PageRank for Product Image Search

WWW 2008 China

Yushi Jing & Shumeet Baluja

CS685: Data Mining

Prof. Arnab Bhattacharya

Presented By –
Ashish Gupta & Akshay Mittal

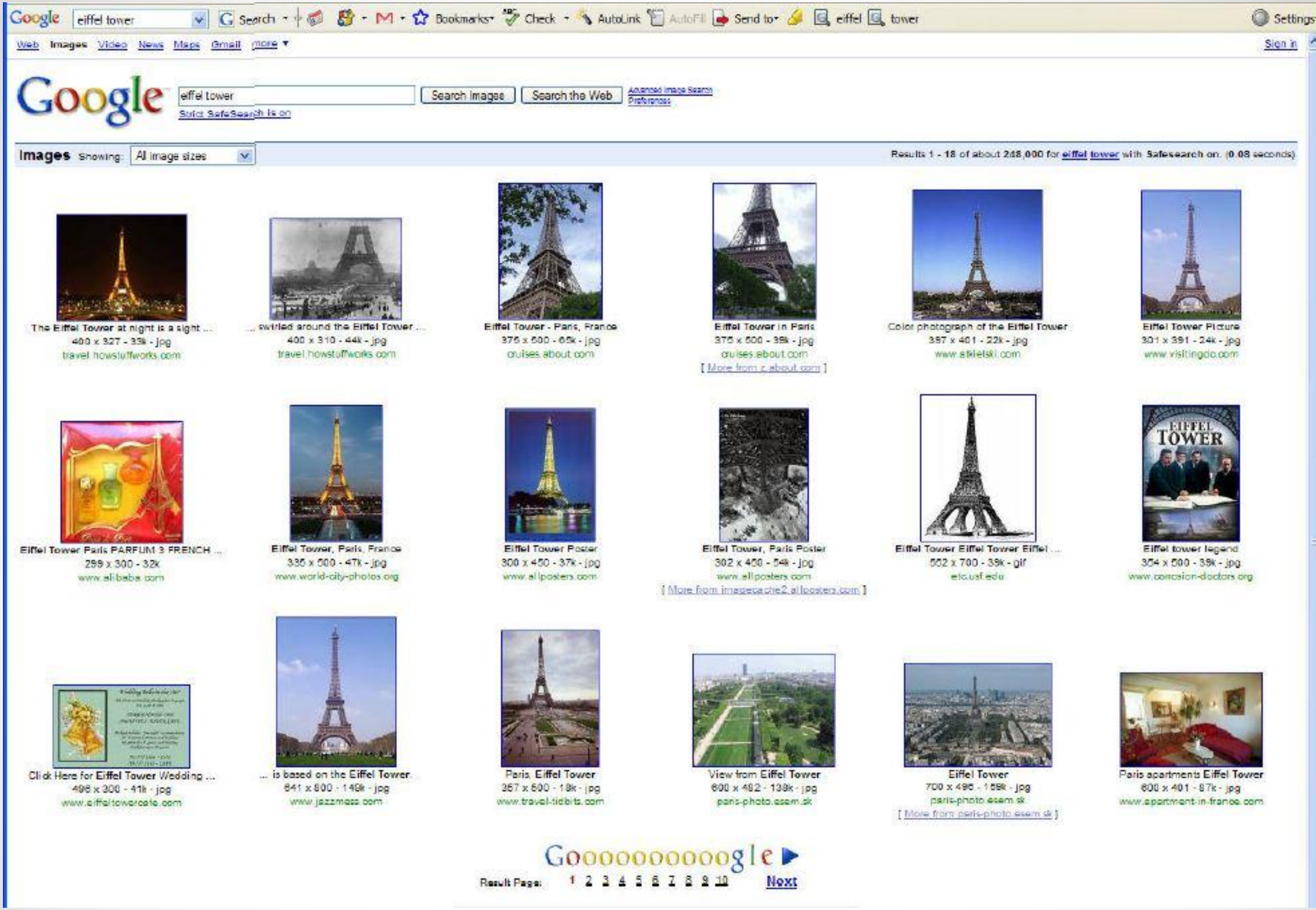
Image Search

Text on Pages

- Well-studied
- Human-recognizable objects
  - Faces
  - Highly textured objects
- Processing

Drawback

- Inconsistent results in terms of quality



Google eiffel tower Search Bookmarks Check AutoLink AutoFill Send to eiffel tower Settings
Web Images Video News Maps Gmail more ▼
Sign in
Google
eiffel tower
Search Images
Search the Web
Advanced Image Search
Preferences
Strict SafeSearch is on
Images Showing: All image sizes
Results 1 - 18 of about 248,000 for eiffel tower with Safesearch on. (0.08 seconds)
The Eiffel Tower at night is a sight ...
400 x 327 - 33k - jpg
travel.howstuffworks.com
... swirled around the Eiffel Tower ...
400 x 310 - 44k - jpg
travel.howstuffworks.com
Eiffel Tower - Paris, France
375 x 500 - 65k - jpg
cruises.about.com
Eiffel Tower in Paris
375 x 500 - 38k - jpg
cruises.about.com
[More from c.about.com]
Color photograph of the Eiffel Tower
397 x 401 - 22k - jpg
www.alkielski.com
Eiffel Tower Picture
301 x 391 - 24k - jpg
www.visitingdc.com
Eiffel Tower Paris PARFUM 3 FRENCH ...
299 x 300 - 32k
www.alibaba.com
Eiffel Tower, Paris, France
335 x 500 - 47k - jpg
www.world-city-photos.org
Eiffel Tower Poster
300 x 450 - 37k - jpg
www.allposters.com
Eiffel Tower, Paris Poster
302 x 450 - 54k - jpg
www.allposters.com
[More from imagecache2.allposters.com]
Eiffel Tower Eiffel Tower Eiffel ...
552 x 700 - 39k - gif
etc.usf.edu
EIFFEL TOWER
Eiffel tower legend
354 x 500 - 38k - jpg
www.concision-doctors.org
Click Here for Eiffel Tower Wedding ...
496 x 300 - 41k - jpg
www.eiffeltowercafe.com
... is based on the Eiffel Tower.
641 x 800 - 149k - jpg
www.jazzmess.com
Paris, Eiffel Tower
357 x 500 - 18k - jpg
www.travel-tidbits.com
View from Eiffel Tower
600 x 482 - 138k - jpg
paris-photo.esem.sk
Eiffel Tower
700 x 496 - 158k - jpg
paris-photo.esem.sk
[More from paris-photo.esem.sk]
Paris apartments Eiffel Tower
600 x 401 - 87k - jpg
www.apartment-in-france.com
Gooooooooooogle ▶
Result Page: 1 2 3 4 5 6 7 8 9 10 Next

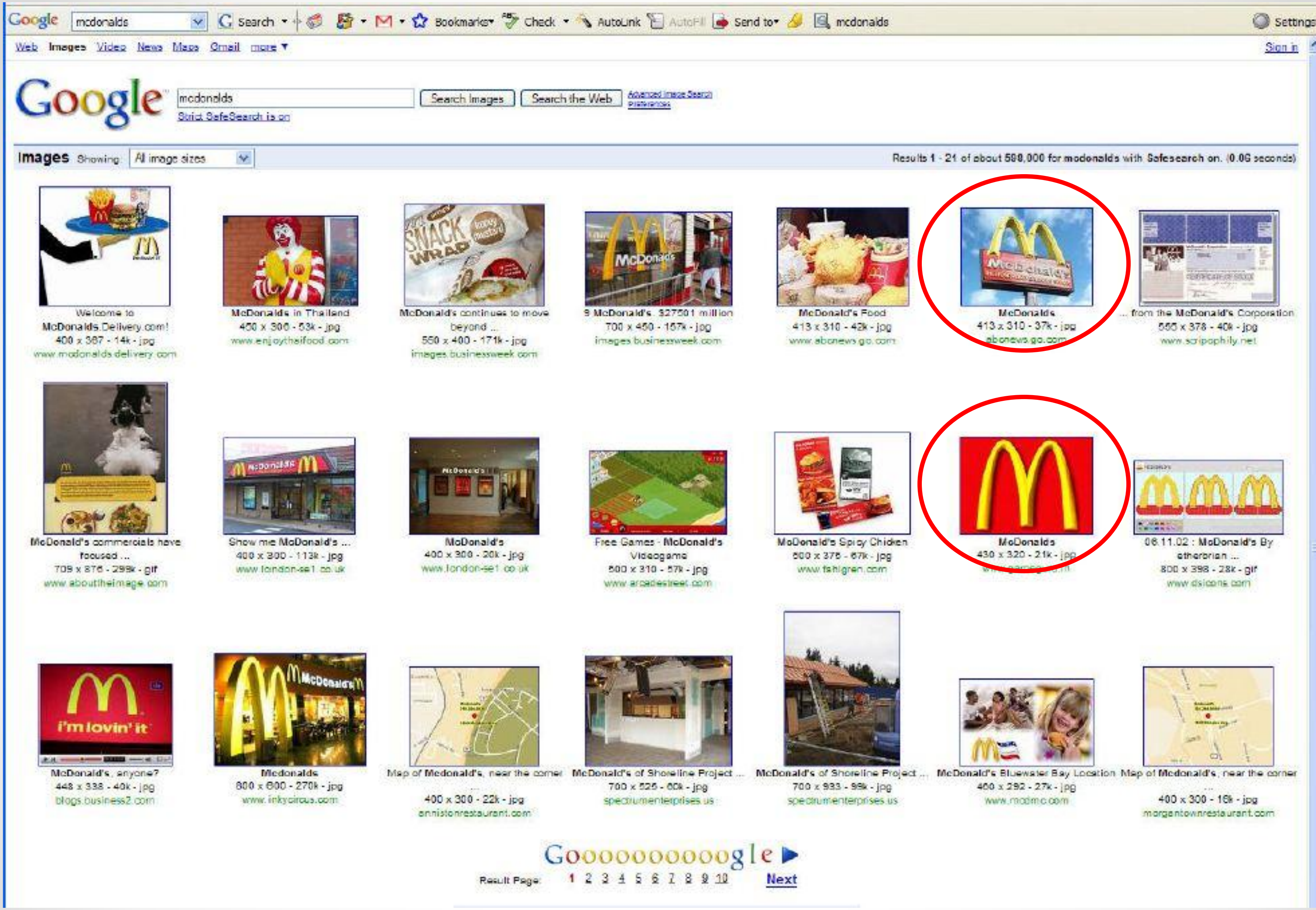

Google mcdonalds Search Bookmarks Check AutoLink AutoFill Send to mcdonalds Settings
Web Images Video News Maps Gmail more
Sign in
Google
mcdonalds
Search Images
Search the Web
Advanced Image Search
Preferences
Strict SafeSearch is on
Images Showing: All image sizes
Results 1 - 21 of about 598,000 for mcdonalds with Safesearch on. (0.06 seconds)
Welcome to McDonalds.Delivery.com! 400 x 367 - 14k - jpg www.mcdonalds.delivery.com
McDonalds in Thailand 450 x 300 - 53k - jpg www.enjoythaifood.com
McDonald's continues to move beyond ... 550 x 400 - 171k - jpg images.businessweek.com
9 McDonald's, $27501 million 700 x 450 - 157k - jpg images.businessweek.com
McDonald's Food 413 x 310 - 42k - jpg www.abcnews.go.com
McDonalds 413 x 310 - 37k - jpg abcnews.go.com
... from the McDonald's Corporation 555 x 378 - 40k - jpg www.scripophily.net
McDonald's commercials have focused ... 709 x 876 - 299k - gif www.abouttheimage.com
Show me McDonald's ... 400 x 300 - 113k - jpg www.london-se1.co.uk
McDonald's 400 x 300 - 20k - jpg www.london-se1.co.uk
Free Games - McDonald's Videogame 500 x 310 - 57k - jpg www.arcadestreet.com
McDonald's Spicy Chicken 500 x 375 - 67k - jpg www.fshlgren.com
McDonalds 430 x 320 - 21k - jpg
06.11.02 : McDonald's By etherbrian ... 800 x 398 - 28k - gif www.dsicons.com
McDonald's, anyone? 448 x 338 - 40k - jpg blogs.business2.com
Mcdonalds 800 x 600 - 270k - jpg www.inkycircus.com
Map of Mcdonald's, near the corner ... 400 x 300 - 22k - jpg annistonrestaurant.com
McDonald's of Shoreline Project ... 700 x 525 - 60k - jpg spectrumenterprises.us
McDonald's of Shoreline Project ... 700 x 933 - 98k - jpg spectrumenterprises.us
McDonald's Bluewater Bay Location 400 x 292 - 27k - jpg www.mcdmc.com
Map of Mcdonald's, near the corner ... 400 x 300 - 16k - jpg morgantownrestaurant.com
Gooooooooooogle
Result Page: 1 2 3 4 5 6 7 8 9 10 Next

Image Search

- Analyzing distribution of "Visual Similarities"
- Symbol "M" – Repetition in a large fraction of images
- Finding multiple visual themes and their relative strengths

Challenges

  - Image Processing (crooked, rotated, non-standard color, etc.)
  - Ranking

Image Processing – Content Based Image Ranking

- Object Category Model
- Homogeneous Object Category
- Limited Scale of Experiment



Query: “nemo”

Solution

- Model *expected user behavior*
- Similarities as probabilistic visual hyperlinks
- Graph:- *vertices* (images) & *edges* (visual hyperlinks)

Google ? PageRank : ImageRank

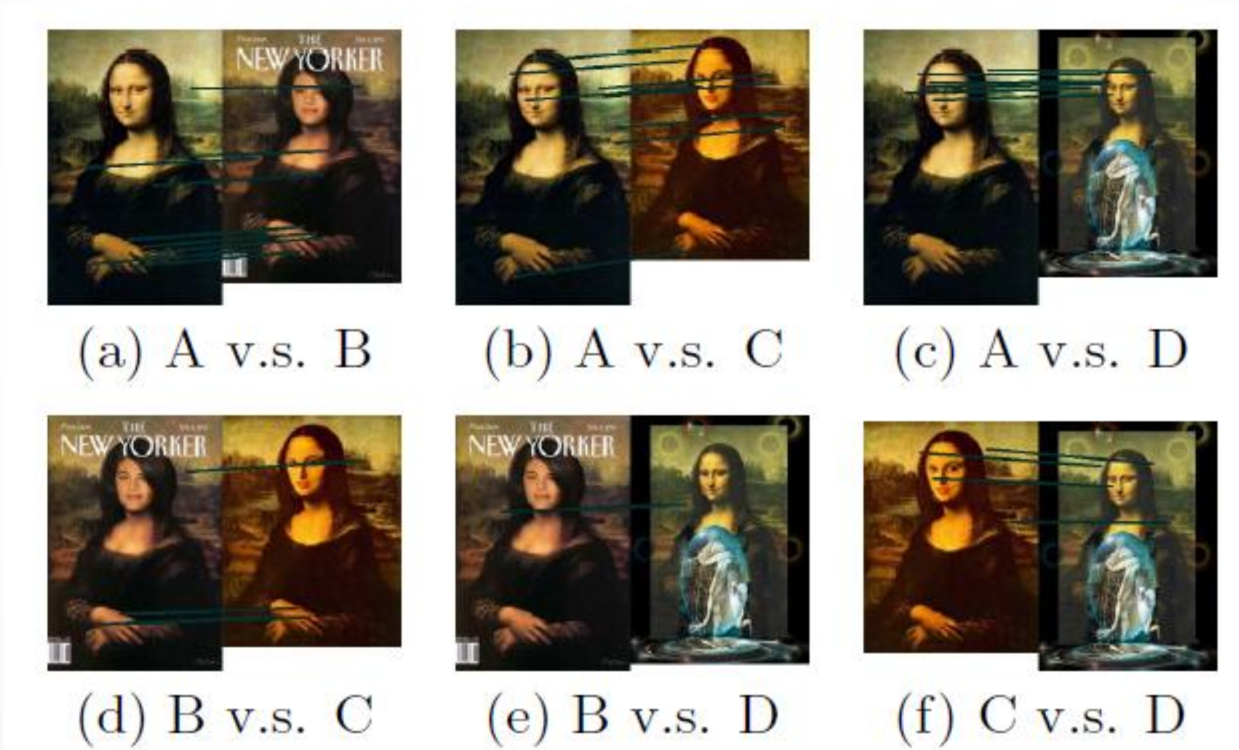
PageRank	ImageRank
$\mathbf{R} = M^* \mathbf{R}$	$\mathbf{IR} = S^* \mathbf{IR}$
M^* - stochastic matrix	S^* - normalized symmetrical adjacency matrix
Damping Effect: $d > 0.8$	
$\mathbf{R} = dM^* \mathbf{R} + (1 - d)\mathbf{1}/N$	$\mathbf{IR} = dS^* \mathbf{IR} + (1 - d)p$ where $p = [1/n](n \times 1)$

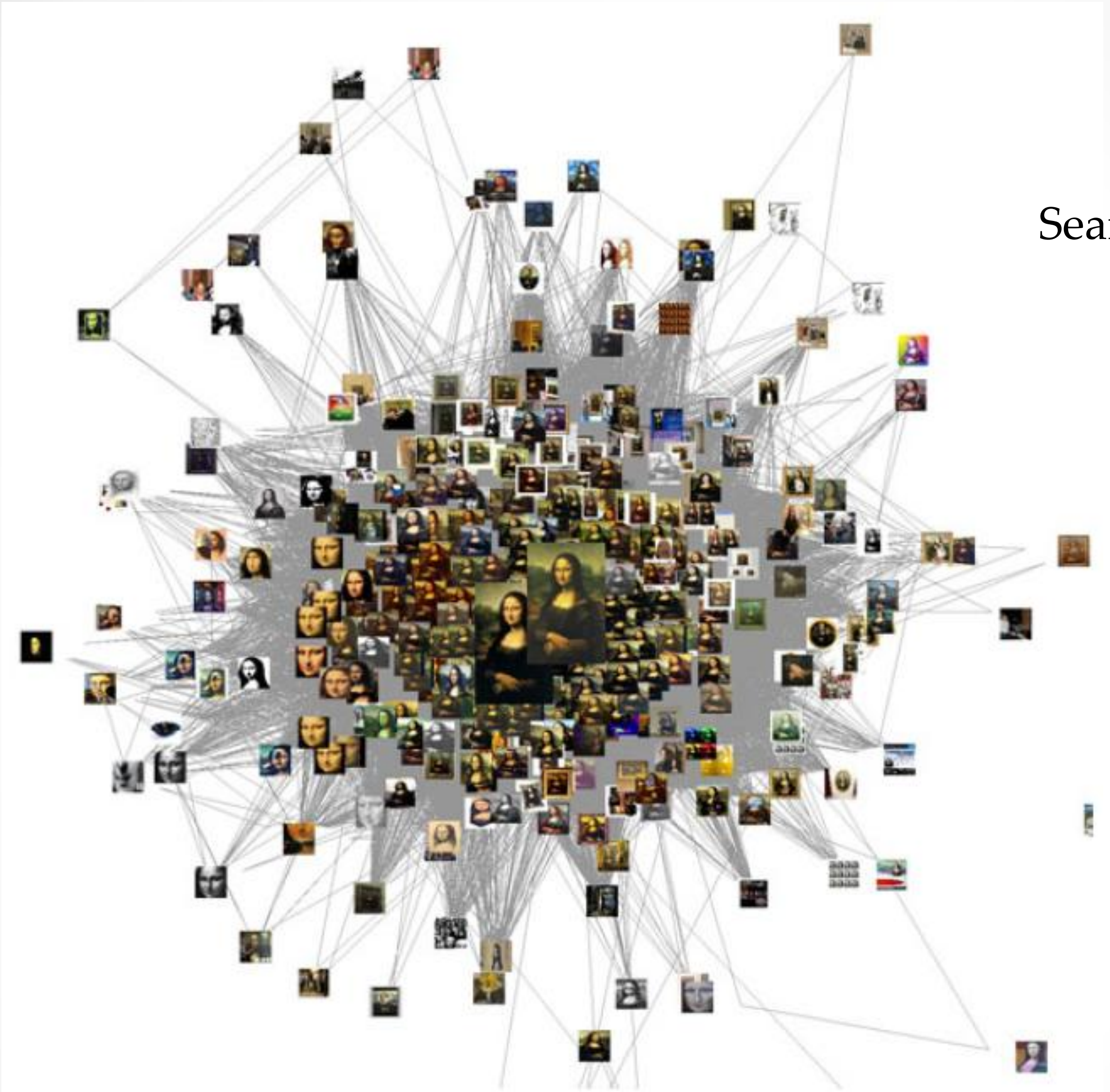
- Global Features versus Local Descriptors
- Scale Invariant Feature Transform (SIFT)
- Difference of Gaussian interest point detector
- Orientation Histogram feature representation

Search: "prius"

A Full Retrieval System

- Queries with homogeneous visual concepts
- "Mona-Lisa" and other comical variations

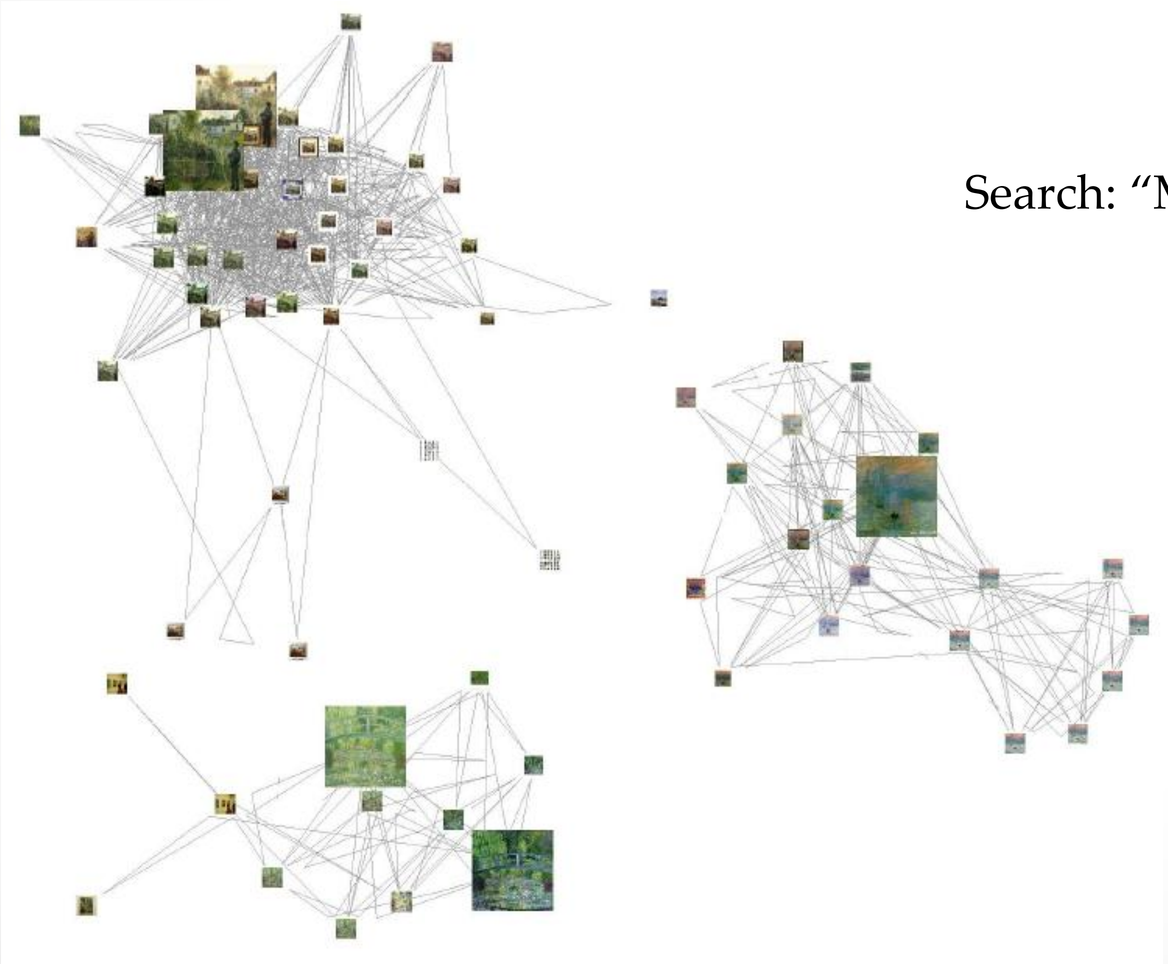




Search: “Mona Lisa”

A Full Retrieval System

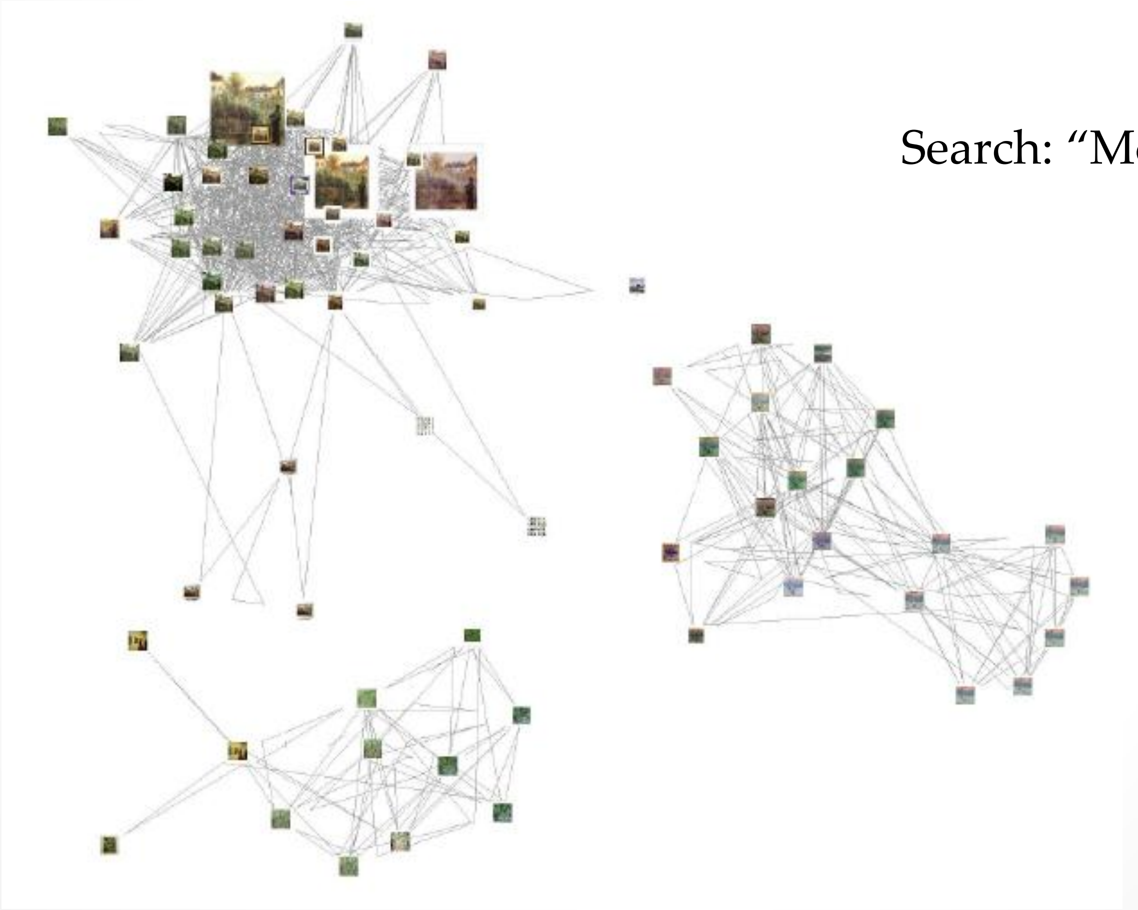
- Queries with heterogeneous visual concepts
- “Apple” (computer and fruit), “Jaguar” (car and animal)



Search: “Monet Paintings”

Why Eigenvector Approach ?

- Alternative: Select high degree nodes in the graph
- Fails to identify the different distinctive visual concepts



Search: “Monet Paintings”

Experimental Results

- Existing Google Search results

- Neglect Sparse Graphs

- Challenge – “quantify the quality of image search results”

- User preference influenced by personal biases

- Comparing the quality of a set of images is difficult

- <u>Example</u>: Five relevant but mediocre images *versus* five mixed with great and bad results

- Two evaluation strategies –
  - Minimizing Irrelevant Images
  - Click Study

Minimizing Irrelevant Images

- Identification of Irrelevant images with 150 volunteers revealed

	ImageRank	Google
Among top 10 results	0.47	2.82
Among top 5 results	0.30	1.31
Among top 3 results	0.20	0.81

- Importance
  - Google Product Search
  - Mixed-Result-Type Search

- Overall Performance
  - ImageRank contains less irrelevant images than Google for 762 queries
  - In only 70 queries, Google performs better
  - In rest 202, both approaches tied

- Drawbacks
  - Inflated logo-score
  - Saved web-pages as images

Click Study

- User satisfaction not purely a function of relevance
- Diversity of images is an important factor
- Experiment: Applied ImageRank to Google's top-1000 images for top 130 queries
- Images received 17.5% more clicks than those in *default ranking*
- Extremely severe bias that favors the default ordering
- Function of the position of image in Google Search result besides relevance and quality

Conclusion

- An effective method to infer a graph in which the images could be embedded
- Outperformed Google ranking on the vast majority of queries tried
- Extremely important is the ability to reduce the number of irrelevant images
- Human-coded information recaptured by
  - Query dependent approach
  - Reliance on the intelligence of crowds
- Customizability of the similarity function

Questions ???