

Towards Semantic-Level Visual Search

Prof. Shih-Fu Chang

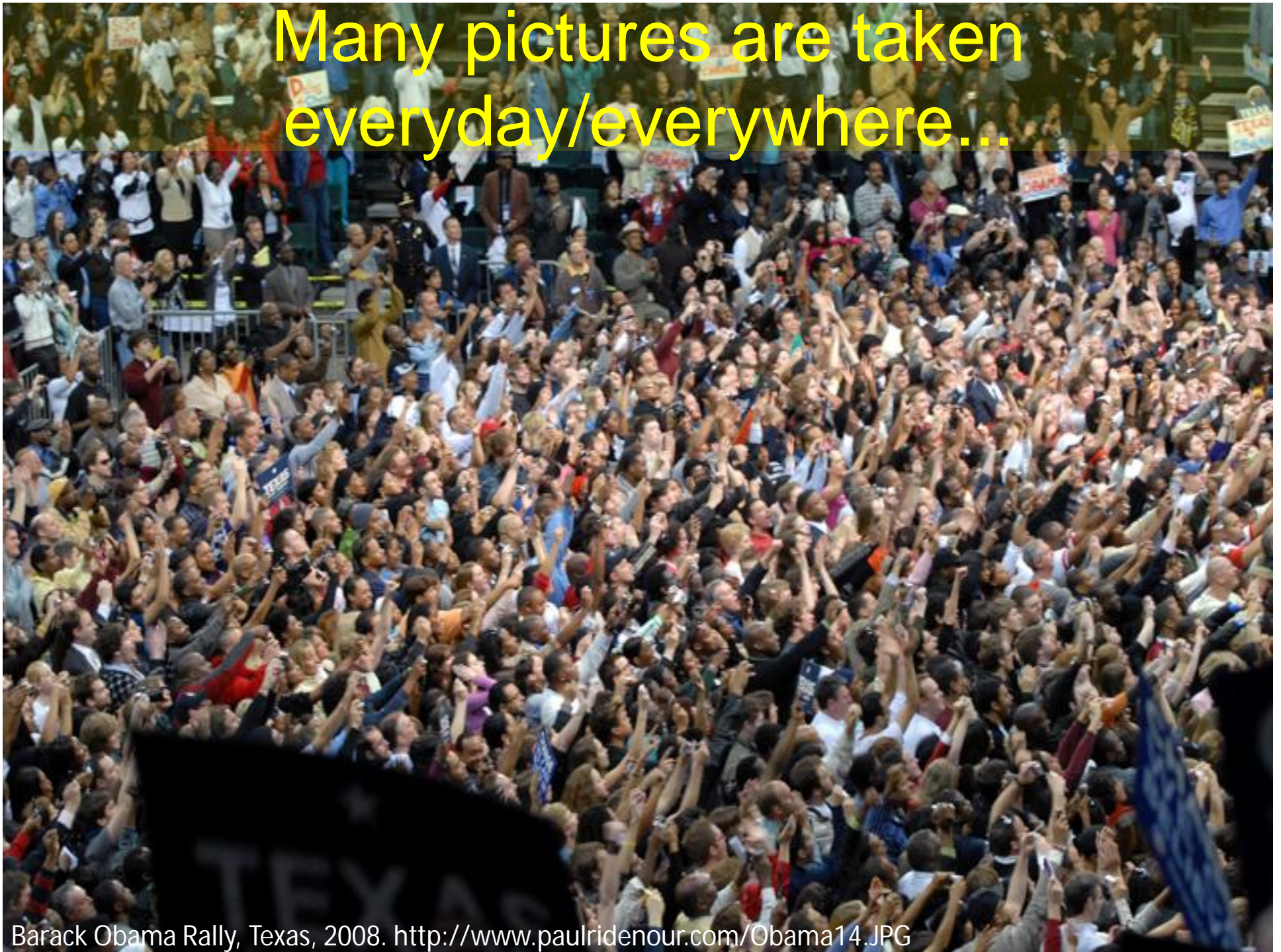
Department of Electrical Engineering

Digital Video and Multimedia Lab

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

Keynote talk at Int. Conf. on Semantic Computing, 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 ...

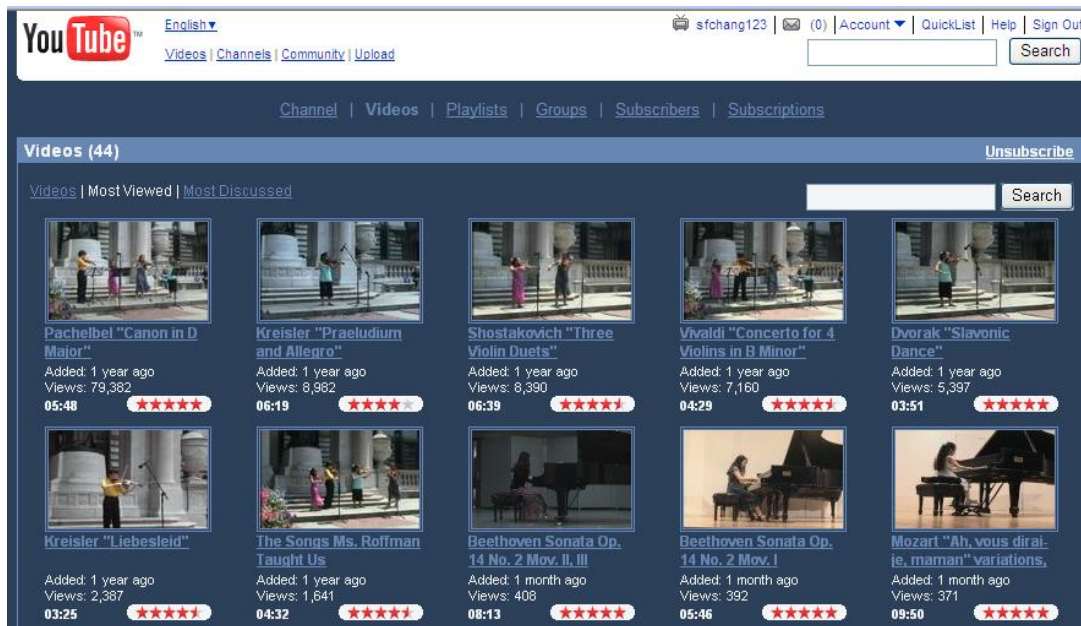


- Most photos and videos remain in shoebox or e-shoebox
- Each uploaded photo has only 0.97 tag on average
- But we love to share them ...

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 with a dark blue header. The channel name is not visible, but the user is logged in as 'sfchang123'. The page displays a grid of 10 video thumbnails, each with a title, duration, and view count. The videos are:

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



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

video

Challenge

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; . . .

- But tagging is boring and hard
- No wonder < 1 tag per uploaded photo



















Scarce Tags → Faulty Search Engines

“Manhattan Cruise”

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

Images Showing: All image sizes Any content All colors Results 1 - 18 of about 676,000 (0.13 seconds)

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

 <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 [corrected] ... 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 Search Images Search the Web [Advanced Image Search](#) [Preferences](#)
Moderate SafeSearch is on

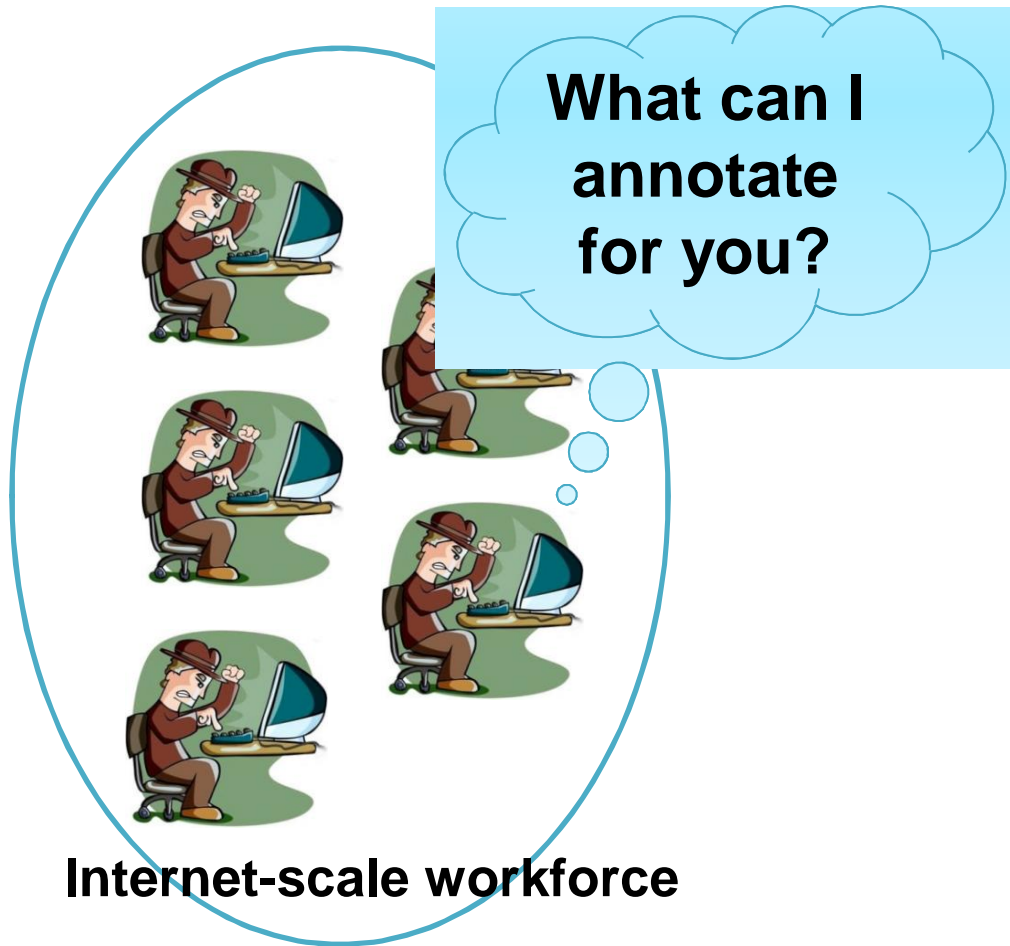
Images Showing: Results 1 - 18 of about 236,000 (0.17 seconds)

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.cruisecrazies.com</p>	 <p>Cruise Ship Crashes Into Manhattan ... 636 x 429 - 20k - jpg www.cruisecrazies.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>

Creative Solutions: 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](#)

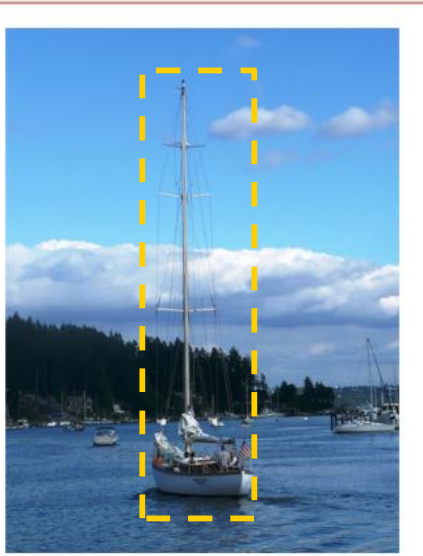
Medical Image Segmentation (OCT) - 1060 - BC
Requester: [Cardiff Turk](#)

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:

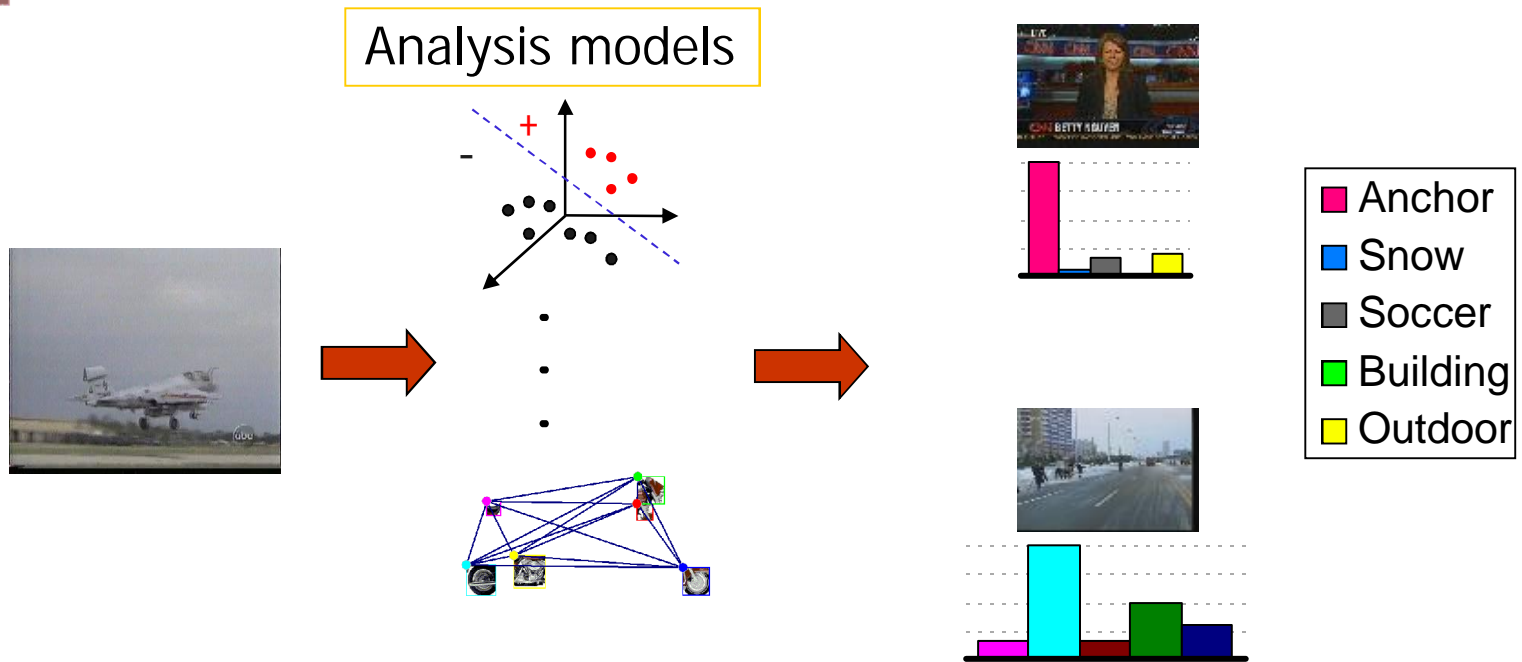
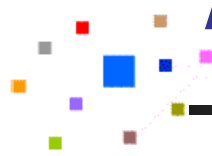
- Include all visible parts and draw as tightly as possible
- **If there are multiple instances, pick only ONE (any one).**
- **Do NOT draw on the instances that already have bounding boxes.**

[SEE INSTRUCTIONS WITH EXAMPLES](#)

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

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

An ideal alternative: Automatic Visual Annotation



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

- Rich semantic labels

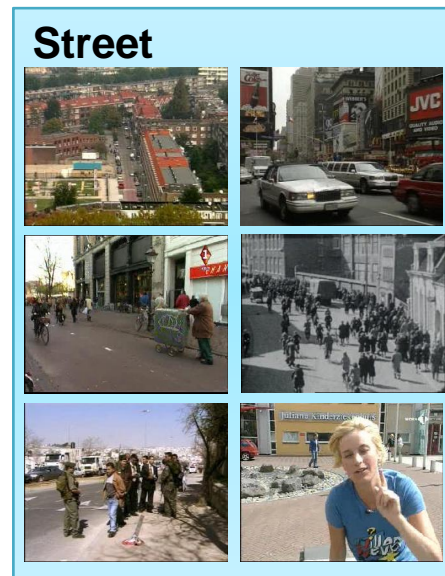
Hot topic ... community fast growing!

(as of Nov. 2009)

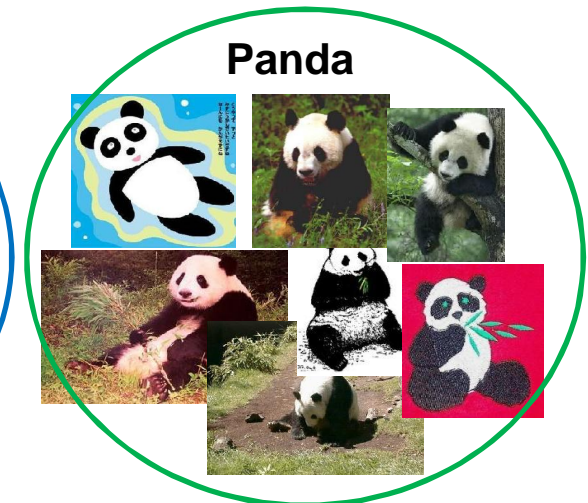
	Data domain	amount	types	Lexicon size
TRECVID	Broadcast news, documentary, Flickr, Youtube...	400 hours Sound & Vision 170 hours Television News 100 hours BBC rushes (130,000+ subshots)	video shots, keyframes	10 (2004, 2005) 39 (2006) 39 (2007) 20 (2008)
LSCOM	Broadcast news video	170 hours Television 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

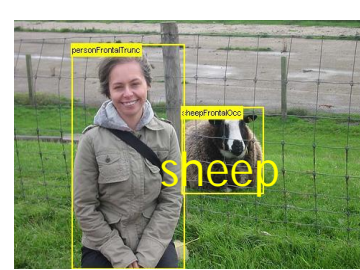
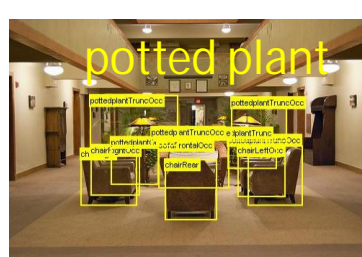
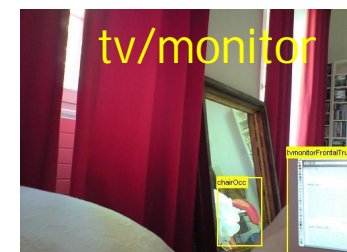
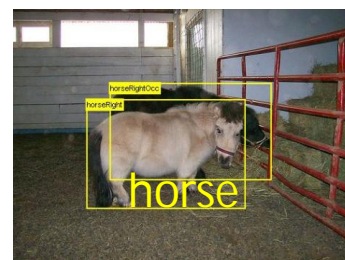
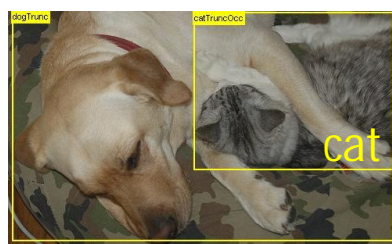
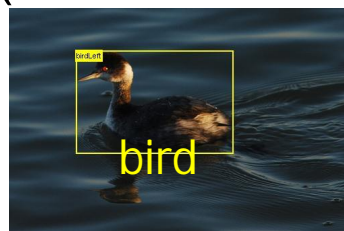
TRECVID



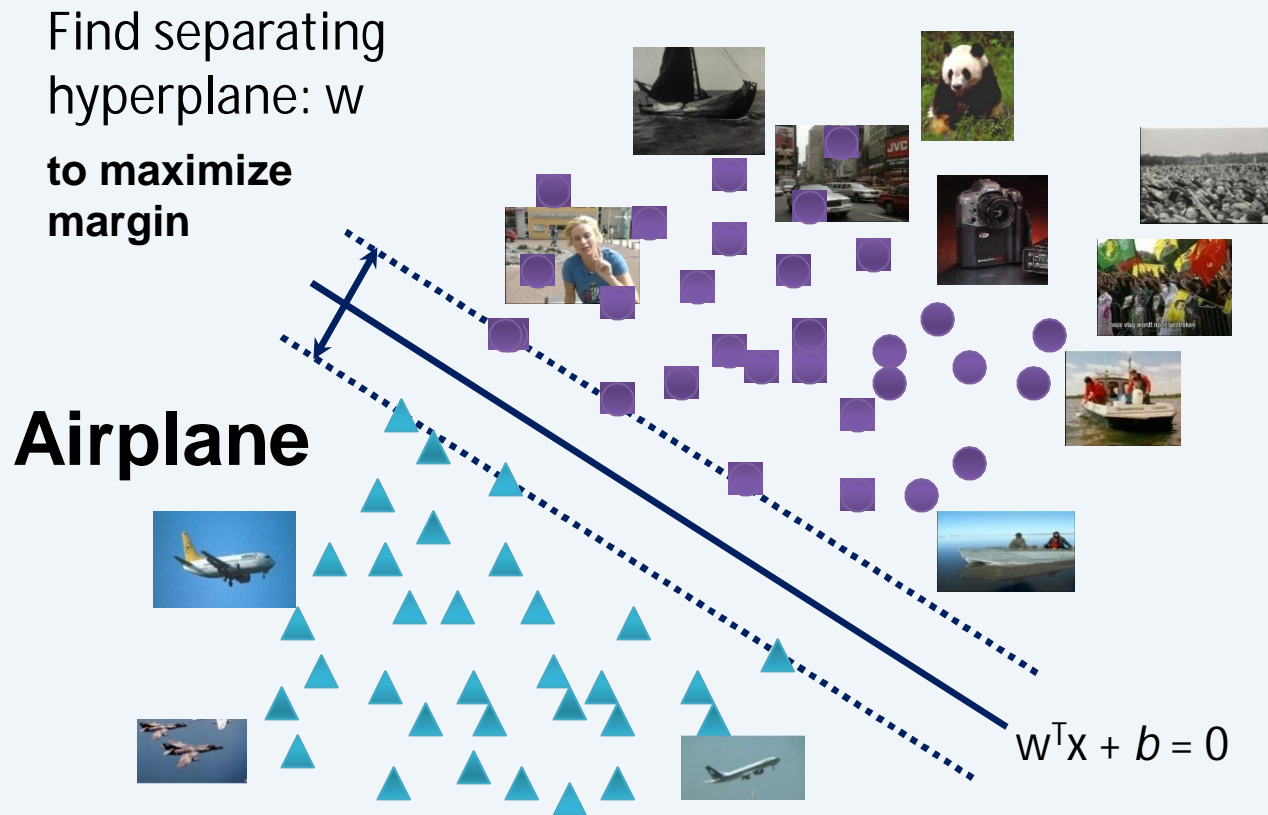
CalTech 101



Object Localization (PASCAL VOC)



Build Classifier



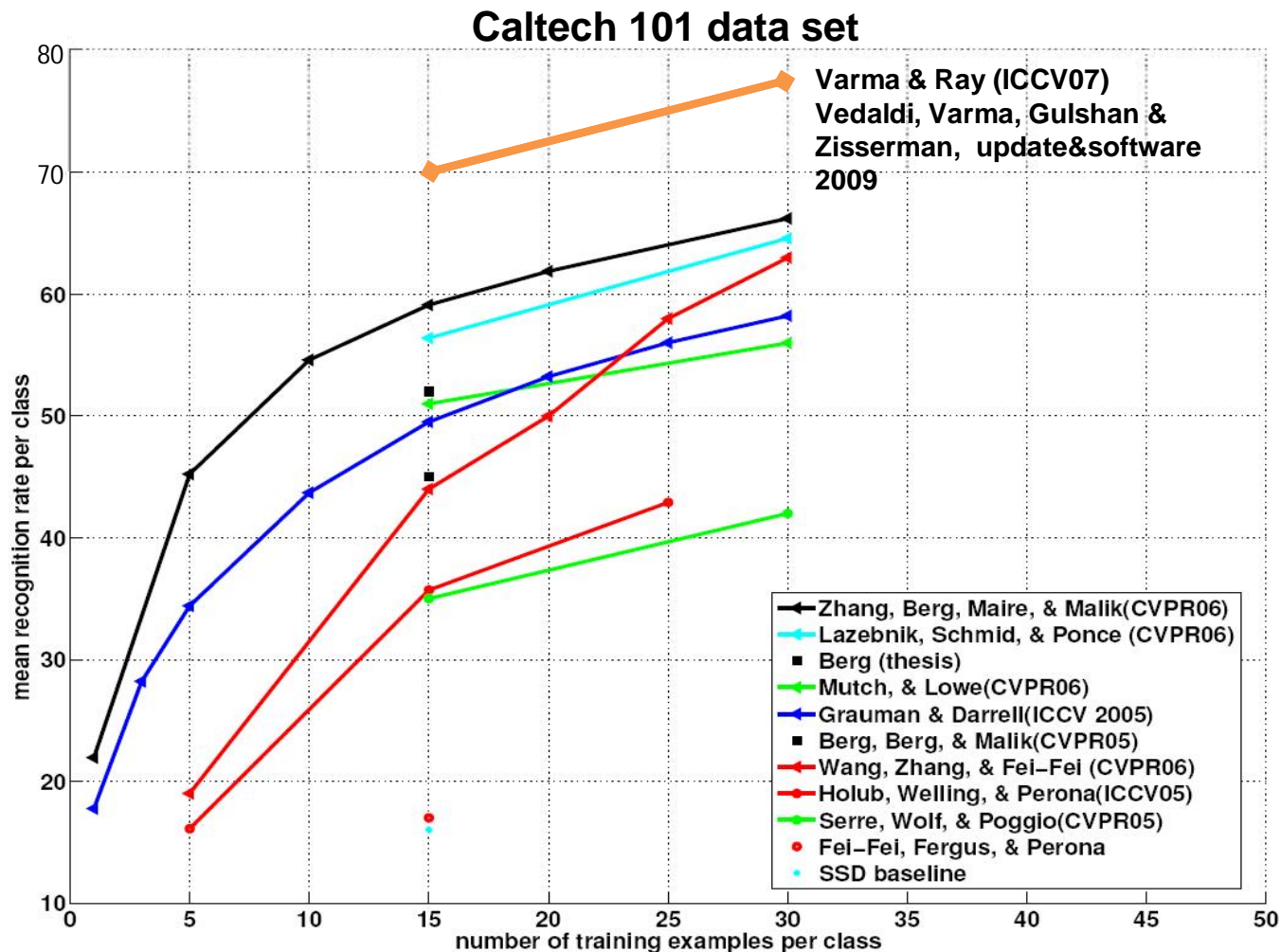
Decision function: $f(x) = \text{sign}(w^T x + b)$

$w^T x_i + b > 0$ if label $y_i = +1$

$w^T x_i + b < 0$ if label $y_i = -1$

Rapid Advances in Image Annotation

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



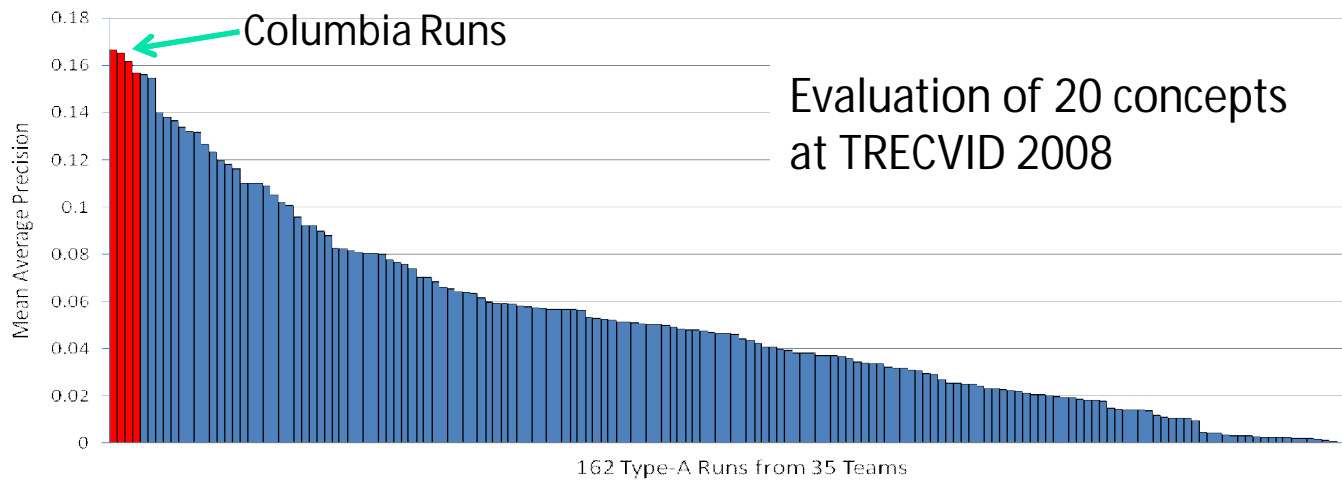
CuZero: tag videos with 400+ classifiers



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



LSCOM Ontology Defines 1000 Concept Classes

(IBM, Columbia, CMU '06)

- Large Scale Concept Ontology for Multimedia
 - Broadcast news video
 - Defined by experts and actual intelligence users
 - Selection criteria: *useful, observable, detectable*
 - 30M+ labels for 449 concepts annotated over 60,000+ video shots
 - Download site
 - <http://www.ee.columbia.edu/dvmm/lscom/>

TRECVID: Detection Examples

- Top five classification results

Classroom



Demonstration Or Protest



Cityscape



Airplane flying



Singing



What can a small recognition engine do?

- Several video search engines incorporate ~1000 visual concepts
 - IBM IMAR
 - U. Amsterdam MediaMill
 - CMU Informedia
 - Columbia CuZero
- But this is still a noisy and relatively small vocabulary

Cope with Small Noisy Vocabulary

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

Find shots of something burning with flames visible

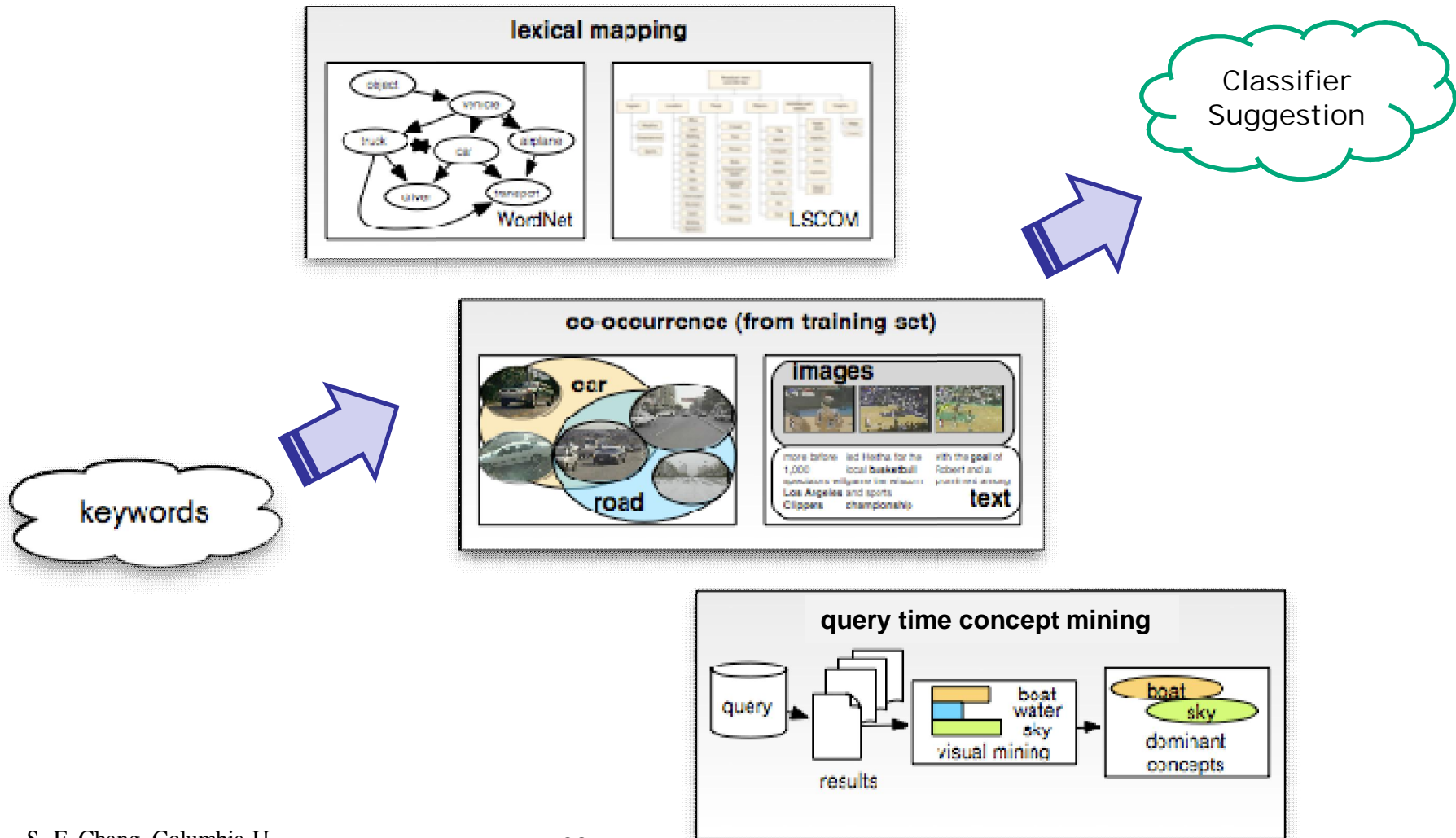
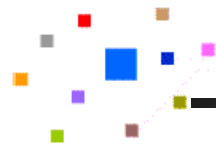


Cannot find matched classifiers!
Which classifiers work?

hundreds of classifiers

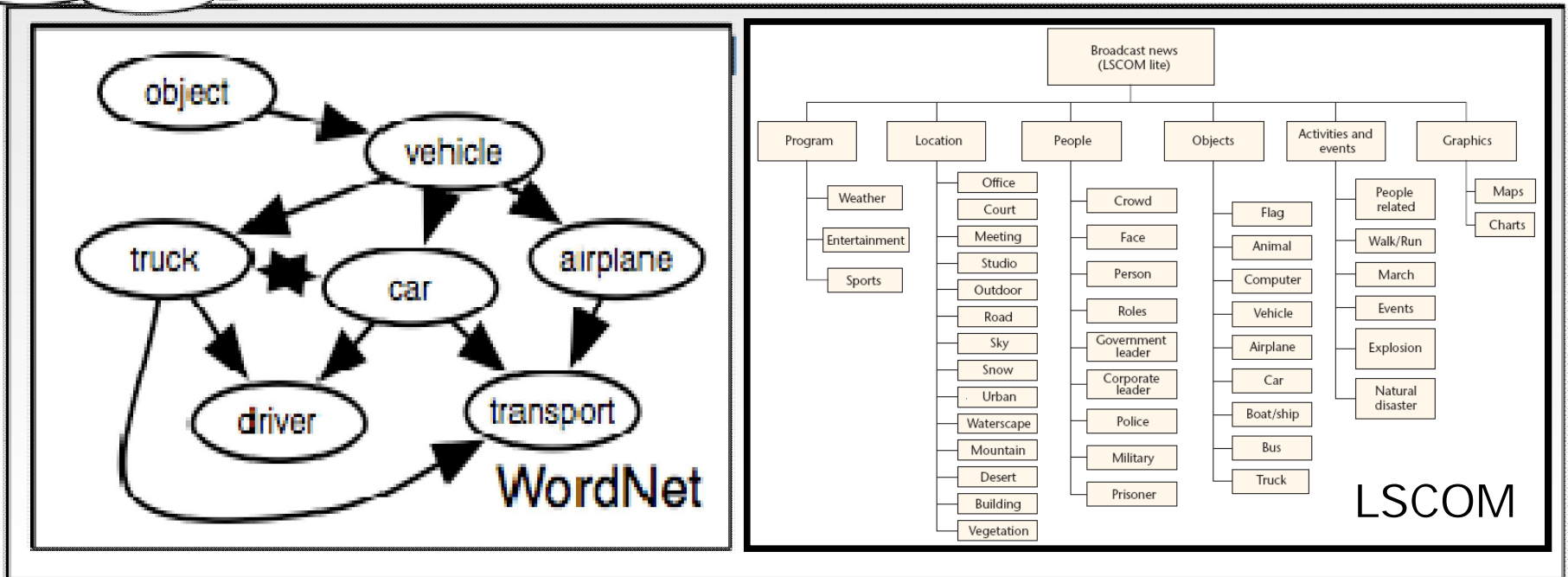
car
urban
fire
outdoor
airplane
road
car crash
building
Explosion
person

Lessons from IR: word-concept query expansion



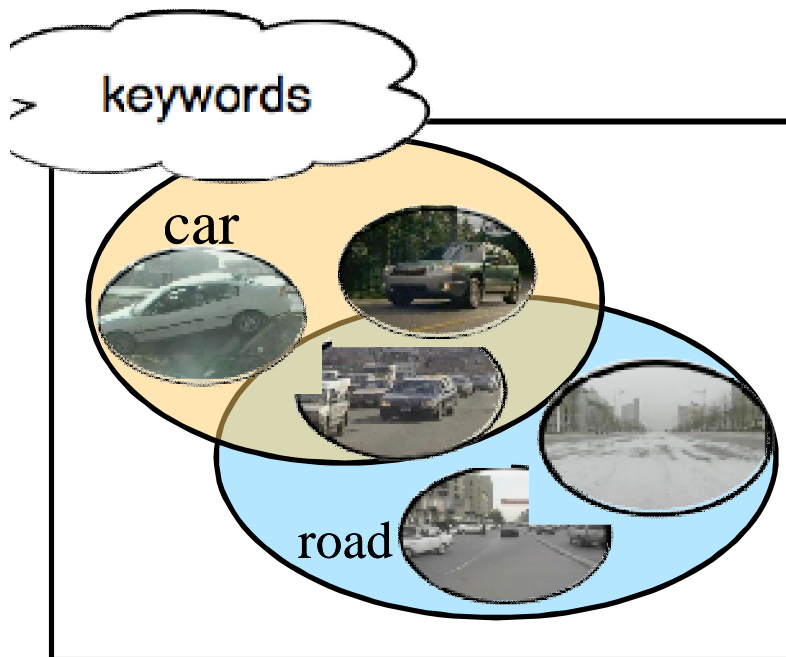
Lexical mapping

keywords



Mapping keywords to concept definition, synonyms, sense context, etc

Co-occurrent concepts



images

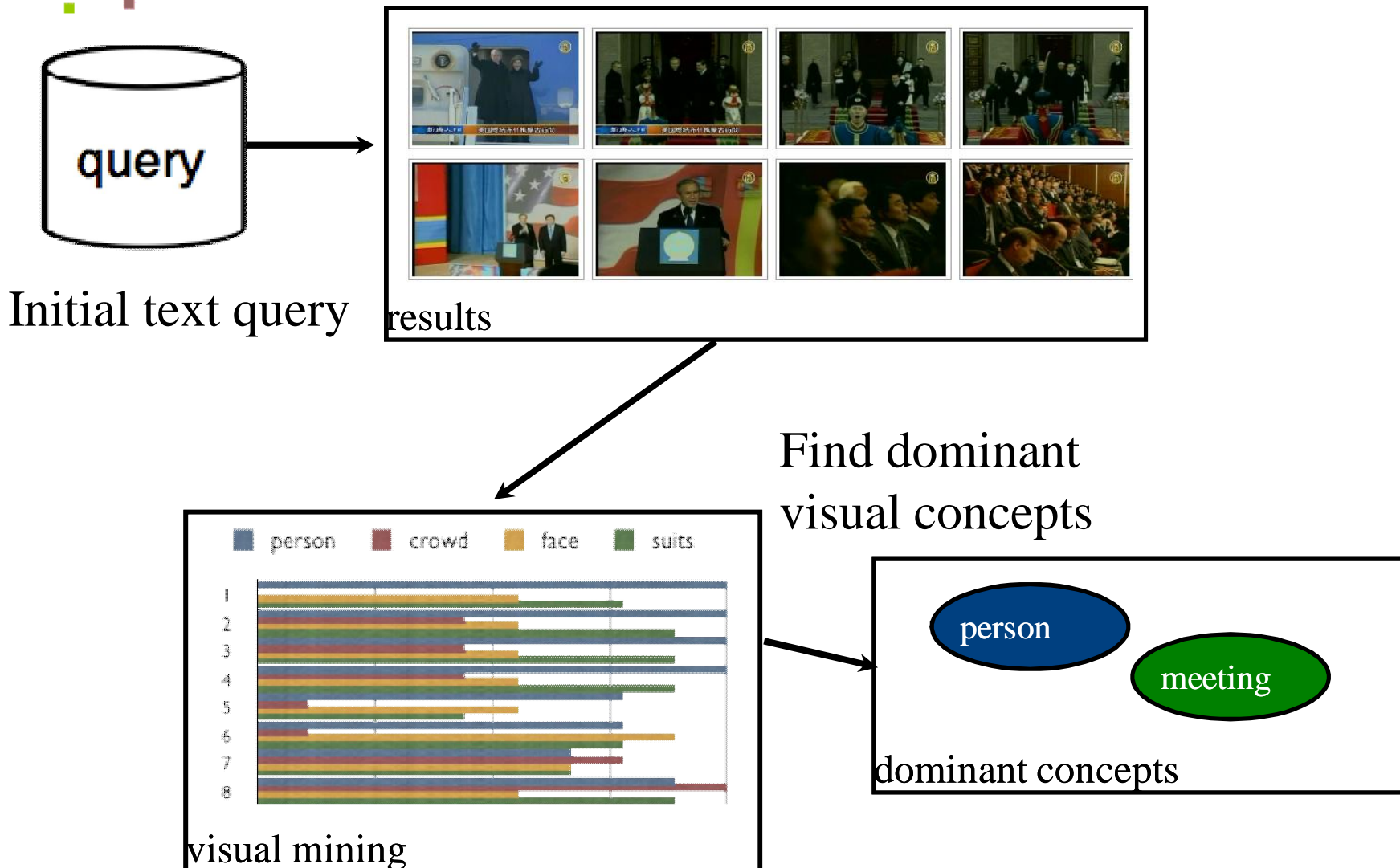
Basketball courts and the American won the Saint Denis on the Phoenix Suns because of the 50 **point** for 19 in their role within the National Association of **Basketball**

George Rizq **led** Hertha for the local **basketball** game the wisdom and sports **championship** of the president

Baghdad to attend the game I see more **goals** and the **players** did not offer great that Beijing Games as the beginning of his brilliance Nayyof 10 this atmosphere the culture of sports **championship**

text

Query-Time Concept Mining



Use visual examples to compensate the deficiency

1

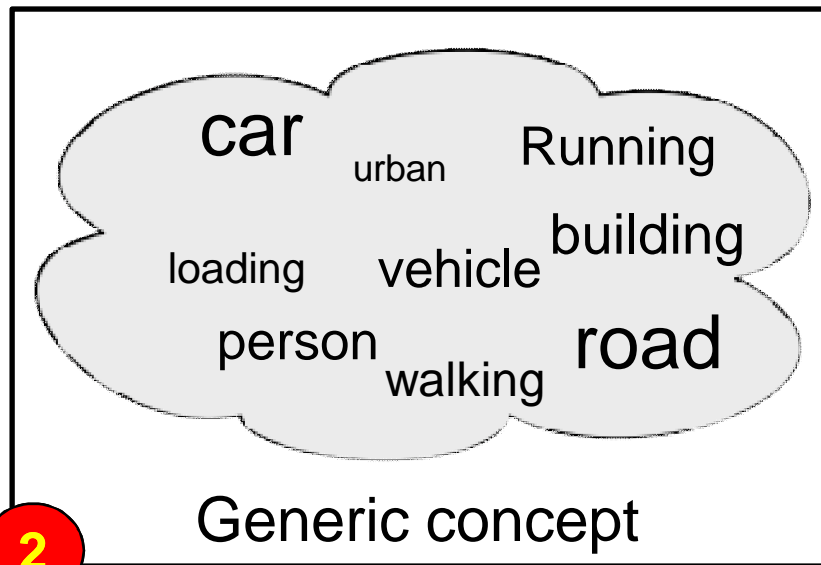
Specific exemplar



Query: “find person running around a building”

2

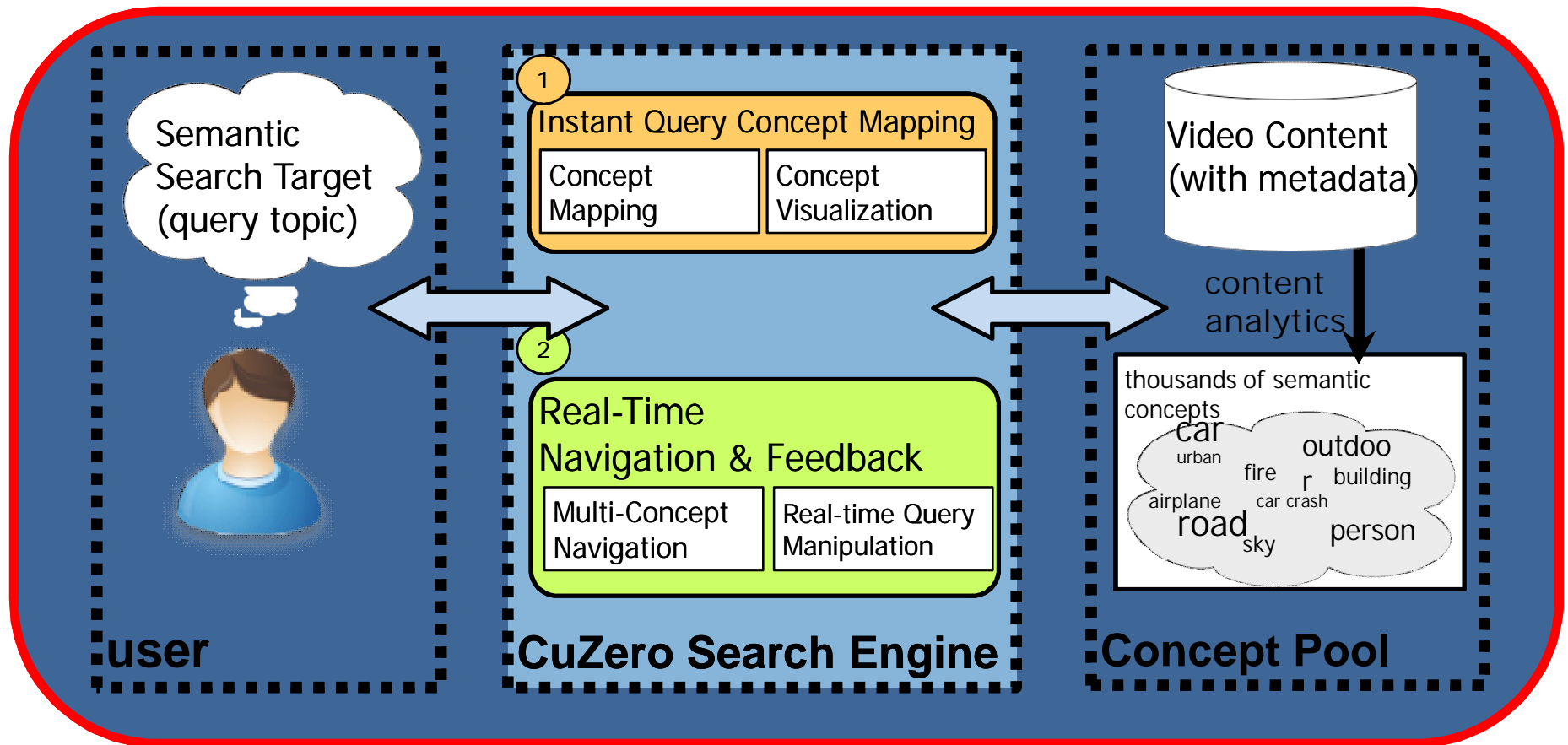
Generic concept



CuZero

Zero-Latency Video Search

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



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

Demos

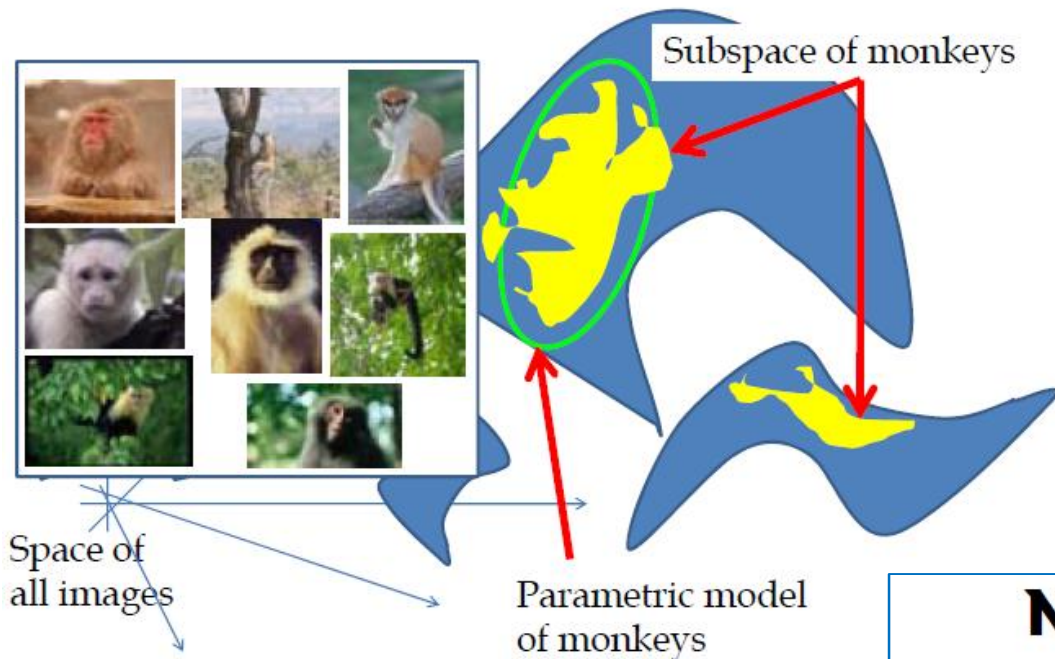
- Find lake front buildings in the park
- Find person walking around building
- Find a car on a road in a snowy condition

Scalability:

Robust recognition is still developing

- Many concept classes are very challenging
- Face, background scene, OCR are relatively mature compared to others

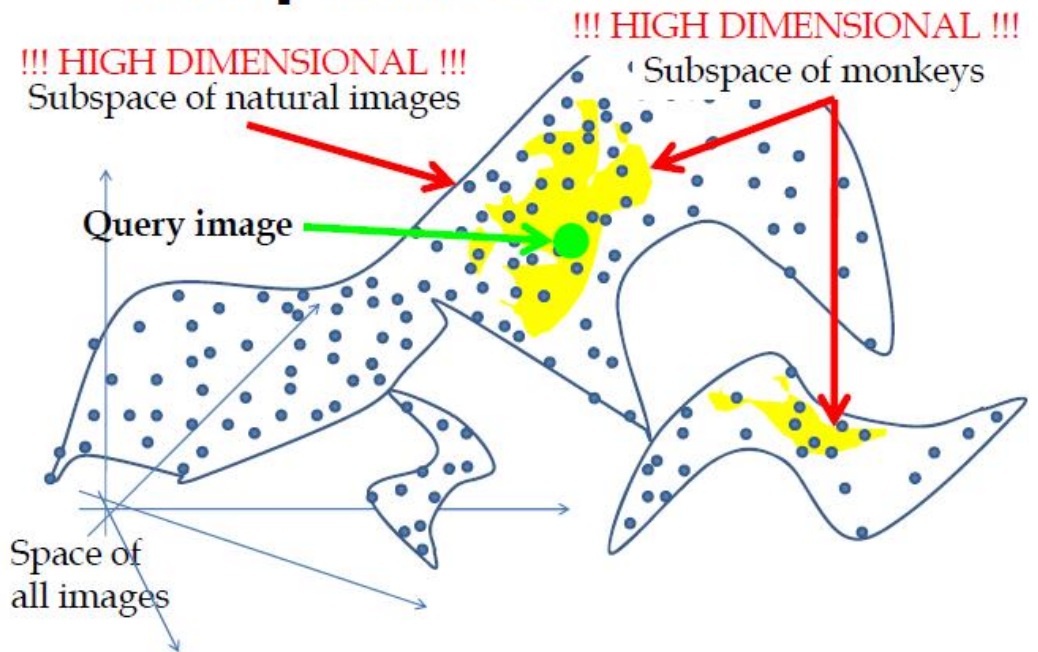
Parametric models



(Slide from Robert Fergus' site)

Instead, try visual matching with A LOT of data

Non-parametric Approach



(Slide of Robert Fergus)

Image matching with local interest points



Matching SIFT points
[Lowe, 1999]

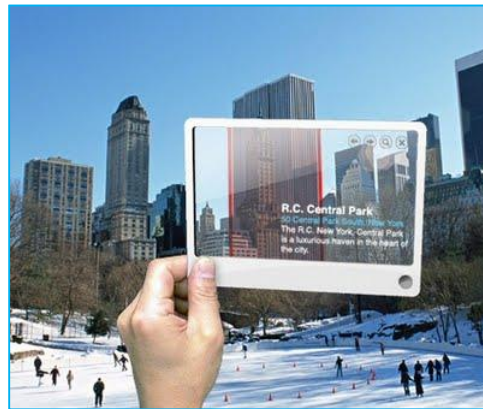
- Measure useful information
 - Image similarity
 - Copy/source identification
 - Discover possible transformation between images
- Can be done efficiently

Visual Matching Spurs Interesting Applications

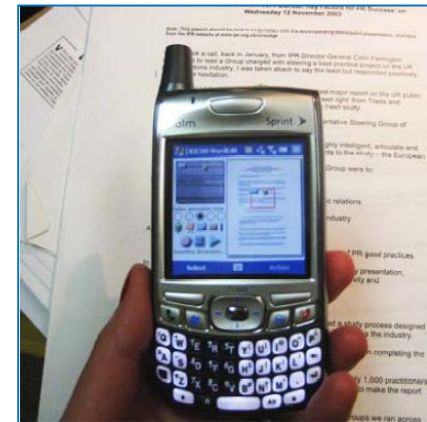
*use image pixels to find additional useful
information*



product search



landmark search



document search

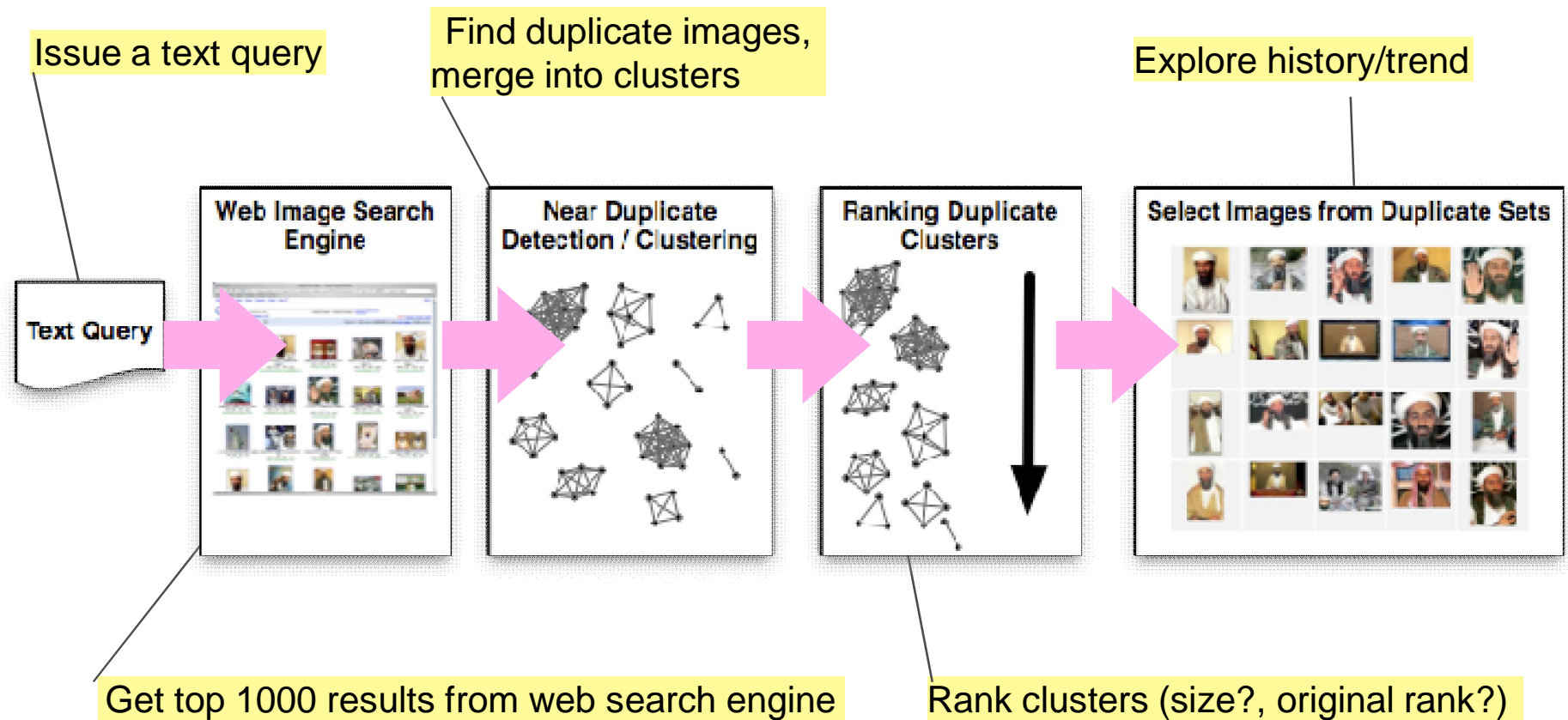
Recent applications:

[iPhone App SnapTel](#)

[Google Goggles](#)

[Demo](#): Flickr Object Retrieval

Image matching for search result summary



Matching Reveals Image Provenance

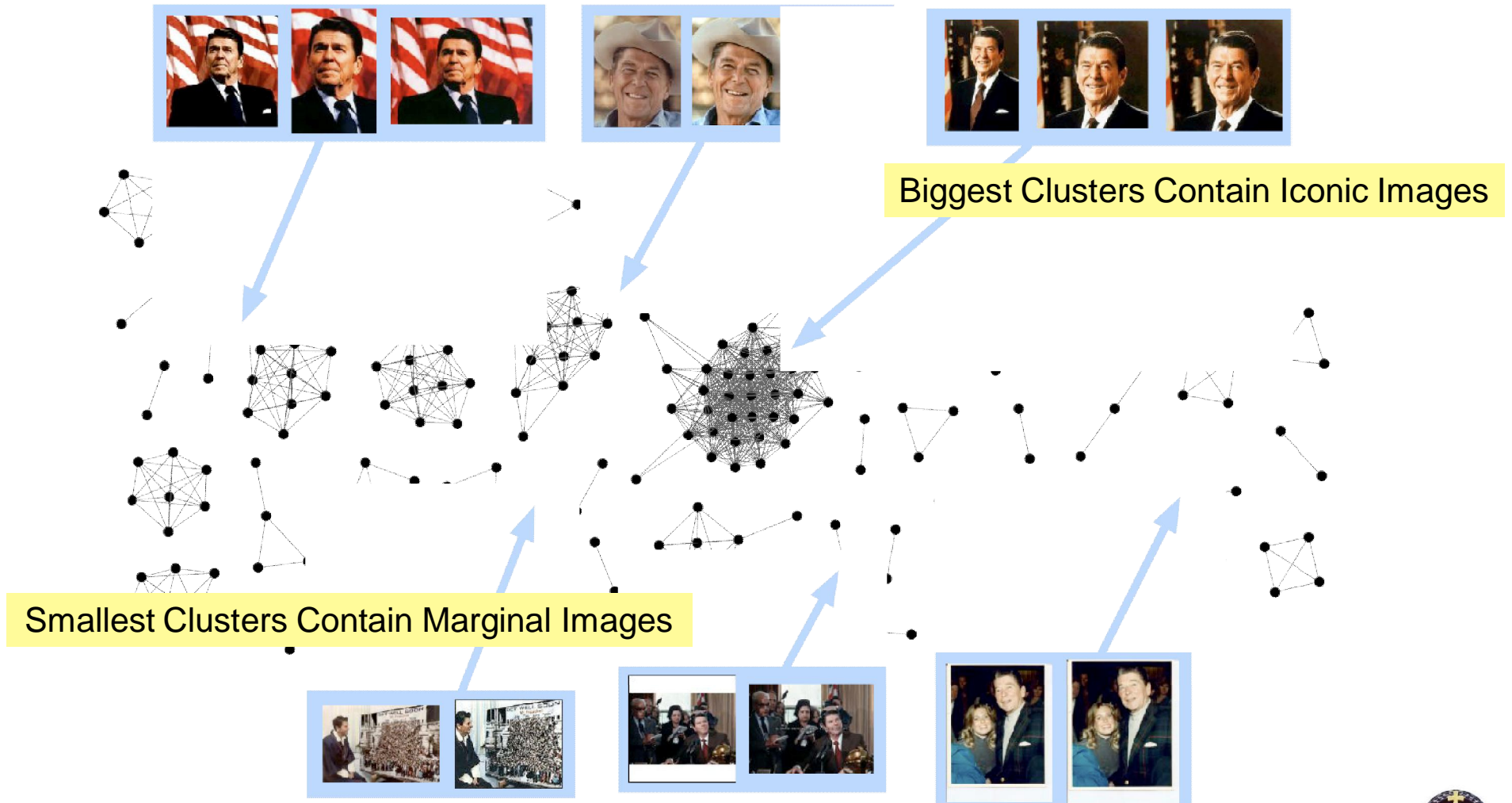


Image Matching Reveals Possible Manipulations



Original



Scaled



Cropped



Gray



Overlay

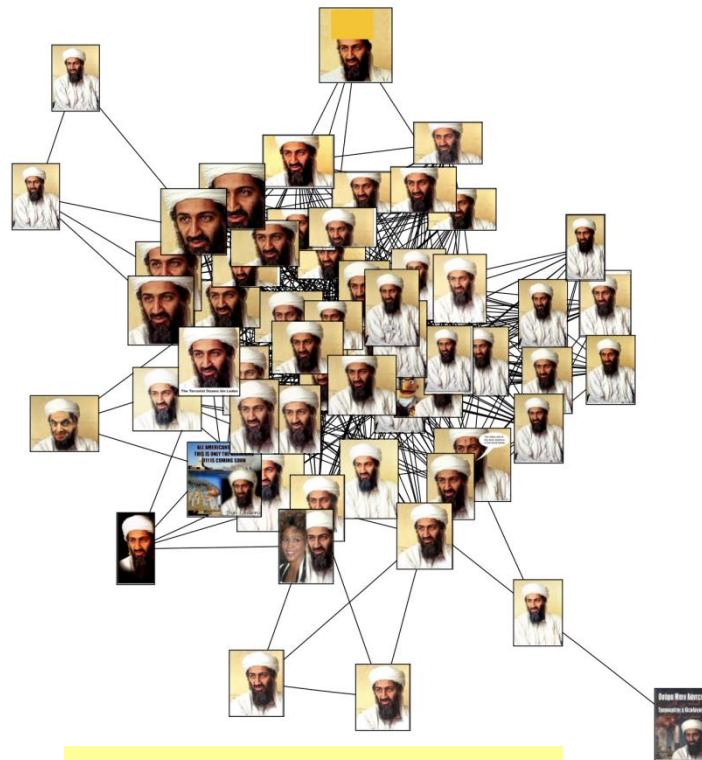


Insertion

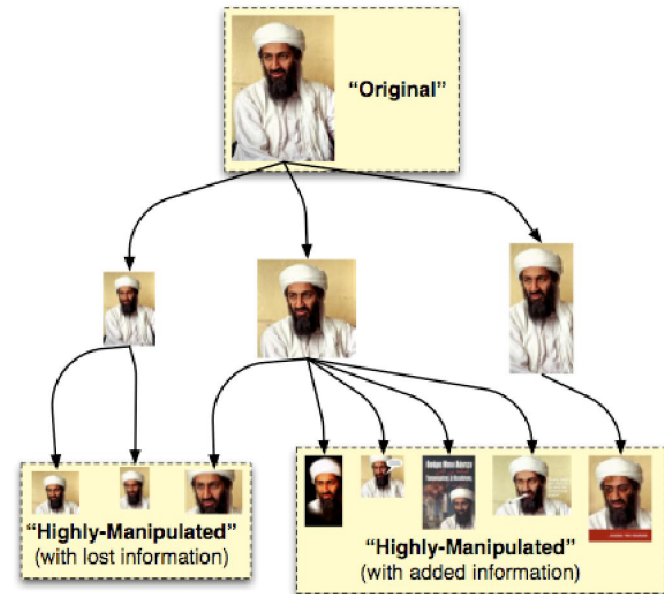
- Given a returned image set, detect possible manipulations
- Each implies editing direction (one image derived from other)
- Other possible manipulations: color correction, multiple compression, sharpening, blurring

Deeper Analysis of Search Results: Visual Migration Map (VMM)

Kenndy and Chang, ACMMM 08

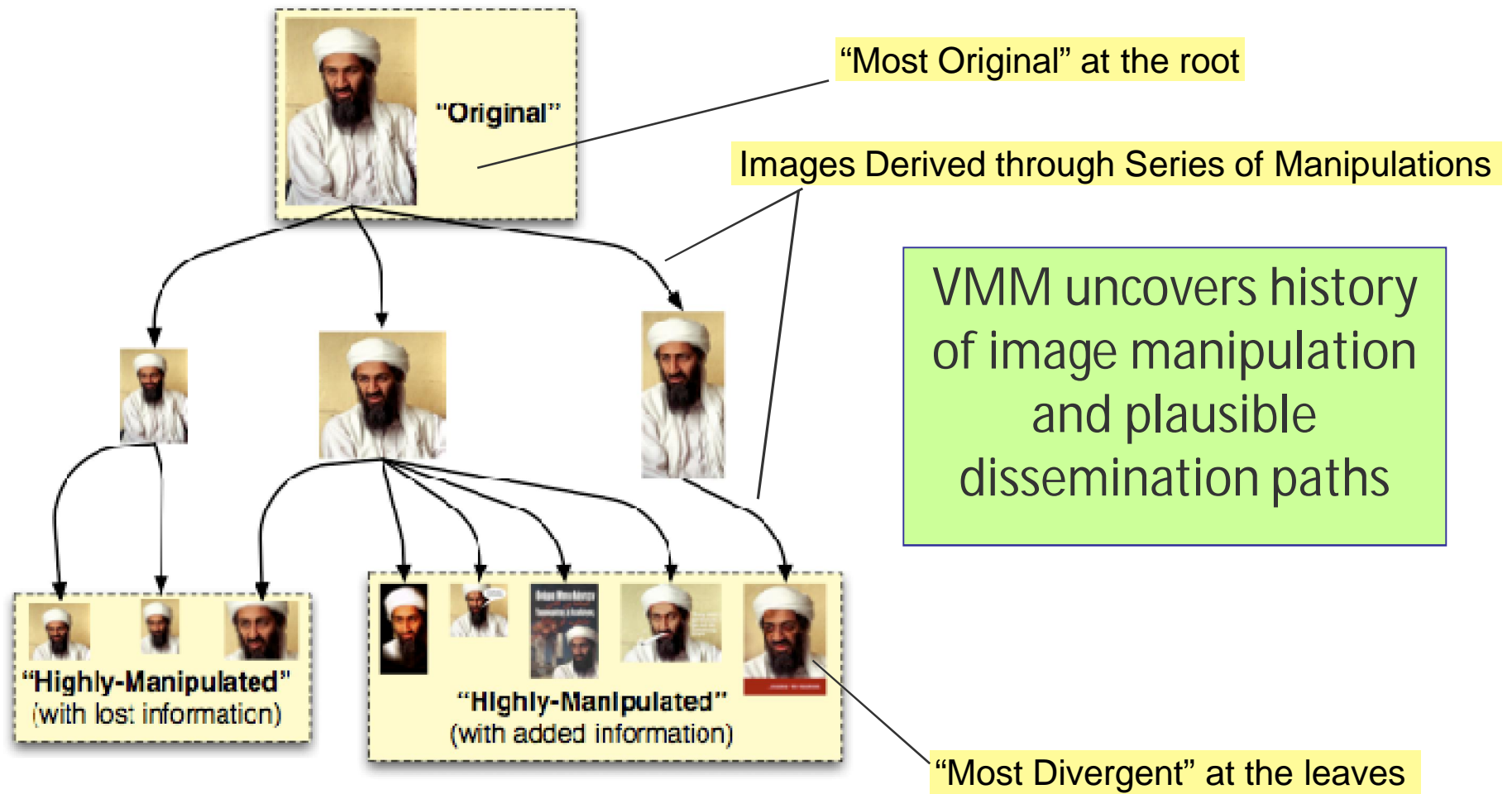


Duplicate Cluster



Visual Migration Map

Visual Migration Map (VMM)



VMM Applications: Geographic/Cultural Dispersion



Manipulations Often Correlate with Perspectives

Joke Website:
 “Every time I get stoned I go and do

Democratic National Committee Site:
 “Capture Osama Bin Laden!”



<http://www.democrats.org/page/petition/osama>

[ml](#)



Daily Excelsior Newspaper:
 “Further Details of Bin Laden Plot
 Unearthed: ABC Report.”

<http://www.dailyexcelsior.com/00jan31/inter.htm>

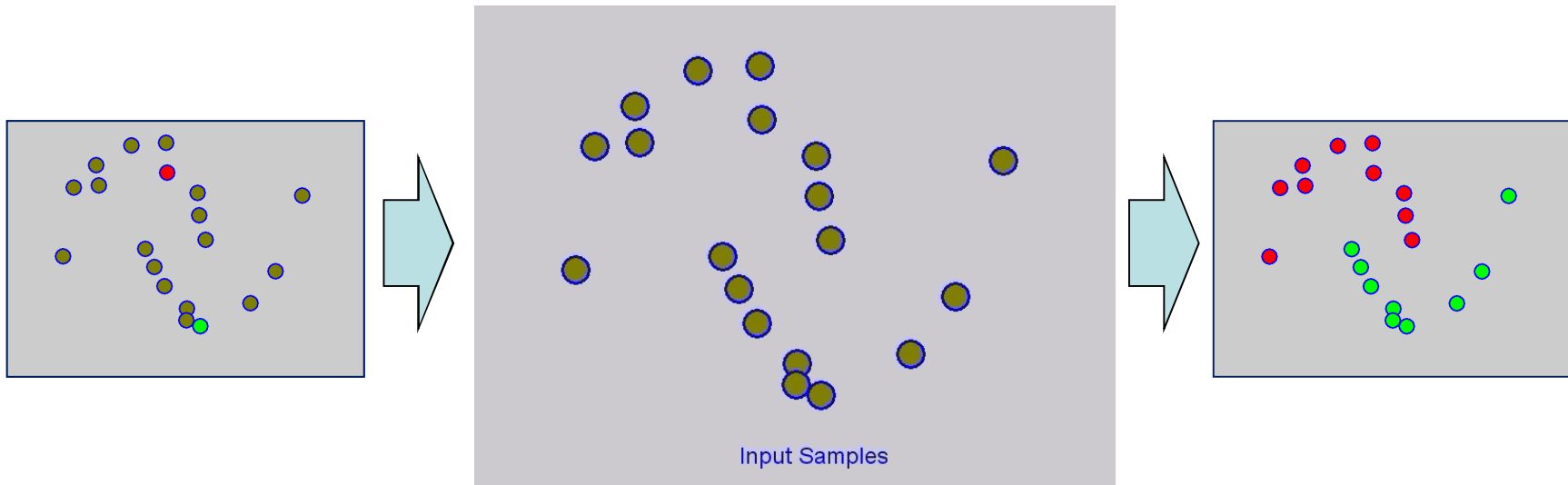


e:
 den - My Idol of All Time!”



http://www.myspace.com/mamu_potnoi


Image Matching + Manifold Graph Allows Useful Propagation of Information



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

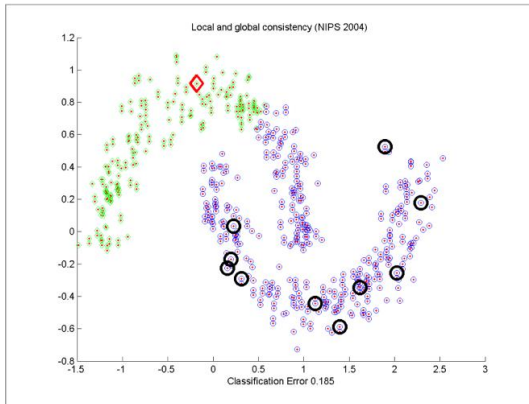
Q -- risk function

f -- classification

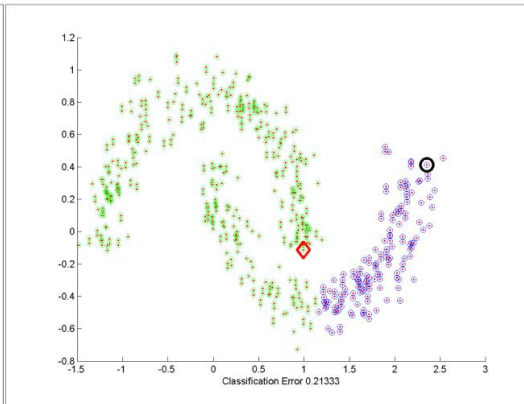
y -- label matrix

Non-Trivial Issues

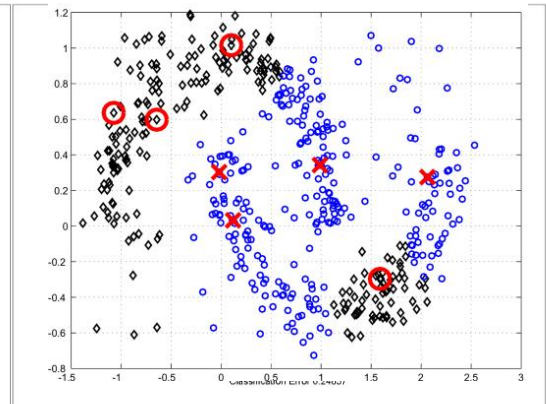
LGC
Method



(a)

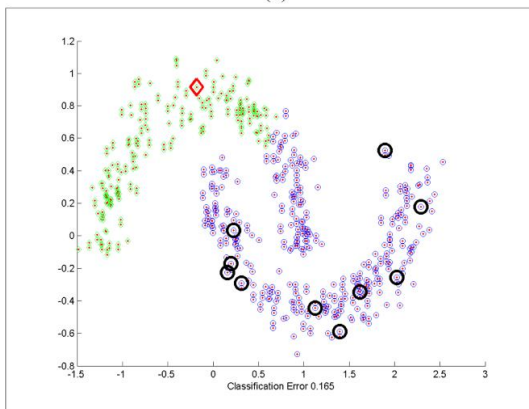


(b)

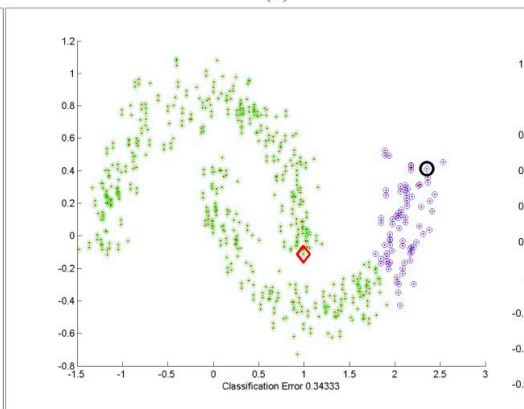


(c)

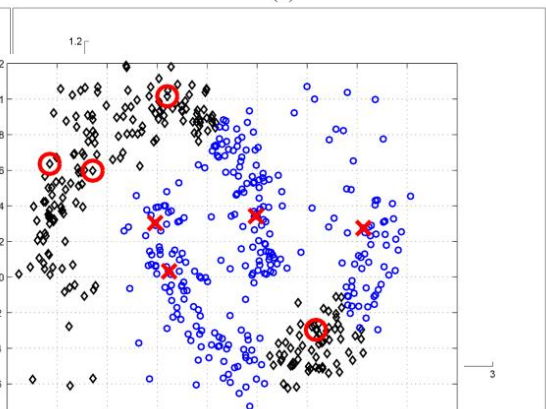
GFHF
Method



(d)



(e)



Unbalanced
Labels

Bad Label
Locations

Noisy
Labels

An active area in Machine Learning

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

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.

Graph Transduction via Alternating Minimization (GTAM)

(Wang, Jebara, Chang, ICML08) (Wang and Chang, CVPR09)

-- Optimization over both Labels (Y) and Prediction (F)

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

- **Propagation Step**

- given label (Y), propagate over graph, predict F

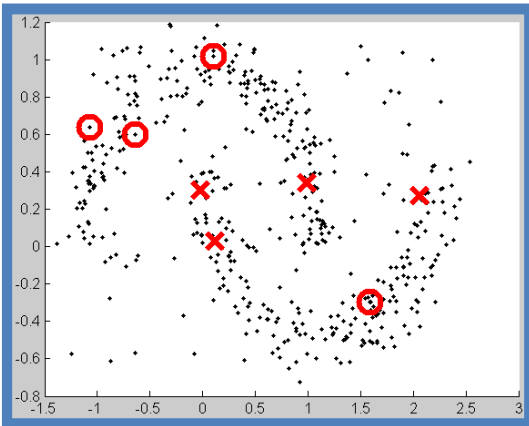
$$\frac{\partial Q}{\partial \mathbf{F}^*} = 0 \Rightarrow \mathbf{F}^* = (\mathbf{L}/\mu + \mathbf{I})^{-1} \mathbf{V} \mathbf{Y} = \mathbf{P} \mathbf{V} \mathbf{Y}$$

- **Label Diagnosis and Selection Step**

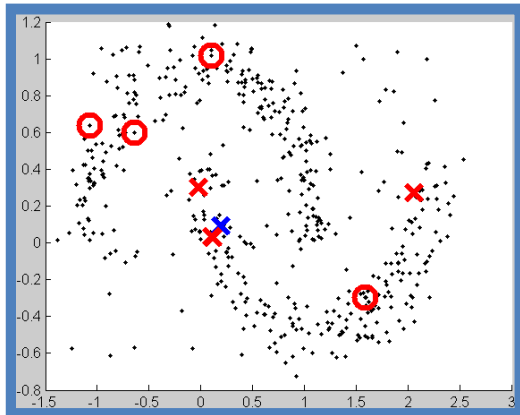
- add good labels or remove bad labels

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

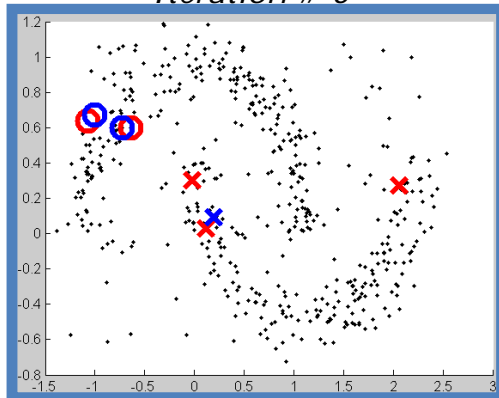
Initial Labels



Iteration # 2

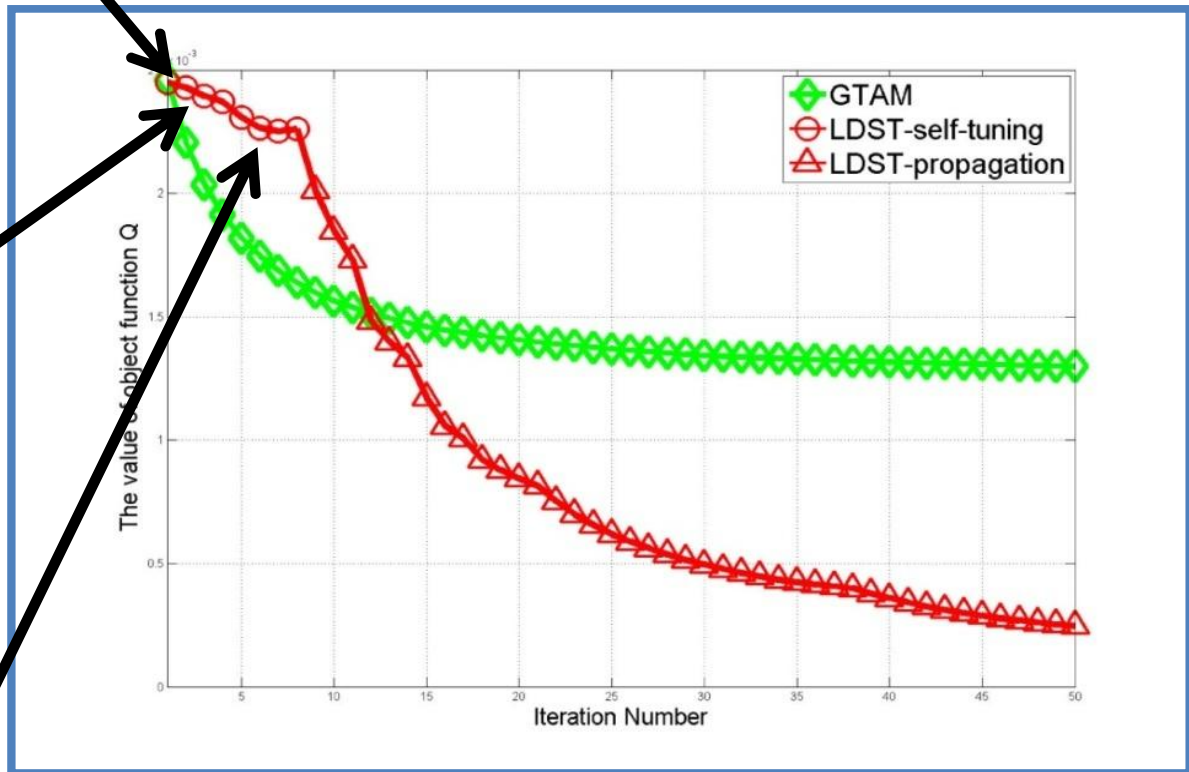


Iteration # 6



Iteratively Tune and Propagate Label Information

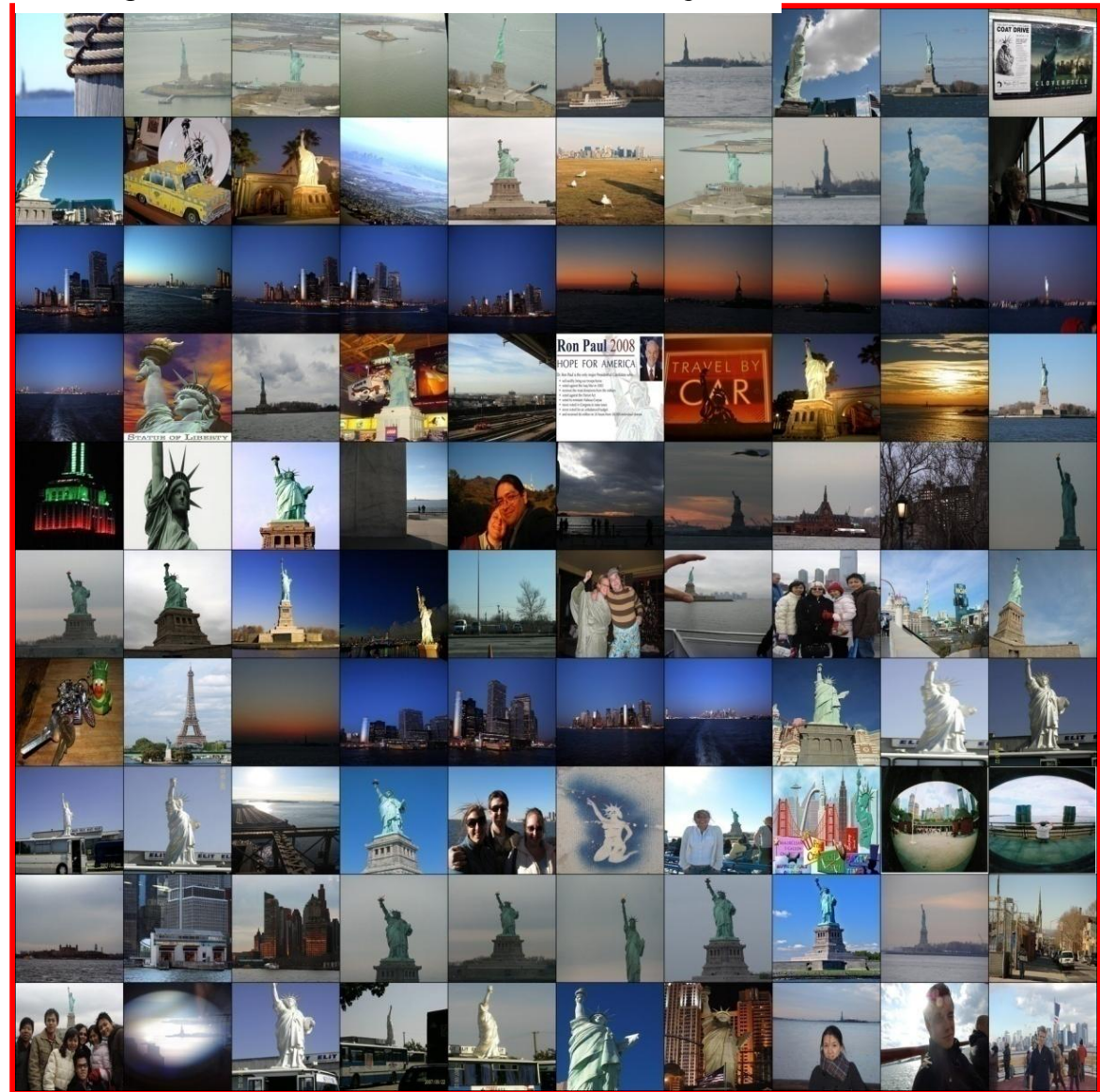
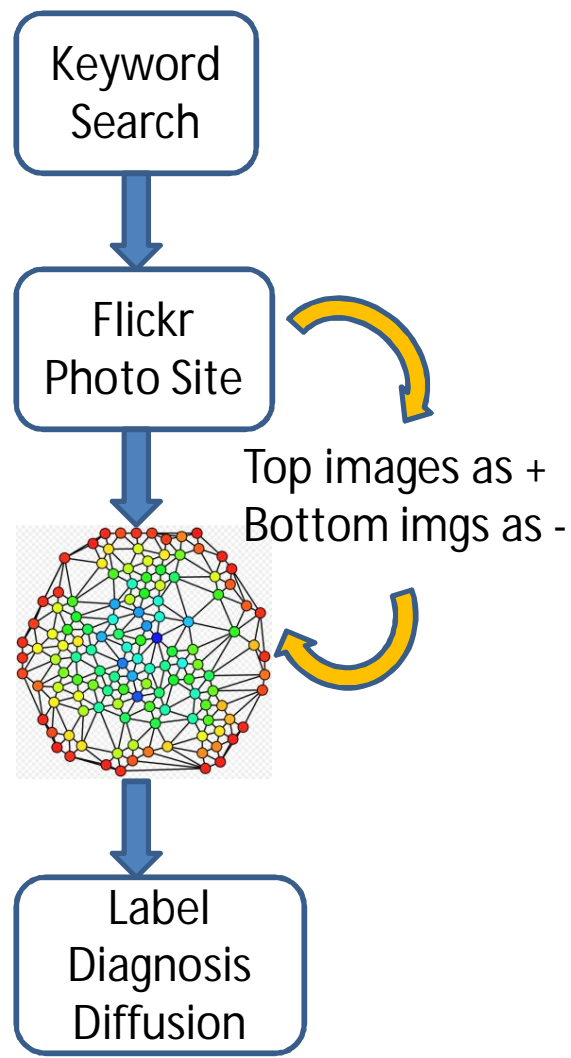
(Wang and Chang CVPR'09)



The values of the cost function Q during optimization procedure of LDST and GTAM methods.

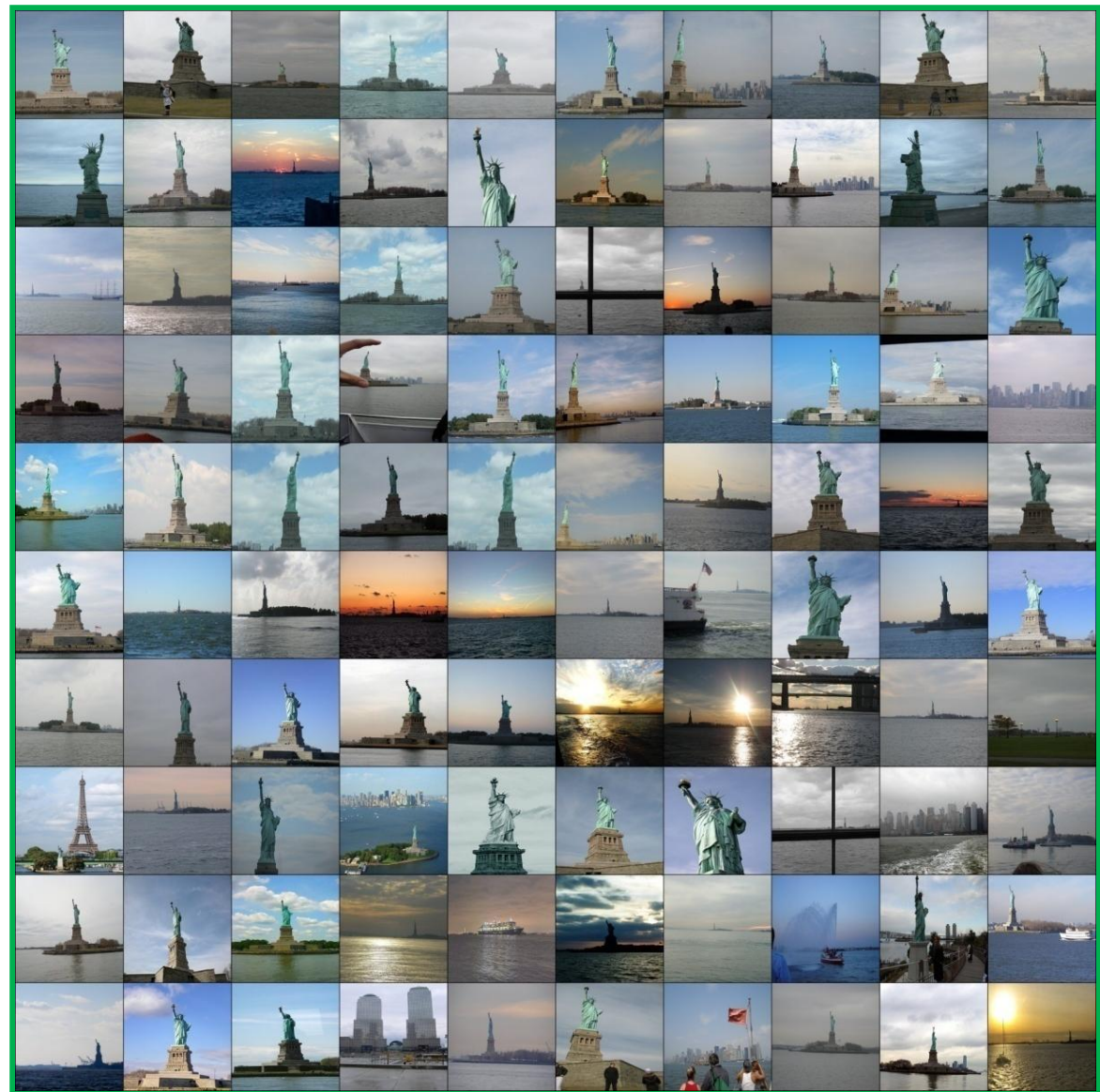
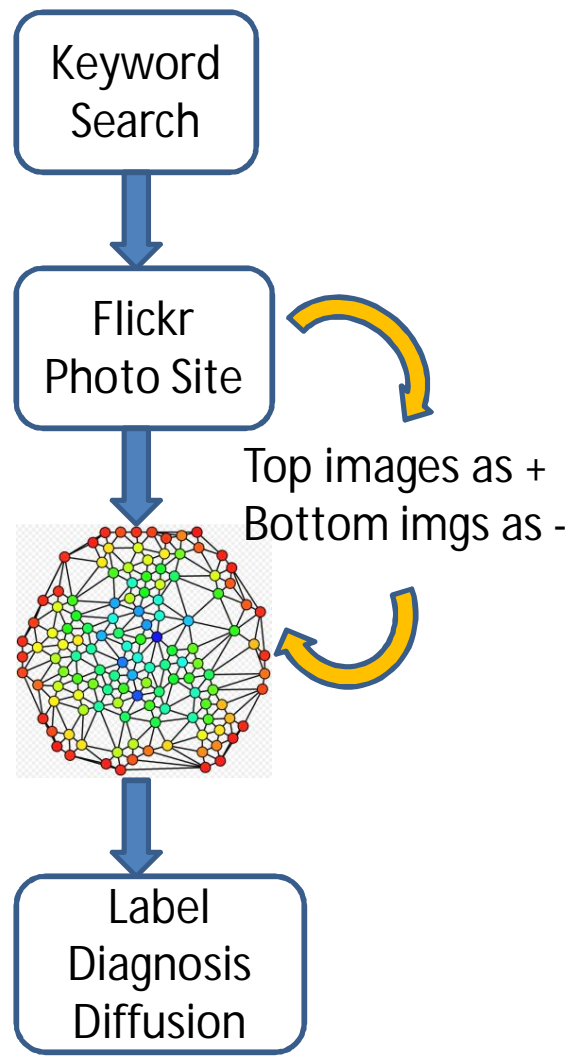
Tune and propagate top scores through image graph

Google Search "Statue of Liberty"



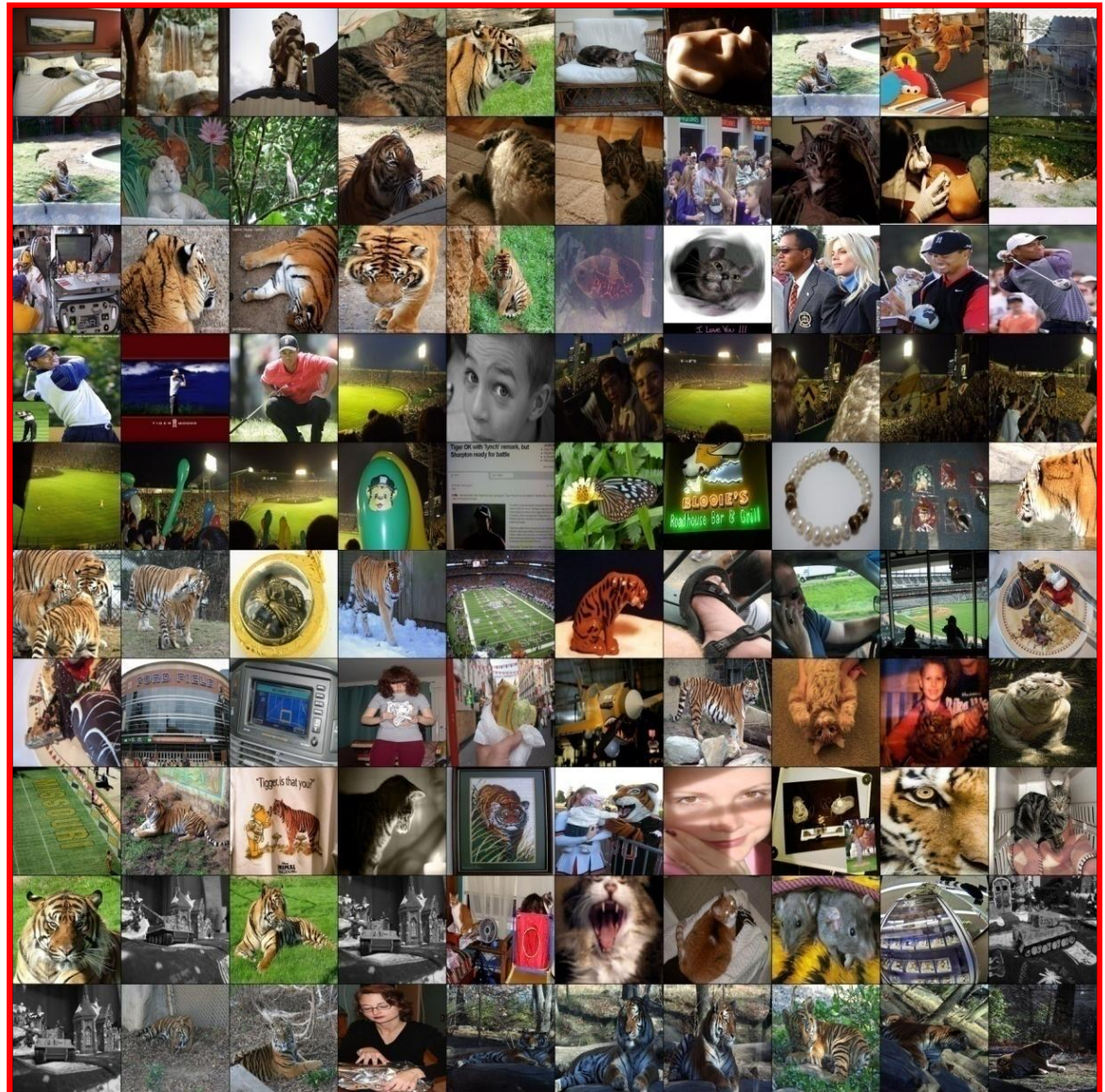
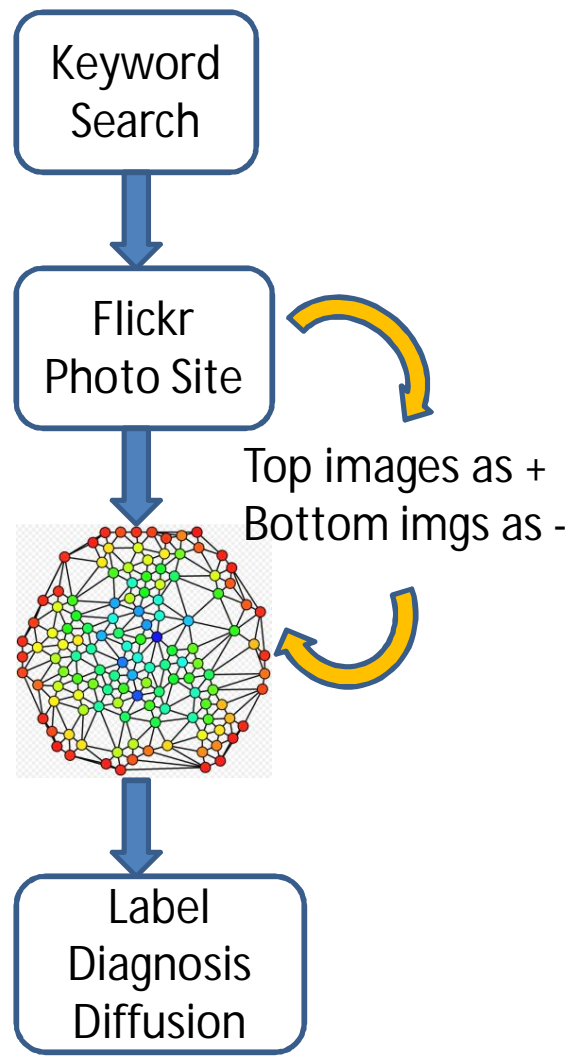
Tune and propagate top scores through image graph

Rerank

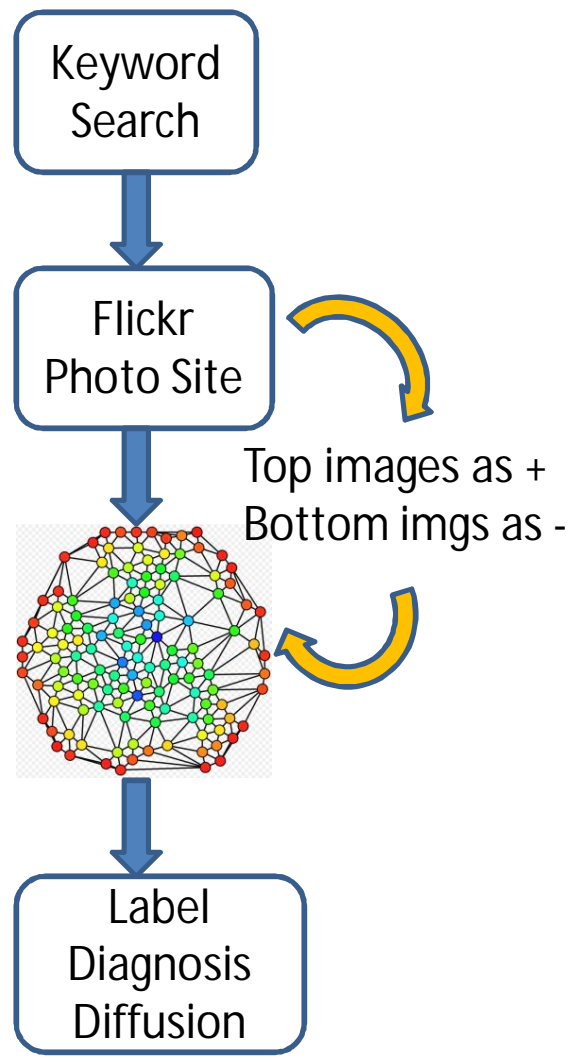


Tune and propagate top scores through image graph

Google Search "Tiger"



Tune and propagate top scores through image graph

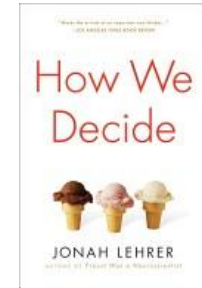
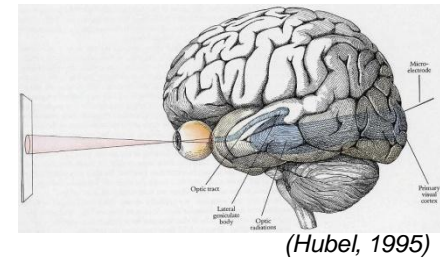


Rerank



Detect Image Semantics via Brain State Decoding

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



Joint work with
Prof. Paul Sajda's group

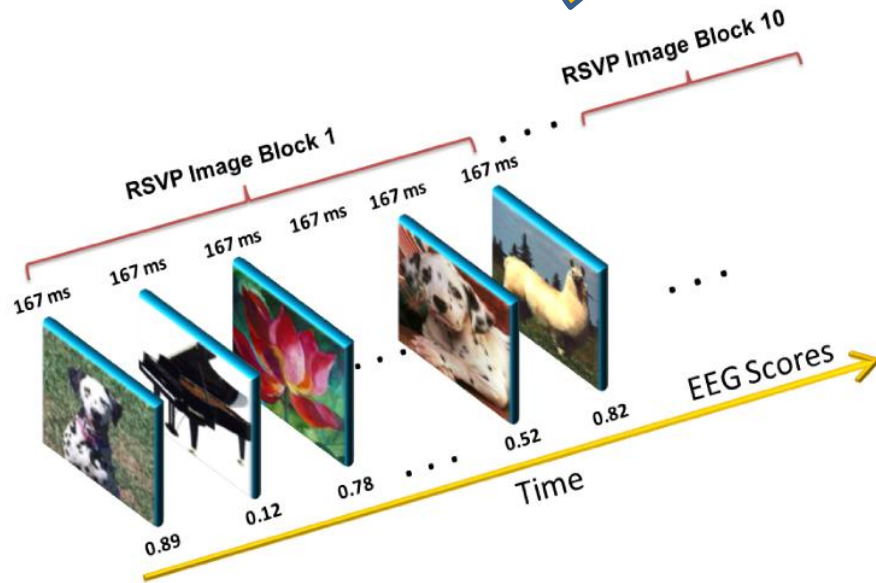


Brain Machine Interface for Image Retrieval

Use EEG brain signals
to detect target of interest
([video](#))

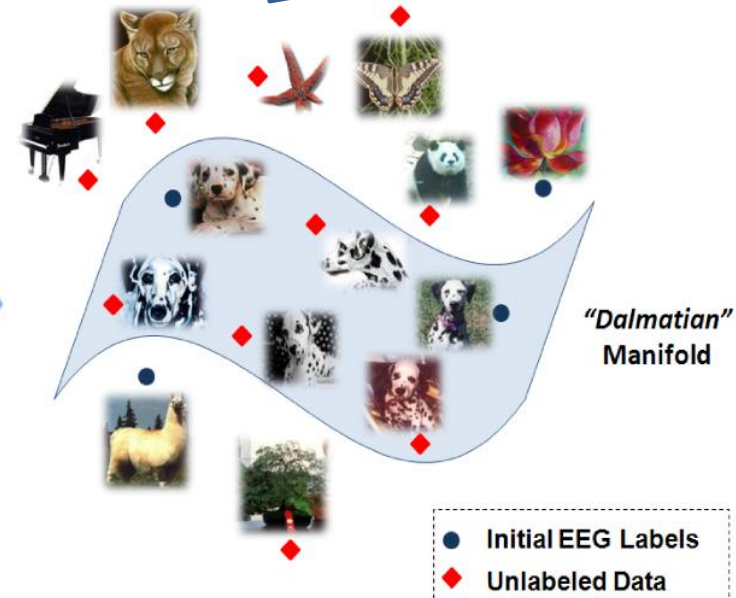


Use image graph to tune
& propagate information



Rapid Serial Presentation of Caltech 101 Images

*C3Vision System
by Sajda et al*

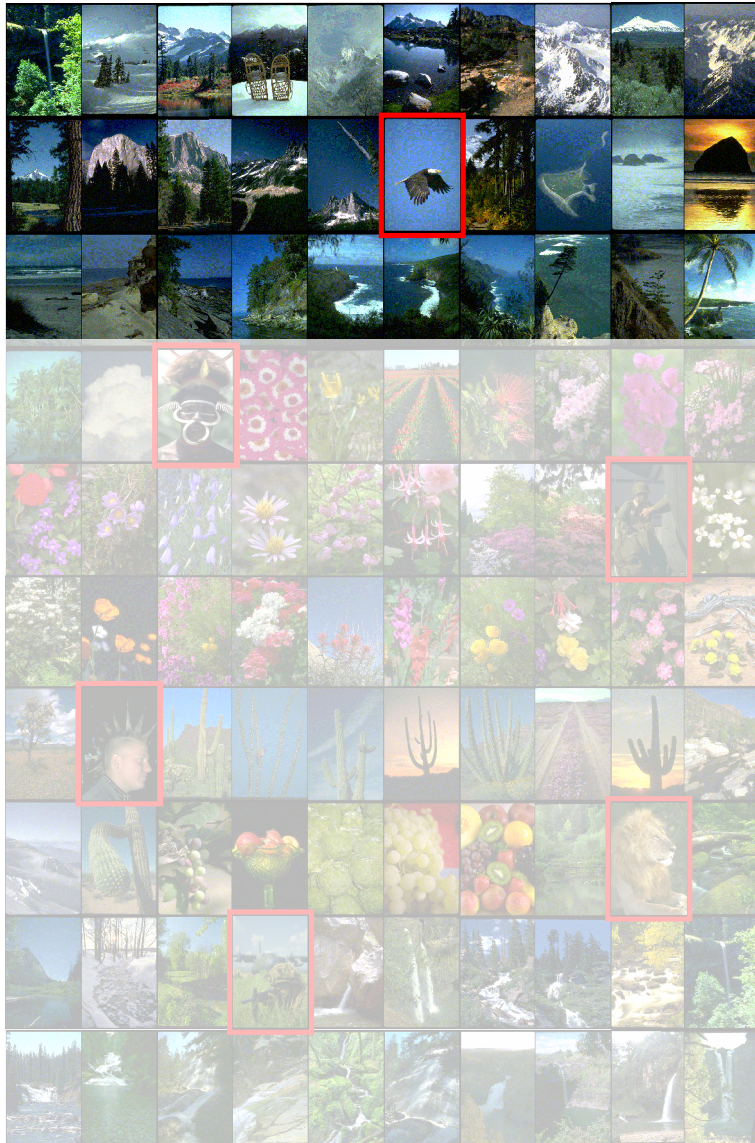


Graph-Based Visual Pattern Discovery

*TAG System
by Wang and Chang*

The Visual Interest Readout Experiment

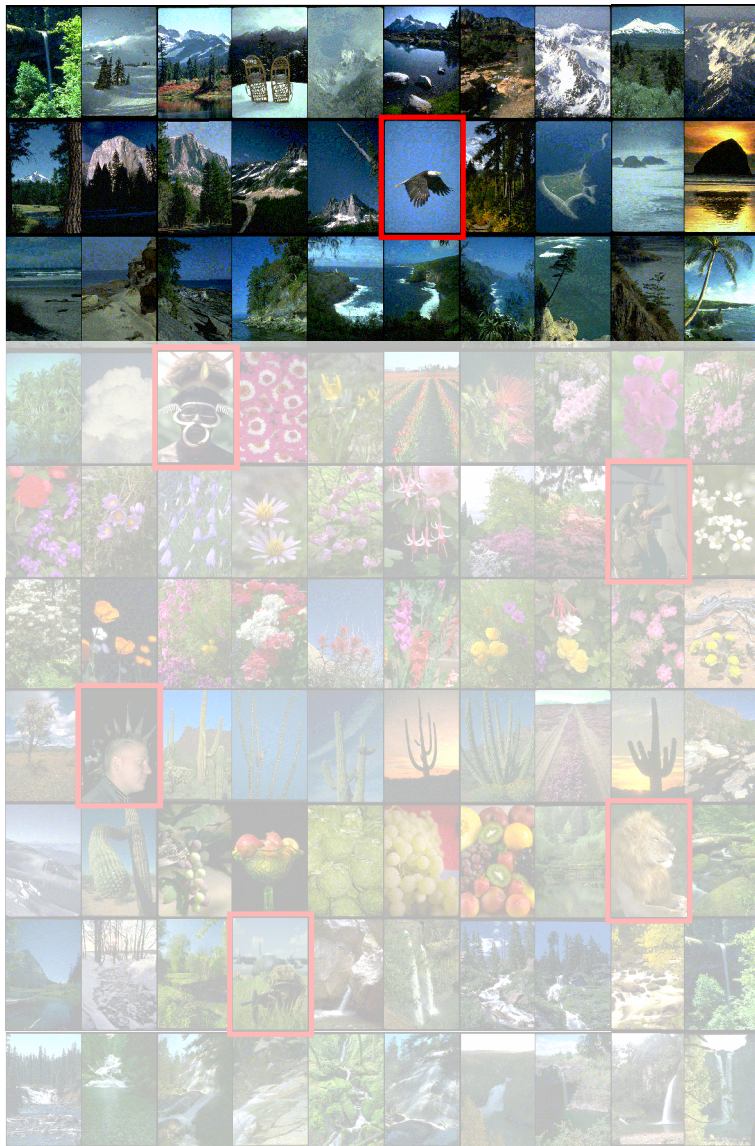
User thinks about what he/she wants to search



Database (any target that may interest users)



The Paradigm



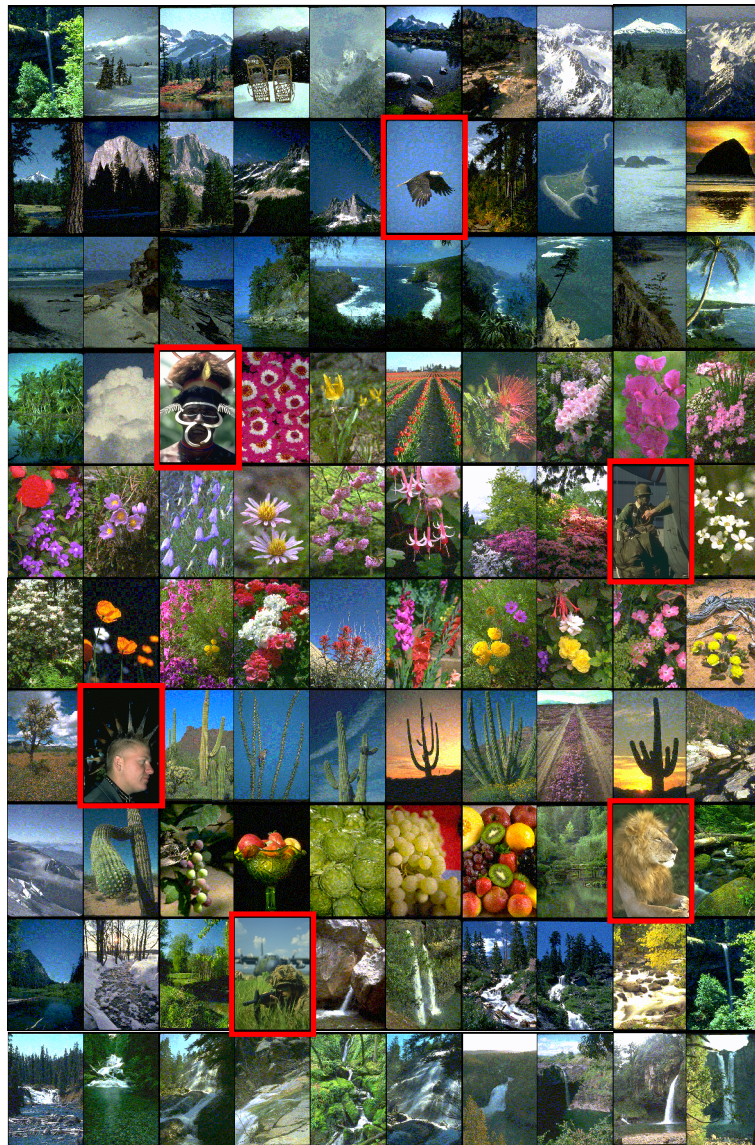
Database



Neural (EEG) decoder

Interest-scores

The Paradigm



Database



Neural (EEG) decoder

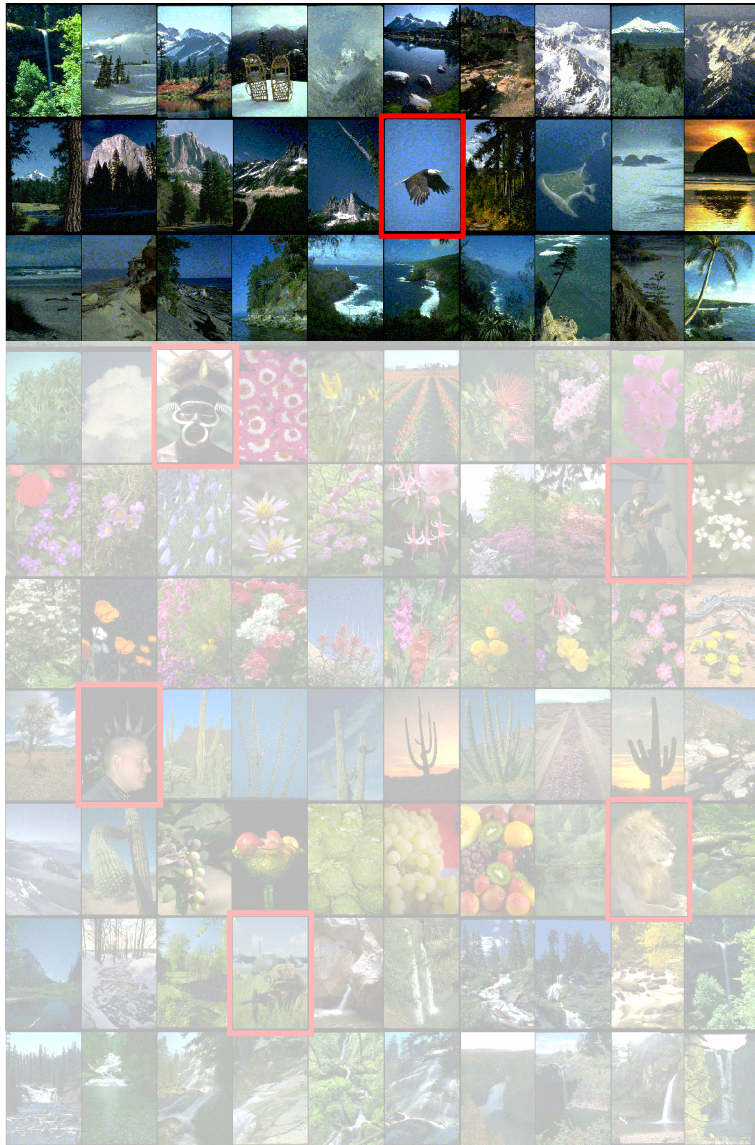
Exemplar labels (noisy)

Semi-supervised
Graph-based propagation

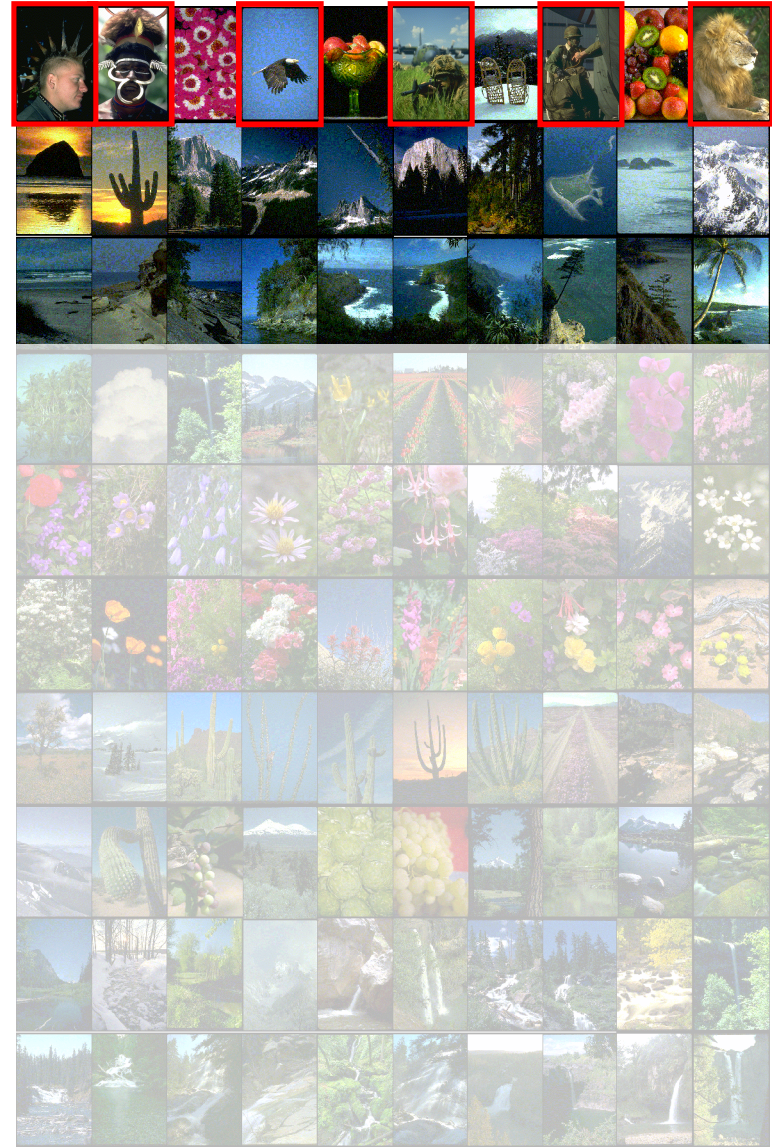
*Features from
the entire DB*

prediction score

The Paradigm

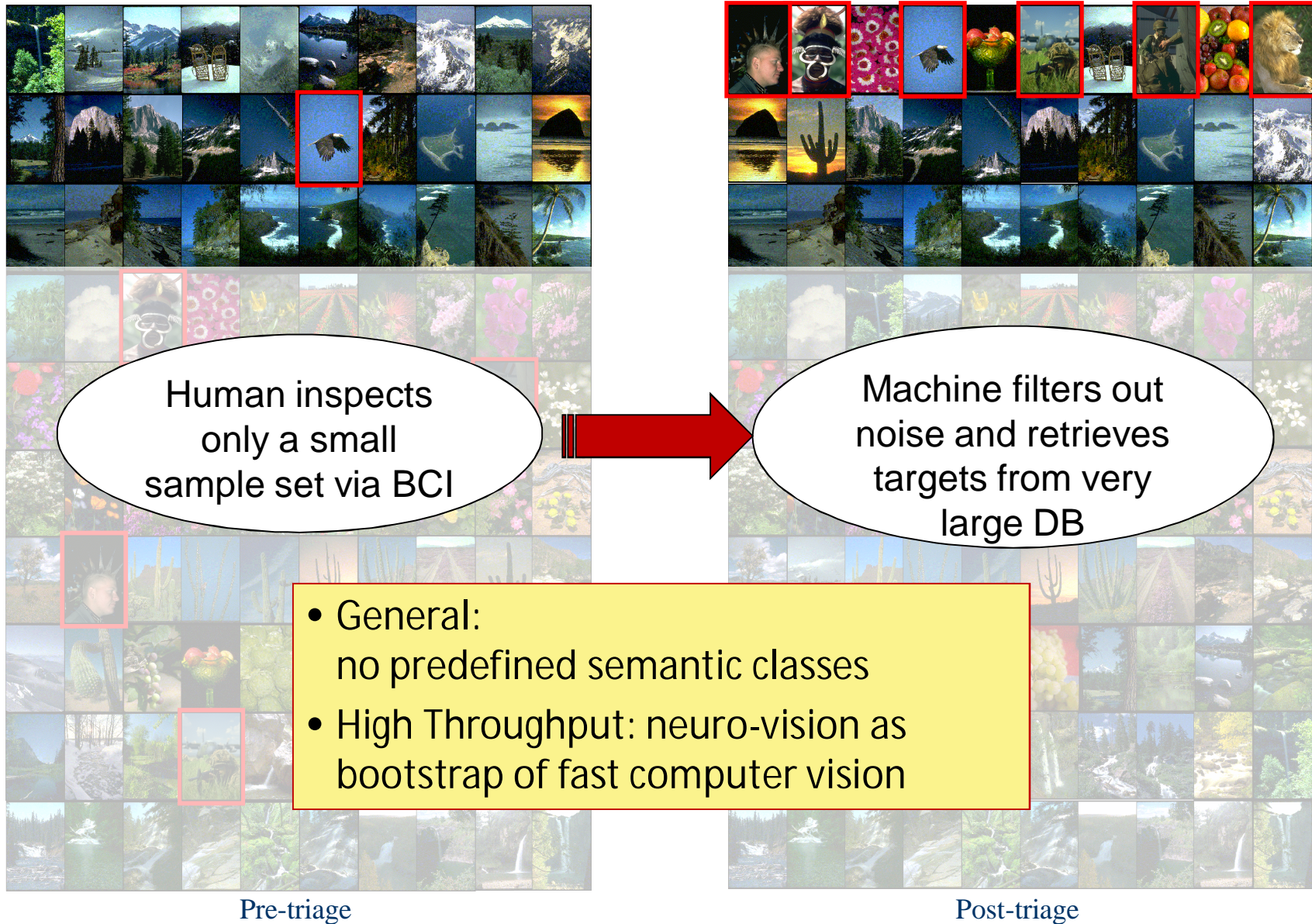


Pre-triage

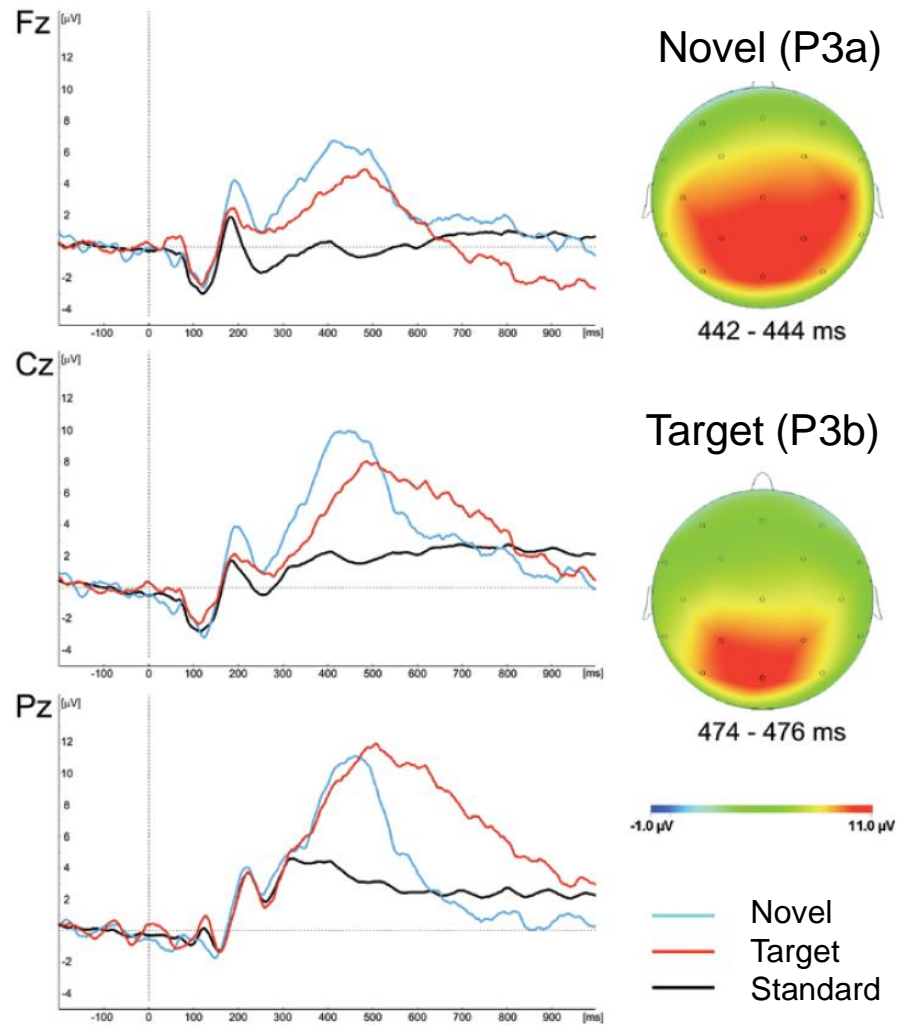
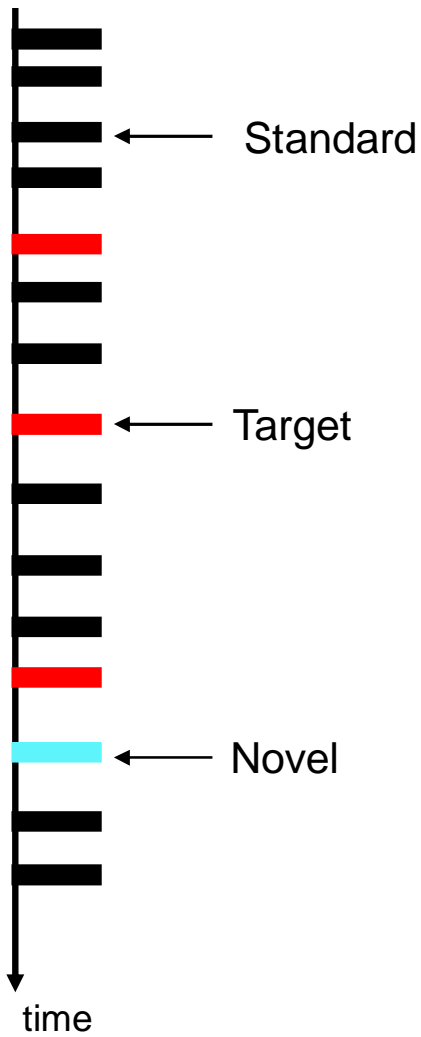


Post-triage

The Paradigm



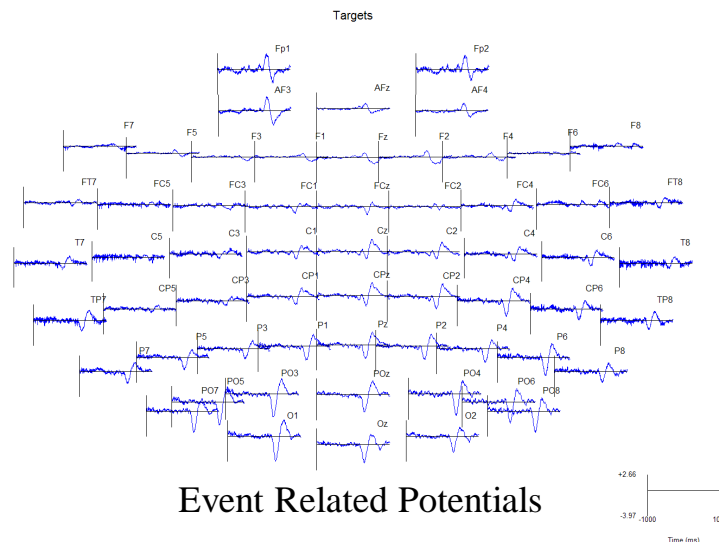
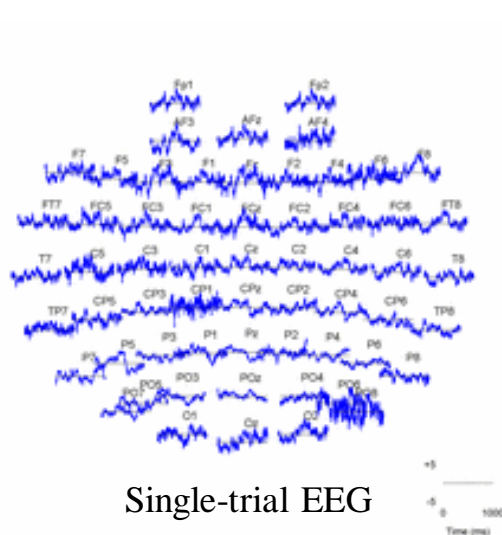
The Neural Signatures of “Recognition”



From D. Linden, 2005

Single-trial EEG Analysis

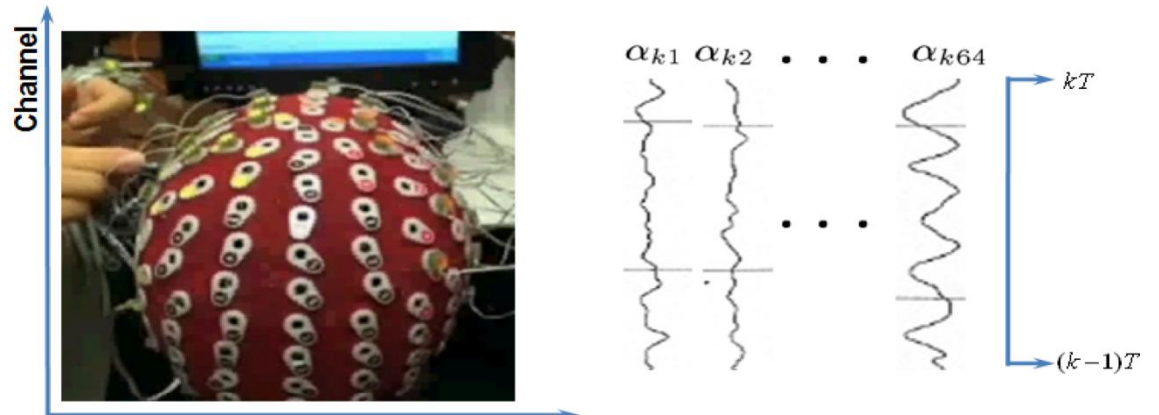
- Typically EEG is averaged over trials to increase the amplitude of the signal correlated with cortical processes relative to artifacts (very low SNR)
- High-density EEG systems were designed without a principled approach to handling the volume of information provided by simultaneously sampling from 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



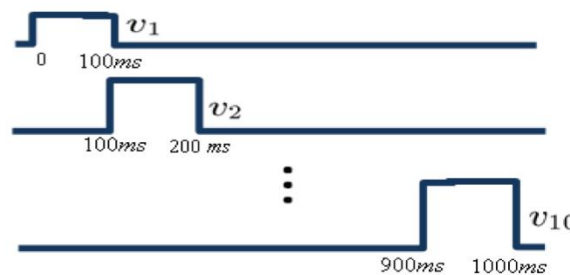
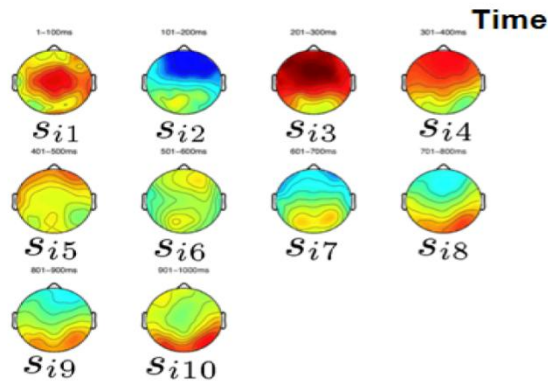
Identifying Discriminative Components in the EEG Using Single-Trial Analysis

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)



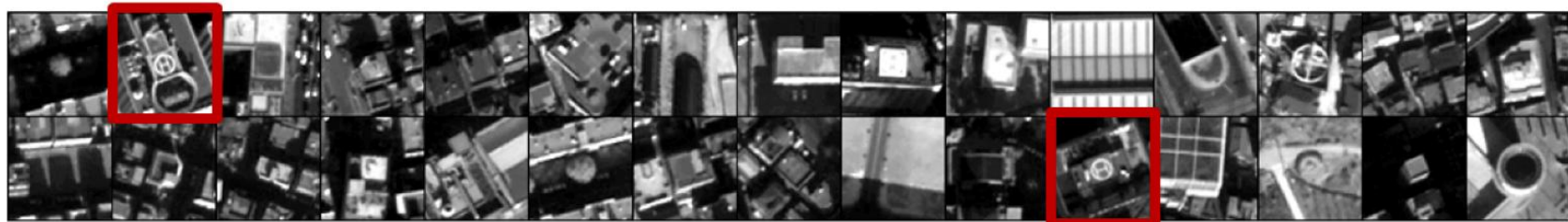
Optimal spatial filtering across electrodes within each short window (e.g., 100ms)



Optimal temporal filtering over time windows after onset

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*)



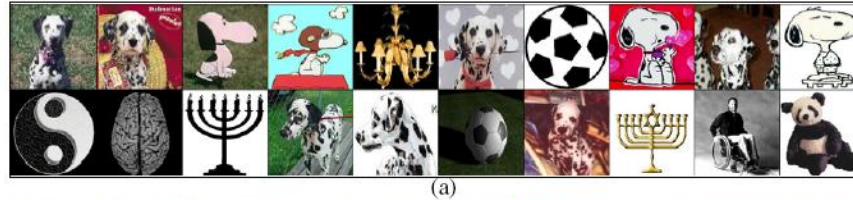
(a)



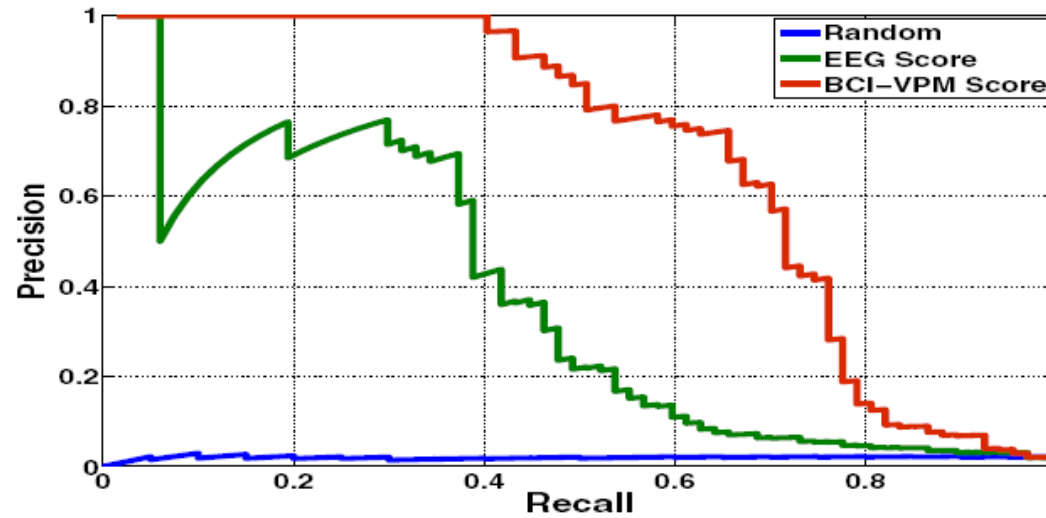
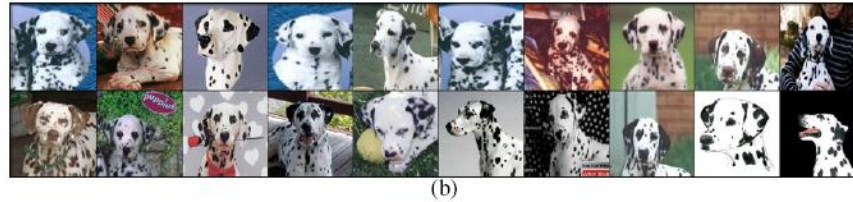
(b)

Example results

Top 20 results of
Neural EEG detection



Top 20 results of
Hybrid System (BCI-VPM)



Application in target searching in satellite images:

Initial EEG neural signal detection:



(a)



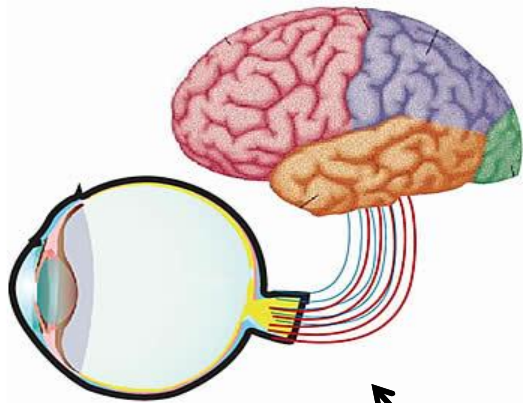
(b)

After graph refinement and diffusion

Images from DigiGlobe)

Summary: Cross Fertilization of Several Fields

Human Vision



Internet/Data Vision

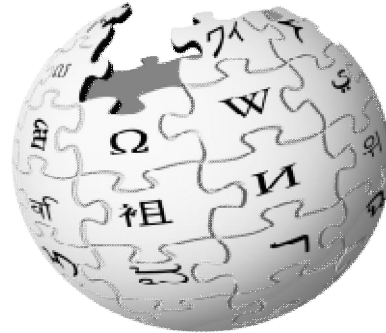


Image Recognition/
Machine Learning

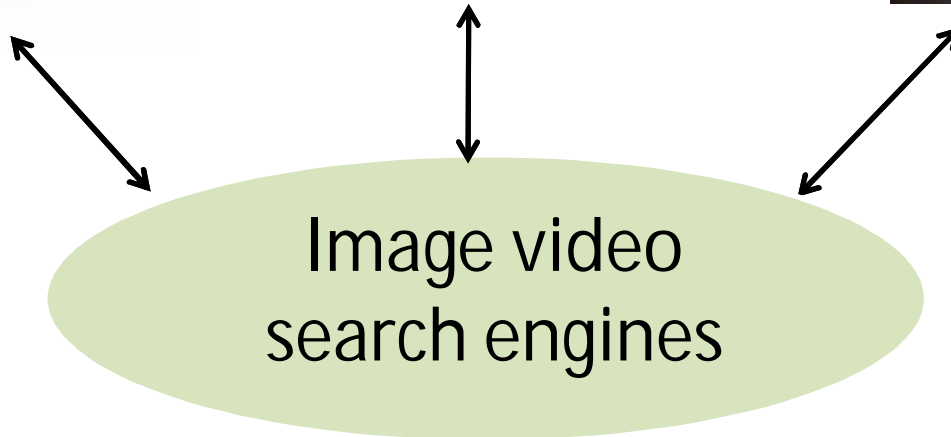
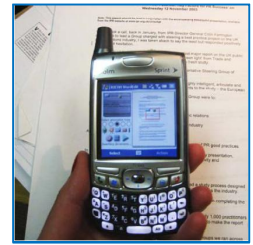


Image video
search engines

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
 - Paul Sajda, Barbara Hanna, Eric Pohlmeier
- 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.