EE 6882
Visual Search Engine

March 5th, 2012
Lecture #7

- Relevance Feedback
- Graph-Based Semi-Supervised Learning
- Application of Image Matching: Manipulation Detection

What We have Learned

- Image Representation and Retrieval Using Global Features
- Local Features and Image Matching
- Image Classification

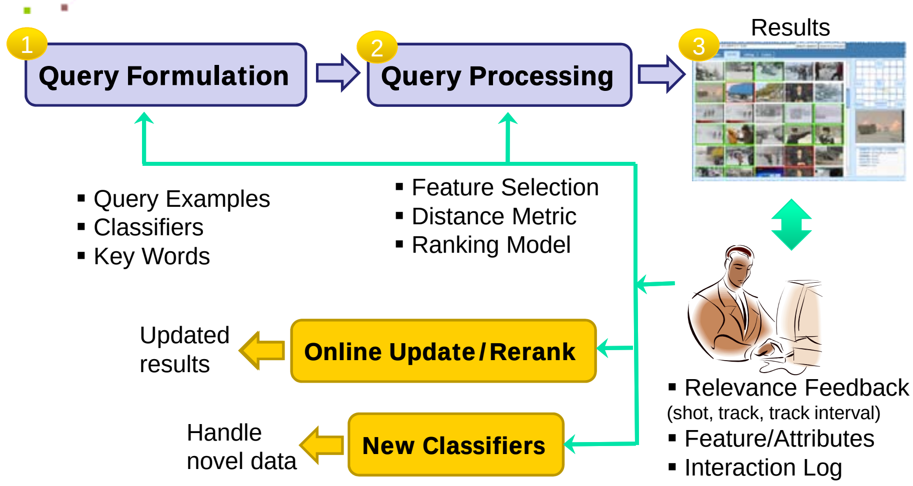
2

What more can be done?

- Image Representation and Retrieval
  - User in the Loop: Relevance Feedback
- Local Features and Image Matching
  - New Features
  - Different Ways of Quantization, Codebook Learning
  - Application: Duplicate Detection
- Image Classification
  - Machine Learning Techniques
  - Semi-Supervised Learning
  - Multi-Modal Fusion
- Others:
  - User interfaces

3

When User in the Loop: Interactive Query Refinement

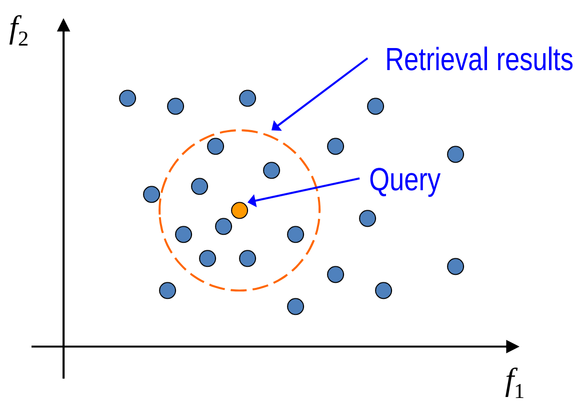


Example:
Columbia TAG Interactive Image Search System

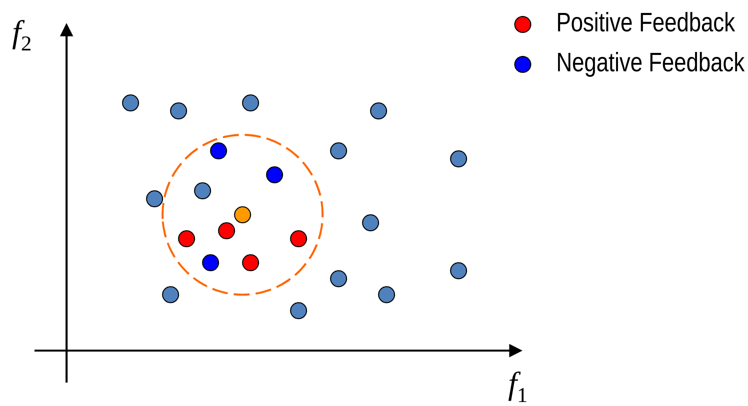
- Demo:
  Rapid Image Annotation with
  User Interaction

A Very Simple Case: Query Update

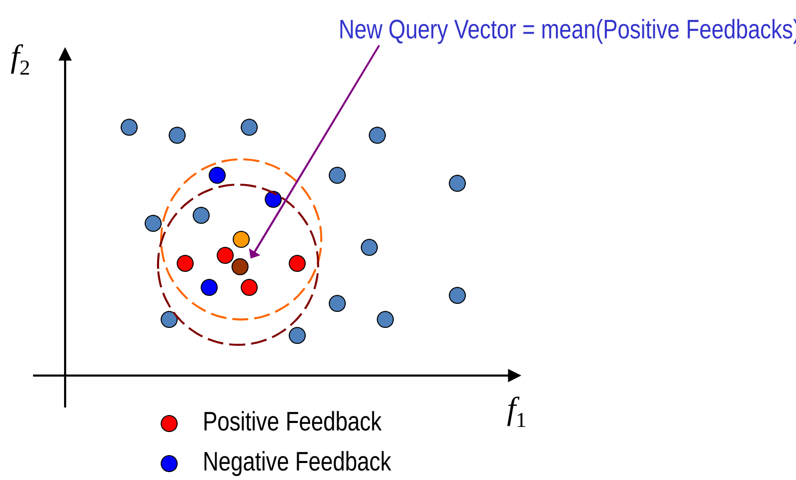
- Automatically update Query Point based on user feedback



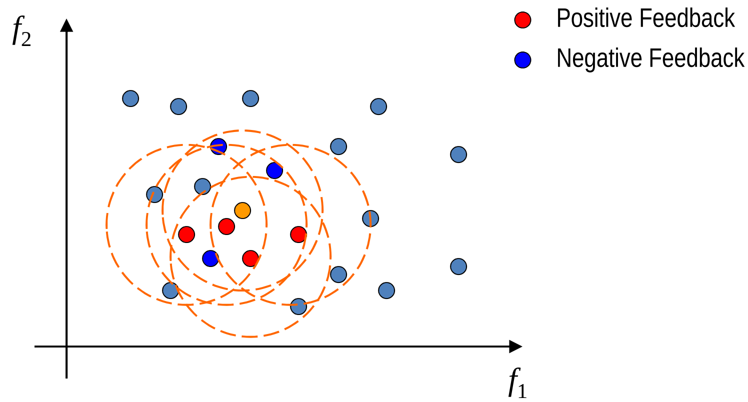
User Provides Feedback



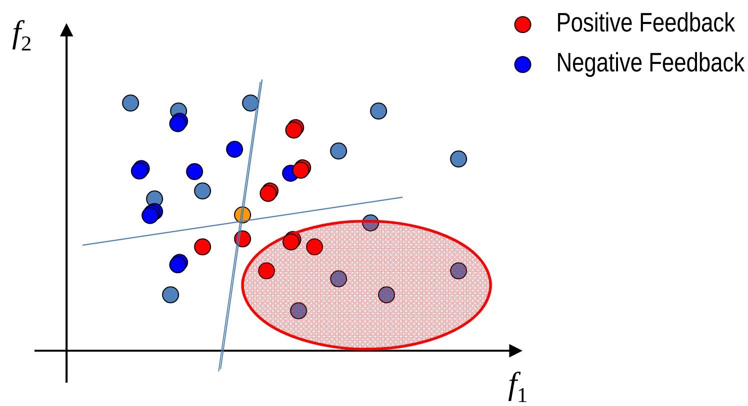
Query Update



Query Expansion

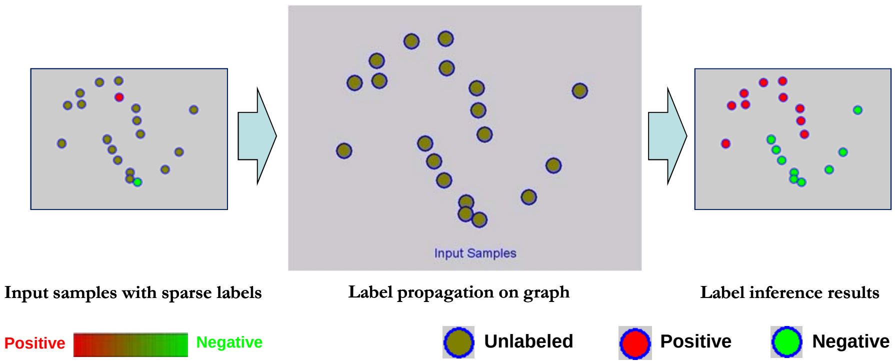


Build Multiple Classifiers



Graph-based Semi-Supervised Learning

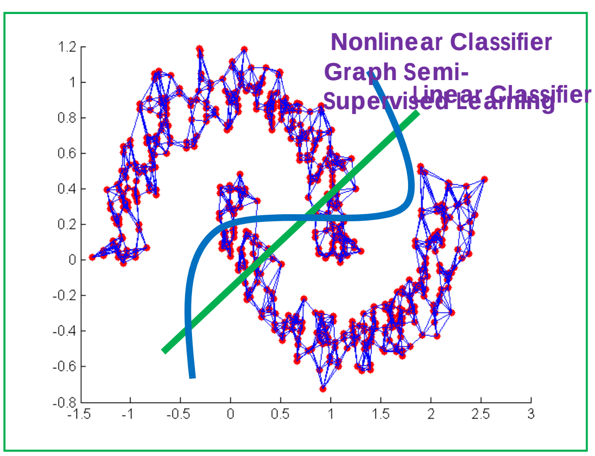
- Given a small set of labeled data and a large number of unlabeled data in a high-dimensional feature space
  - Build sparse graphs with **local** connectivity
  - Propagate information over graphs of large data sets
  - Hopefully robust to noise and scalable to gigantic sets



11

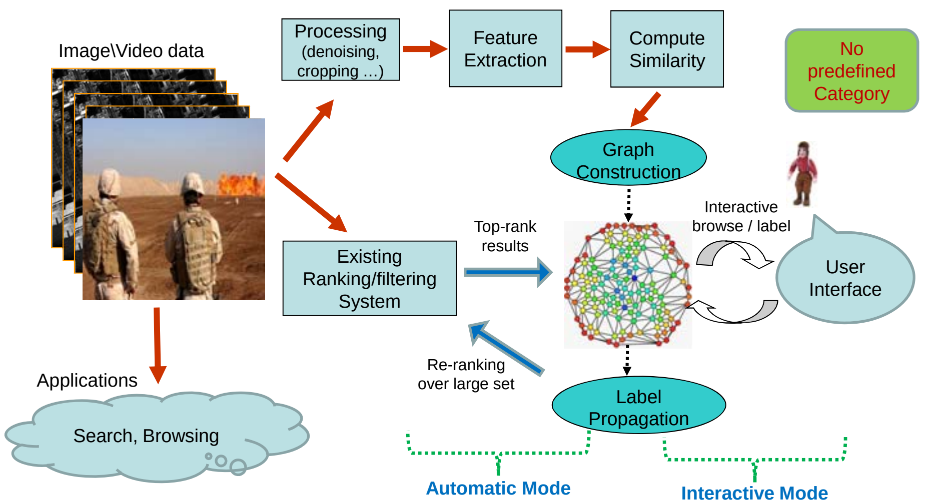
Intuition

- Capture local structures via sparse graph



Through Spare Graph Construction (e.g., kNN)

Possible Applications: Propagating Labels in Interactive Search & Auto Re-ranking



S.-F. Chang, Columbia U.

13

Background Review

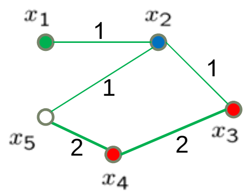
- Given a dataset $\mathcal{X} = (\mathcal{X}_l, \mathcal{X}_u)$ of labeled samples $\mathcal{X}_l$, and unlabeled samples $\mathcal{X}_u$
- *undirected* graph $\mathcal{G} = \{\mathcal{X}, \mathcal{E}\}$of samples $\mathcal{X}$ as vertices and edges $\mathcal{E}$ weighted by sample similarity

$$w_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$$

- Define weight matrix $\mathbf{W} = \{w_{ij}\}$; vertex degree $\mathbf{D} = diag\left([d_1, \cdots, d_n]\right)$

$$d_i = \sum_j w_{ij}$$

Example



Weight matrix

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & 2 & 0 \end{bmatrix}$$

Node degree

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

classes

samples

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ ? & ? & ? \end{bmatrix}$$

Label matrix

$$\Rightarrow \quad \mathcal{G} = \{\mathcal{X}, \mathcal{E}\} \quad \{\mathbf{W}, \mathbf{D}, \mathbf{Y}\} \quad \Rightarrow$$

Graph-based SSL

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0.1 & 0.2 & 0.9 \end{bmatrix}$$

Label prediction

Some Options of Constructing Sparse Graph

- Distance Threshold
- K-Nearest Neighbor Graph

$$\max_{\hat{P}} \sum_{ij} \hat{P}_{ij} W_{ij} \qquad \hat{P}_{ij} = 1 \; if \; x_i \; and \; x_j \; connect$$

$$s.t. \; \sum_j \hat{P}_{ij} = k, \hat{P}_{ii} = 0, \; \forall i, j \; \in 1, \ldots, n$$

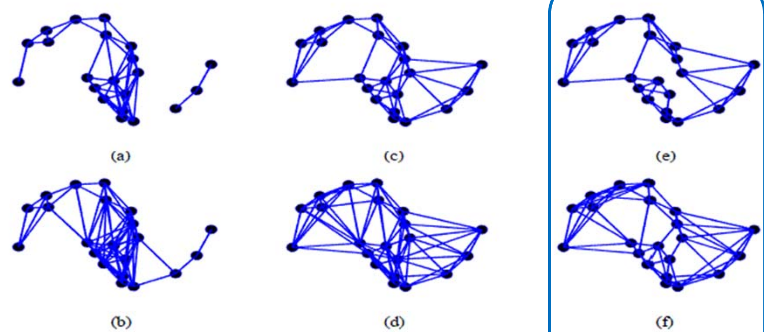
- B-Matched Graph

(Huang and Jebara, AISTATS 2007)
(Jebara, Wang, and Chang, ICML 2009)

$$\max_{P} \sum_{ij} P_{ij} W_{ij}$$

$$s.t. \; \sum_j P_{ij} = b, P_{ii} = 0, P_{ij} = P_{ji}, \; \forall i, j \; \in 1, \ldots, n$$

Several Ways of Constructing Sparse Graphs

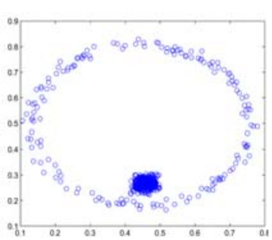


k,b=4

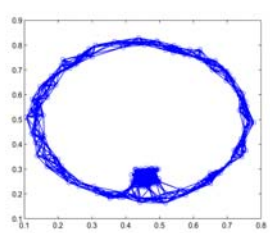
k,b=6

Distance threshold | Rank threshold (kNN) | B-Match

Examples of Graph Construction

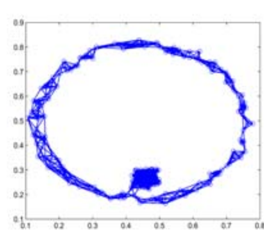


(KNN)



k = 4

(B-Matching)



b = 4

Graph Construction – Edge Weighting

- Binary Weighting

$$\mathbf{W} = \mathcal{P}$$

- Gaussian Kernel Weighting

$$\mathbf{W}_{ij} = \mathcal{P}_{ij} \exp\left(-\frac{d(\mathbf{x}_i, \mathbf{x}_j)}{2\sigma^2}\right)$$

- Locally Linear Reconstruction Weighting

$$\min_{W} \sum_{i} \|\mathbf{x}_i - \sum_{j=1}^{n} \mathcal{P}_{ij} w_{ij} \mathbf{x}_j\|^2$$
$$s.t. \ \sum_{j} w_{ij} = 1, w_{ij} \geq 0$$

Measure Smoothness: Graph Laplacian

- Graph Laplacian $\Delta = \mathbf{D} - \mathbf{W}$, and normalized Laplacian $\mathbf{L} = \mathbf{D}^{-1/2} \Delta \mathbf{D}^{-1/2}$
- smoothness of function *f* over graph

$$< f, \mathbf{L}f > = f^T \mathbf{L} f$$
$$= \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \left\| \frac{f(x_i)}{\sqrt{D_{ii}}} - \frac{f(x_j)}{\sqrt{D_{jj}}} \right\|^2$$

Multi-class $< \mathbf{F}, \mathbf{LF} > = tr(\mathbf{F}^{\mathrm{T}} \mathbf{LF})$

Classical Methods:

(Zhu et al ICML03, Zhou et al NIPS04, Joachim ICML03)

- Predict a graph function (F) via cost optimization

prediction function function smoothness empirical loss

$$\mathbf{F}^* = \arg\min_{\mathbf{F}} Q(\mathbf{F}) = \arg\min_{\mathbf{F}} \left\{ Q_{smooth}(\mathbf{F}) + Q_{fit}(\mathbf{F}) \right\}$$

➢ Local and Global Consistency - **LGC** (Zhou et al, NIPS 04)

$$\min_{\mathbf{F}\in\mathbb{R}^{|V|\times c}} \operatorname{tr}\{\mathbf{F}^\top \mathbf{L}\mathbf{F} + \mu(\mathbf{F}-\mathbf{Y})^\top(\mathbf{F}-\mathbf{Y})\} \quad \rightarrow \quad \mathbf{F}^* = (\mathbf{L}/\mu + \mathbf{I})^{-1}\mathbf{Y} = \mathbf{P}\mathbf{Y}$$

➢ Gaussian Random Fields – **GRF** (Zhu et al, ICML03)

$$\min_{\mathbf{F}\in\mathbb{R}^{|V|\times c}} \operatorname{tr}(\mathbf{F}^\top \triangle \mathbf{F})$$
$$s.t. \quad \mathbf{F}_l = \mathbf{Y}_l$$
$$\nabla_{F_u}(Q) = 0$$

Empirical Observations

(Jebara, Wang, and Chang, ICML 2009)

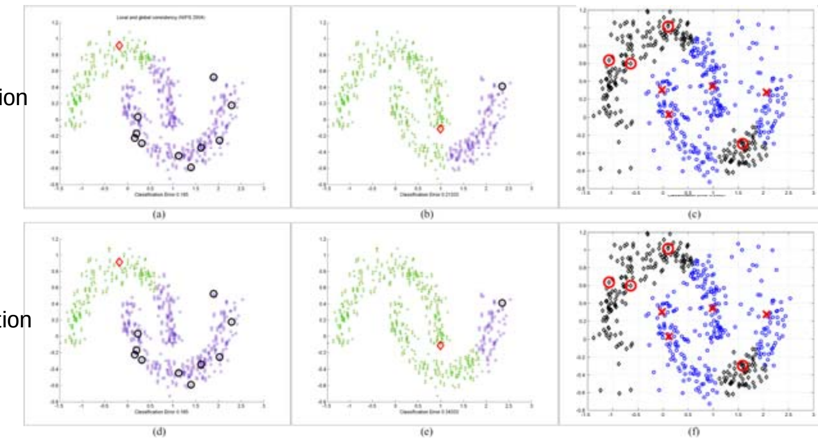
- Compare **method-graphs-weights**
- B-matching tends to outperform kNN
- B-Matching particularly good for GTAM + local linear (LLR) weight

Data set	USPS	COIL	BCI	TEXT
QC + CMN	13.61	59.63	50.36	40.79
LDS	25.2	67.5	49.15	31.21
Laplacian	17.57	61.9	49.27	27.15
Laplacian RLS	18.99	54.54	48.97	33.68
CHM (normed)	20.53	-	46.9	-
LGC-KNN-BN	14.7	59.18	48.94	48.79
LGC-KNN-GK	12.42	57.3	48.42	48.09
LGC-KNN-LLR	15.8	56.75	48.65	40.28
LGC-BM-BN	14.4	59.31	48.34	40.44
LGC-BM-GR	11.89	58.17	48.17	37.39
LGC-BM-LLR	14.44	58.69	48.08	39.83
GRF-KNN-BN	19.11	64.45	48.77	47.65
GRF-KNN-GK	12.94	61.31	48.98	47.65
GRF-KNN-LLR	19.20	61.19	48.46	47.14
GRF-BM-BN	18.98	60.63	48.44	43.16
GRF-BM-GR	12.82	60.87	48.77	42.88
GRF-BM-LLR	18.95	60.84	48.25	42.94
GTAM-*KNN-BN*	6.42	29.70	47.56	49.36
GTAM-*KNN-GK*	4.77	16.69	47.29	49.13
GTAM-*KNN-LLR*	6.69	15.35	45.54	41.48
GTAM-*BM-BN*	5.2	25.83	47.92	17.81
GTAM-*BM-GR*	**4.31**	13.65	47.48	28.74
GTAM-*BM-LLR*	5.45	**12.57**	**43.73**	**16.35**

Noisy Label and other Challenges

LGC Propagation

GRF Propagation



Unbalanced Labels | Ill Label Locations | Noisy Data and Labels

Label Unbalance - A Quick Fix

- Normalize labels within each class based on node degrees

Example:

classes

samples

$$\mathbf{Y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Label matrix

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix} \longrightarrow \mathbf{V} = \begin{bmatrix} \frac{1}{1+3} & 0 & 0 & 0 & 0 \\ 0 & \frac{3}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{1+3} & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{4} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Node degree matrix

Dealing with Noisy Labels
-- Graph Transduction via Alternate Minimization

*(**GTAM**, Wang, Jebara, & Chang, ICML, 2008) (**LDST**, Wang, Jiang, & Chang, CVPR, 2009)*

- Change *uni-variate* optimization to *bi-variate* formulation:

$$\min_{\mathbf{F}\in\mathbb{R}^{|V|\times c}} \mathrm{tr}\{\mathbf{F}^\top\mathbf{L}\mathbf{F} + \mu(\mathbf{F}-\mathbf{Y})^\top(\mathbf{F}-\mathbf{Y})\}$$

$$\Downarrow$$

$$\min_{\mathbf{F},\mathbf{Y}} \tfrac{1}{2}\mathrm{tr}\left\{\mathbf{F}^T\mathbf{L}\mathbf{F} + \mu(\mathbf{F}-\mathbf{V}\mathbf{Y})^T(\mathbf{F}-\mathbf{V}\mathbf{Y})\right\}$$
$$s.t.\quad \mathbf{Y}_{ij}\in\{0,1\}, \textstyle\sum_j \mathbf{Y}_{ij} = 1$$

Alternate Optimization

- First, given Y solve continuous valued $\mathbf{F}$

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

- Then, search optimal integer Y given F*

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

Gradient decent search

Alternate Minimization for Label Tuning

Example:

$$\mathbf{Y}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \nabla_Y Q = \begin{bmatrix} 0.8 & 0.1 \\ -0.23 & -0.25 \\ -0.31 & 0.07 \\ -0.17 & -0.04 \end{bmatrix}$$

$$\text{Add label:} \quad (i^+, j^+) = \min_{i,j} \nabla_{(vy_u)} Q; \quad y_{i^+j^+} = 1$$

$$\text{Delete label:} \quad (i^-, j^-) = \max_{i,j} \nabla_{(vy_l)} Q; \quad y_{i^-j^-} = 0$$

add label (3,1)

delete label (1,1)

$$\Rightarrow \quad \mathbf{Y}^T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

- Iteratively repeat the above procedure

Label Diagnosis and Self Tuning

(LDST, Wang, Jian, & Chang, CVPR, 2009)

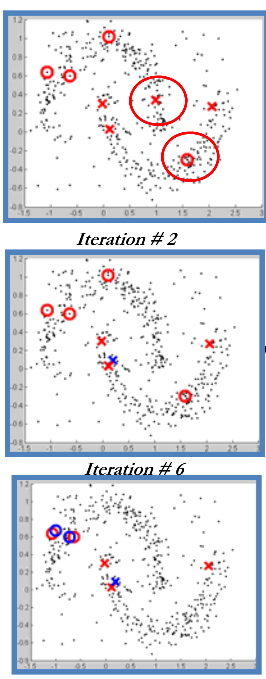
$$\text{Add label:} \quad (i^+, j^+) = \min_{i,j} \nabla_{(vy_u)} Q; \quad y_{i^+j^+} = 1$$

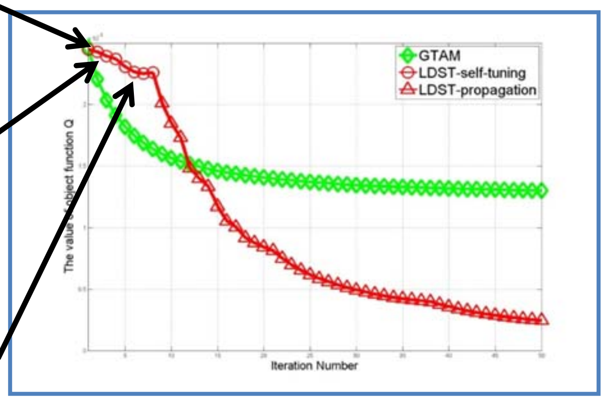
$$\text{Delete label:} \quad (i^-, j^-) = \max_{i,j} \nabla_{(vy_l)} Q; \quad y_{i^-j^-} = 0$$

Initial Labels

Iteration # 2

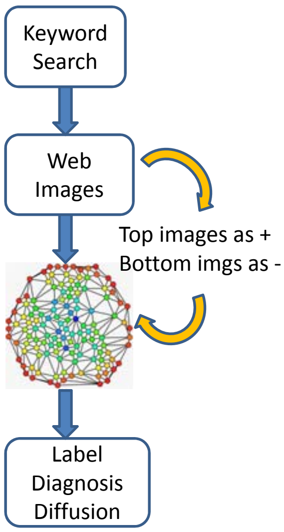
Iteration # 6





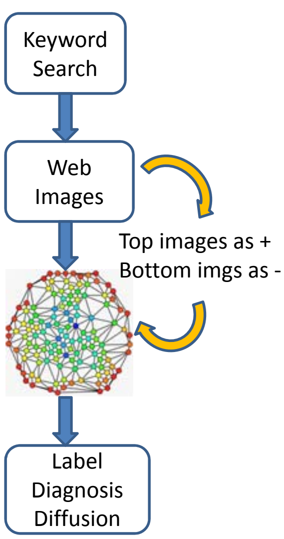
Decline of the cost function Q over iterations (with vs. without label tuning)

Application: Web Search Reranking



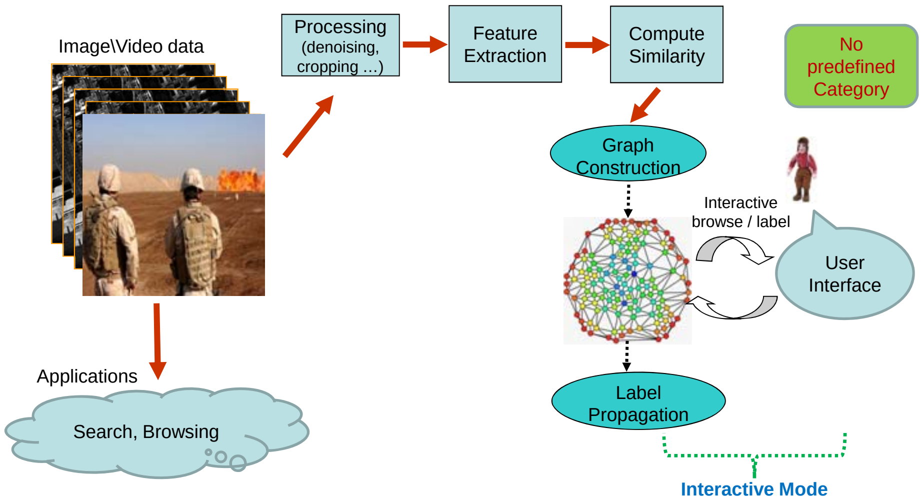
Google Search "Tiger"

Application: Web Search Reranking



Rerank

Possible Applications: Propagating Labels in Interactive Search & Auto Re-ranking



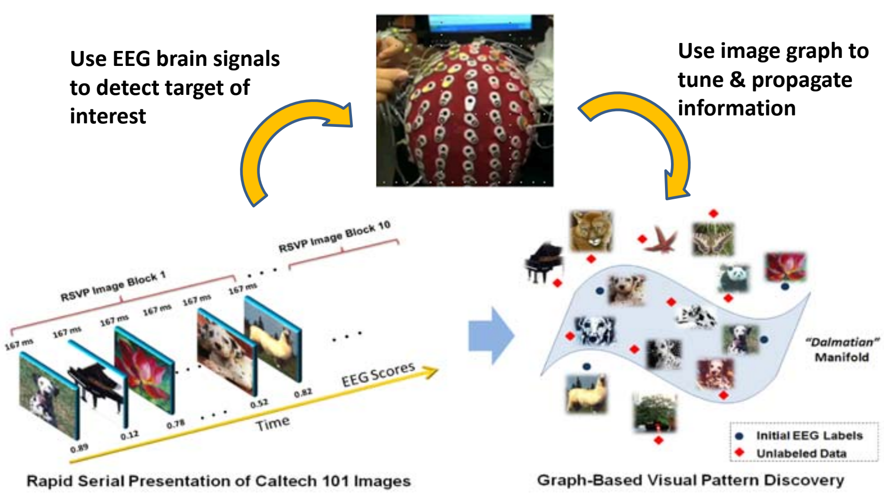
S.-F. Chang, Columbia U.

31

Application: Brain Machine Interface for Image Retrieval

-- denoise unreliable labels from brain signal decoding

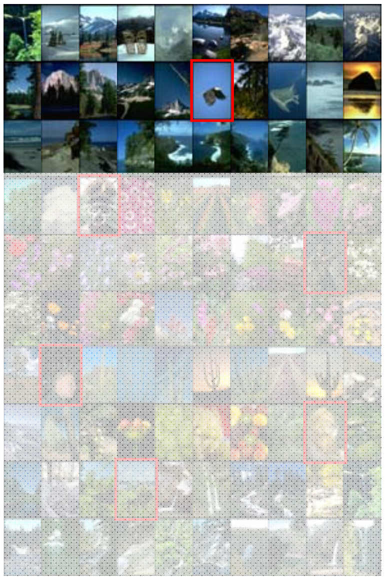
(joint work with Sajda et al, ACMMM 2009, J. of Neural Engineering, May 2011)



COLUMBIA UNIVERSITY

The Paradigm

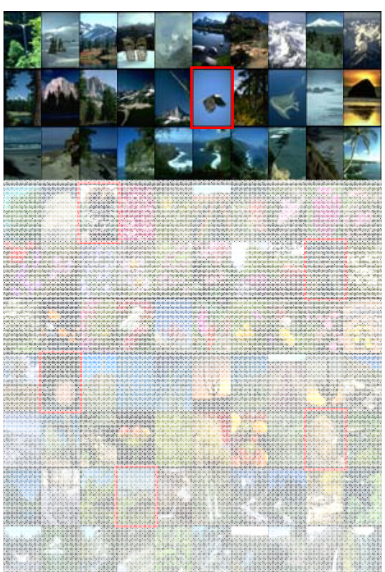
Database (any target that may interest users)

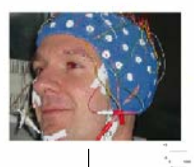


33

COLUMBIA UNIVERSITY

The Paradigm





Neural (EEG) decoder

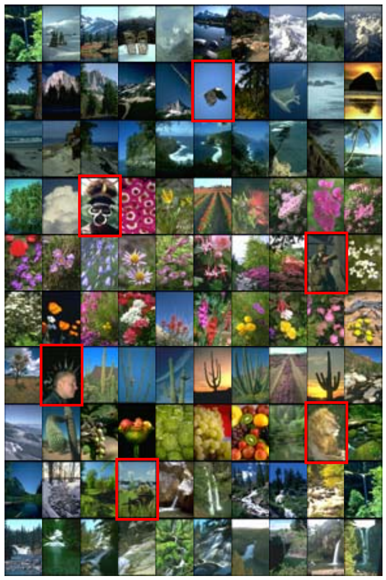
EEG-scores

Database

34

COLUMBIA UNIVERSITY

The Paradigm



Database

Neural (EEG) decoder

Exemplar labels (noisy)

image features

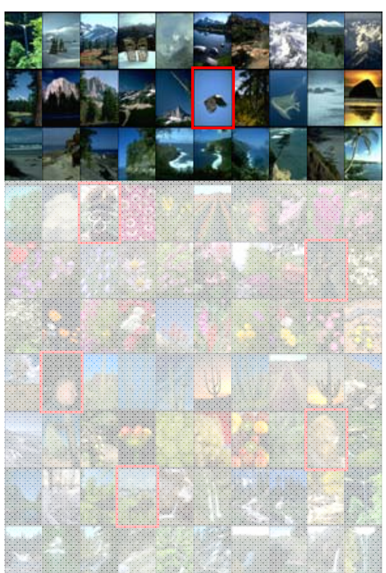
Graph-based Semi-Supervised Learning

prediction score

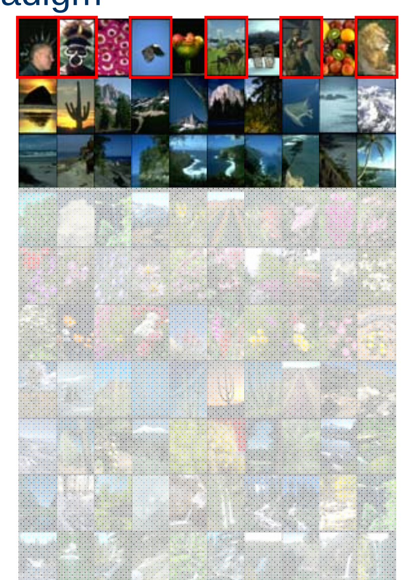
35

COLUMBIA UNIVERSITY

The Paradigm



Pre-triage

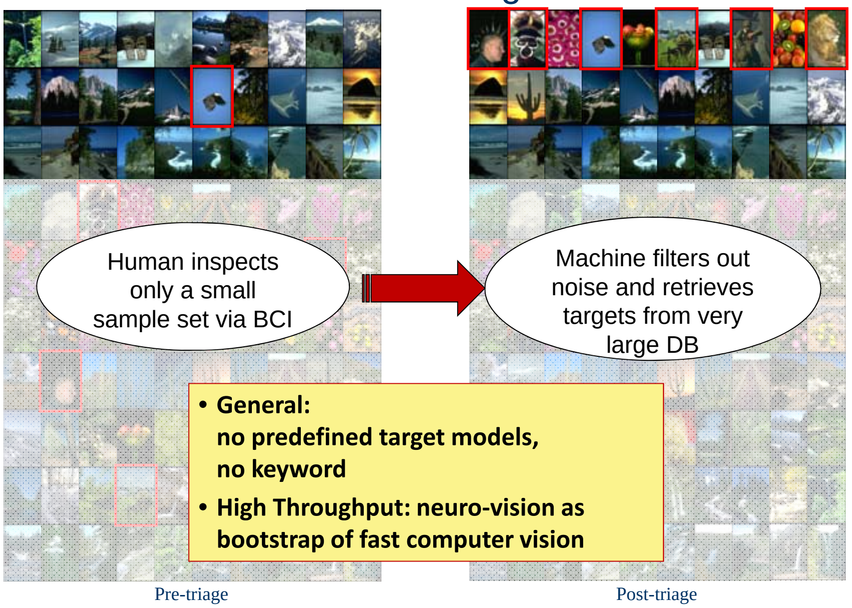


Post-triage

36

COLUMBIA UNIVERSITY

The Paradigm



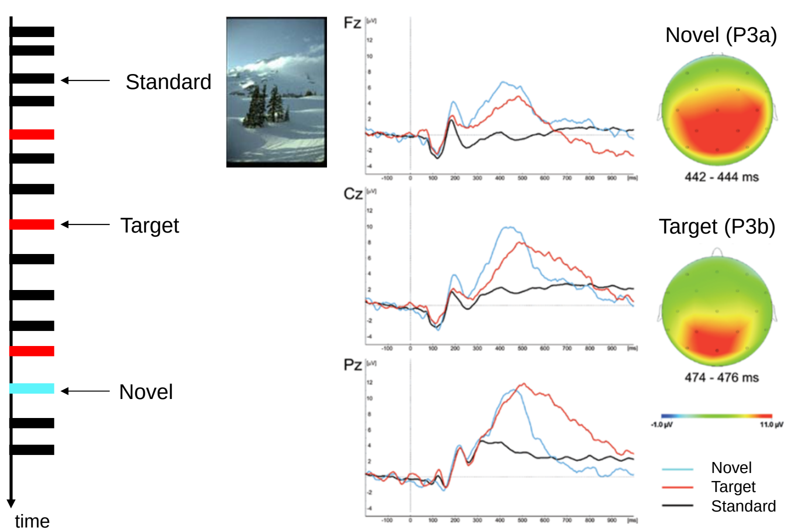
- **General:** **no predefined target models, no keyword**
- **High Throughput: neuro-vision as bootstrap of fast computer vision**

37

COLUMBIA UNIVERSITY

The Neural Signatures of "Recognition"

D. Linden, Neuroscientist, 2005, ***the Oddball Effect***

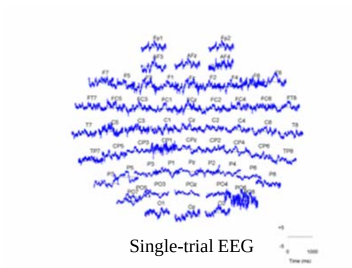


38

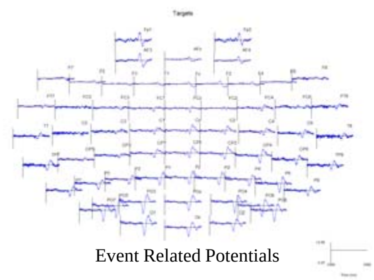
COLUMBIA UNIVERSITY

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



Single-trial EEG



Event Related Potentials

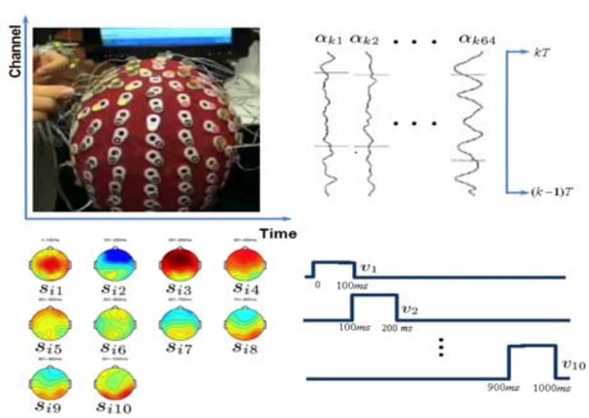
NSF HNCV10 39

COLUMBIA UNIVERSITY

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)

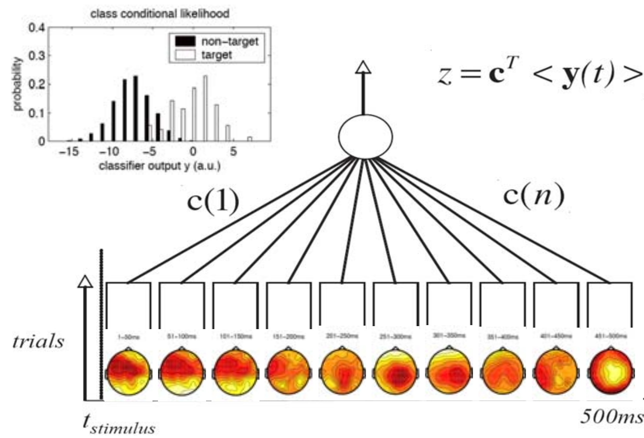
Optimal temporal filtering over time windows after onset

NSF HNCV10 40

COLUMBIA UNIVERSITY

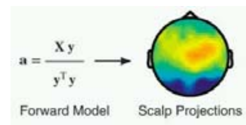
Hierarchical Discriminant Components

… use factorization to greatly reduce the number of parameters (100K -> 100)…



$$z = \mathbf{c}^T < \mathbf{y}(t) >$$

Discover distribution of discriminative components



$$\mathbf{a} = \frac{\mathbf{X}\,\mathbf{y}}{\mathbf{y}^T\mathbf{y}}$$

$t_{stimulus}$

$$\mathbf{y}(t) = \mathbf{w}_\tau^T \mathbf{x}(t) \qquad \tau = \left\{ t_k - \frac{\delta}{2} \ldots t_k + \frac{\delta}{2} \right\}$$

NSF HNCV10

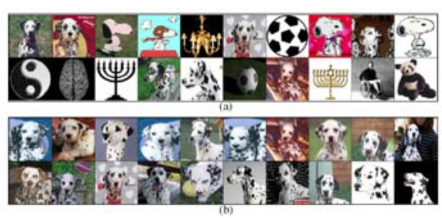
41

COLUMBIA UNIVERSITY

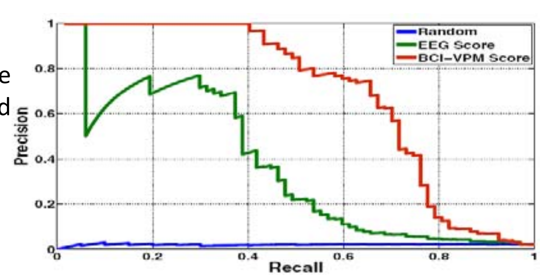
Effect of graph-based reranking (BCI test)

Top (noisy) results of Brain EEG signal detection

Top results after graph-based label denoising & propagation



P-R curve significantly improved



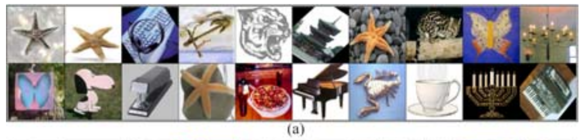
42

COLUMBIA UNIVERSITY

More Example Results

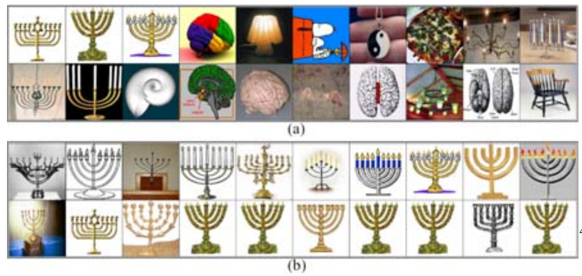
Top 20 results of EEG detection

Top 20 results of Hybrid System (BCI-VPM)



Top 20 results of EEG detection

Top 20 results of Hybrid System (BCI-VPM)



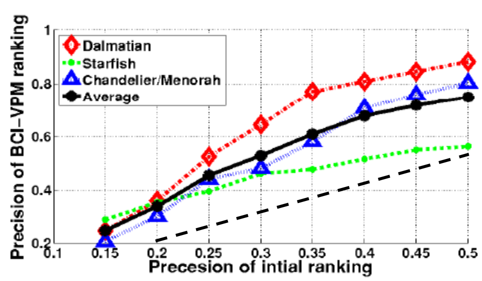
43

Dependency of Neuro & CV Components

… not every case improves …

- Among 12 cases (4 subjects & 3 targets), 8 cases are clearly improved. When the EEG decoder fails, the hybrid system also fails.

Question:
- what's the required EEG accuracy for the hybrid system to work?
- are some categories more difficult?



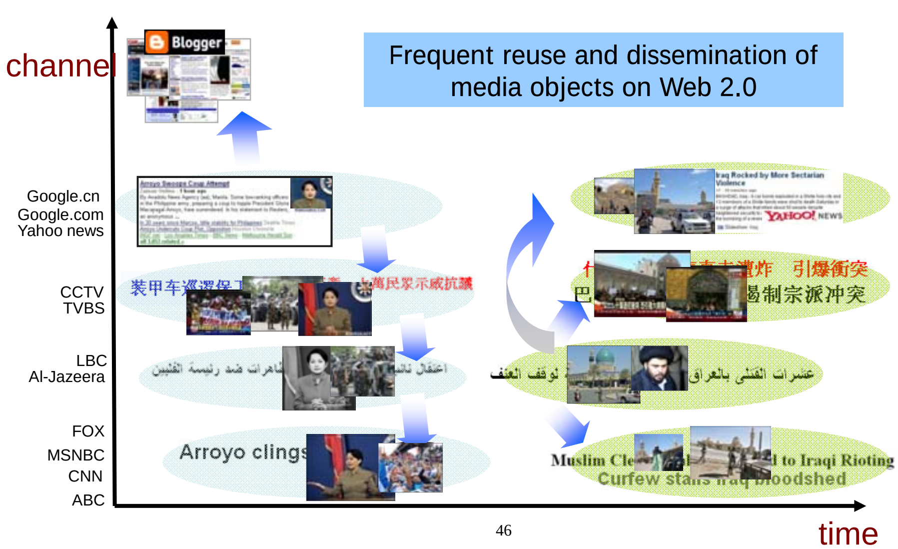
EEG accuracy

44

Part II: Application of Image Matching

Social Media Tracking

Application of Image Matching: Tracking Media Objects on the Web



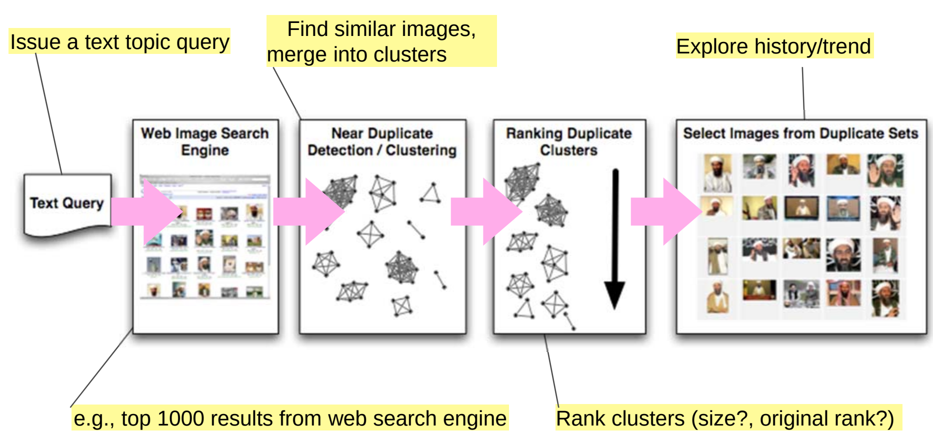
Manipulation correlated with *Perspective*

Raising the Flag on Iwo Jima
Joe Rosenthal, 1945

Anti-Vietnam War,
Ronald and Karen Bowen, 1969

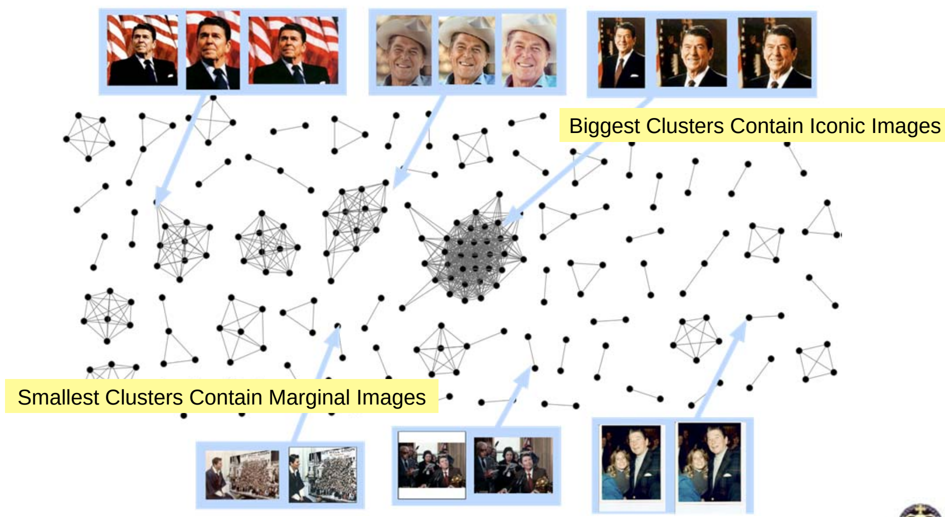
DVMM Lab, Columbia University 47

Web Scenario: Search, Cluster, Insights

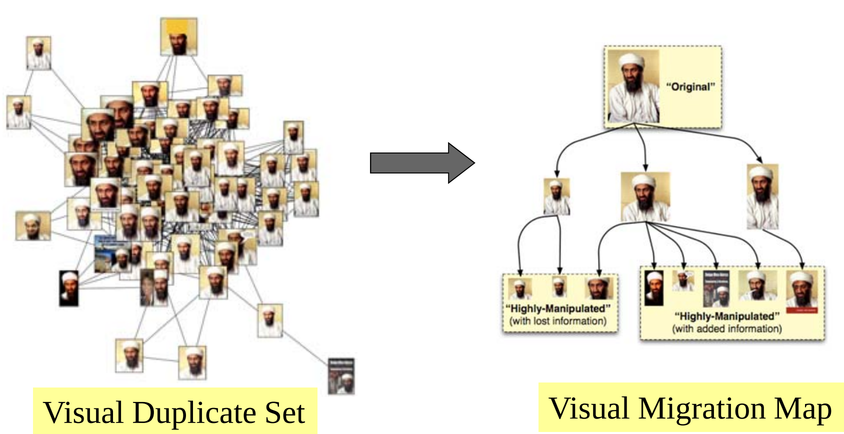


-48- digital video | multimedia lab

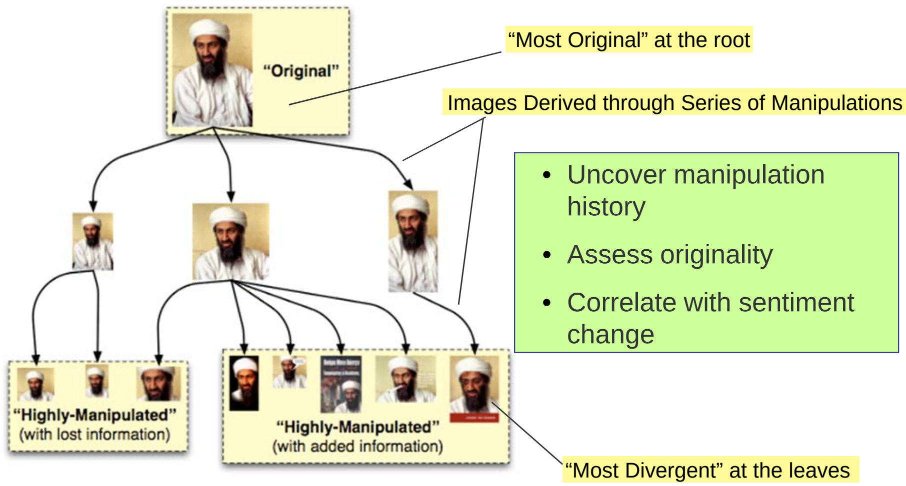
Duplicate Clusters Reveal Image Provenance



Deep Analysis of Visual Data: Visual Migration Map (VMM)

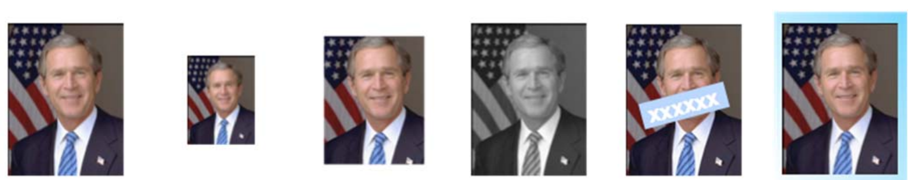


Visual Migration Map (VMM)



- Uncover manipulation history
- Assess originality
- Correlate with sentiment change

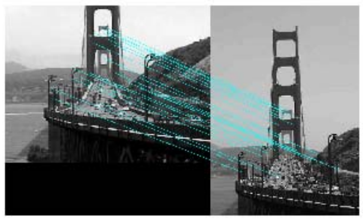
How to automate VMM construction? Start with Basic Manipulations

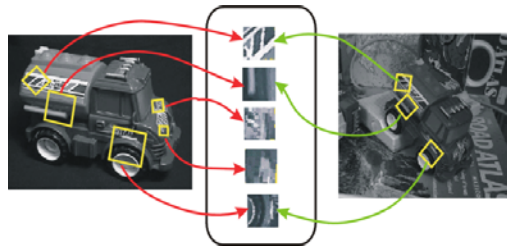


Original Scaled Cropped Gray Overlay Insertion

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

Image Matching by Local Features





[Lowe, 1999]

- Find local interest points in images
- Match interest points to determine image copies of the same scenes or objects
- Estimate geometrical transforms from matched points

-53- »digital video | multimedia lab»

More Challenging: Overlay Detection?

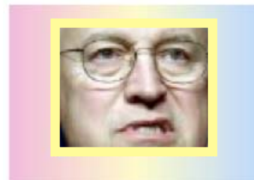
- Given two images, we can observe that a region is different between the two
- But how do we know which is the original?

DVMM Lab, Columbia University

Cropping or Insertion?

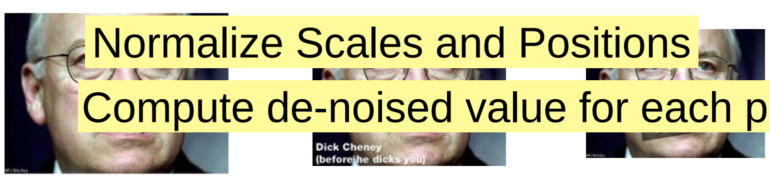

Original


Cropping


Insertion

- Can find differences in image area
- But is the smaller-area due to a crop *or* is the larger area due to an insertion?

Use Context from Many Duplicates

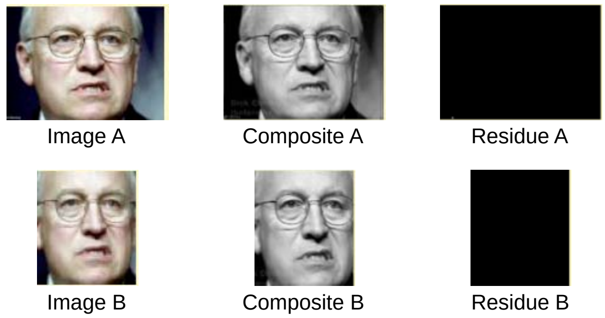


Normalize Scales and Positions

Compute de-noised value for each pixel

“Composite” image

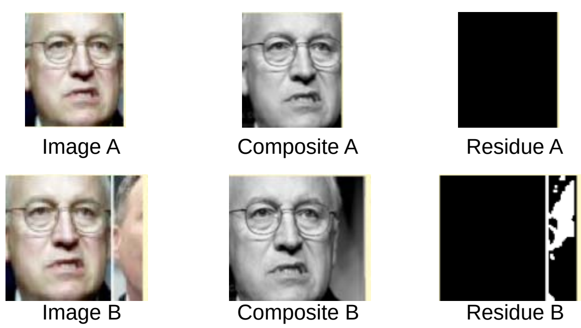
Cropping Detection w/ Context



- In cropping, we expect the content outside the crop area to be consistent with the composite image

DVMM Lab, Columbia University

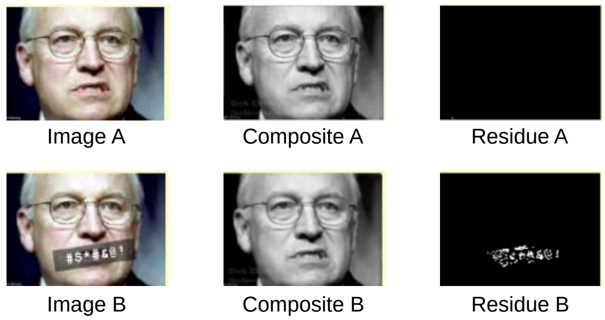
Insertion Detection w/ Context



- In insertion, we expect the area outside the crop region to be different from the typical content

DVMM Lab, Columbia University

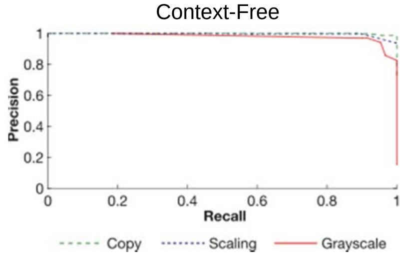
Overlay Detection w/ Context



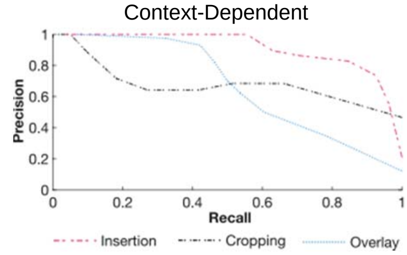
- Comparing images against composite image reveals portions that differ from typical content
- Image with divergent content may have overlay

Evaluation: Manipulation Detectio

Context-Free

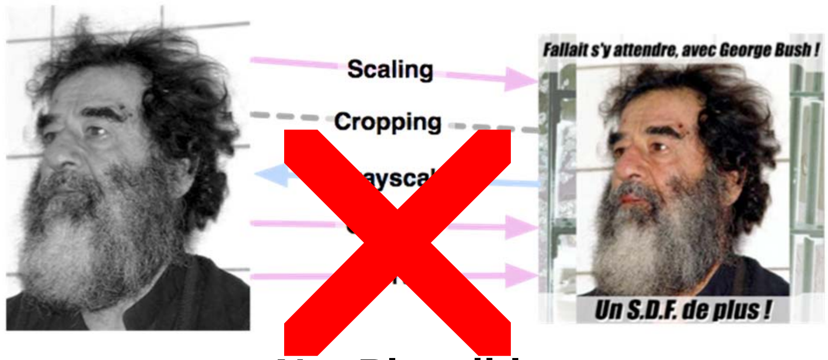


Context-Dependent



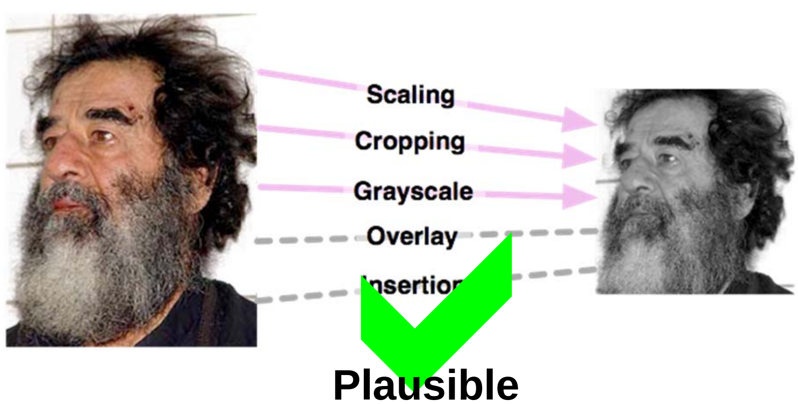
- Context-Free detectors have near-perfect performance
- Context-Dependent detectors still have errors
- Consistency checking can further improve the accuracy
- Are these error-prone results sufficient to build manipulation histories?

Inferring Direction from Consistency



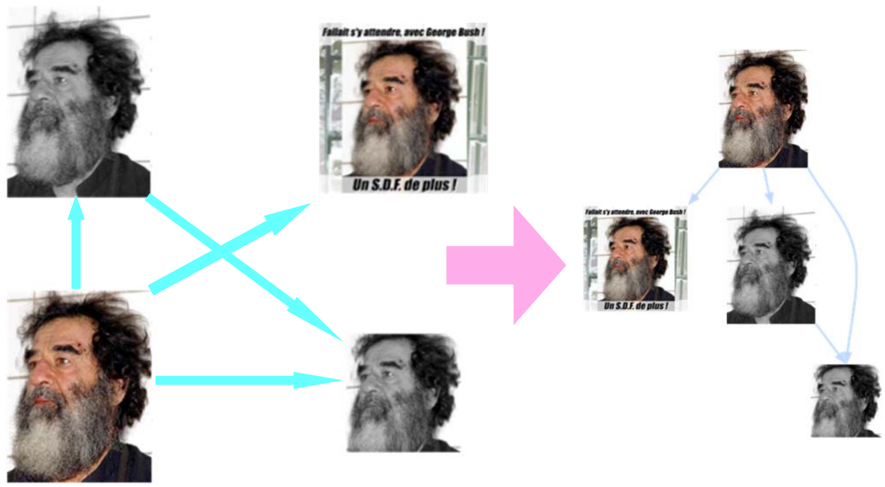
Not Plausible

Manipulation Direction from Consistency



Plausible

Derive Manipulation among Multiple Images



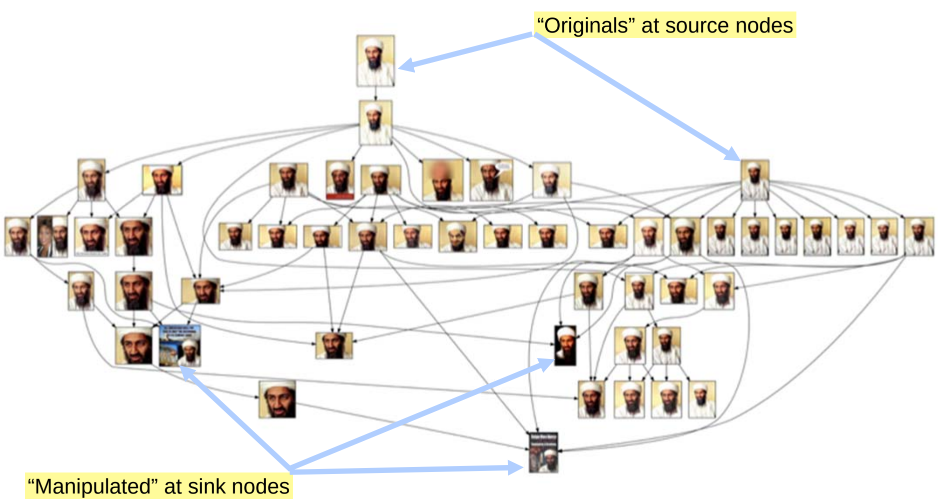
Emerging Migration Map

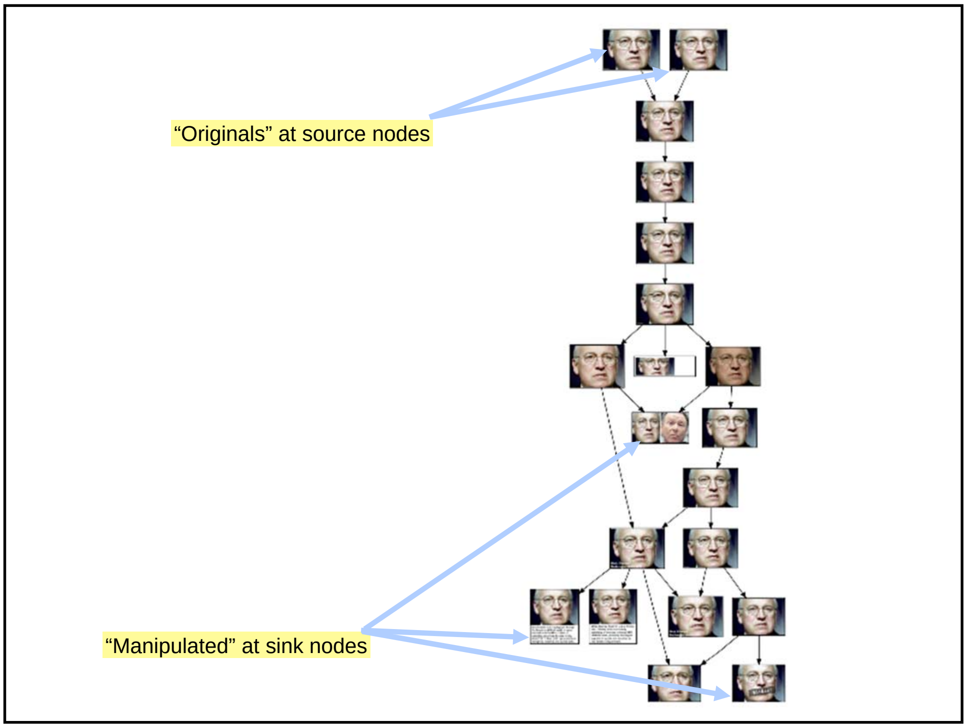
- Individual parent-child relationships give rise to a manipulation history
- Relationships are only *plausible* (we don't know for sure)
- Absences of relationships are more concrete (we can be more certain)
- Redundancy: plausible derivations from parents and ancestors of parents

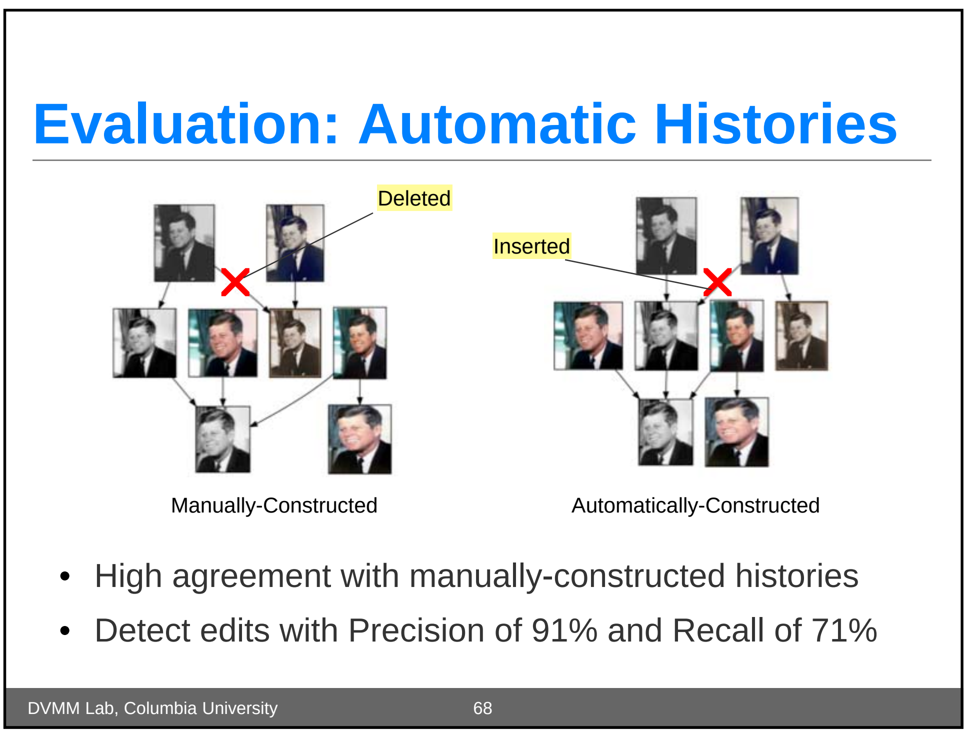
Experiments

- Select 22 iconic images
- Mostly political figures, culled from Google Zeitgeist and TRECVID queries
- Generate manipulation histories:
  - fully-automatic mechanisms
  - pseudo groundtruth through manual annotation

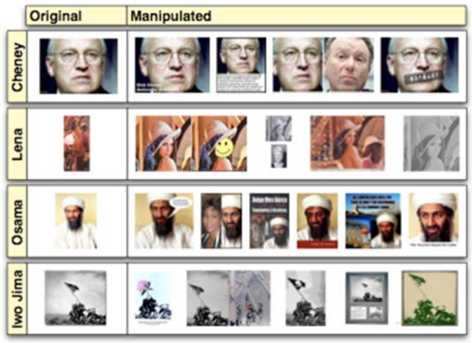
Automatic Visual Migration Map




"Originals" at source nodes
"Manipulated" at sink nodes

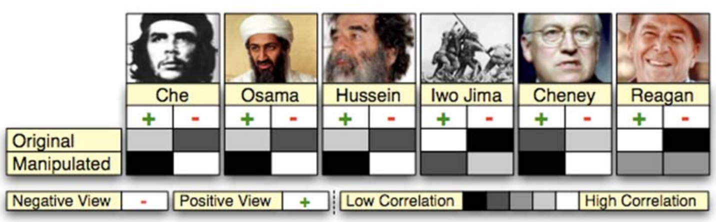

Evaluation: Automatic Histories
Deleted
Inserted
Manually-Constructed
Automatically-Constructed
• High agreement with manually-constructed histories
• Detect edits with Precision of 91% and Recall of 71%
DVMM Lab, Columbia University
68

Application: Summarizing Changes



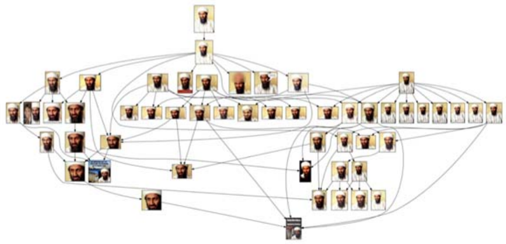
- Analyze manipulation history graph structure to extract most-original and most highly-manipulated images

Application: Finding Perspective

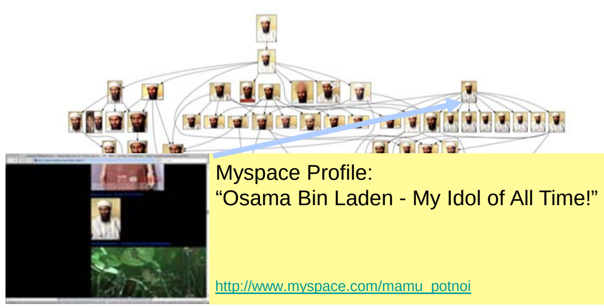


- Survey image type and corresponding perspective across many examples
- Find correlation between high manipulation and negative/critical opinion

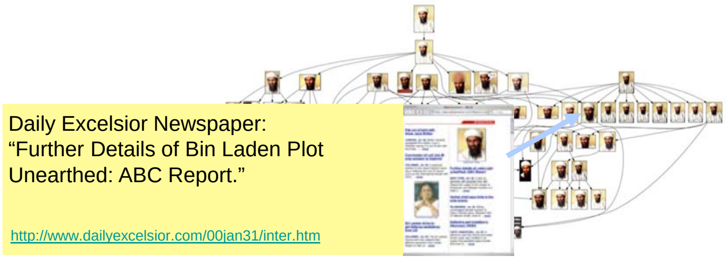
Application: Finding Perspective



Application: Finding Perspective



Application: Finding Perspective



DVMM Lab, Columbia University 73

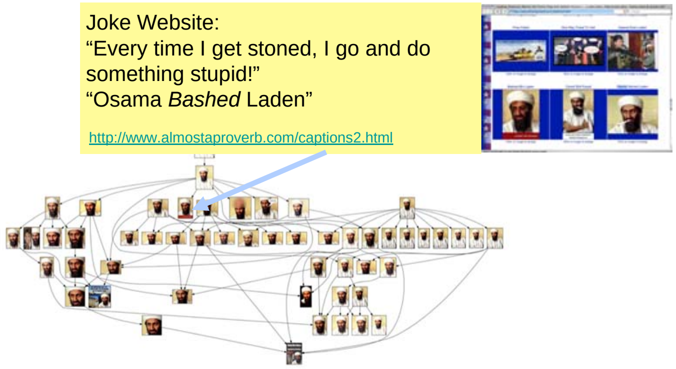
Application: Finding Perspective

DVMM Lab, Columbia University 74

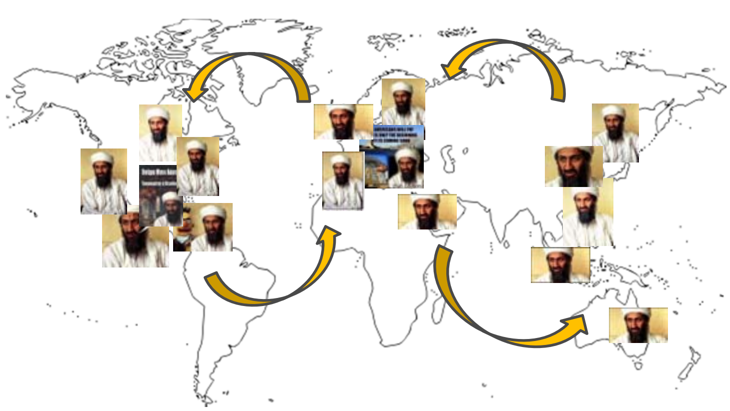
Application: Finding Perspective

Joke Website:
"Every time I get stoned, I go and do something stupid!"
"Osama *Bashed* Laden"

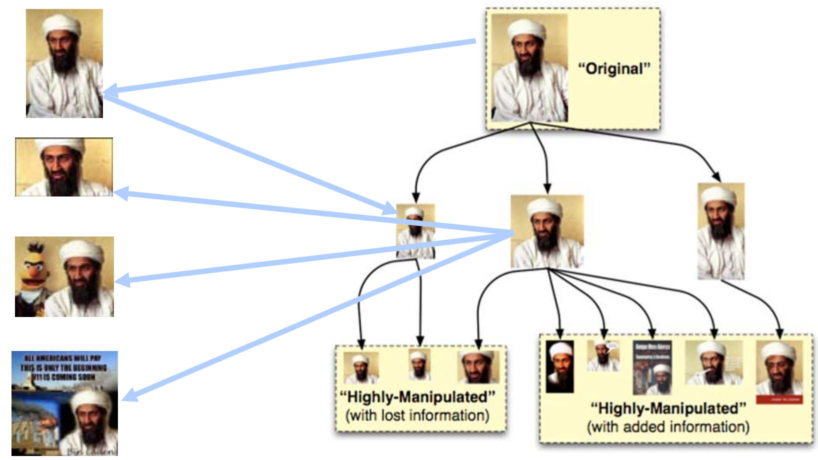
http://www.almostaproverb.com/captions2.html



VMM Applications: Geographic/Cultural Dispersion



VMM Applications: Reverse Profiling



DVMM Lab, Columbia University 77

References

- Graph-Based Relevance Feedback
  - J. Wang, Y.-G. Jiang, and S.-F. Chang. **Label diagnosis through self tuning** for web image search. *CVPR*, 2009.
  - T. Jebara, J. Wang, and S.-F. Chang, Graph Construction and b-Matching for Semi-Supervised Learning, ICML 2009.
  - X. Zhu, Z. Ghahramani, and J. D. Lafferty. Semi-supervised learning using Gaussian fields and harmonic functions. *ICML*, 2003.
  - D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. *NIPS*, 2004.
- Brain Machine Interface
  - J. Wang, E. Pohlmeyer, B. Hanna, Y.-G. Jiang, P. Sajda, and S.-F. Chang, "**Brain State Decoding** for Rapid Image Retrieval," ACM Multimedia Conference, 2009.
- Web Image Tracking
  - L. Kennedy and S.-F. Chang, "Internet Image Archaeology: Automatically Tracing the Manipulation Histories of Images on the Web," ACM Multimedia 2008, Vancouver, Canada, October 2008.
  - L. Xie, et al, "Tracking Visual Memes in Rich-Media Social Communities," International AAAI Conference on Weblogs and Social Media, 2011.

78