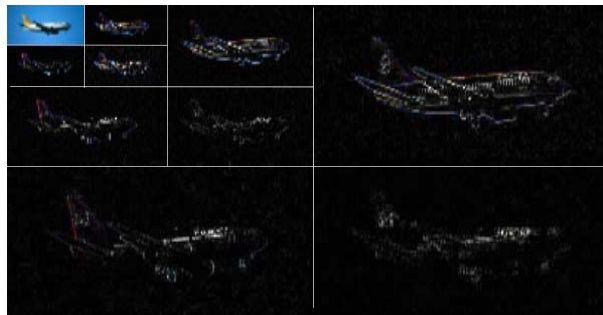
Grid Color Moment



mean
variance
...



Wavelet Texture



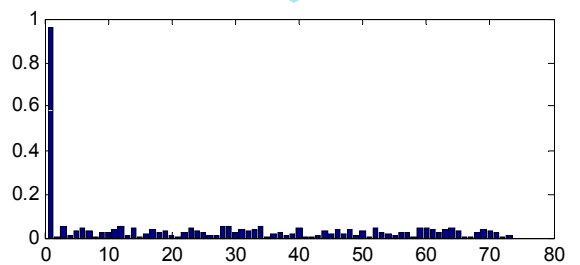
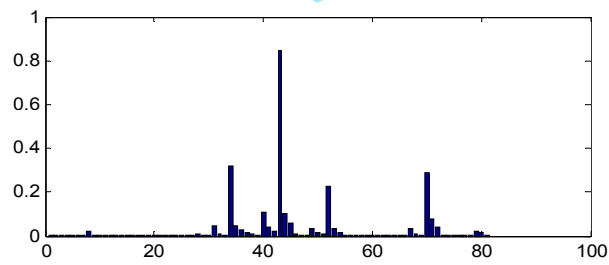
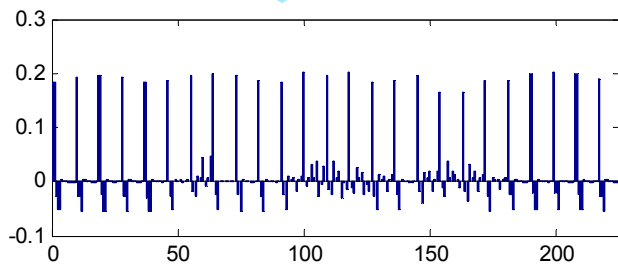
energy in different filter banks



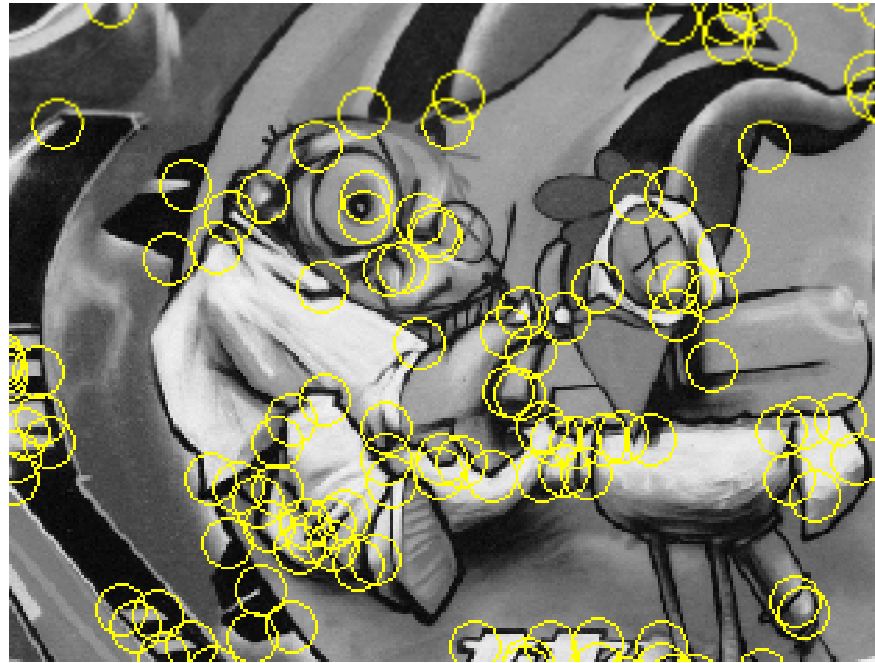
Canny Edge



edge direction counts



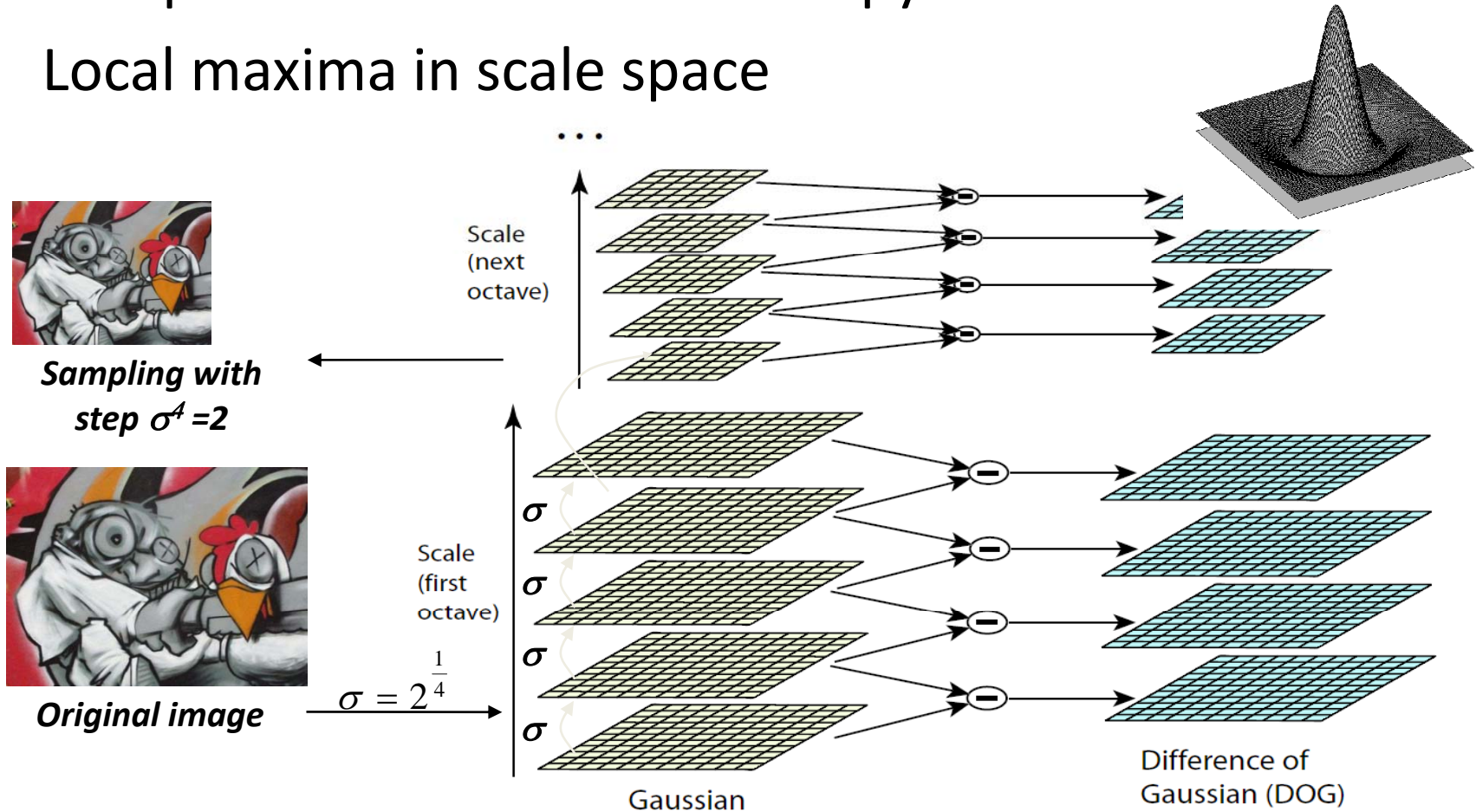
Recent Popular Features: Sample keypoints in images



- Keypoint properties:
 - Interesting content
 - Precise localization
 - Repeatable detection under variations of scale, rotation, etc

Example: Scale-Space

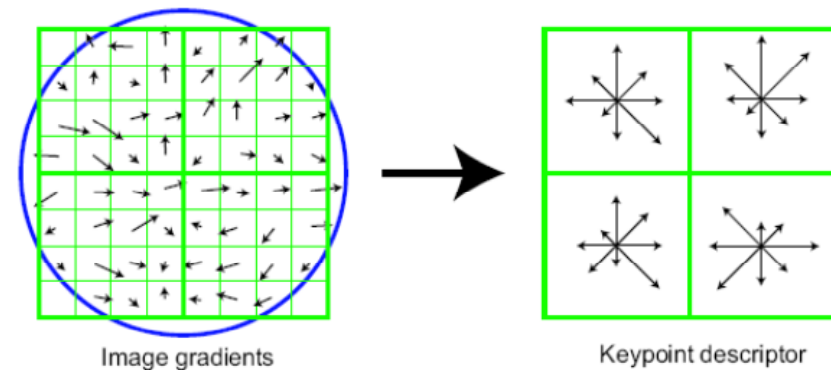
- Computation in Gaussian scale pyramid
- Local maxima in scale space



Extract Appearance Descriptor from Keypoints



Compute gradient
in a local patch

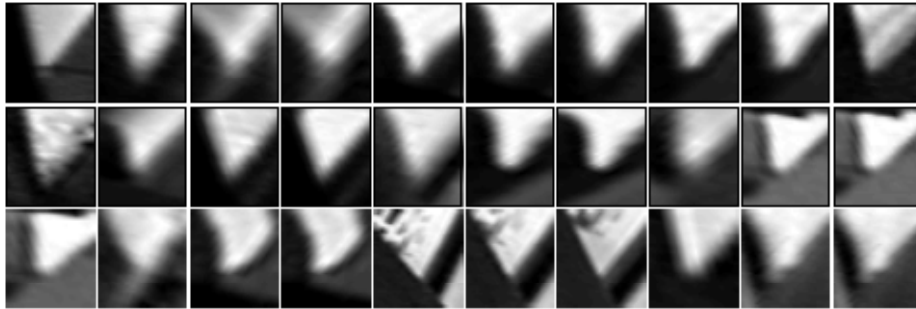


SIFT: Histogram of oriented gradients over local grids

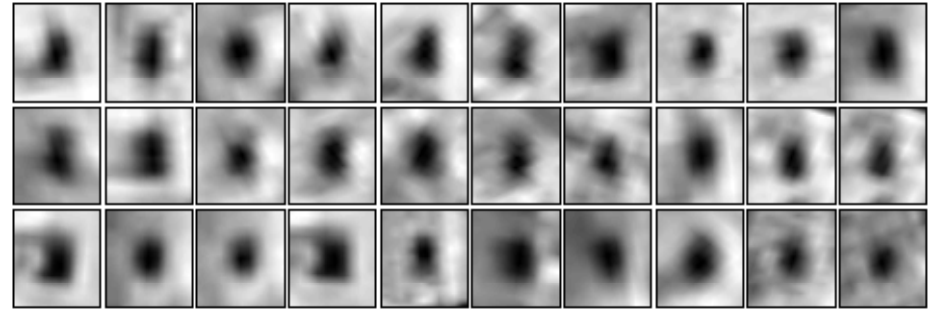
- e.g., 4x4 grids and 8 directions
-> $4 \times 4 \times 8 = 128$ dimensions
- Scale invariant

[Lowe, ICCV 1999]

Clustering of Image Patch Patterns



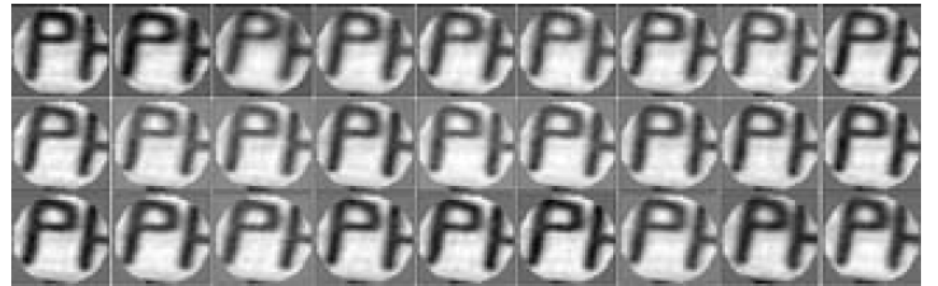
Corners



Blobs



eyes



letters

A New Representation: Visual Words

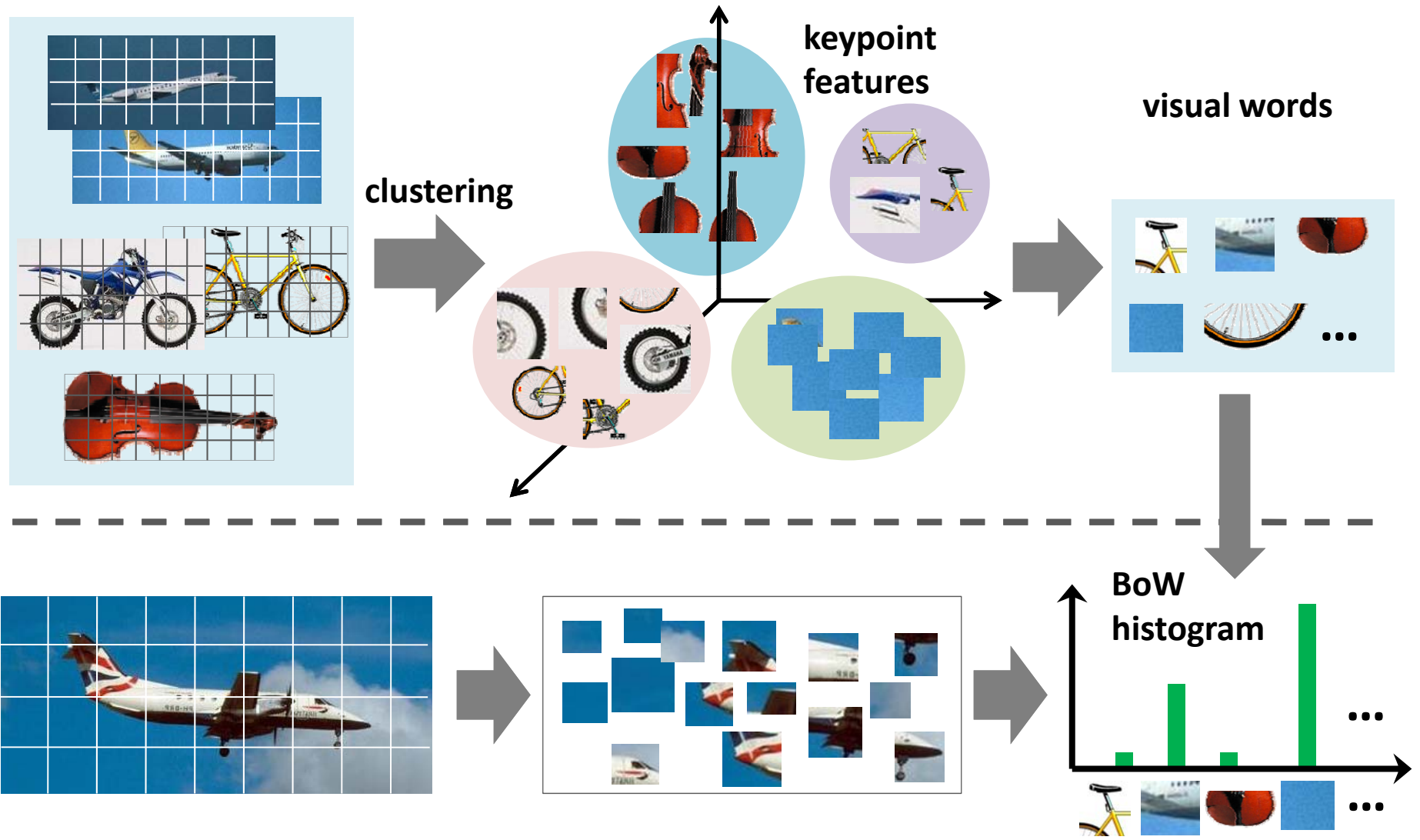
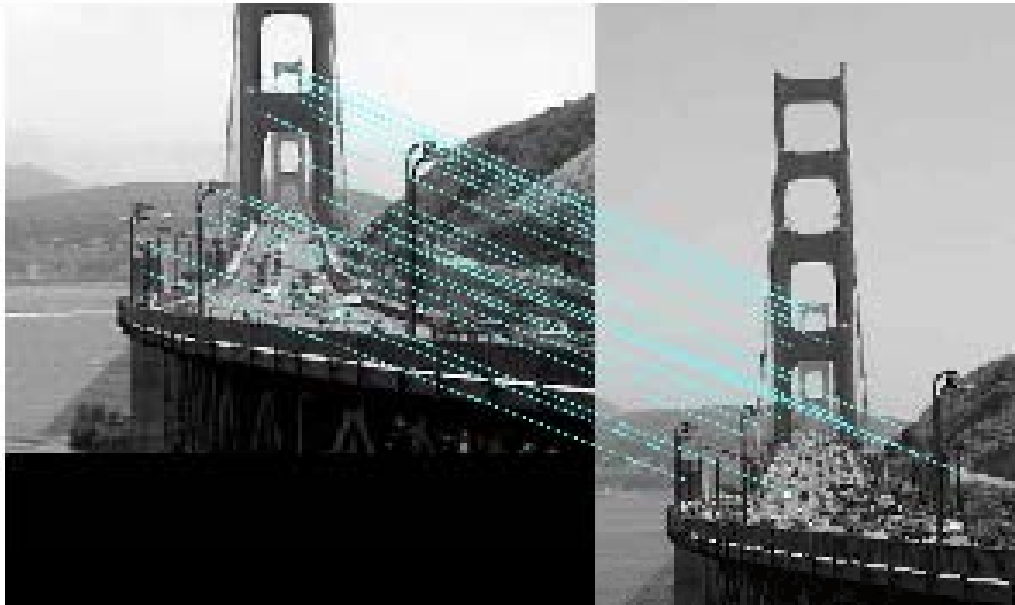
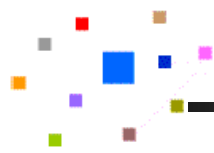


Image matching with local interest points



Matching SIFT points
[Lowe, 1999]

- Measure useful information
 - Image similarity
 - Copy/source identification
 - Discover possible transformation between images
- Can be done efficiently



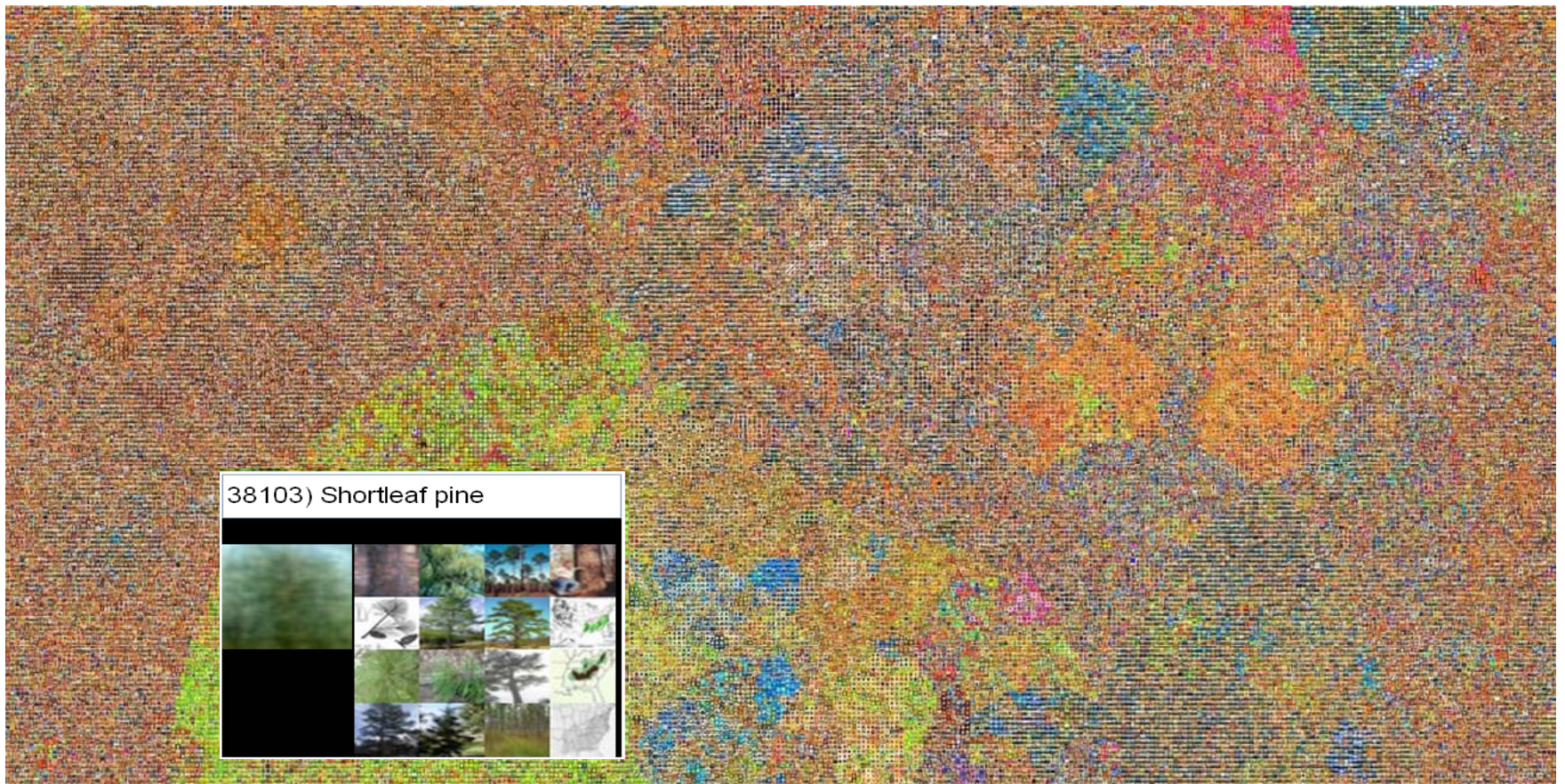
Content Based Image Search

- Demo: Object Retrieval
- Demo 2: Flickr Image Search

Scale Up: Find similar images over Internet

- Billions of images online as dense sampling of the world
- For every image taken, likely to find images that look alike

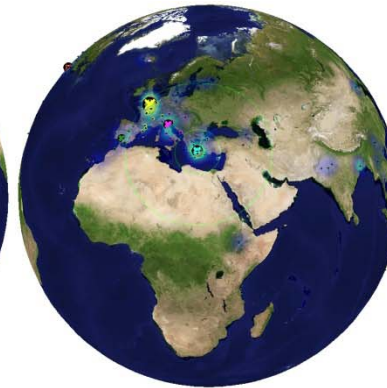
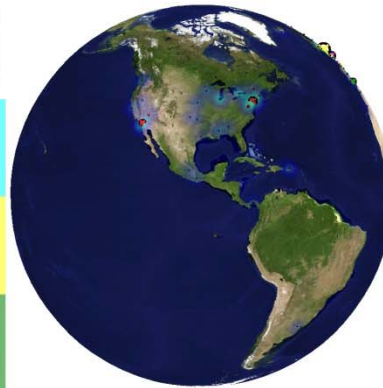
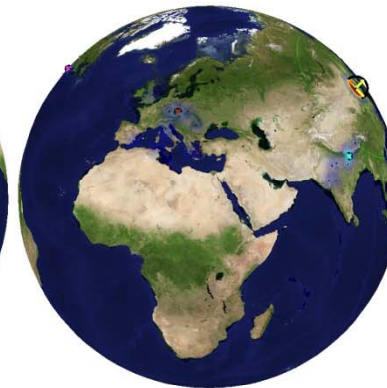
80 Million Tiny Images, Torralba, Fergus & Freeman, PAMI 2008



IM2GPS: where is this photo taken? (Hays & Efros, 2008)

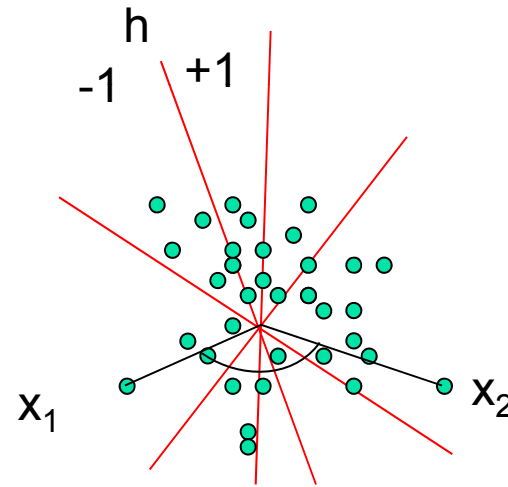
Similar images

Most likely locations



Search over Billions: Scalability is a Big Issue

- Similarity Search: traditional tree-based methods (e.g., kd-tree) not suitable in high dimension
- Need accurate, sublinear solutions ($o(N)$, $O(\log(N))$, $O(1)$)
- Recent trends:
projection based hashing
 - Random projection:
Locality Sensitive Hash (LSH)
[Indyk & Motwani 98, Charikar 02]
 - Principal projection:
Spectral Hashing [Weiss et al 08]
 - Restricted boltzman machines
[Hinton et al. 06, Torralba et al. 08]
 - Kernel LSH
[Kulis et al. 09 & Mu et al. 10]

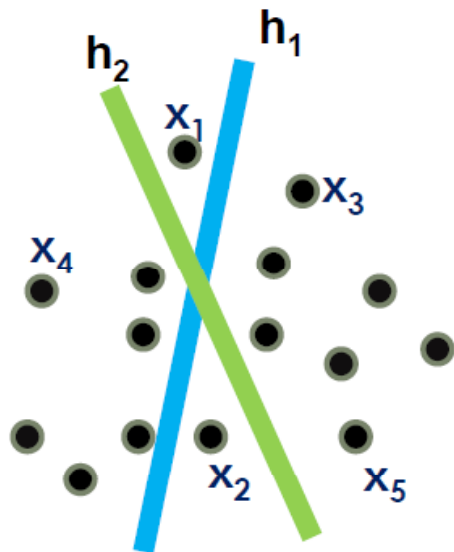


$$P(h(x_1) = h(x_2)) = 1 - \cos^{-1}(x_1 \cdot x_2) / \pi$$
$$= \text{Sim}(x_1, x_2)$$

random projection h with $N(0, 1)$

Binary Codes

Linear projection (hyperplane) based partitioning



X	x ₁	x ₂	x ₃	x ₄	x ₅
y ₁	0	1	1	0	1
y ₂	1	0	1	0	1
...
y _k

010... 100... 111... 001... 110...

Linear Projection based hashing

$$h_k(\mathbf{x}) = \text{sgn}(f(\mathbf{w}_k^T \mathbf{x} + b_k))$$
$$y_k(\mathbf{x}) = (1 + h_k(\mathbf{x}))/2$$

Very efficient training and testing

Probabilistic guarantee of finding true targets within ϵ distance range

[Indyk & Motwani 98]

Beyond Randomness: Semi-Supervised Learned Projection

- Given Pair-Wise Relations

$(\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{M}$: neighbor pair
 $(\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{C}$ nonneighbor-pair

- Measure empirical fit of hash bits

$$J(\mathbf{H}) = \sum_k \left\{ \sum_{(\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{M}} h_k(\mathbf{x}_i) h_k(\mathbf{x}_j) - \sum_{(\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{C}} h_k(\mathbf{x}_i) h_k(\mathbf{x}_j) \right\}$$

- Are the partitions balanced?
Measure hash bit variance

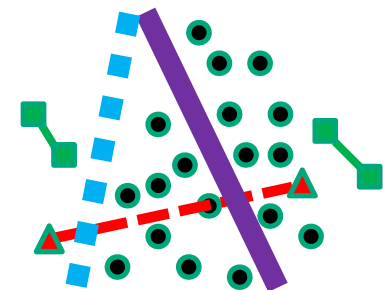
$$\sum_k E[\|h_k(\mathbf{x}) - \mu_k\|^2]$$

- Elegant eigen-decomposition solution

$$J(\mathbf{W}) = \frac{1}{2} \text{tr} [\mathbf{W}^\top \mathbf{X}_l \mathbf{S} \mathbf{X}_l^\top \mathbf{W}] + \frac{\eta}{2} \sum_k E[\|\mathbf{w}_k^\top \mathbf{x}\|^2]$$

$$S_{ij} = \begin{cases} 1 & : (\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{M} \\ -1 & : (\mathbf{x}_i, \mathbf{x}_j) \in \mathcal{C} \\ 0 & : \textit{otherwise.} \end{cases}$$

[Wang, Kumar, Chang
CVPR10, ICML10]



- Incremental learning via AdaBoosting

 Similar

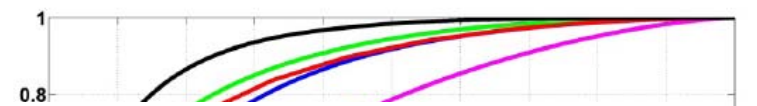
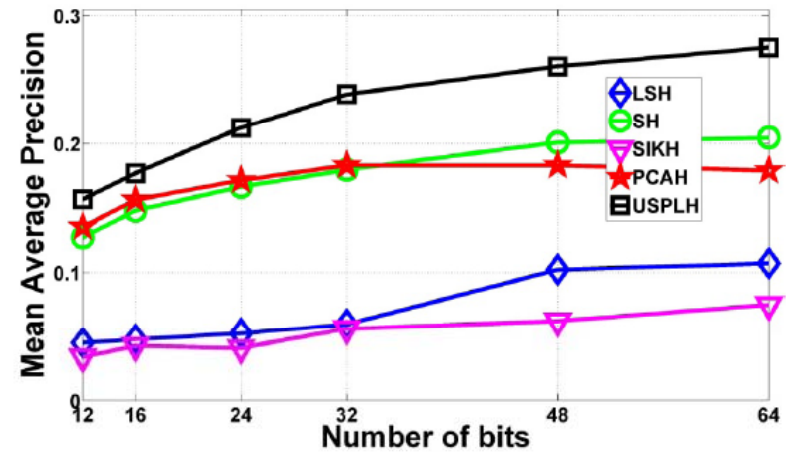
 Dissimilar

 Poor Projection

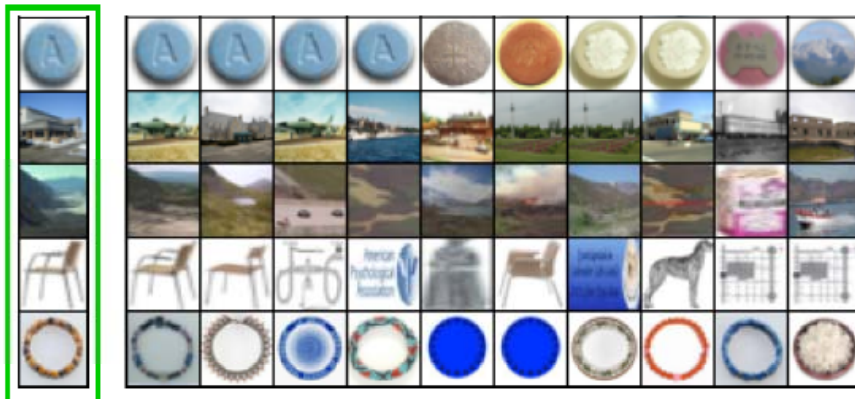
 Good Projection

- Learned projection hash increases accuracy > 2X
- Query time: a few seconds
- Compact code - 48 bits vs. 128 bytes per sample)
- **Challenge: scale to billions?**

SIFT-1M



Tiny Image – 80M



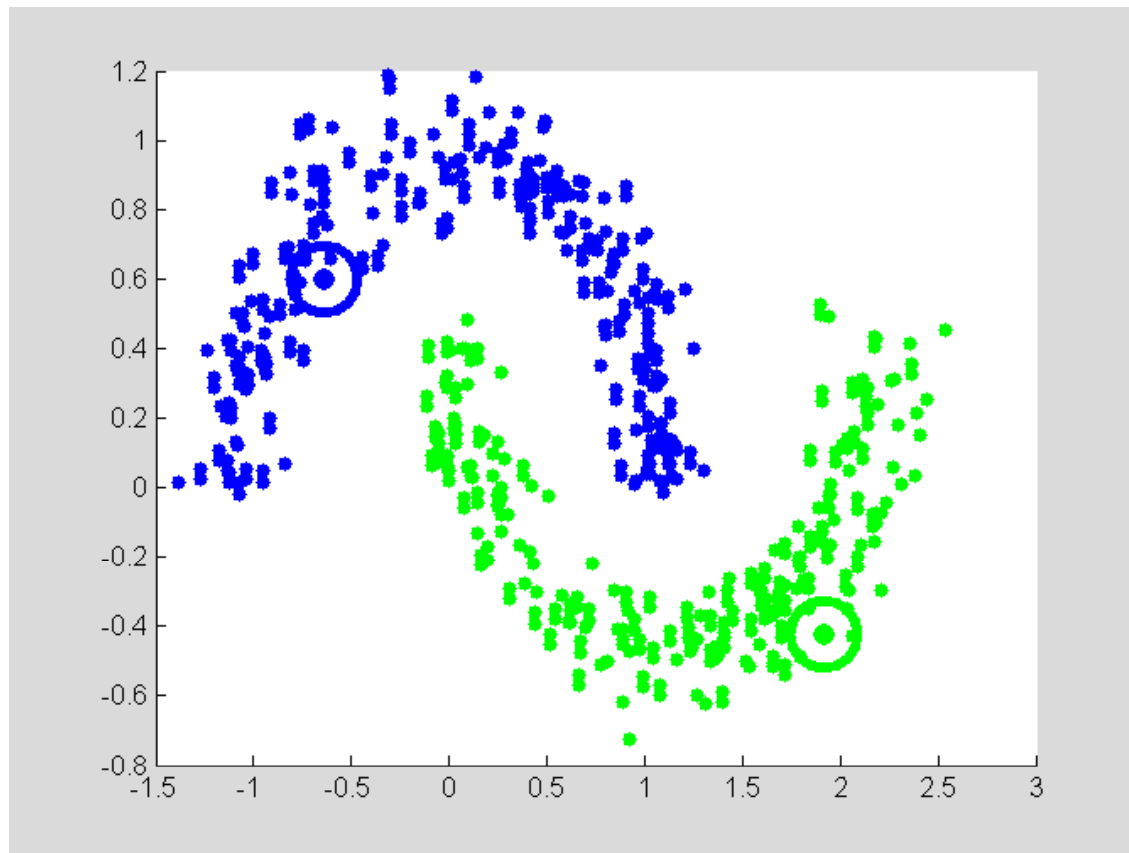
Binary Reconstructive Embedding



Learned Projection Hash

Beyond pair-wise similarity: Image Manifold

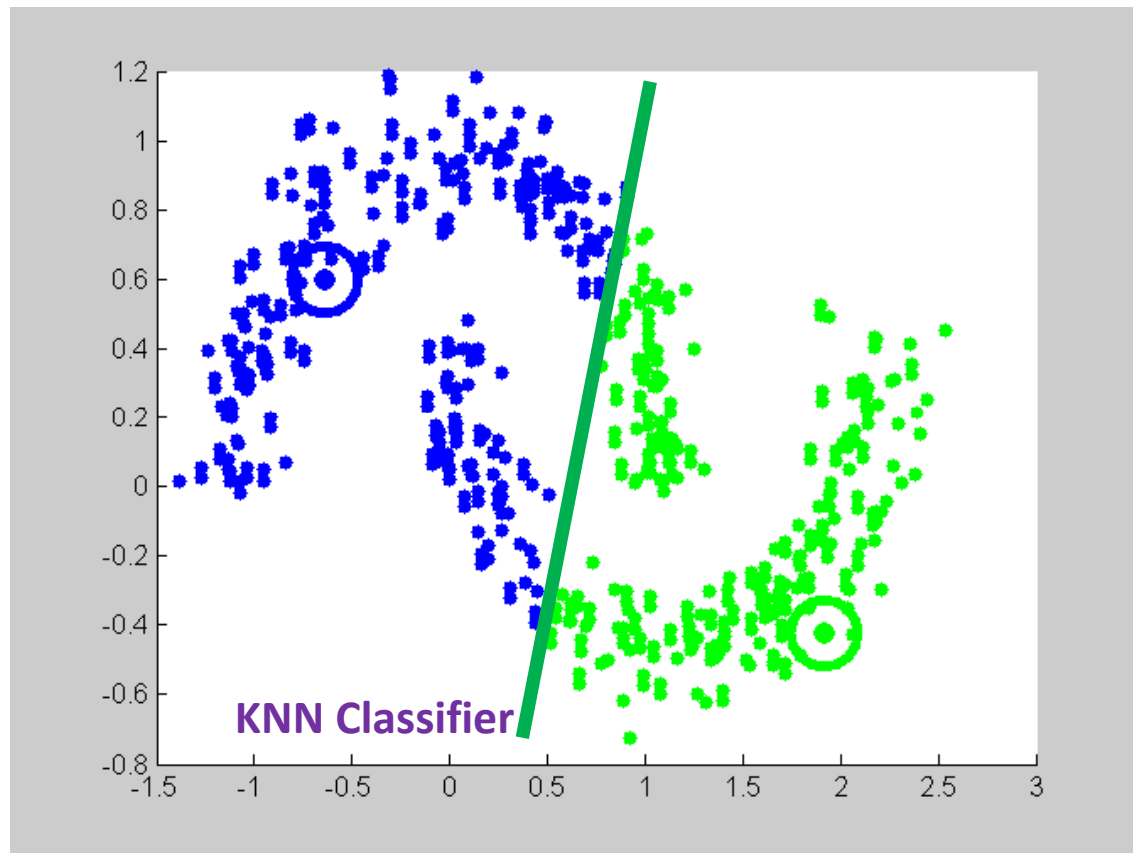
- Recover manifold distributions in high-dimensional space
- Each point represents an image, construct sparse graph
- Graph for information propagation, like PageRank



consider pair-wise similarity graph and data

Manifold vs. Supervised Learning

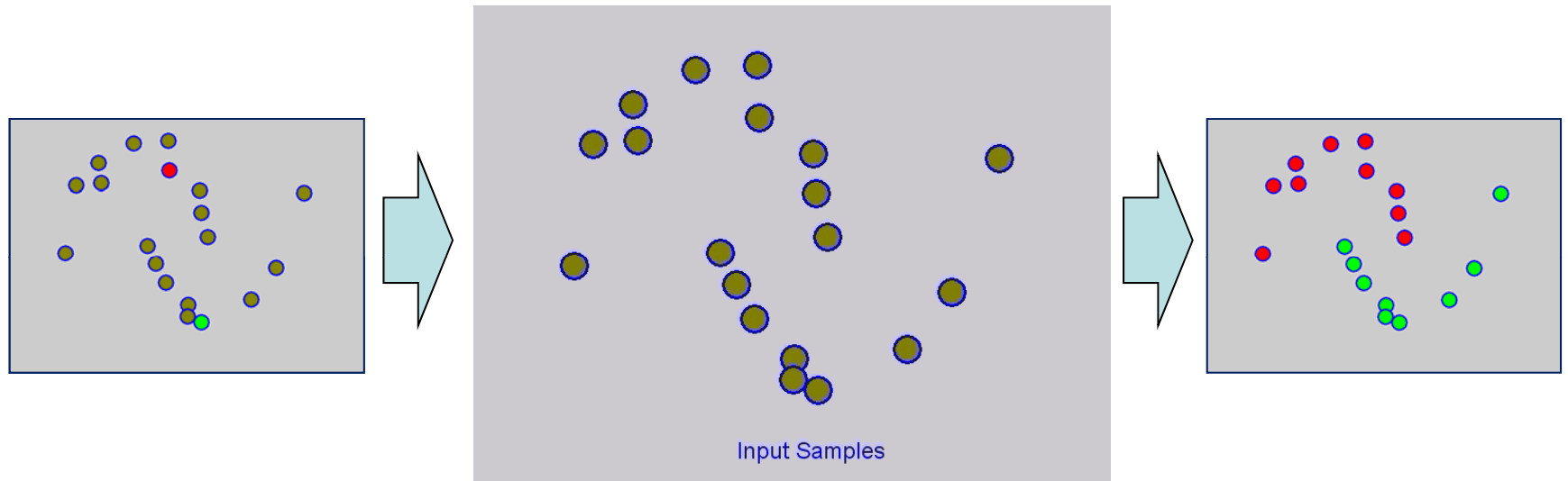
- Poor classifier when ignoring manifold distribution



Only consider based data

Graph-based Semi-Supervised Learning

- Construct Sparse for Label Propagation



Input samples with sparse labels

Label propagation on graph

Label inference results

Positive  Negative

 Unlabeled

 Positive

 Negative

$$f^* = \min_f Q(f, y, \mathcal{G}(V, W))$$

\mathcal{G} -- graph

V -- graph node

W -- weight matrix

Q -- risk function

f -- classification

y -- label matrix

An active area in Machine Learning

- Given initial labels, Y , find classification F over graph

Label
smoothness

Fit known
labels

$$\begin{aligned} Q(F) &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} \left\| \frac{F_{i\cdot}}{\sqrt{D_{ii}}} - \frac{F_{j\cdot}}{\sqrt{D_{jj}}} \right\|^2 + \mu \sum_{i=1}^l \|F_{i\cdot} - Y_{i\cdot}\|^2 \\ &= \text{tr}\{F^\top L F + \mu(F - Y)^\top (F - Y)\} \end{aligned}$$

(Zhou, et al NIPS04)

- Gaussian fields & Harmonic functions (Zhu et al ICML03)

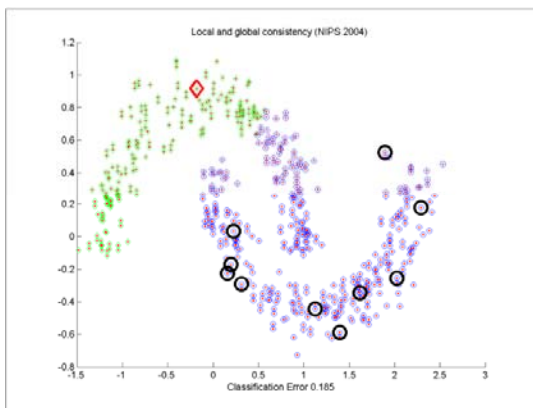
$$Q(F) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n w_{ij} \|F_{i\cdot} - F_{j\cdot}\|^2$$

1) $\Delta F = 0$ on unlabeled data, where $\Delta = D - W$ is the graph Laplacian;

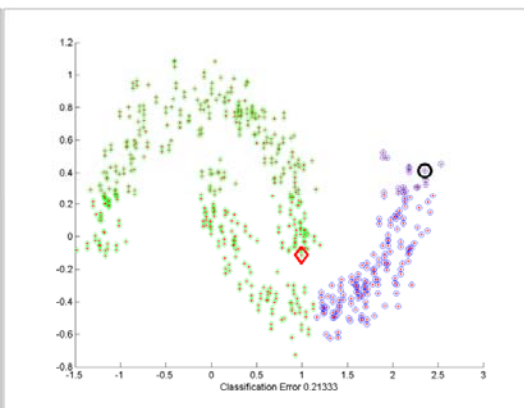
2) $F_{i\cdot} = Y_{i\cdot}$ on labeled data.

Non-Trivial Issues

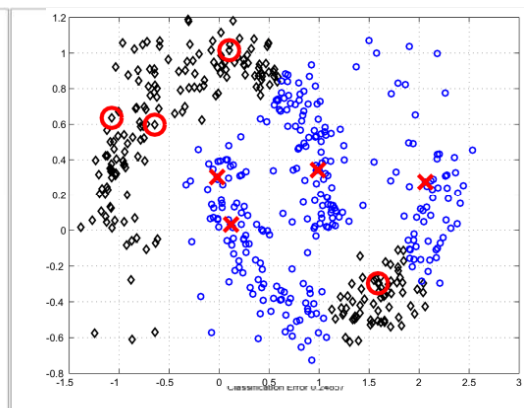
LGC
Method



(a)

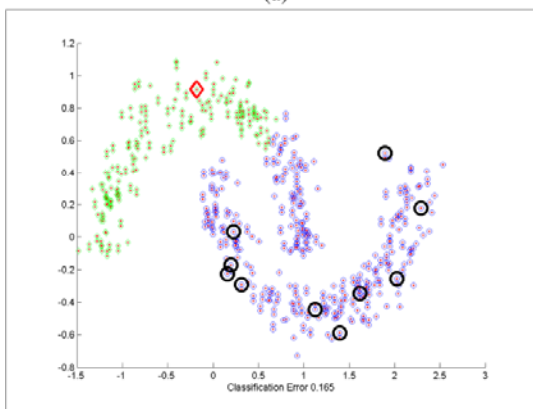


(b)

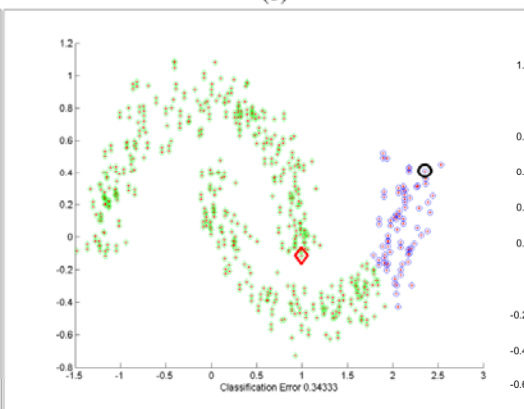


(c)

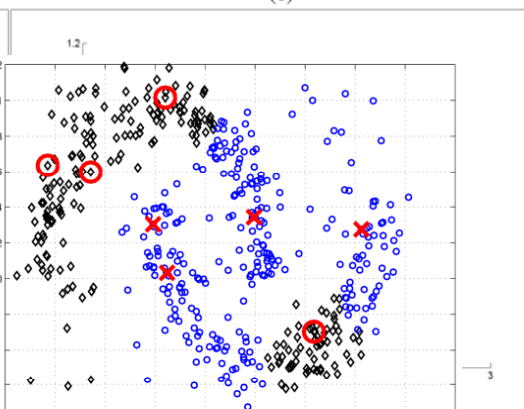
GFHF
Method



(d)



(e)

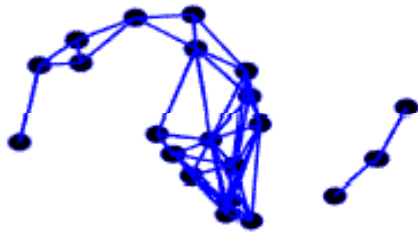


Unbalanced
Labels

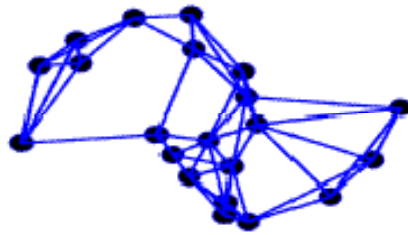
Bad Label
Locations

Noisy
Labels

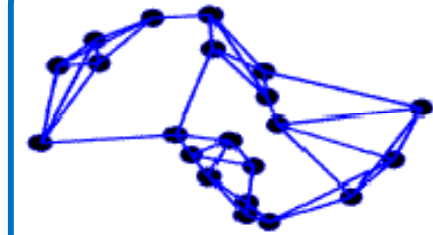
A Key Problem: How to Construct a Good Sparse Graph from Massive Data?



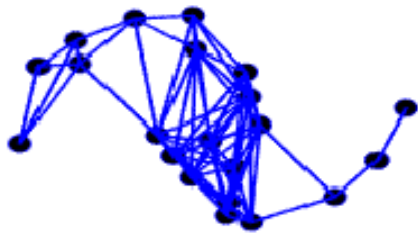
(a)



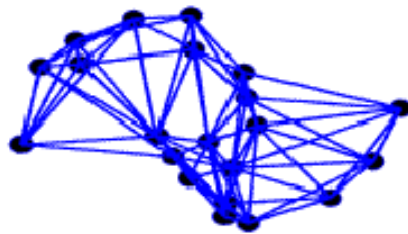
(c)



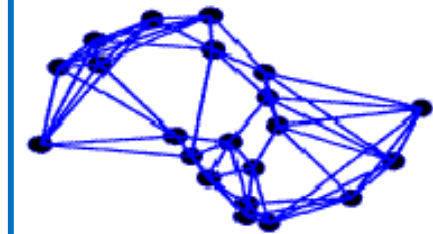
(e)



(b)



(d)



(f)

Distance threshold

Rank threshold (kNN)

B-Match

- **B-Match preserves sparse structures**
- **Efficient method exists much faster than $O(N^3)$**

$$\min_{P \in \mathbb{B}} \sum_{ij} P_{ij} D_{ij}$$

P: connectivity

D: distance

$$s.t. \sum_j P_{ij} = b, P_{ii} = 0, P_{ij} = P_{ji},$$

(Jebara, Wang, Chang ICML 2009)

Another Key Issue: Noisy Labels

→ Treat both labels & prediction as optimization variables

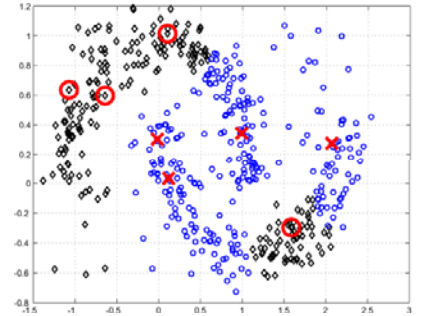
$$\mathcal{Q}(\mathbf{F}, \mathbf{Y}) = \frac{1}{2} \text{tr} \left\{ \mathbf{F}^T \mathbf{L} \mathbf{F} + \mu (\mathbf{F} - \mathbf{V} \mathbf{Y})^T (\mathbf{F} - \mathbf{V} \mathbf{Y}) \right\}$$

- Diagnose and add/remove labels
 - Select the most reliable and informative label

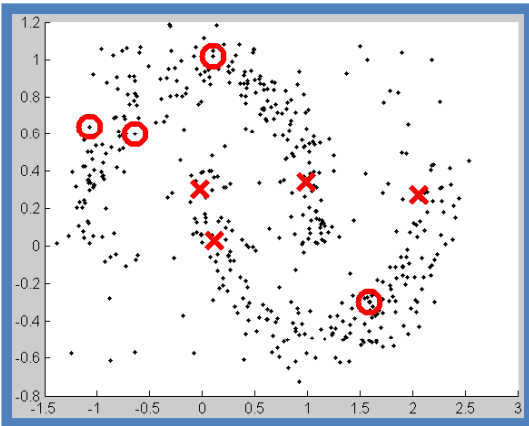
$$\mathcal{Q}(\mathbf{Y}) = \frac{1}{2} \text{tr} \left(\mathbf{Y}^T \mathbf{V}^T \left[\mathbf{P}^T \mathbf{L} \mathbf{P} + \mu (\mathbf{P}^T - \mathbf{I})(\mathbf{P} - \mathbf{I}) \right] \mathbf{V} \mathbf{Y} \right)$$

- Propagate trusted labels to the whole graph
 - given label (Y), propagate over graph, predict F

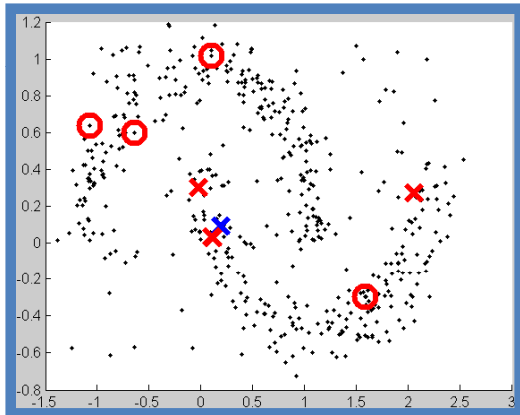
$$\frac{\partial \mathcal{Q}}{\partial \mathbf{F}^*} = 0 \Rightarrow \mathbf{F}^* = (\mathbf{L} / \mu + \mathbf{I})^{-1} \mathbf{V} \mathbf{Y} = \mathbf{P} \mathbf{V} \mathbf{Y}$$



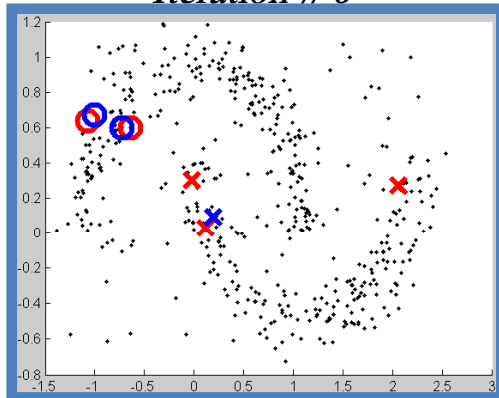
Initial Labels



Iteration # 2

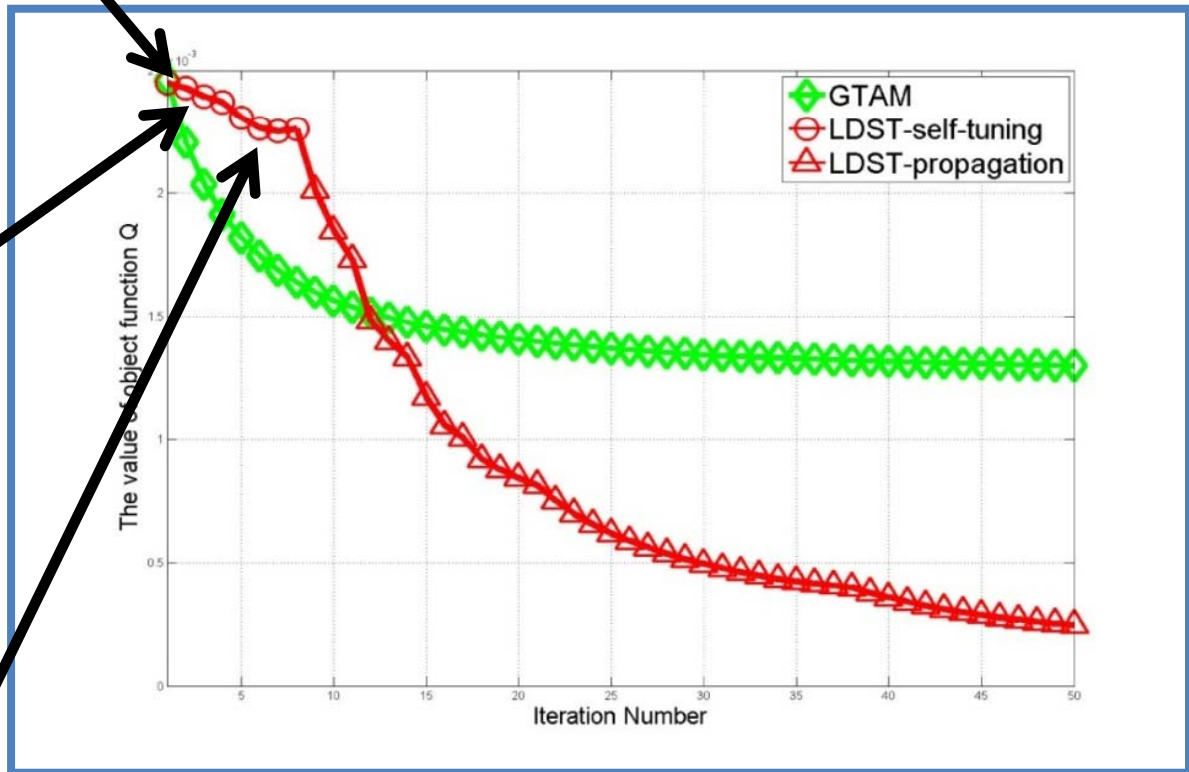


Iteration # 6



Iteratively Tune and Propagate Label Information

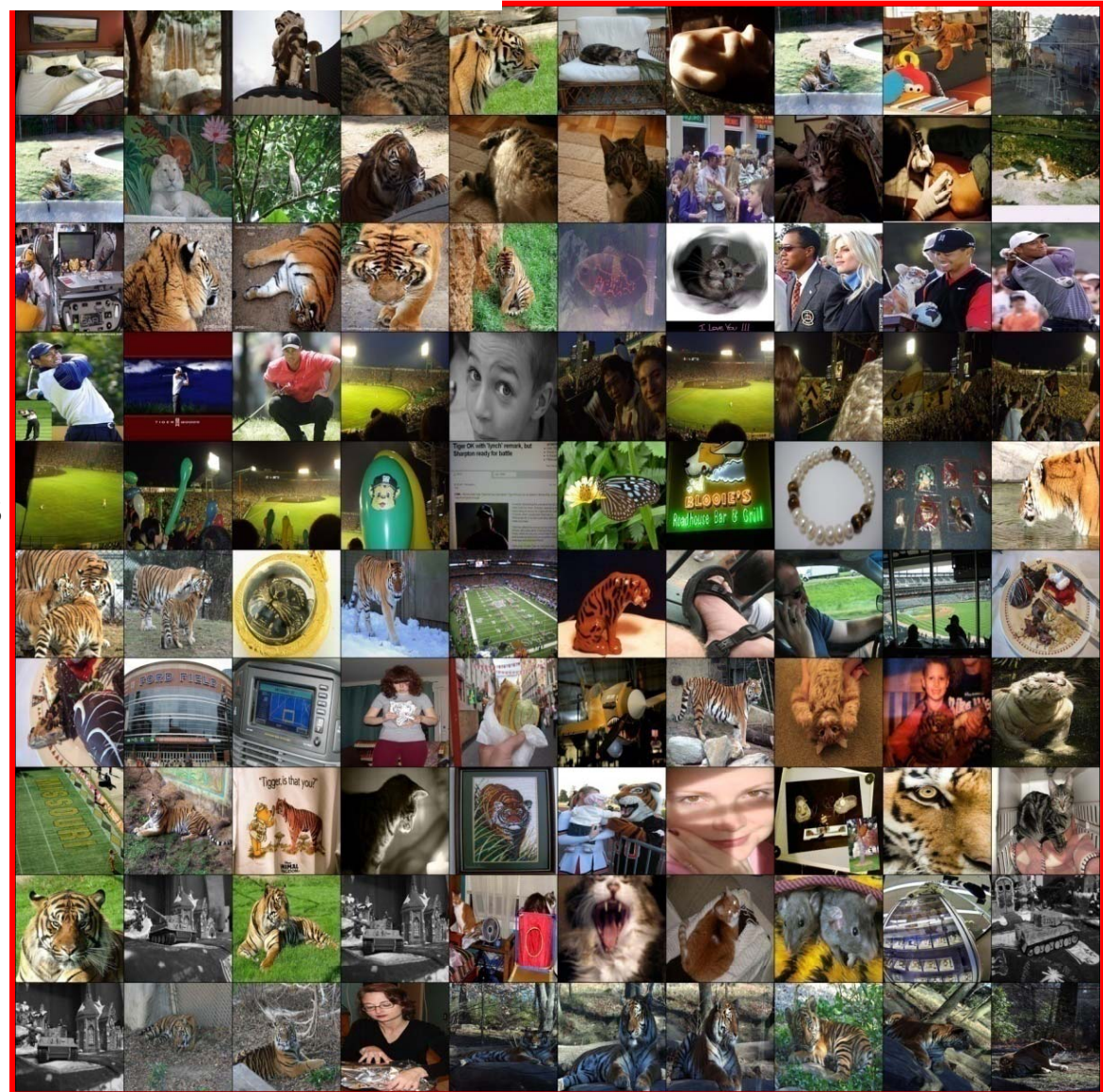
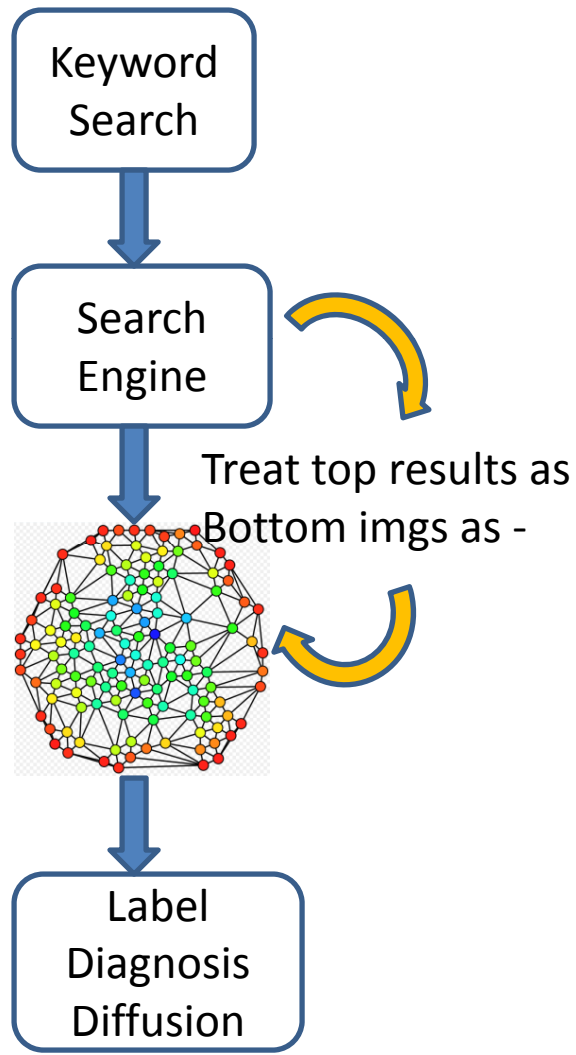
(Wang and Chang CVPR'09)



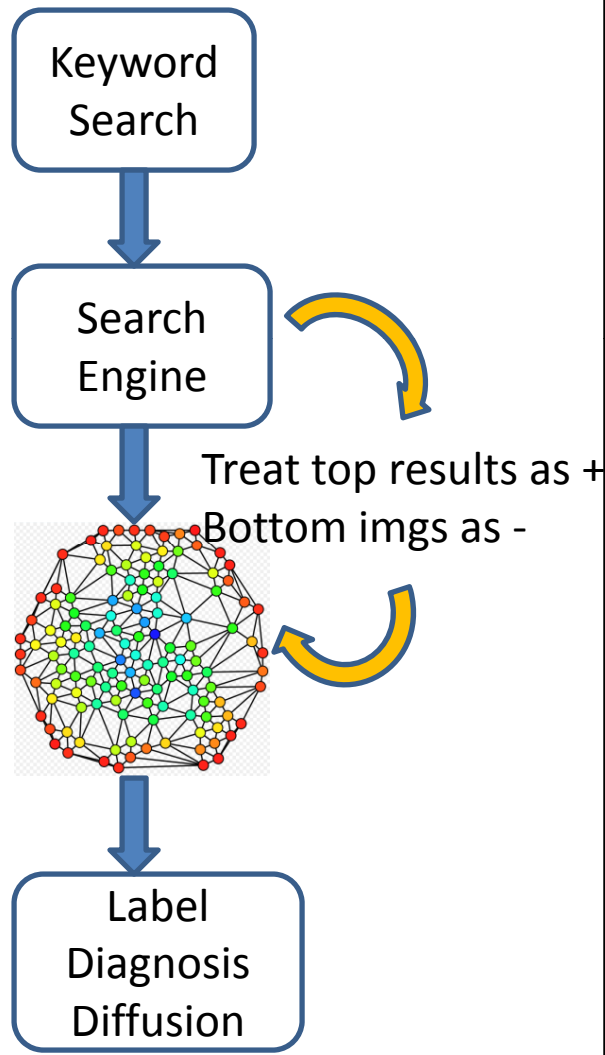
Decline of the cost function Q over iterations (with vs. without label tuning)

Application: Internet Search Result Reranking

Google Search "Tiger"



Application: Internet Search Result Reranking

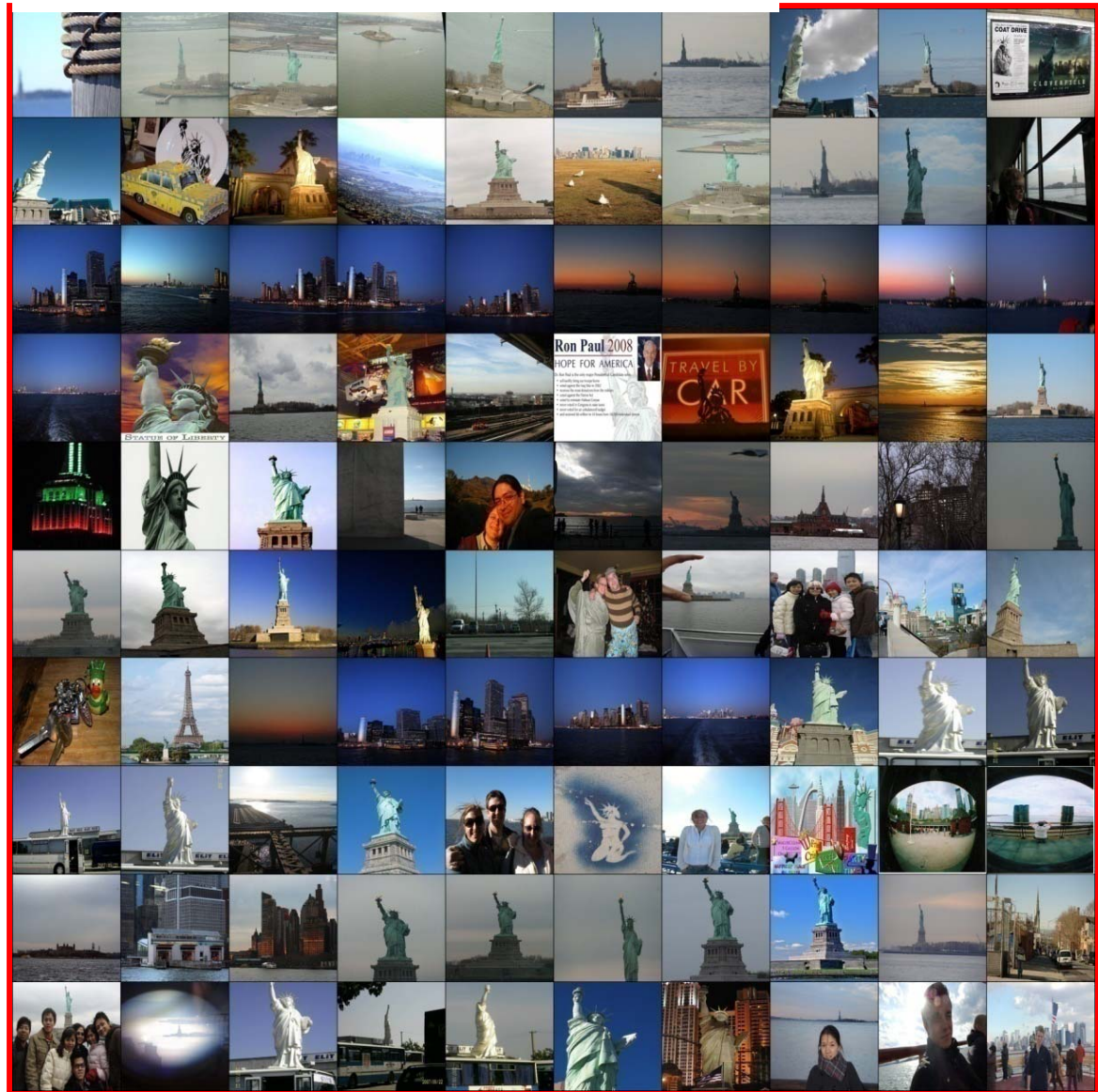
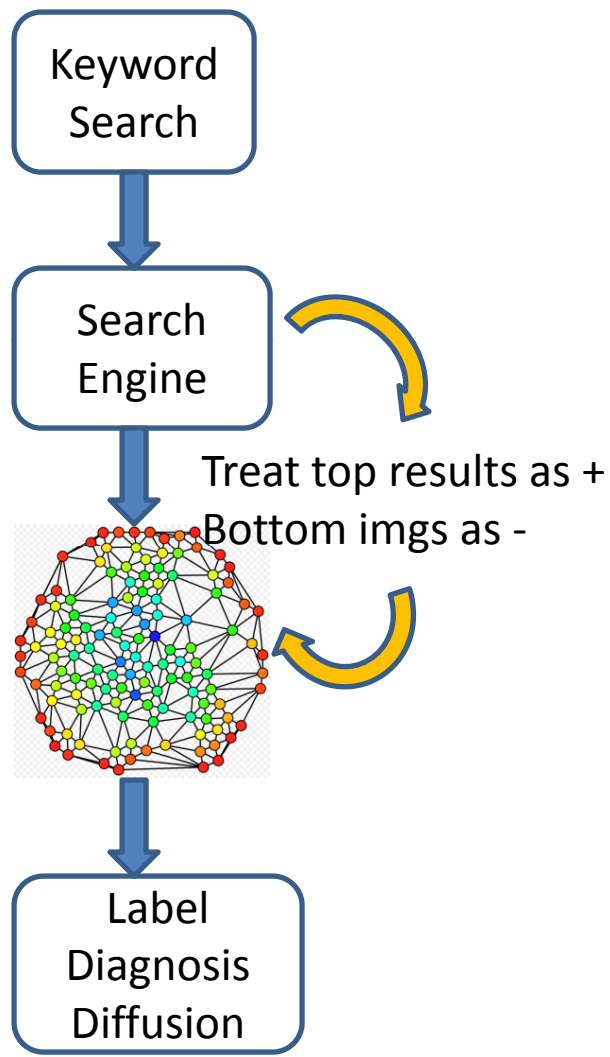


Reranked Results



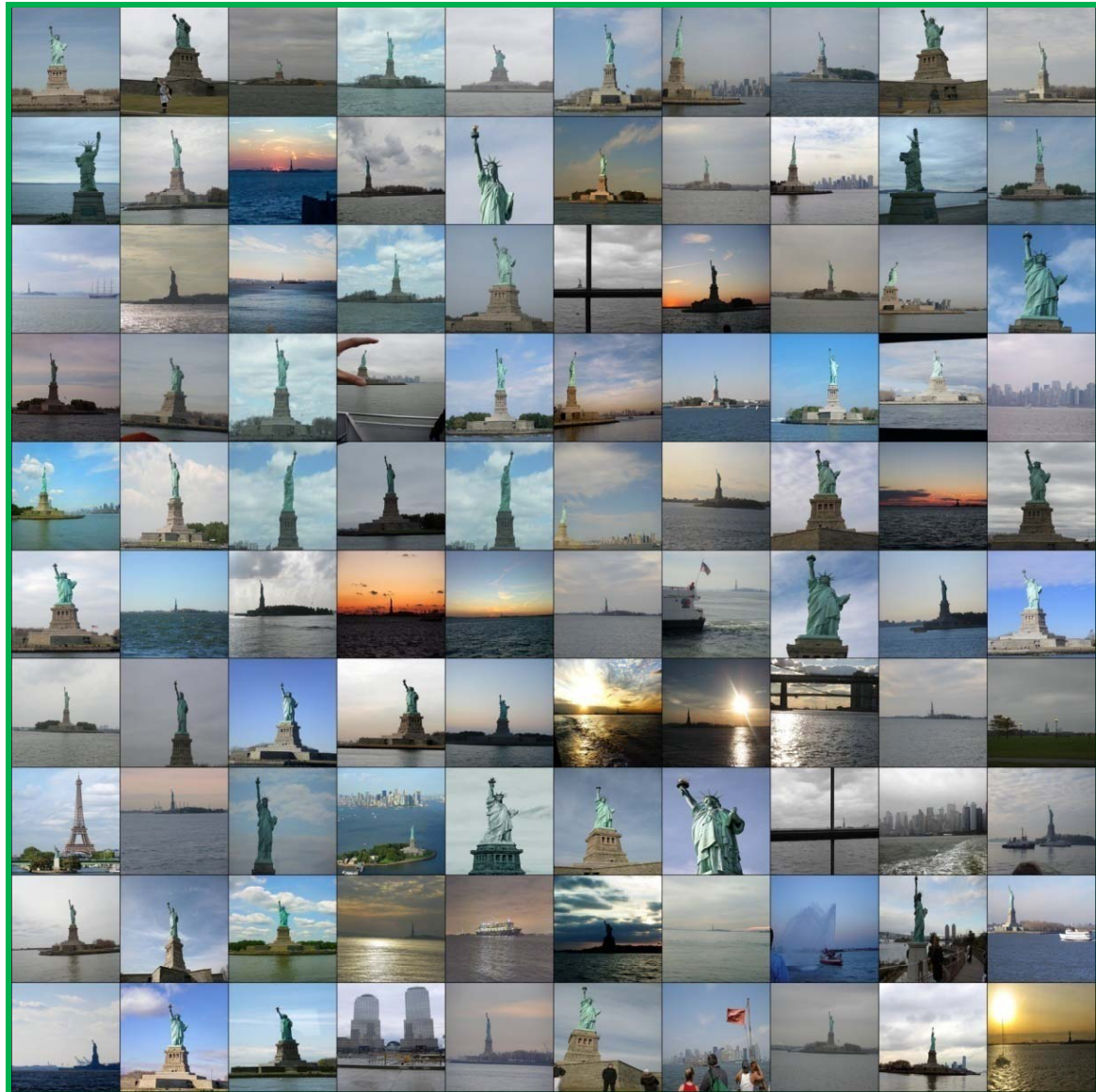
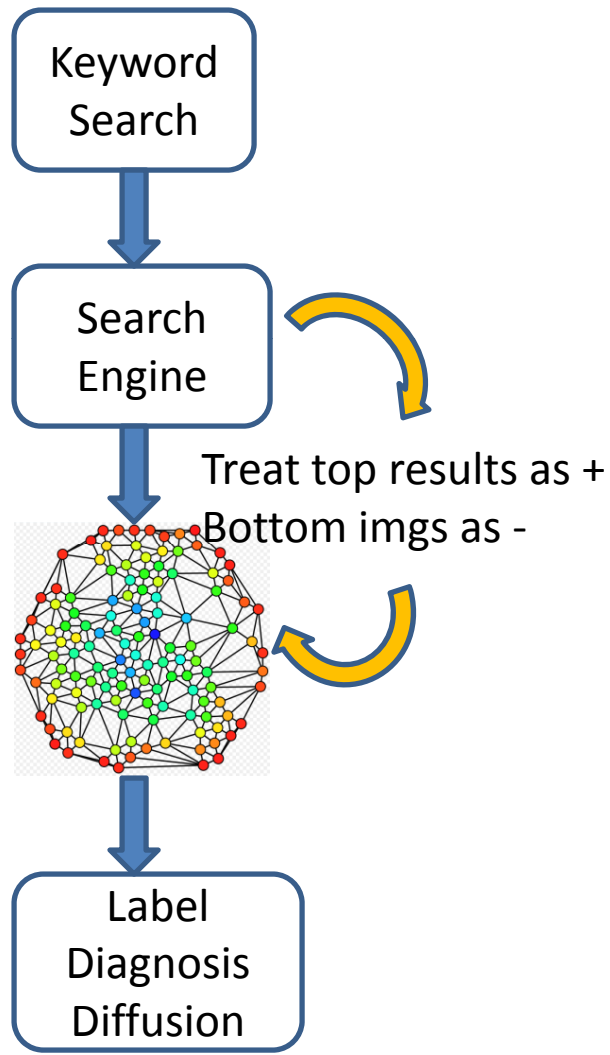
Application: Internet Search Result Reranking

Google Search “Statue of Liberty”

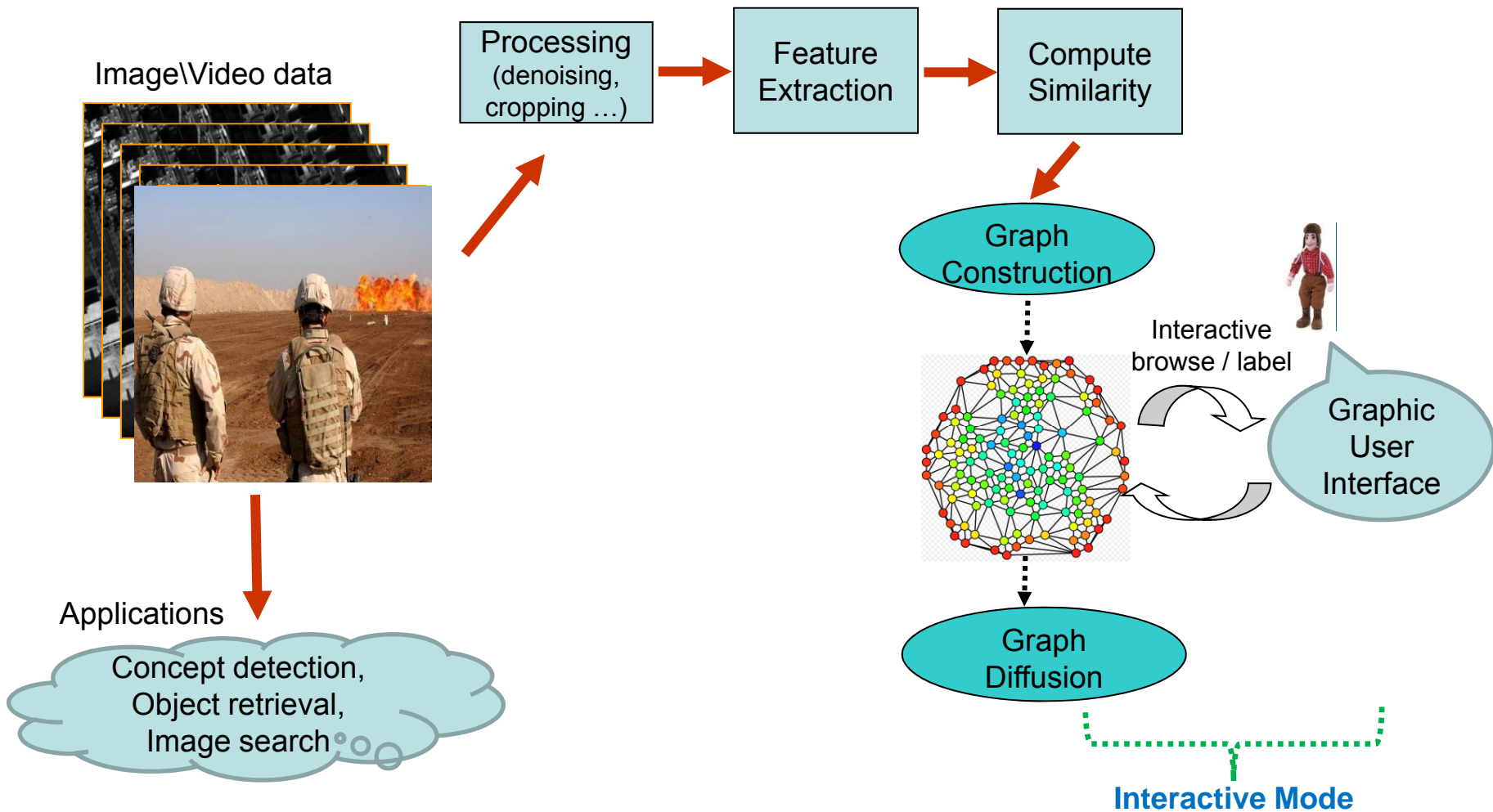


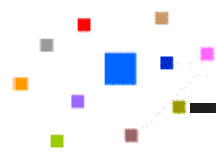
Application: Internet Search Result Reranking

Reranked Results



Interactive Image Annotation: Columbia TAG System

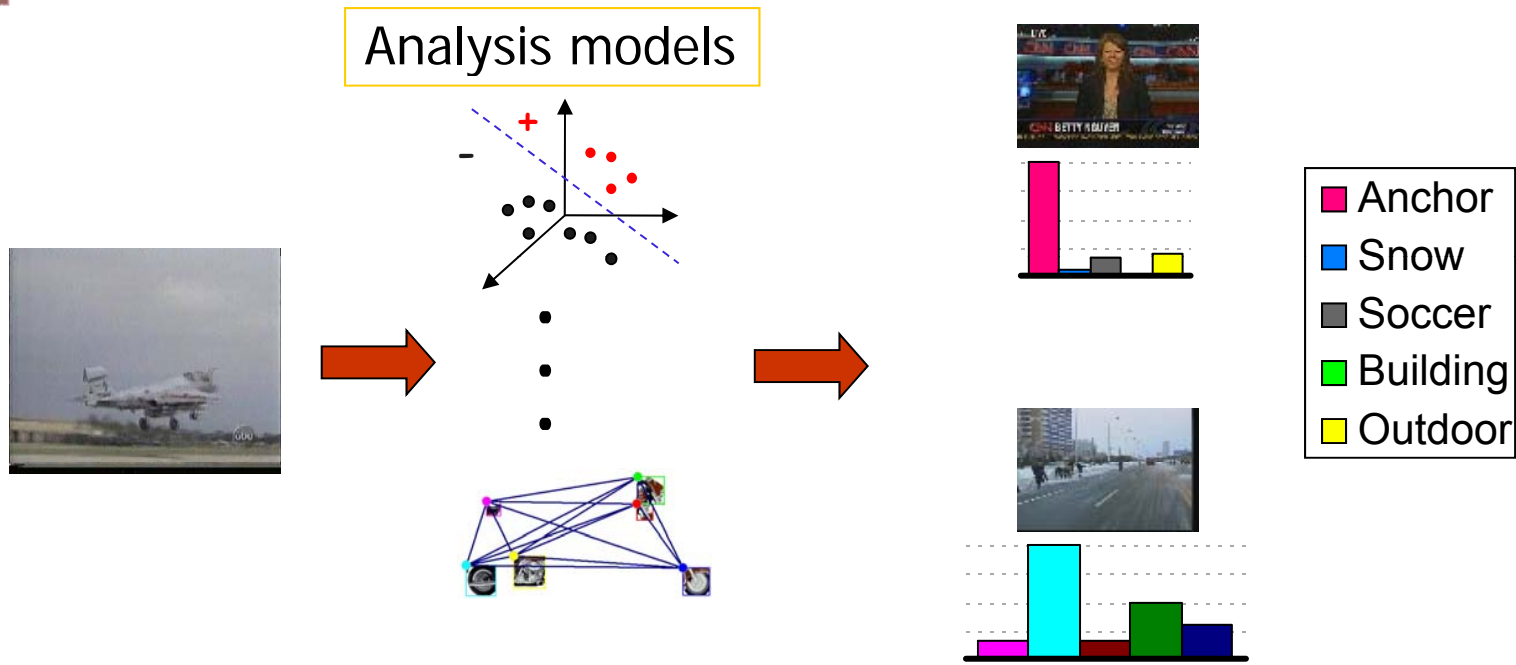




Columbia TAG Image Search System

- Demo:
Rapid Image Annotation with
User Interaction

Beyond Matching: Automatic Visual Recognition



- Audio-visual features
- Geo, time, camera metadata
- User context

- Rich semantic labels

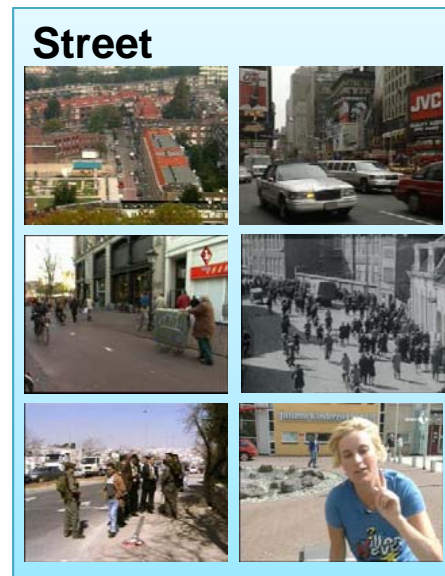
Hot topic ... community fast growing!

(as of Nov. 2009)

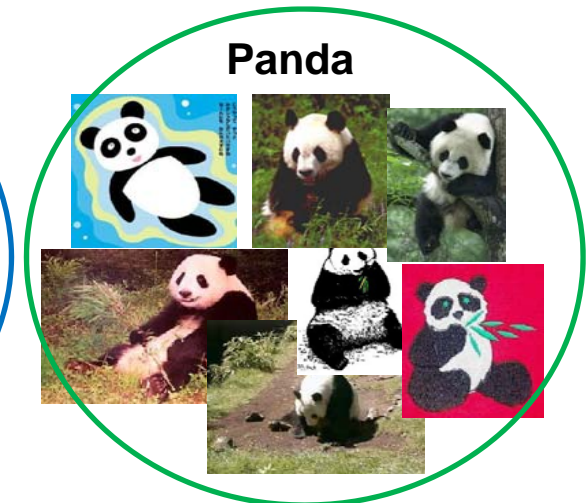
	Data domain	amount	types	Lexicon size
TRECVID	Broadcast news, documentary, Flickr, Youtube...	400 hours Sound & Vision 170 hours Television News 100 hours BBC rushes (130,000+ subshots)	video shots, keyframes	10 (2004, 2005) 39 (2006) 39 (2007) 20 (2008)
LSCOM	Broadcast news video	170 hours Television News 61901 subshots	video subshots	1000+ concepts
CalTech256	Internet Images	30,607 images	images	256 classes
PASCAL	Internet Images	9,963 images 24,640 annotated objects	images, objects	20 classes
Tiny Image	Internet Images	80,000,000 tiny images (32x32)	images	75,378 WordNet nouns
LabelMe	Internet and user uploaded images and videos	30,369 images from 183 folders	images, keyframes	111,490 object labels
ImageNet	Internet images	9,386,073 images	images	14,847 WordNet synsets
Lotus Hill Dataset	Internet Images	500,000+ images and keyframes	images, keyframes	280 object classes

Image/Video Classification

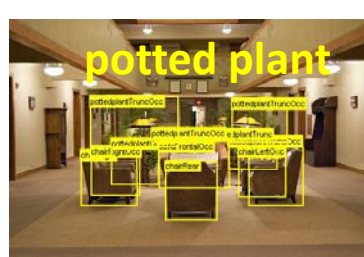
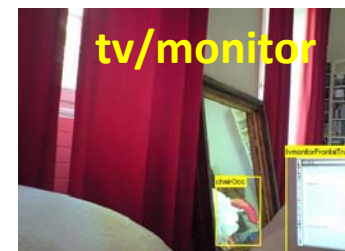
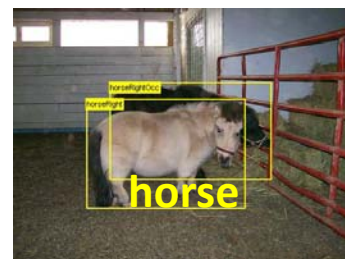
TRECVID



CalTech 101

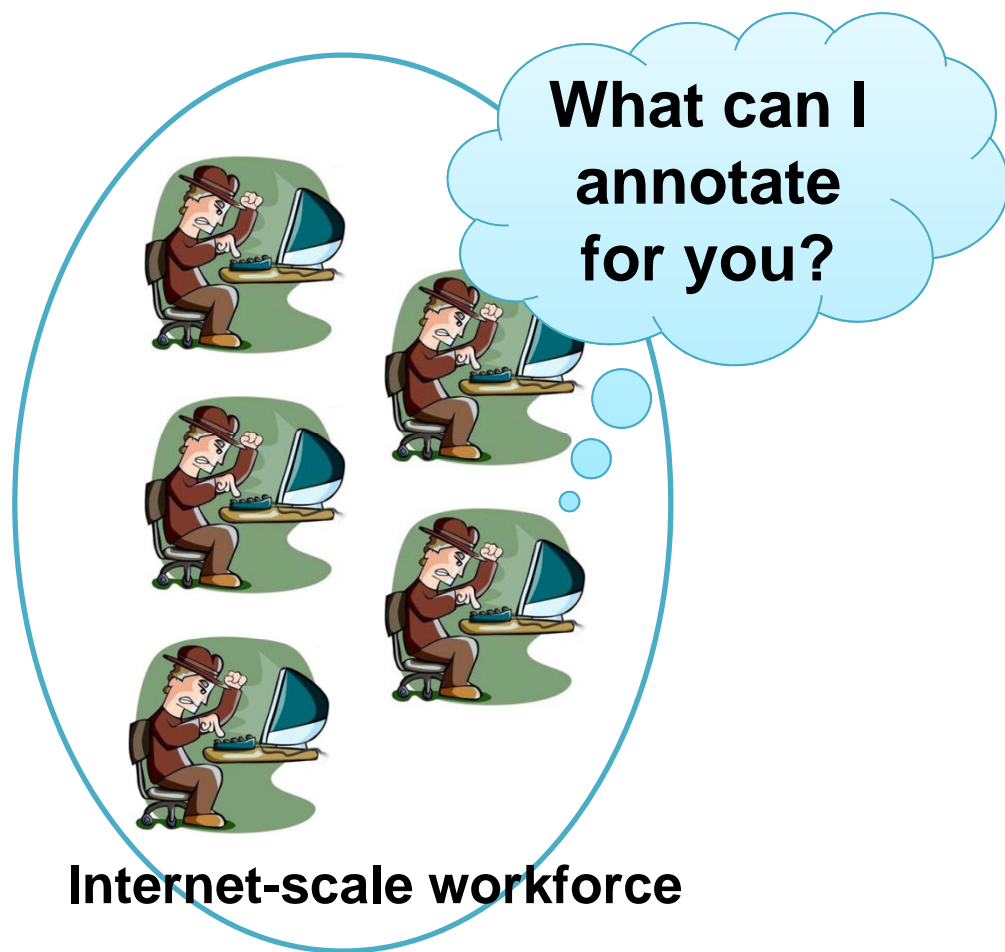


Object Localization (PASCAL VOC)



Create Large Corpora: Crowd Sourcing

- Amazon Mechanical Turk
- Web Open Market for Human Computing



earn \$0.03 per image tag

Crowd sourcing for image annotation

amazonmechanical turk
beta Artificial Intelligence

Search for containing

HITs containing 'annotation'
1-2 of 2 Results
Sort by:

Draw bounding boxes around objects in images
Requester: [mlabel-dolores](#)

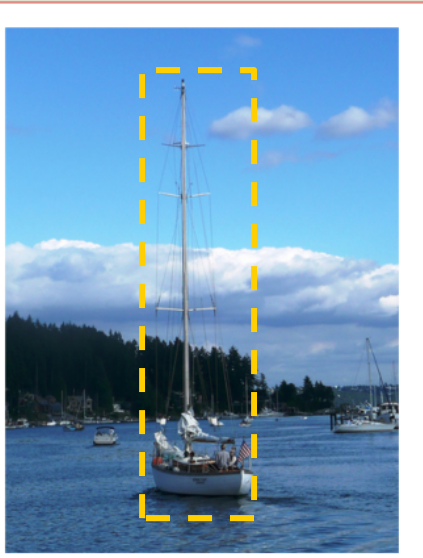
Medical Image Segmentation (OCT) - 1060 - BC
Requester: [Cardiff Turk](#)

Draw bounding boxes around objects in images
Requester: mlabel-dolores
Reward: \$0.05 per HIT
HITs Available: 1655
Duration: 60 minutes
Qualifications Required: None

Sign In

Main | Instructions with examples | Look up "gig" in Wikipedia | in Google

Draw a box around **gig**: *long and light rowing boat; especially for racing*



Draw a bounding box around the following object in the image:

gig: long and light rowing boat; especially for racing

Instructions:

- Include all visible parts and draw as tightly as possible
- **If there are multiple instances, pick only ONE (any one).**
- **Do NOT draw on the instances that already have bounding boxes.**

[SEE INSTRUCTIONS WITH EXAMPLES](#)

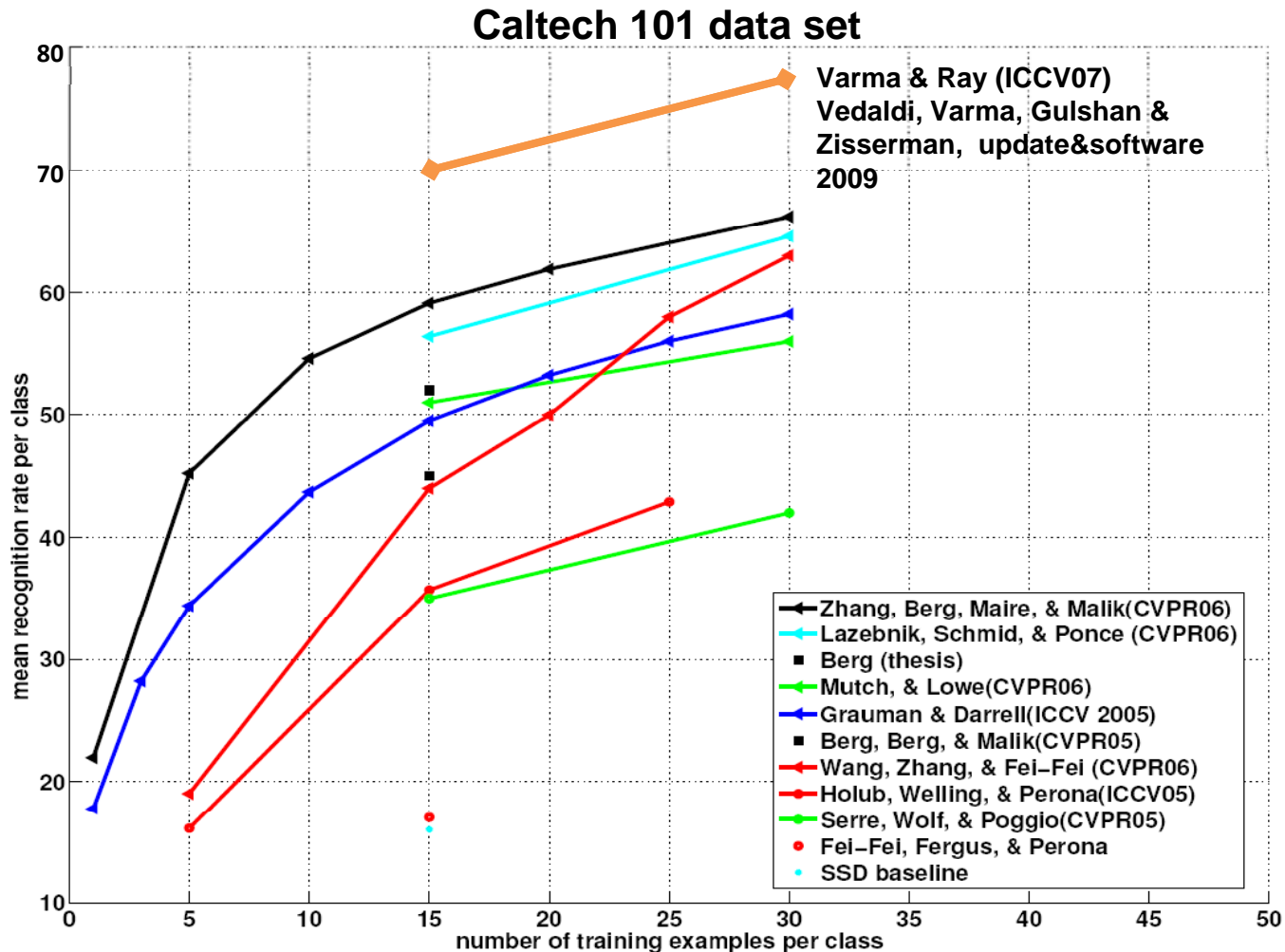
Check here if there's NO gig in this image or if every instance already has a bounding box.

roup
roup

Such task is nicely called Human Intelligence Task (HIT)!

Rapid Advances in Image Annotation

- “Moore’s Law”: accuracy doubles in about 2 years



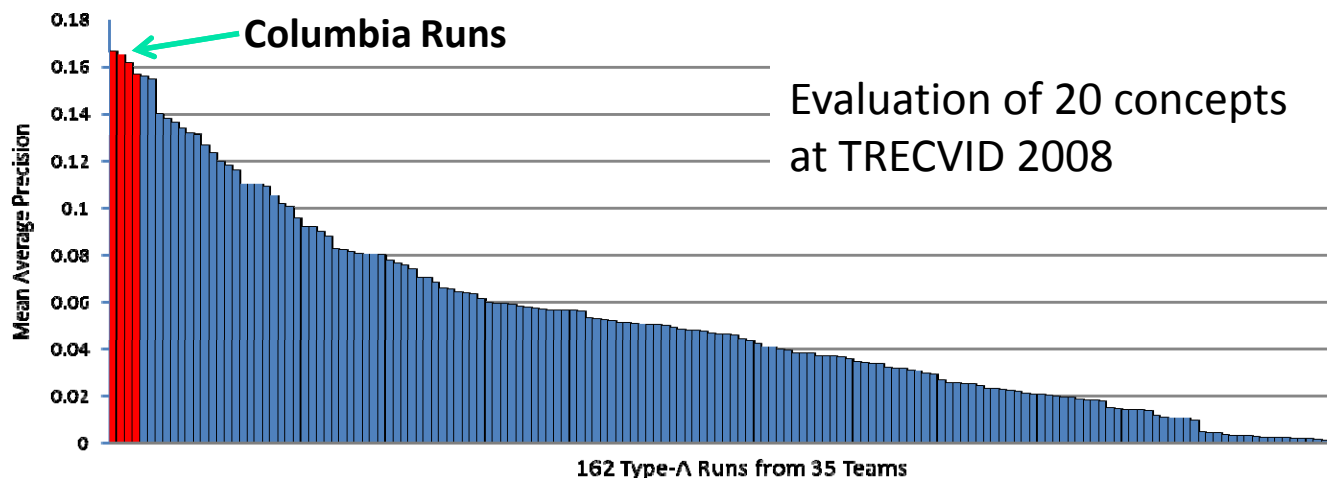
CuZero: tag videos with 400+ classifiers



concept detection models:
objects, people, location, scenes,
events, etc

airplane airplane_takeoff airport_or_airfield armed_person building car cityscape crowd
desert dirt_gravel_road entertainment explosion_fire forest highway hospital insurgents
landscape maps military military_base military_personnel mountain nighttime people-
marching person powerplants riot river road rpg shooting smoke tanks urban
vegetation vehicle waterscape_waterfront weapons weather

TRECVID 2008 High-Level Feature Extraction



TRECVID: Detection Examples

- Top five classification results

Classroom



Demonstration Or Protest



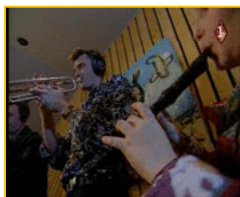
Cityscape



Airplane flying



Singing



What can a small recognition engine do?

- How to leverage a noisy and small visual vocabulary?
- State of the Art video search engines have ~1000 visual concepts
 - IBM IMAR
 - U. Amsterdam MediaMill
 - CMU Informedia
 - Columbia CuZero

Obvious Problems with Small Vocabulary

- Given a search topic, users often have difficulty in choosing matched concept classifiers

Find shots of something burning with flames visible



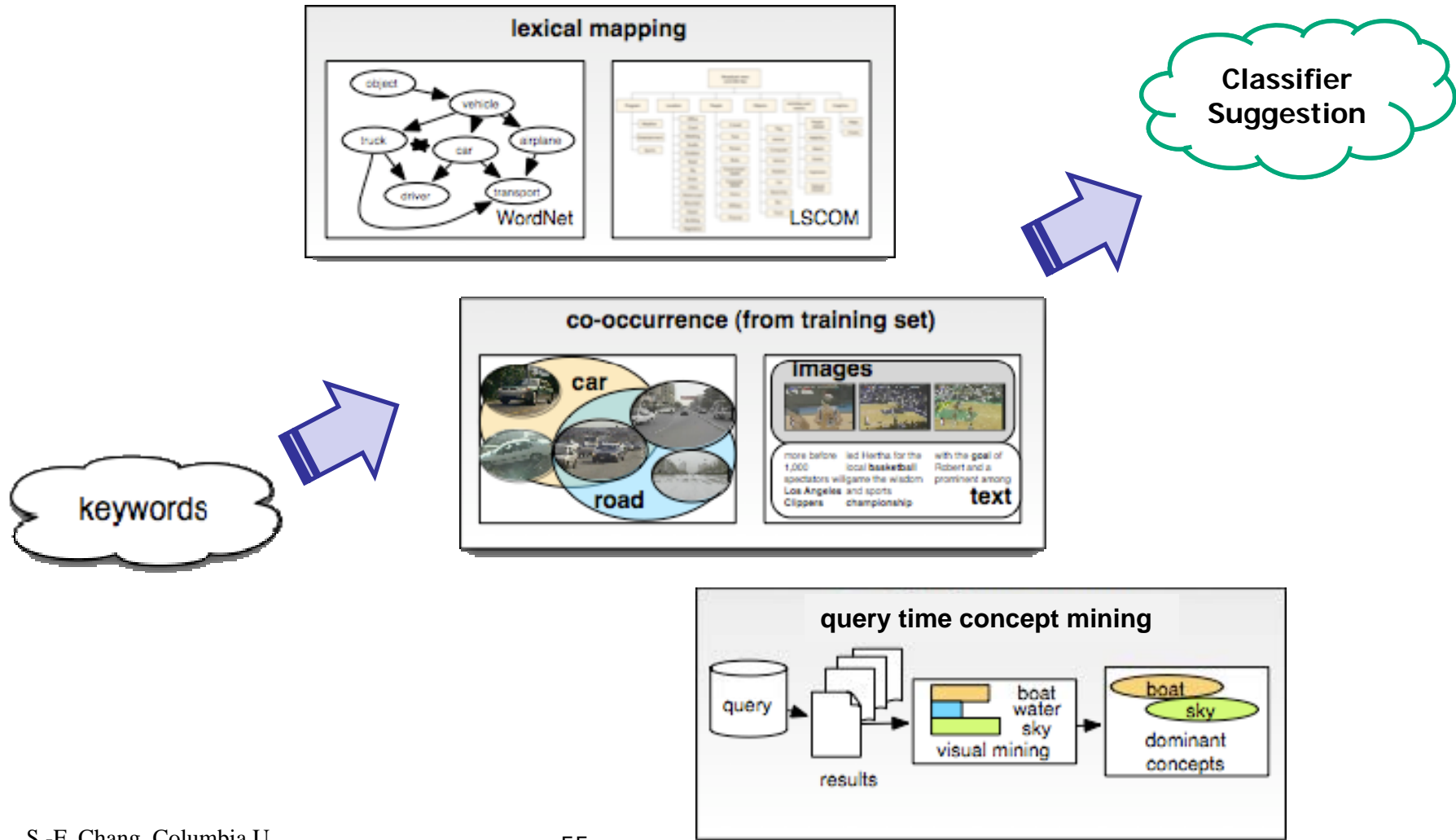
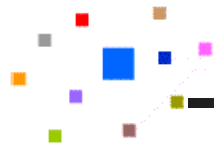
Cannot find matched classifiers!
Which classifiers work?

hundreds of classifiers

A cloud-shaped graphic containing several classifier terms: car, urban, fire, outdoor, airplane, road, Explosion, car crash, building, and person.

car	fire	outdoor
urban		building
airplane	car crash	person
road	Explosion	

Ideas from Information Retrieval: Query Expansion



Use visual examples to compensate the deficiency

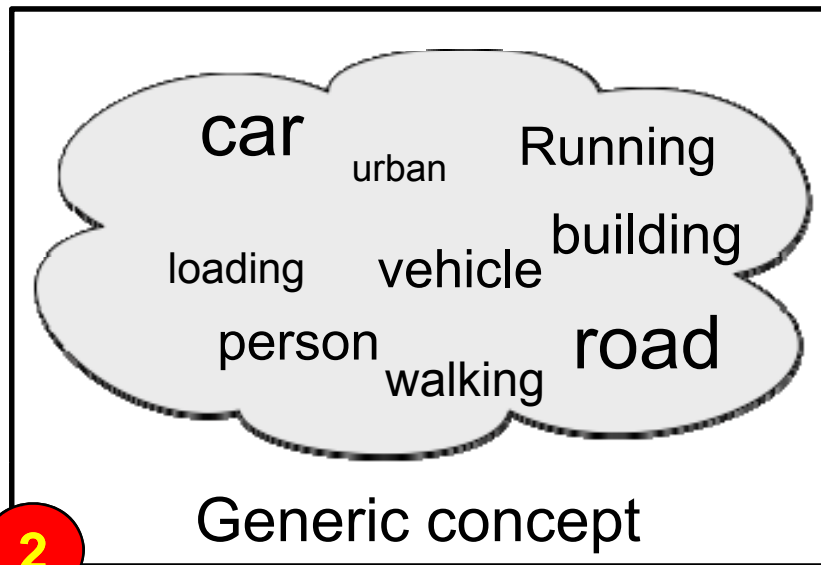
1

Specific exemplar



Query: “find person running around a building”

2

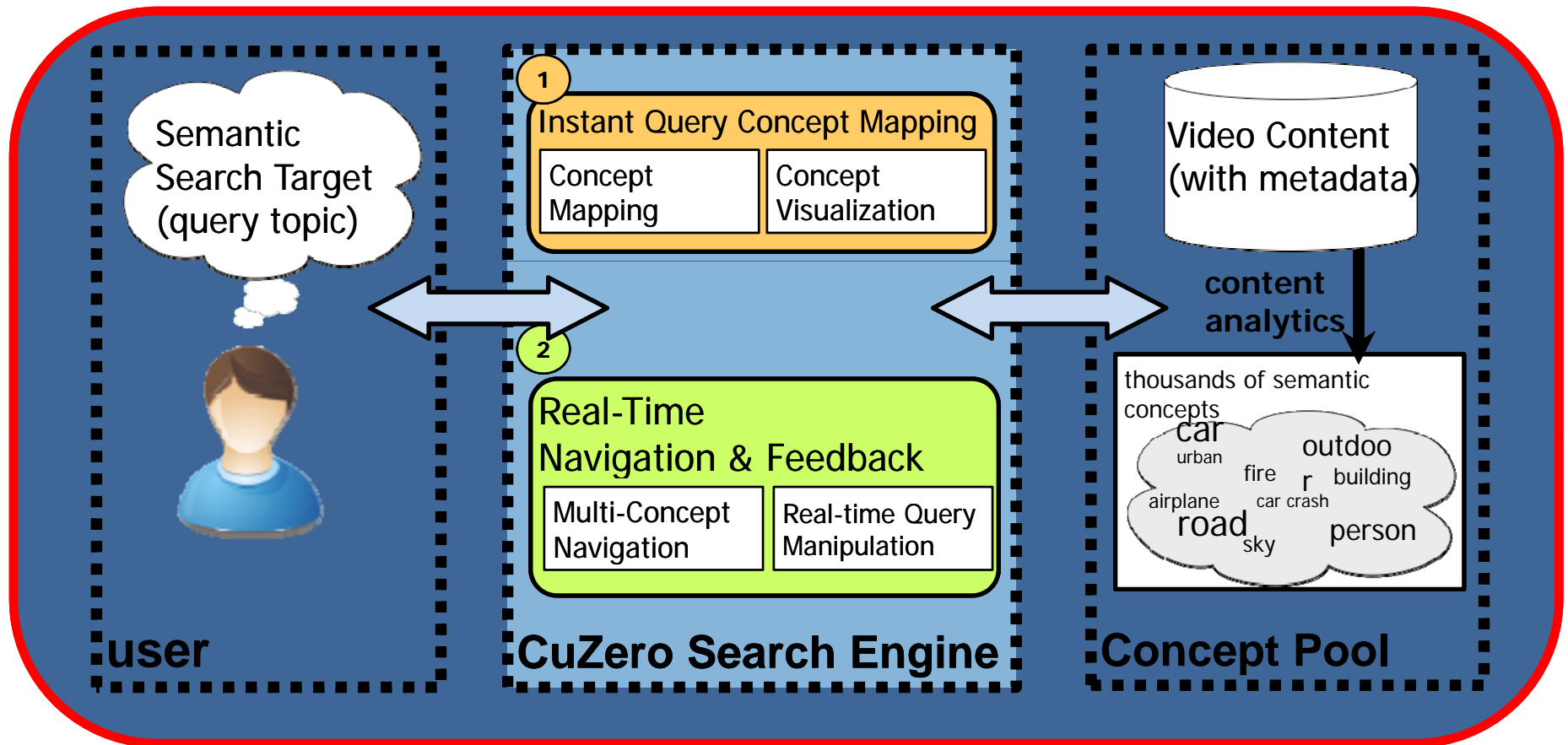


Generic concept

CuZero

Zero-Latency Video Search

<http://www.ee.columbia.edu/cuzero>



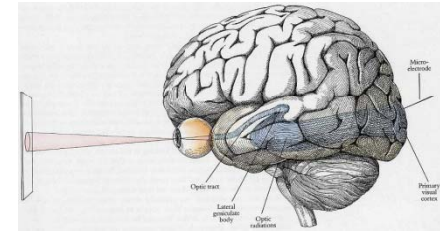
(Zavesky and Chang ,
Multimedia Info Retrieval MIR '08)

Demos

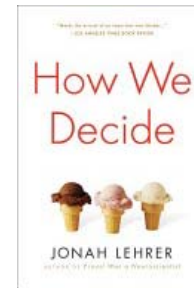
- Find lake front buildings in the park
- Find person walking around building
- Find a car on a road in a snowy condition

Visual Search via Brain State Decoding

- Human Vision is Superb by quick “gist” in the “Blink of an Eye”



(Hubel, 1995)



Joint work with Paul Sajda's group



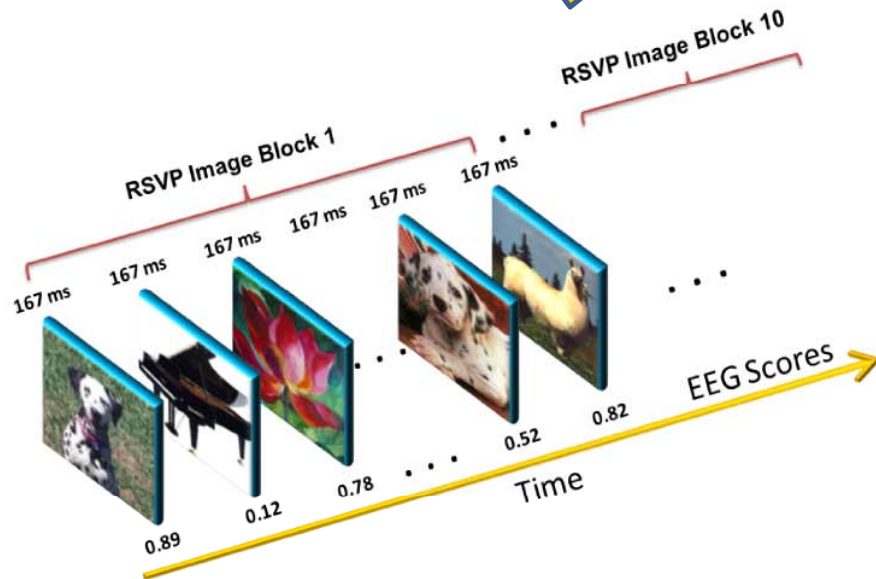
Brain Machine Interface for Image Retrieval

Use EEG brain signals
to detect target of interest

([video](#))

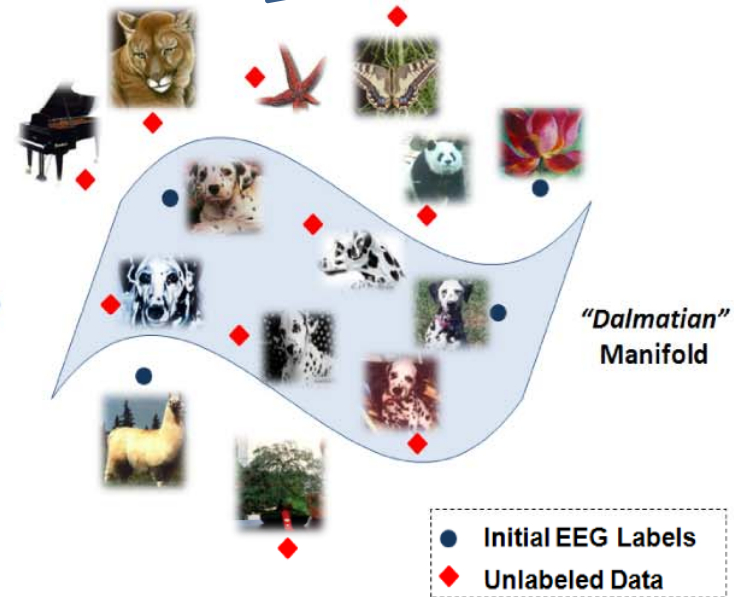


Use image graph to tune
& propagate information



Rapid Serial Presentation of Caltech 101 Images

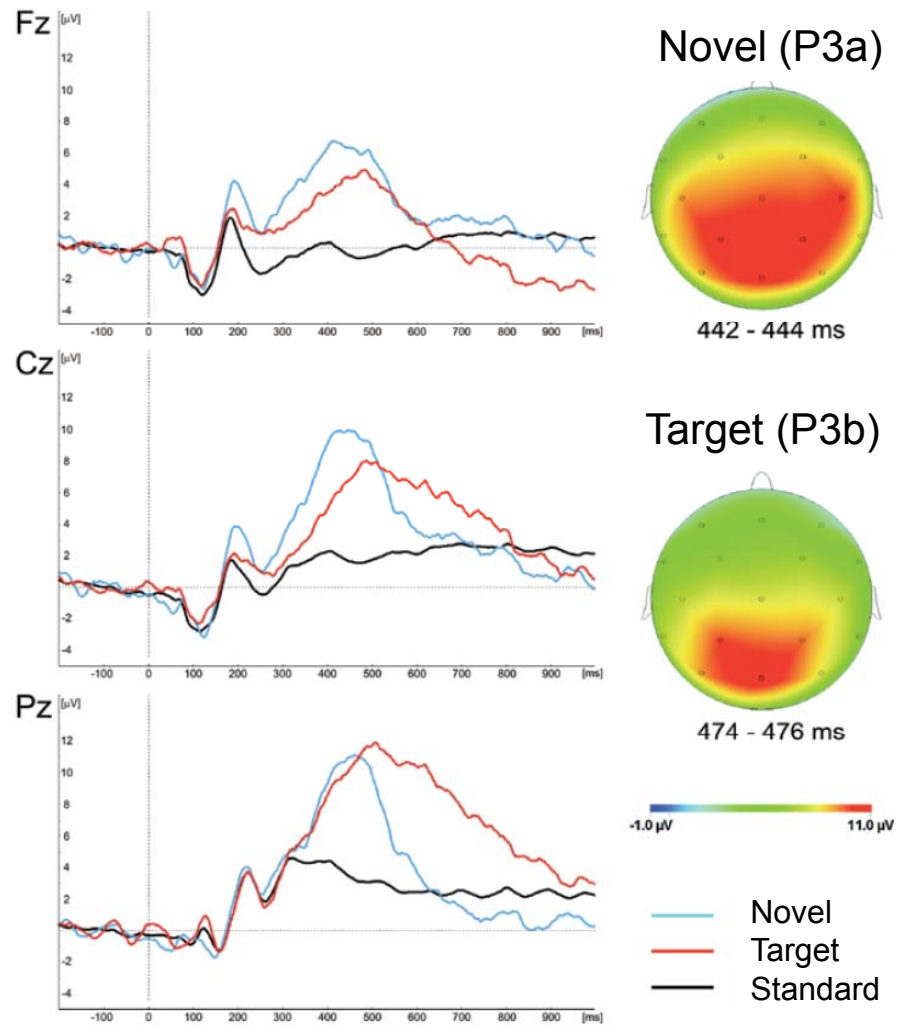
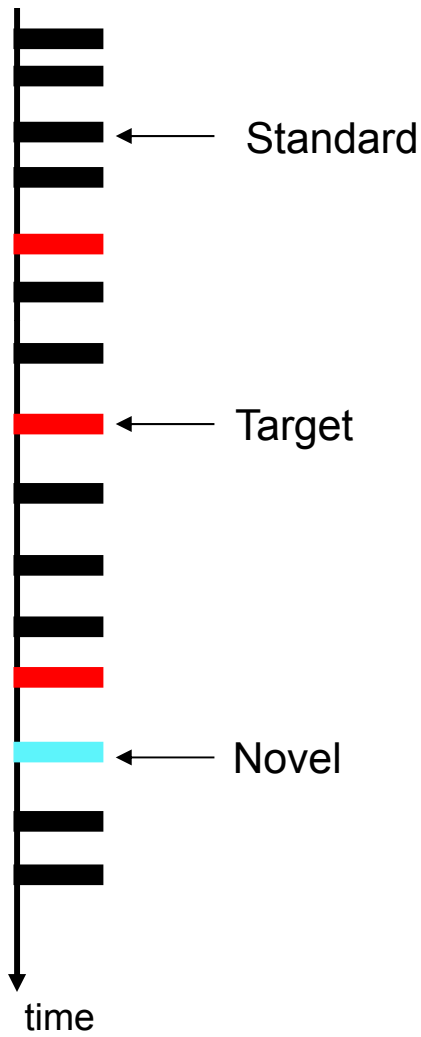
*C3Vision System
by Sajda et al*



Graph-Based Visual Pattern Discovery

*TAG System
by Wang and Chang*

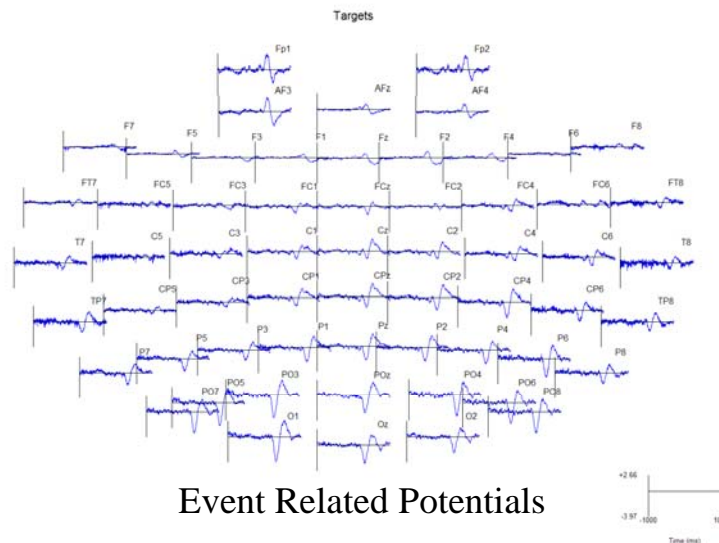
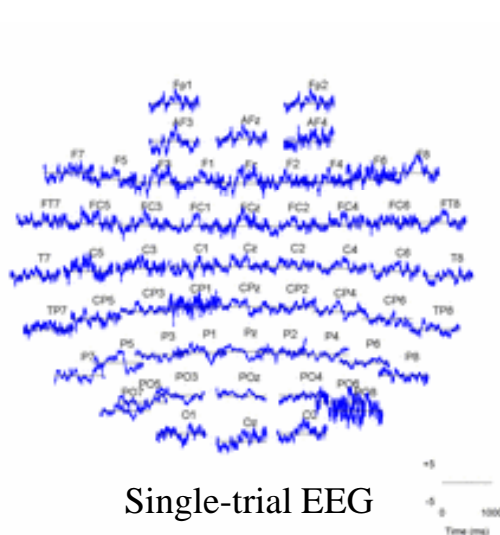
The Neural Signatures of “Visual Attention”



From D. Linden, 2005

Single-trial EEG Analysis

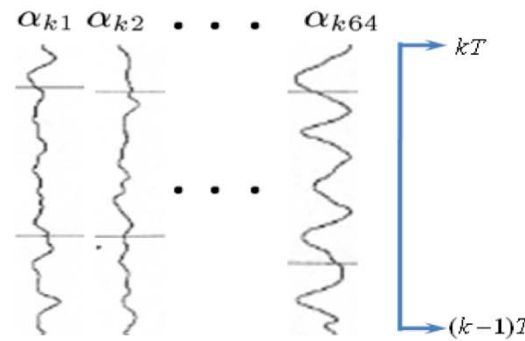
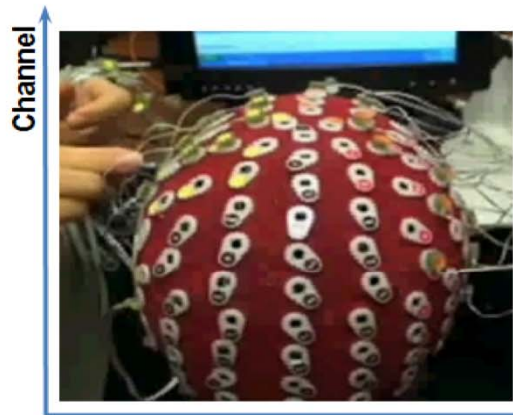
- Typically EEG is averaged over trials to increase the amplitude of the signal correlated with cortical processes relative to artifacts (very low SNR)
- High-density EEG systems were designed without a principled approach to handling the volume of information provided by simultaneously sampling from large electrode arrays.
- Our solution: identifying neural correlates with individual stimuli via single trial EEG analysis.
- We apply principled methods to find optimal ways for combining information over electrodes and moments in time contained in individual trials



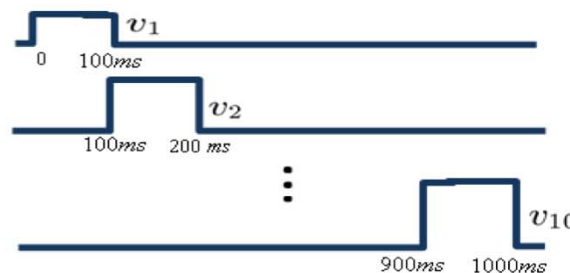
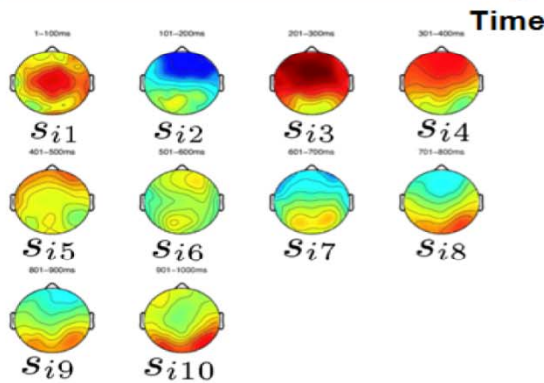
Identifying Discriminative Components in the EEG Using Single-Trial Analysis

LDA or Logistic Regression is used to learn the contributions of EEG signal components at different spatial-temporal locations

(Parra, Sajda et al. 2002, 2003)



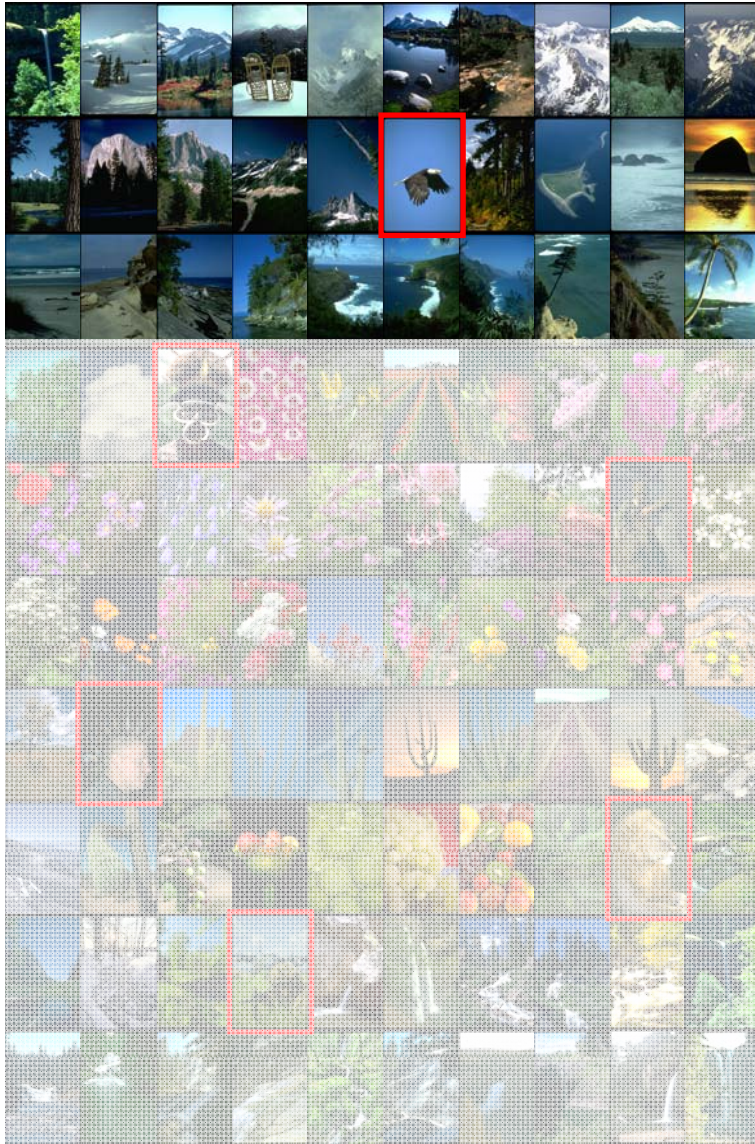
Optimal spatial filtering across electrodes within each short window (e.g., 100ms)



Optimal temporal filtering over time windows after onset

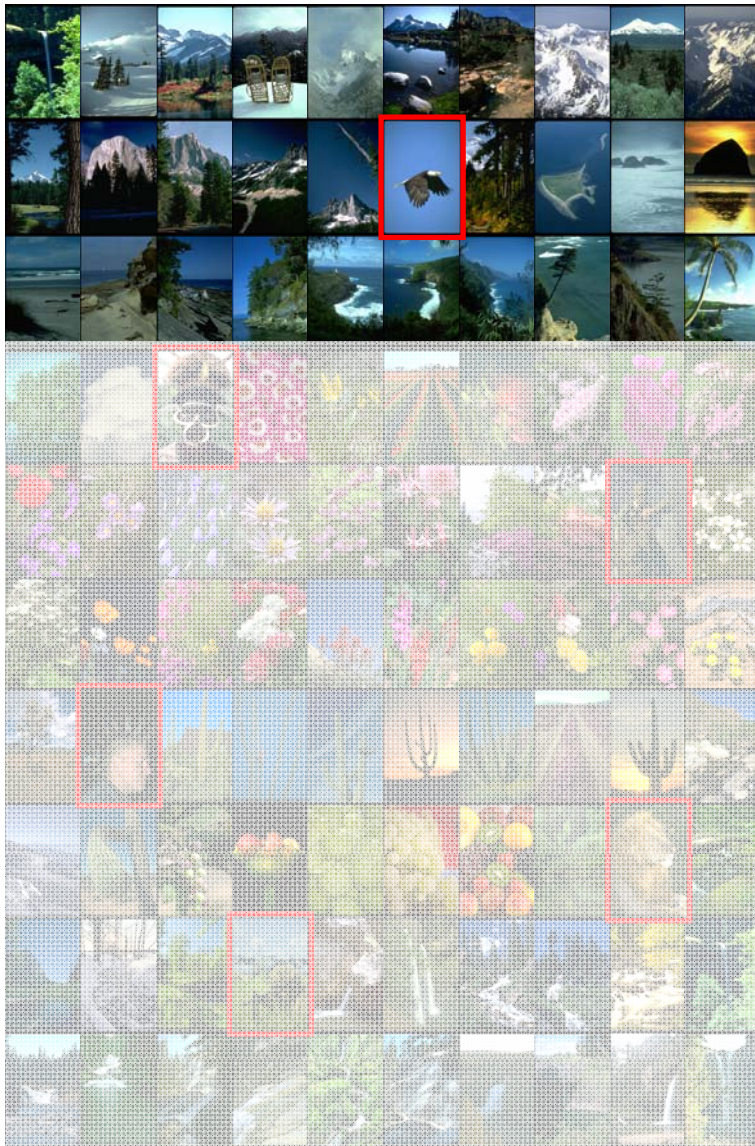
The Visual Interest Readout Experiment

User thinks about what he/she wants to search

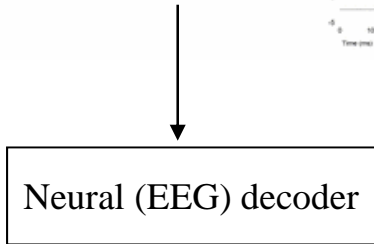


Database (any target that may interest users)

The Paradigm



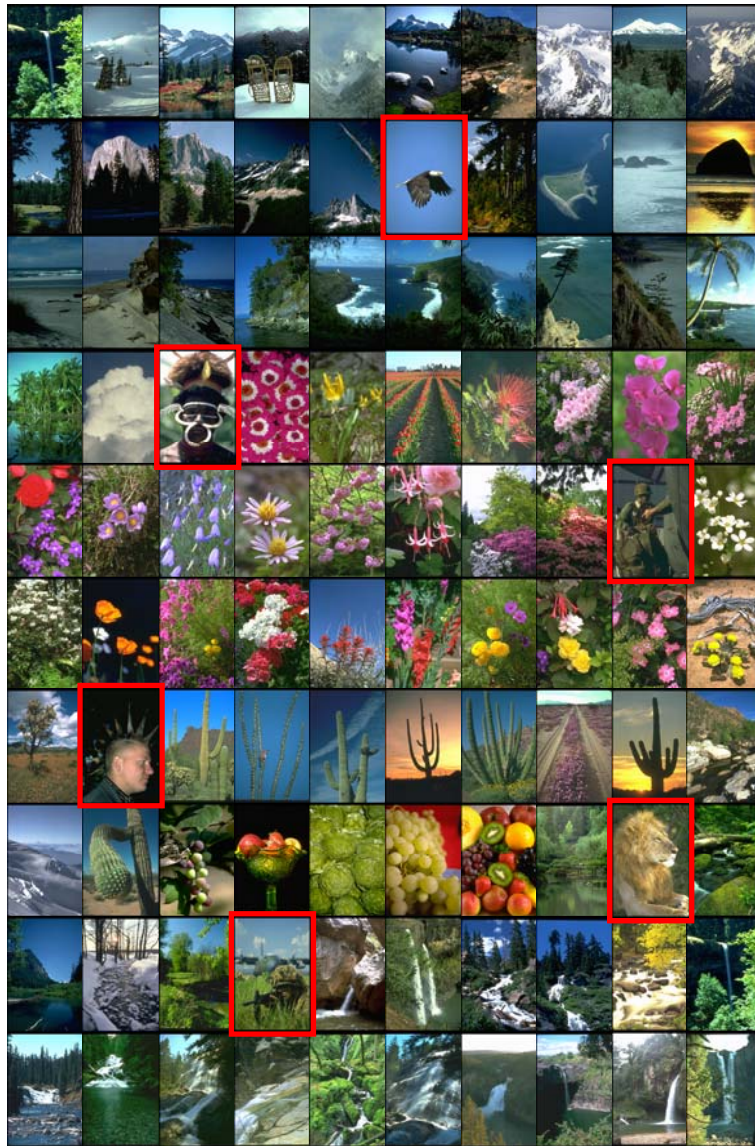
Database



↓

Interest-scores

The Paradigm



Database



Neural (EEG) decoder

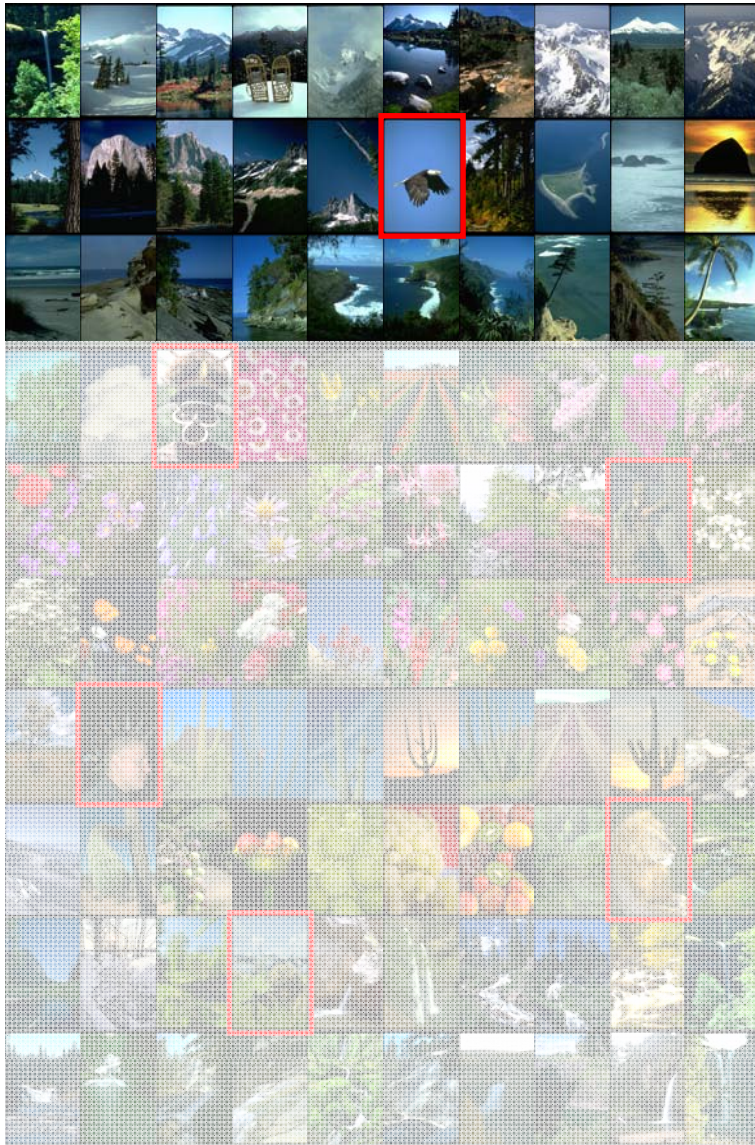
Exemplar labels (noisy)

Semi-supervised
Graph-based propagation

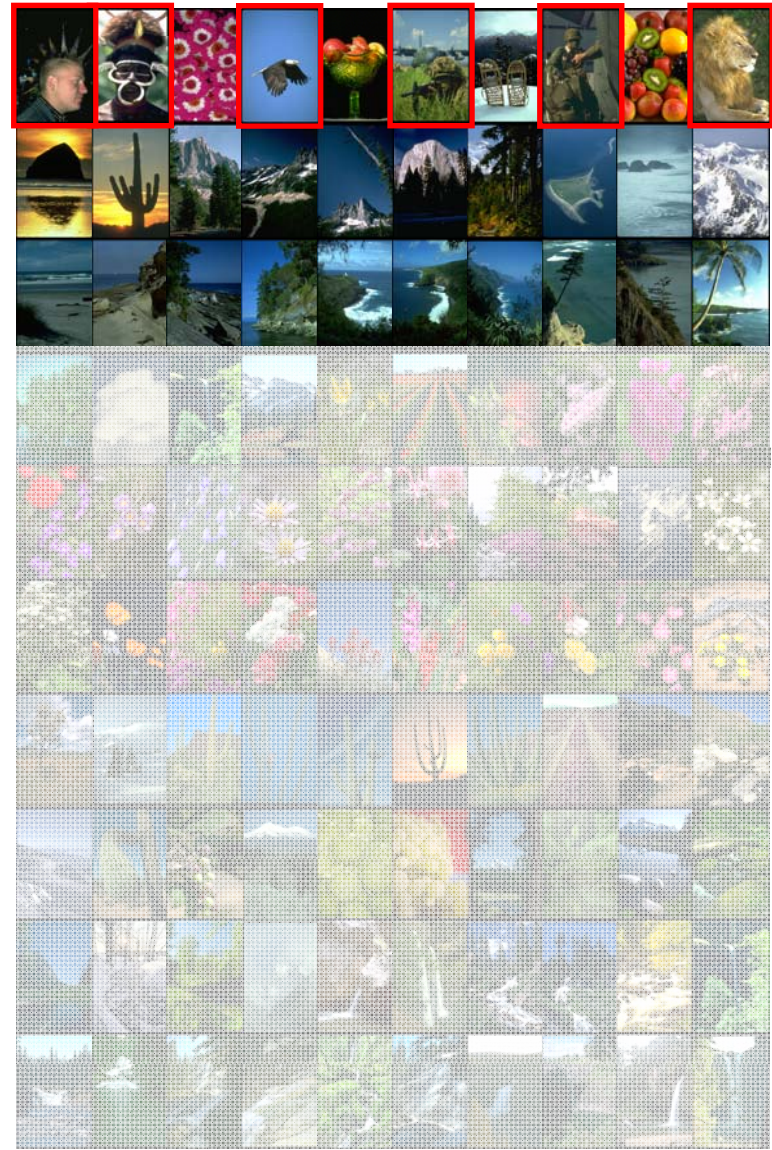
*Features from
the entire DB*

prediction score

The Paradigm

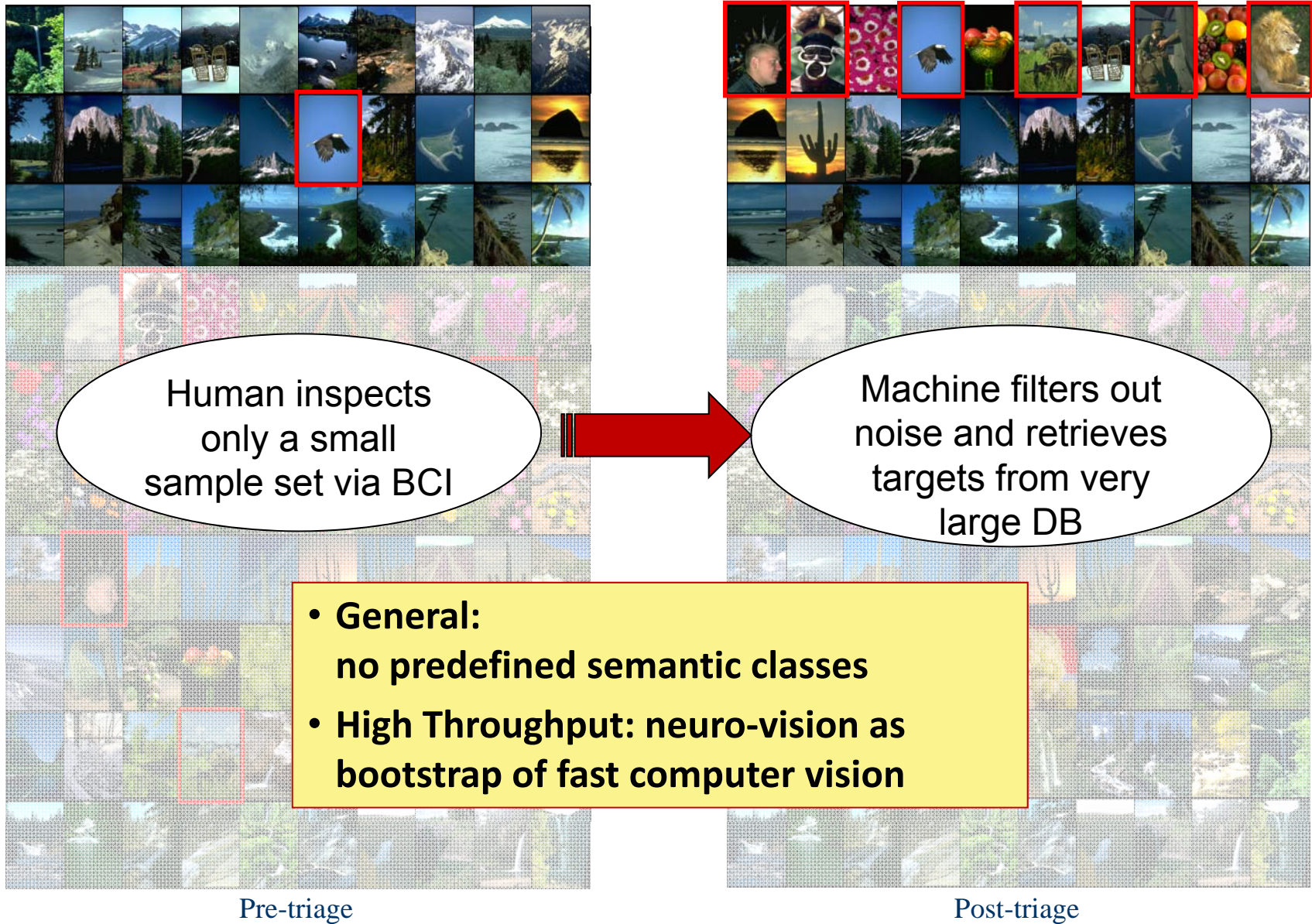


Pre-triage



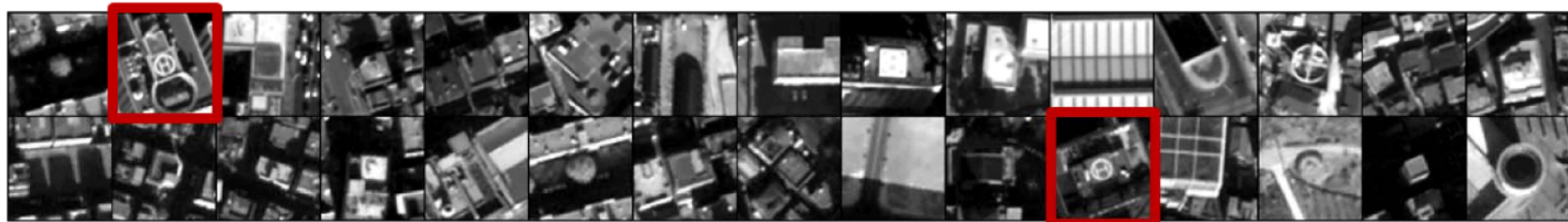
Post-triage

The Paradigm



Experiments

- CalTech101: 3798 images from 62 categories
Satellite images
- Generic neural decoder trained per user using images (*Soccer Ball* or *Baseball Gloves*) from Caltech256
- A subset images randomly sampled to construct 6-Hz RSVP sequence
- Initial Trials: 4 subjects, 3 targets (*Dalmatian*, *Chandelier/Menorah*, & *Starfish*)



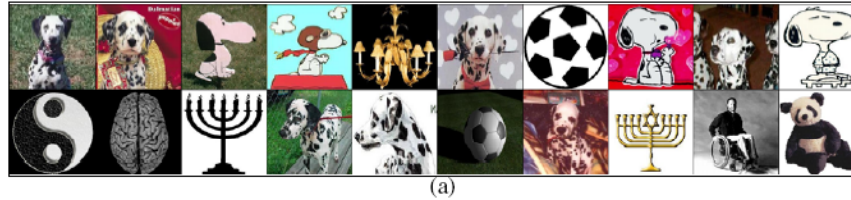
(a)



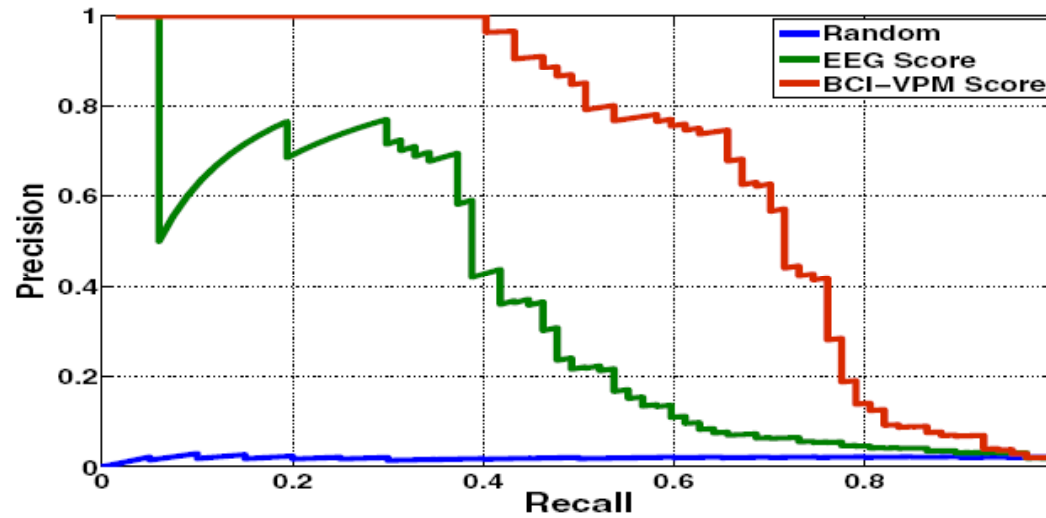
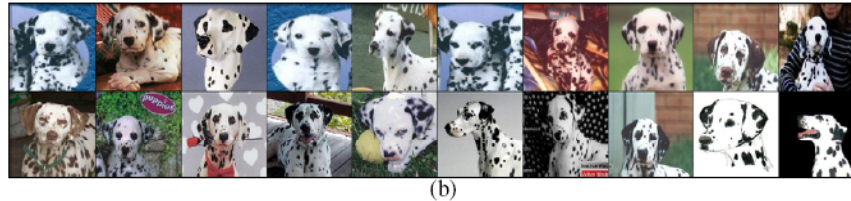
(b)

Example results

Top 20 results of
Neural EEG detection



Top 20 results of
Hybrid System (BCI-VPM)



DARPA NIA Program: Remote Sensing Target Search

Initial EEG neural signal detection:



(a)



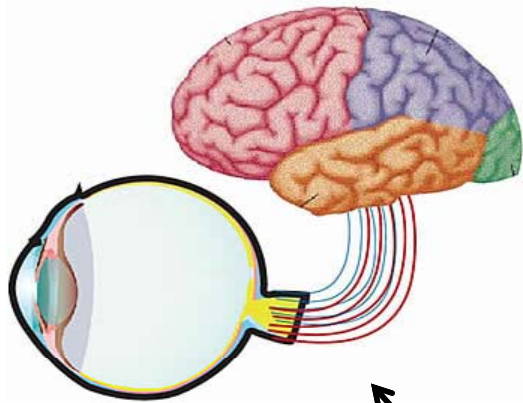
(b)

After graph refinement and diffusion

Images from DigiGlobe)

Summary: Cross Fertilization of Several Fields

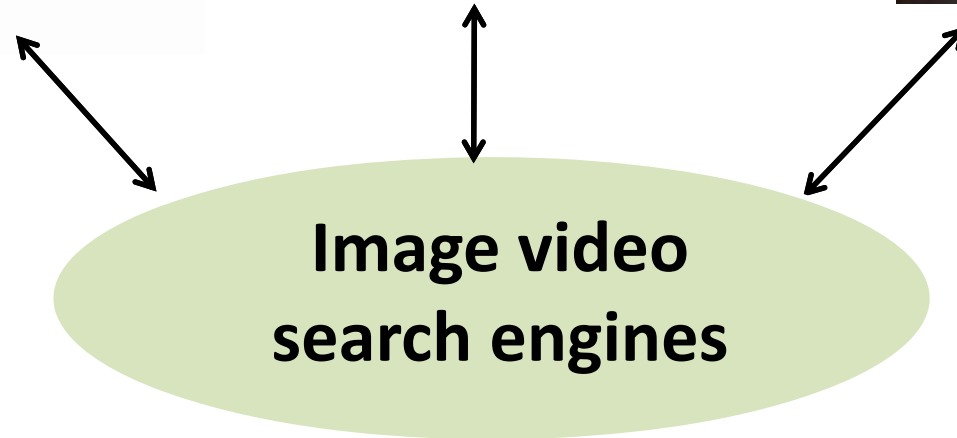
Human Vision



Internet/Data Vision



Image Recognition/
Machine Learning



Conclusions

- Great opportunity for video search research
- Exciting topics
 - **Semantic Search:**
Large-scale visual ontology and intuitive search
 - **Machine Learning and Computer Vision:**
Robust classification and image understanding
 - **Matching of Billions of Images or More**
Robust features and fast matching
 - **Internet Vision:**
Explore new applications on Internet
 - **Neuro-Computer Vision:**
Synergistic integration with neural vision systems



Acknowledgments

- Columbia University
 - Eric Zavesky, Yu-Gang Jiang, Jun Wang, Junfeng He, Wei Liu, Wei Jiang, Akira Yanagawa
- Yahoo! Research
 - Lyndon Kennedy
- City University, Hong Kong
 - Chong-Wah Ngo
- IRIT, France
 - Elie El Khoury

References

(many papers can be found at <http://www.ee.columbia.edu/dvmm>)

- E. Zavesky, S.-F. Chang, “**CuZero**: Embracing the Frontier of Interactive Visual Search for Informed Users,” ACM Multimedia Information Retrieval, 2008.
- Y.-G. Jiang, A. Yanagawa, S.-F. Chang, C.-W. Ngo, “**CU-VIREO374**: Fusing Columbia374 and VIREO374 for Large Scale Semantic Concept Detection,” ADVENT Technical Report #223-2008-1 Columbia University, August 2008. (<http://www.ee.columbia.edu/ln/dvmm/CU-VIREO374/>)
- J. Wang, E. Pohlmeier, B. Hanna, Y.-G. Jiang, P. Sajda, and S.-F. Chang, “**Brain State Decoding** for Rapid Image Retrieval,” ACM Multimedia Conference, 2009.
- Jun Wang, Yu-Gang Jiang, Shih-Fu Chang, “**Label Diagnosis** through Self Tuning for Web Image Search,” CVPR 2009.
- Jun Wang, Sanjiv Kumar, Shih-Fu Chang, “**Semi-Supervised Hashing** for Scalable Image Retrieval”, CVPR 2010.
- Junfeng He, Wei Liu, Shih-Fu Chang, “Scalable Similarity Search with **Optimized Kernel Hashing**,” ACM SIGKDD, 2010.
- Wei Liu, Junfeng He, Shih-Fu Chang, “**Large Graph Construction** for Scalable Semi-Supervised Learning,” ICML 2010.
- M. Naphade, J.R. Smith, J. Tesic, S.-F. Chang, W. Hsu, L. Kennedy, A. Hauptmann, J. Curtis, “**Large-Scale Concept Ontology for Multimedia**,” *IEEE MultiMedia* magazine, 2006.
- Lyndon Kennedy, Shih-Fu Chang, “Internet image archaeology: automatically tracing the manipulation history of photographs on the web,” ACM Multimedia, 2008.