Leveraging Internet Data
IM2GPS: Estimating Geographic Information from a Single Image (by James Hays and Alexei Efros)

Adriana Kovashka
CS PhD Student

Where is this?

… and this?

Overview of IM2GPS

- Intuition
  - “What is it like?” vs. “What is it?”
- Data
  - 6 million geo-tagged images from Flickr
- Method
  - Represent images in 6 ways, compare
- Result
  - *Estimated* image location

Representations in IM2GPS

- Tiny Images
- Color histograms
- Texton histograms
- Line features
- Gist descriptor with color
- Geometric context

IM2GPS Results

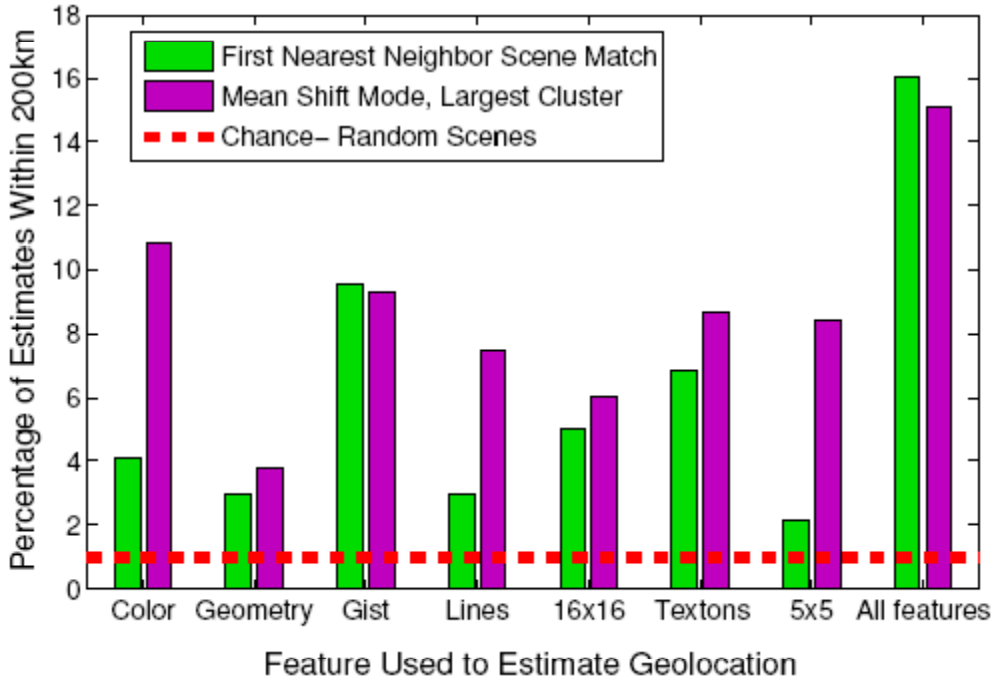


Figure 5. *Geolocation performance across features.* Percentage of test cases geolocated to within $200km$ for each feature. We compare geolocation by 1-NN vs. largest mean-shift mode.

Hays 2008

Note on the Task

- This is *not* scene categorization
  - Specific locations used
  - “Urban vs. natural” insufficient
  - Can think of current task as *place* recognition*

Demo Overview

- Data
  - 50096 images (incl. 237 test images)
  - 100 most populated cities in the world
- Representations
  - Gist, color, Tiny Images
- Comparison
  - K-nn

Procedure

- Use code by Hays to query/download Flickr images
  - about 3 days
- Download, modify, run Gist code
  - about 30 hours
- Test
  - about 6 hours for 7000 images
  - 10 min for 237 test images

Representations

- Gist (512 dim)
  - Used Torralba's scene recognition code
- Color (32 dim)
  - Computed histograms in L*a*b* color space
  - 4 bins for L, 14 for a and b
- Tiny Images (768 dim)
  - Resized images to 16x16x3
  - Vectors of color pixels

Comparison Methods

- Method One
  - Sim(x, y) = inner product between concatenation of three representations of x and y
- Method Two*
  - Sim(x, y) = $\exp(-dist_A/\sigma_A)*\exp(-dist_B/\sigma_B)*\exp(-dist_C/\sigma_C)$
  - $dist_A$ = Euclidian distance between representations A of x and y
  - σ_A = mean of distances for representation A

Note on the Computation of σ

- Current computation
  - X – matrix of *n*-dim features for all *m* images
  - Subtract mean(X) from all rows of X
  - Square result
  - Sum rows
  - Take square roots of sums
  - Take mean of resulting column
- Better computation
  - Average of Euclidian distance between i and j for each pair of images (i, j)
  - Computationally very expensive

Dataset

- Queried for 104 city tags
- Negative tags to remove duplicates, noise
- Downloaded images uploaded over 2 weeks
- 50096 images from Flickr (237 test)
  - 6M in IM2GPS (more tags, time)
- Disproportionate image set sizes per city!

'Abidjan' [0]
'Ahmedabad' [3]
'Alexandria' [152]
'Ankara' [10]
'Athens' [213]
'Atlanta' [843]
'Baghdad' [3]
'Bandung' [114]
'Bangalore' [477]
'Bangkok' [428]
'Barcelona' [2221]
'Beijing' [658]
'BeloHorizonte' [3]
'Berlin' [1655]
'Bogota' [404]
'Bombay' [16]
'Boston' [1631]
'Brasilia' [97]
'BuenosAires' [132]
'Busan' [0]
'Cairo' [107]
'Calcutta' [4]
'Chengdu' [225]
'Chennai' [114]
'Chicago' [2796]
'Chittagong' [0]
'Chongqing' [37]
'Dallas' [459]
'Delhi' [169]
'Detroit' [263]
'Dhaka' [55]
'Dongguan' [0]
'Guadalajara' [71]
'Guangzhou' [68]
'Guiyang' [0]
'Hanoi' [158]
'Harbin' [76]
'HoChiMinhCity' [9]
'HongKong' [835]
'Houston' [461]
'Hyderabad' [19]
'Istanbul' [681]
'Jakarta' [50]
'Johannesburg' [300]
'Karachi' [9]
'Khartoum' [6]
'Kinshasa' [0]
'Kolkata' [91]
'KualaLumpur' [56]
'Lagos' [25]
'Lahore' [8]
'Lima' [97]
'London' [2891]
'LosAngeles' [1442]
'Madras' [1]
'Madrid' [1822]
'Manila' [230]
'Medellin' [0]
'Melbourne' [529]
'MexicoCity' [59]
'Miami' [1280]
'Milan' [362]
'Monterrey' [26]
'Montreal' [0]
'Moscow' [291]
'Mumbai' [270]
'NYC' [2383]
'Nagoya' [23]
'Nanjing' [17]
'NewYorkCity' [483]
'Osaka' [222]
'Paris' [3052]
'Philadelphia' [883]
'Phoenix' [504]
'PortoAlegre' [69]
'Pune' [5]
'Pyongyang' [13]
'Recife' [221]
'RiodeJaneiro' [1135]
'Riverside' [215]
'Riyadh' [1]
'Rome' [1328]
'Ruhr' [53]
'Saigon' [252]
'SaintPetersburg' [44]
'Salvador' [867]
'SanFrancisco' [2204]
'Santiago' [365]
'SaoPaulo' [229]
'Seoul' [364]
'Shanghai' [118]
'Shenyang' [0]
'Shenzhen' [12]
'Singapore' [1118]
'Surat' [0]
'Sydney' [1541]
'Taipei' [546]
'Tehran' [19]
'Tianjin' [8]
'Tokyo' [1992]
'Toronto' [2009]
'WashingtonDC' [2031]
'Wuhan' [18]
'Yangon' [3]

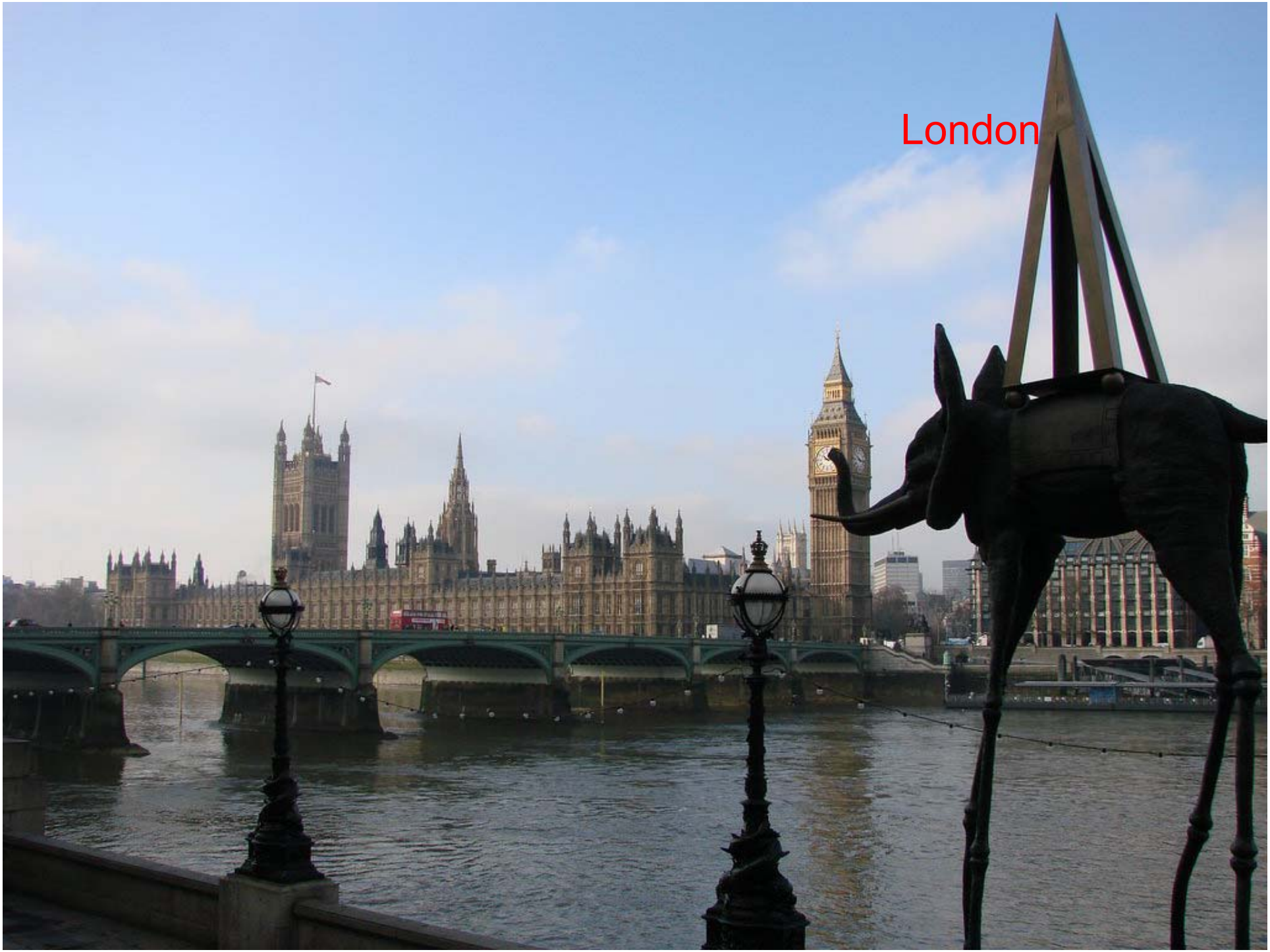

Bangalore


Boston
DAVIS AVE.


Boston


Cairo


Istanbul


London


London
POST
NO
BILLS

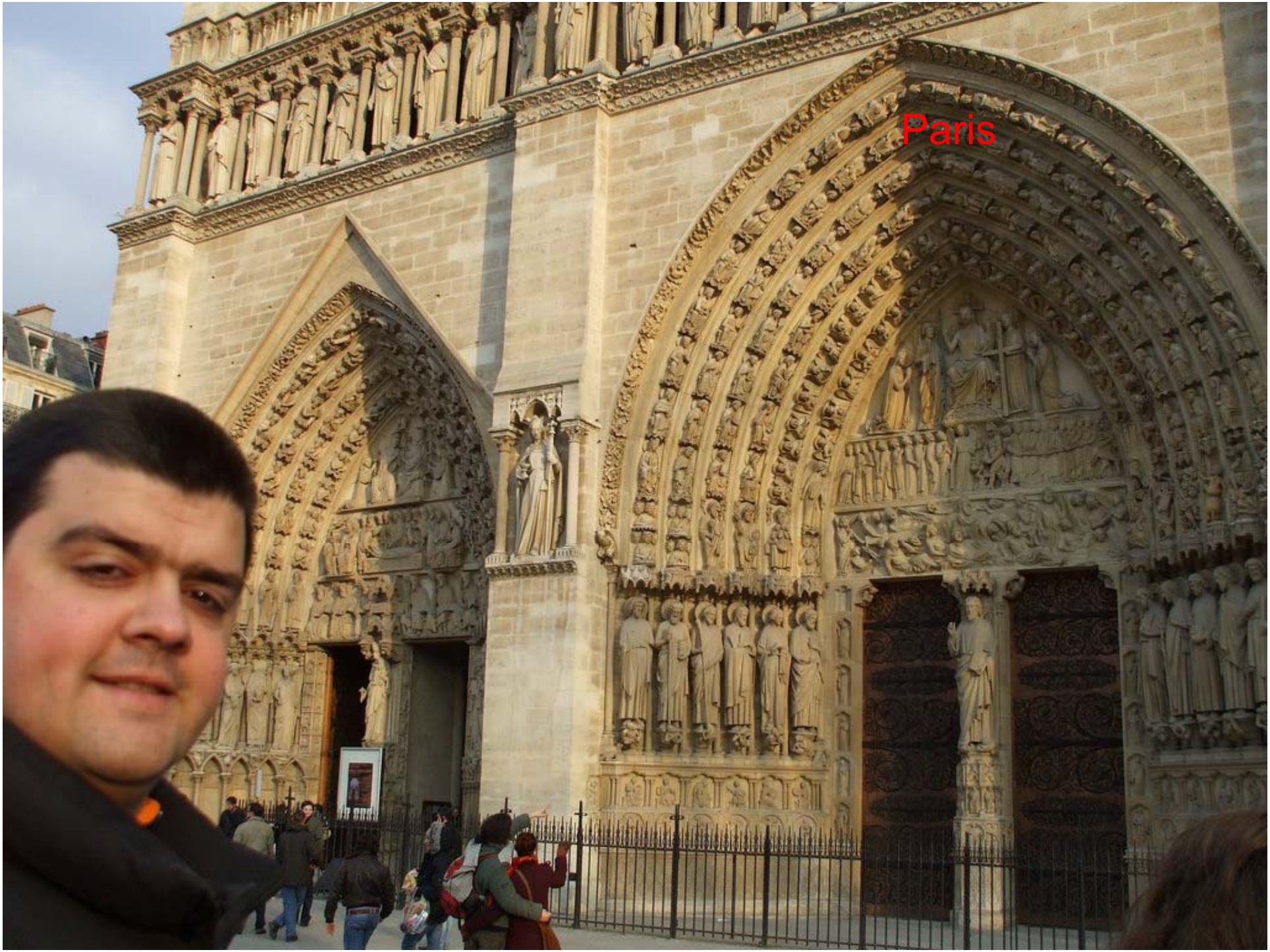

Los Angeles
UNION BANK OF CALIFORNIA


Madrid
TABERNA
Madrid
Desde
1908
LA
DOLORES
LICORES


Milan
Volkswagen. Das Auto.


Moscow

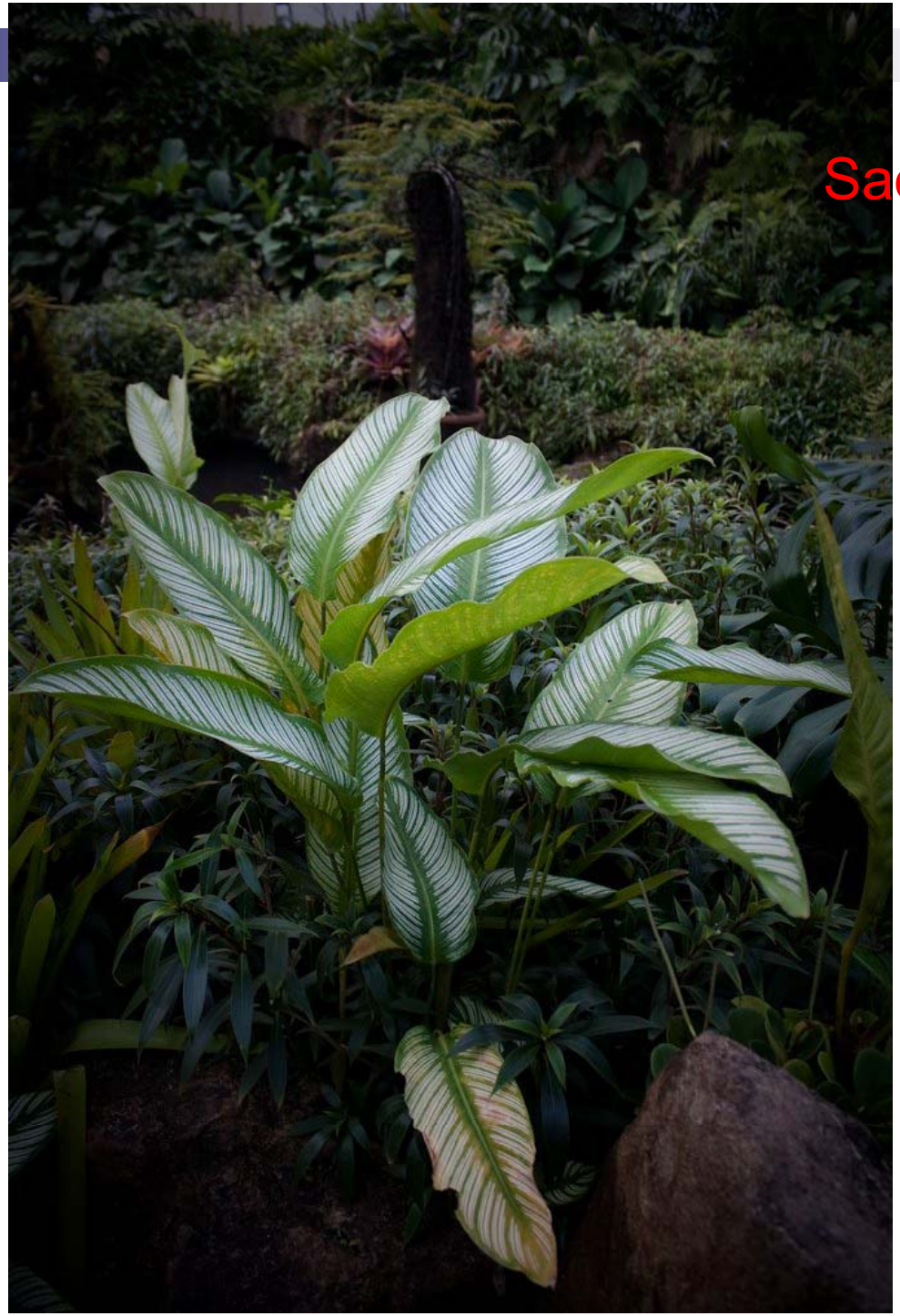

Mumbai


Paris


Rome


San Francisco

San Francisco



Sao Paolo


Tokyo

Tokyo


Query 1 - Greece

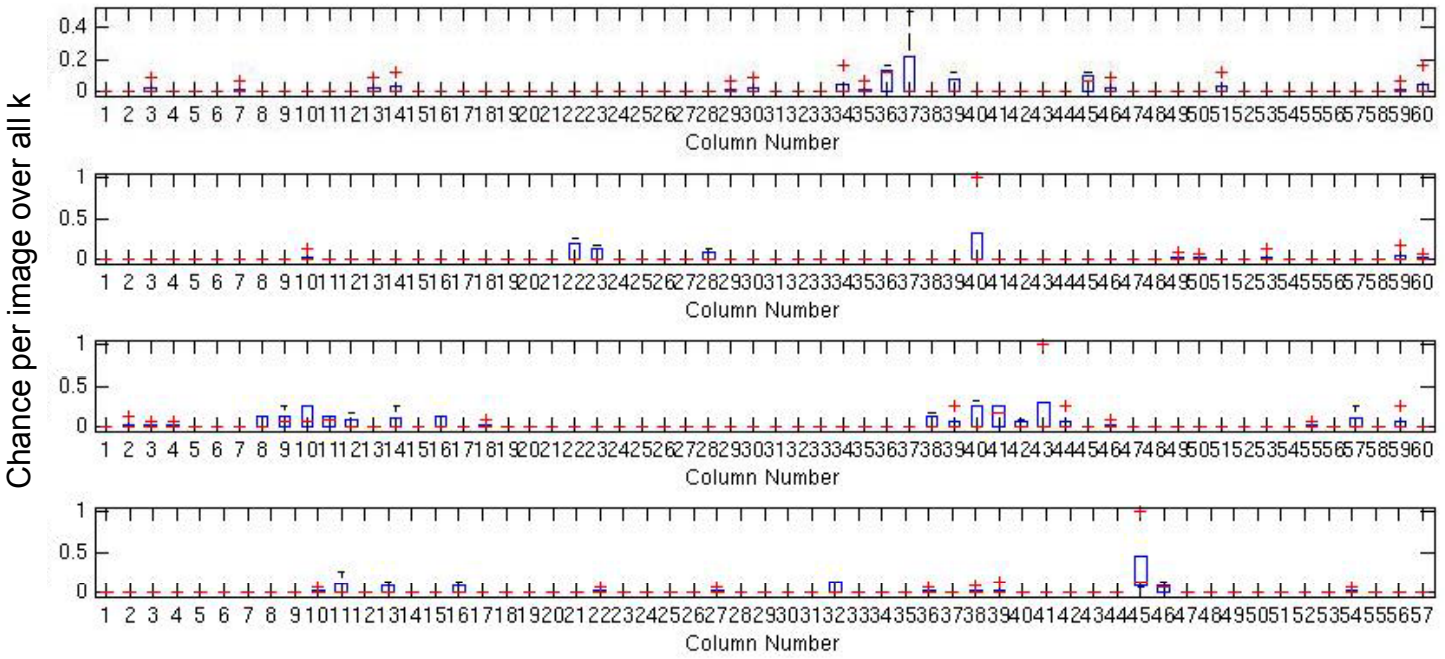

Query 2 - Arizona


Query 3 - Switzerland

Overview of Results

- Evaluation
  - Percentage of correct classifications
  - Percentage of top *m* neighbors within *n* km of query image
  - Average distance of neighbors
- Tests
  - on 237 test images
  - on 7000 images from dataset

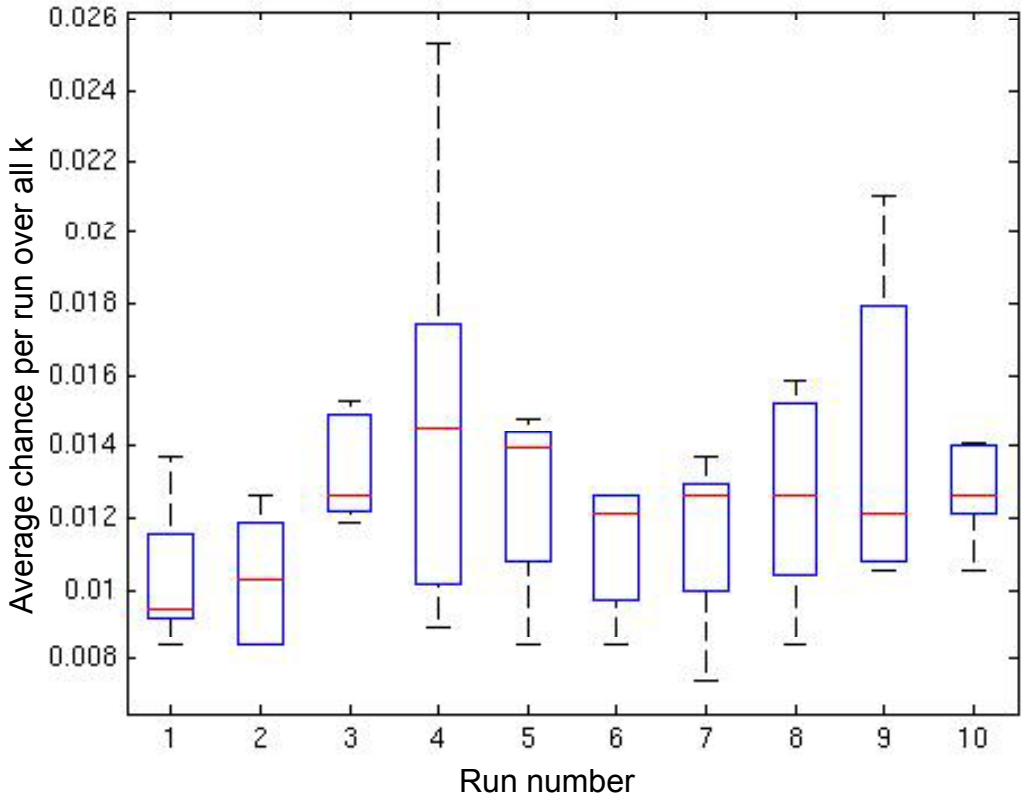
Chance for Test Images (200km)



Images 1 to 237

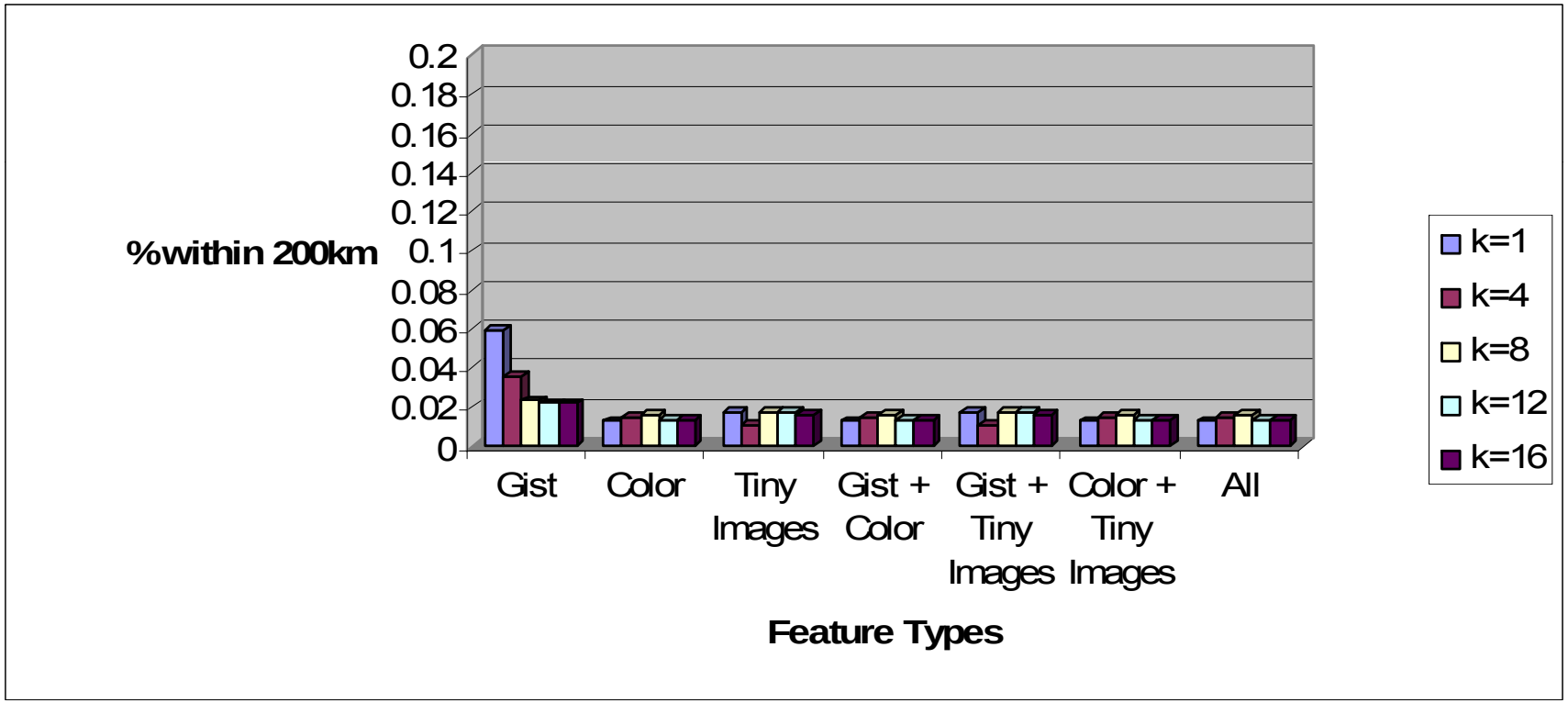
Chance is pretty low for this data.

Chance for Test Images (cont'd)



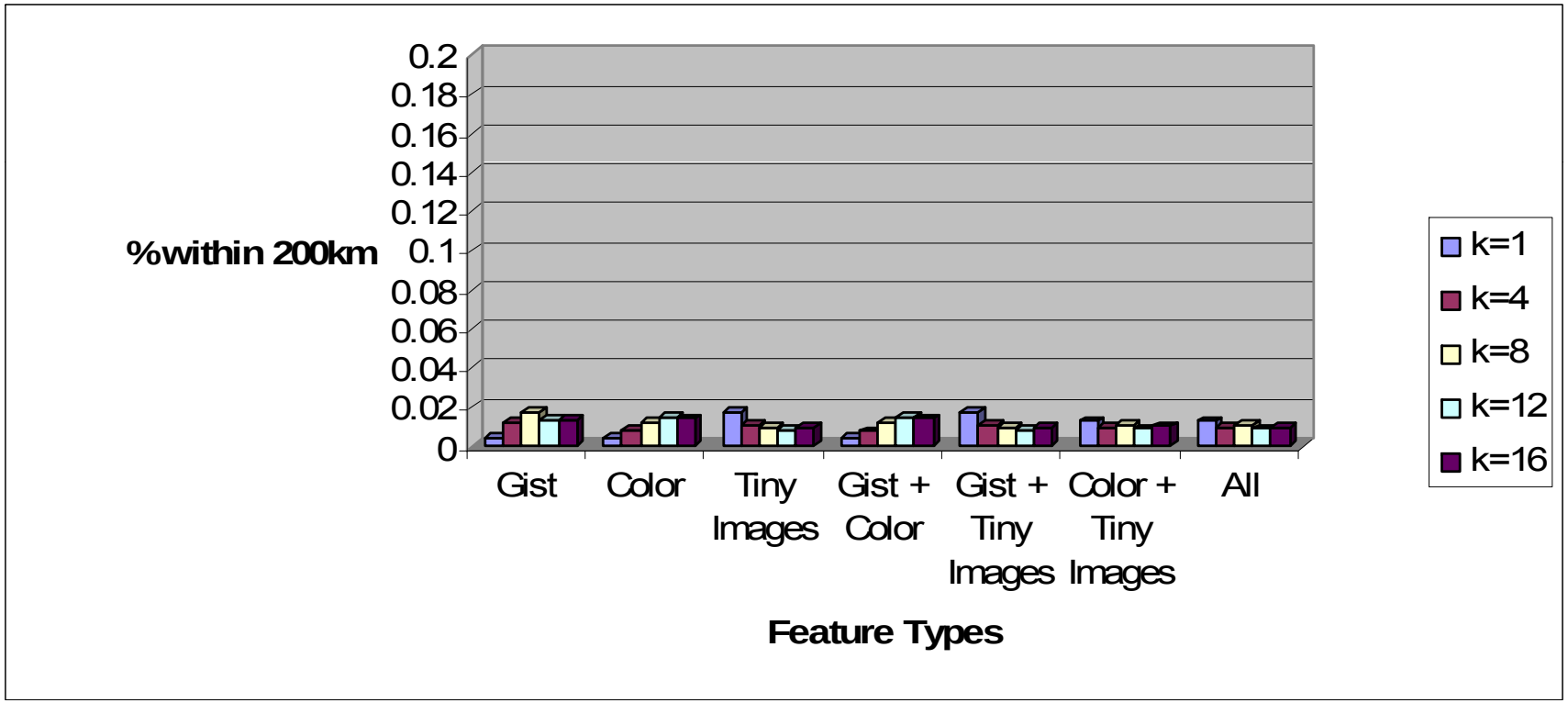
Chance is pretty low for this data.

Test Images, % w/in 200km, M1



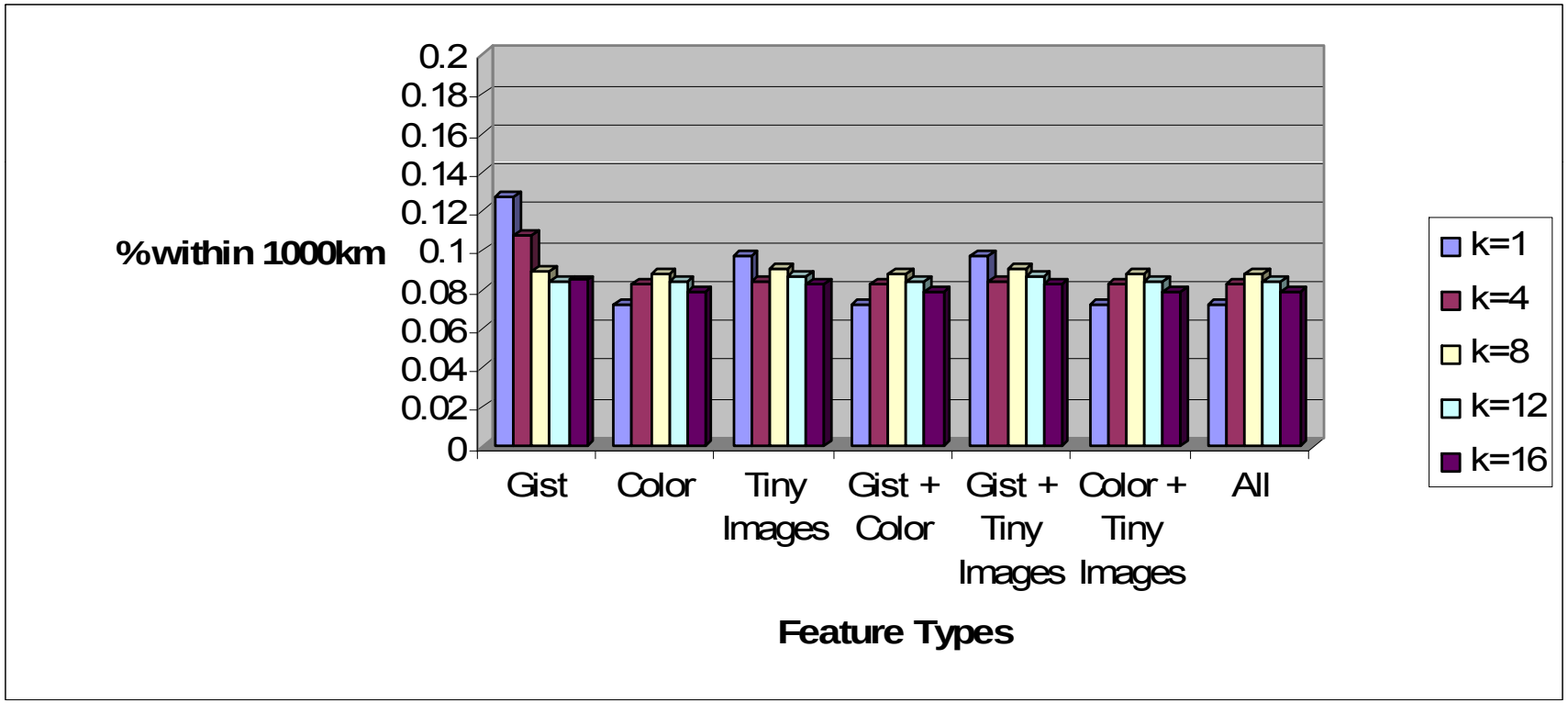
Gist seems to perform best with M1.

Test Images, % w/in 200km, M2



M2 works worse than M1.

Test Images, % w/in 1000km, M1



Results are naturally much better with larger distance allowed.

IM2GPS Results

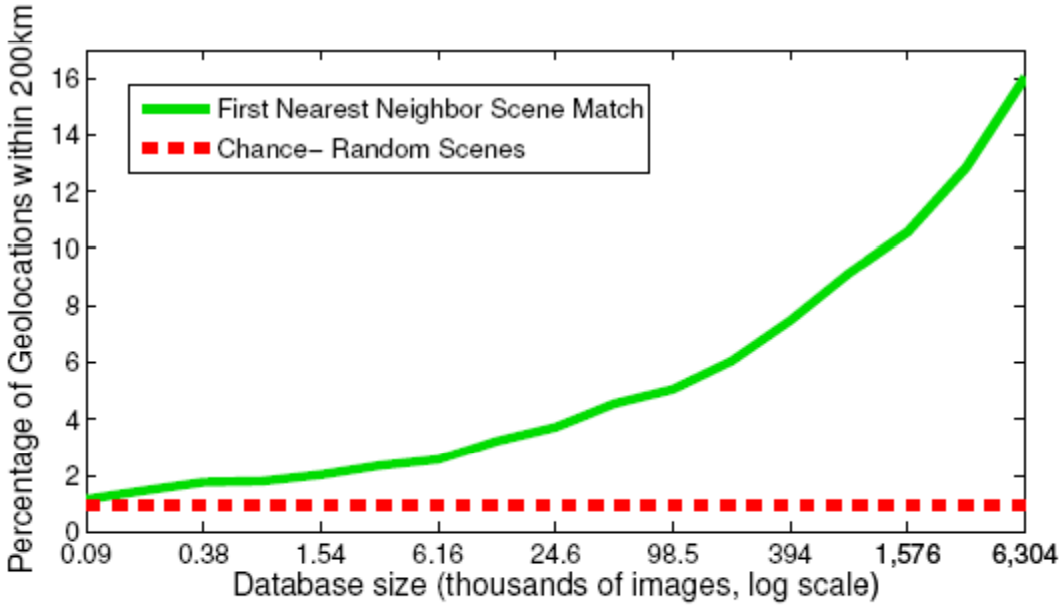
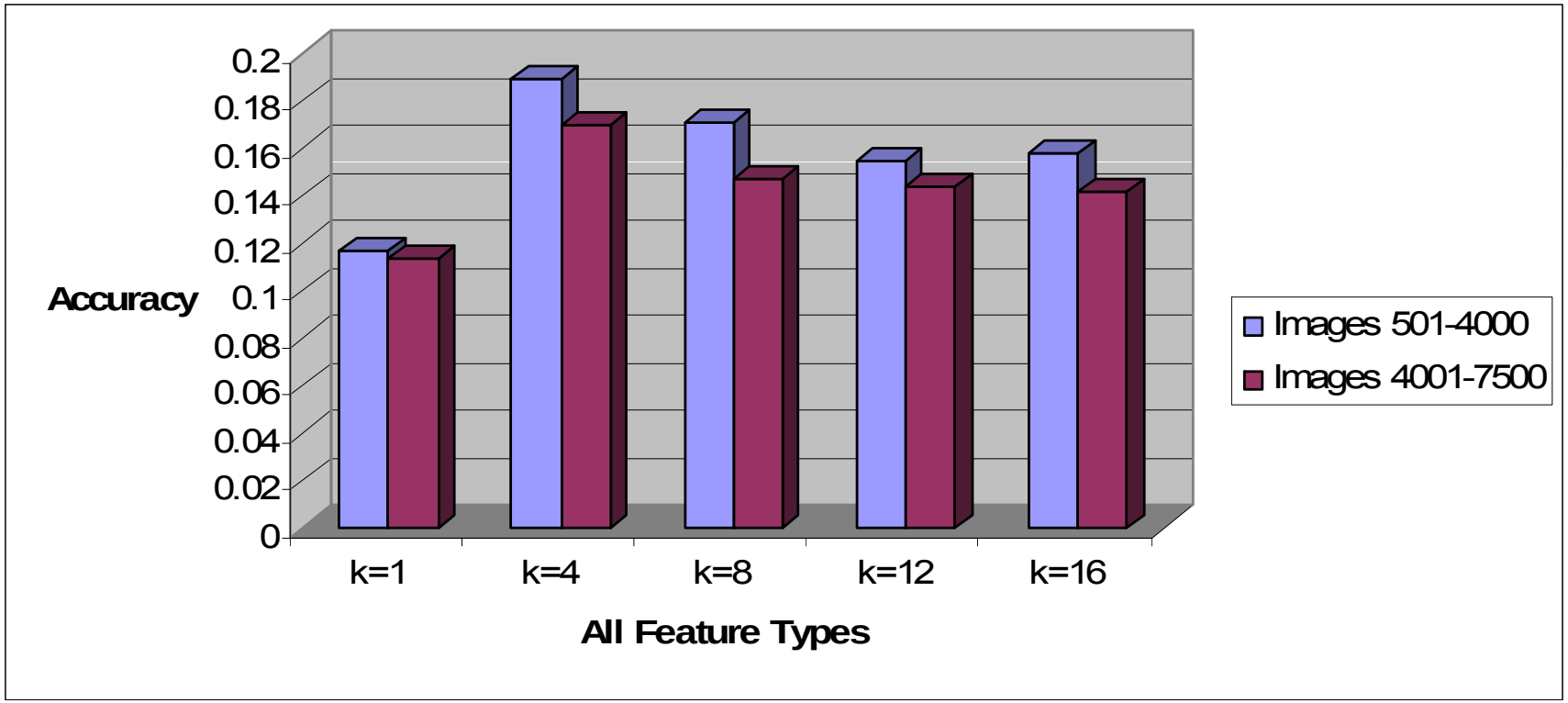


Figure 4. *Geolocation performance across database sizes.* Percentage of test set images that were correctly geolocated within 200*km* of ground truth as function of dataset size using 1-NN. As the database shrinks the performance converges to chance.

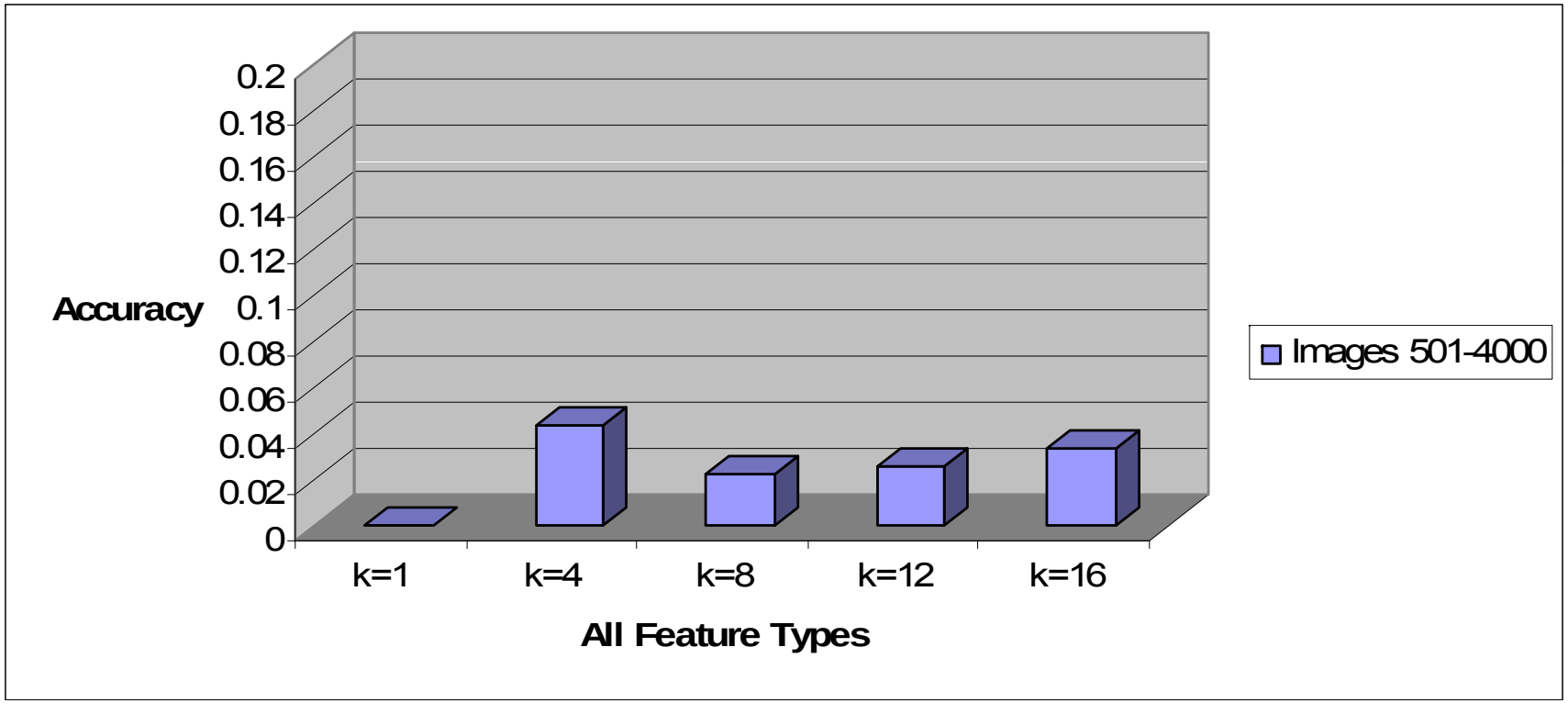
Hays 2008

Dataset, Accuracy, M1



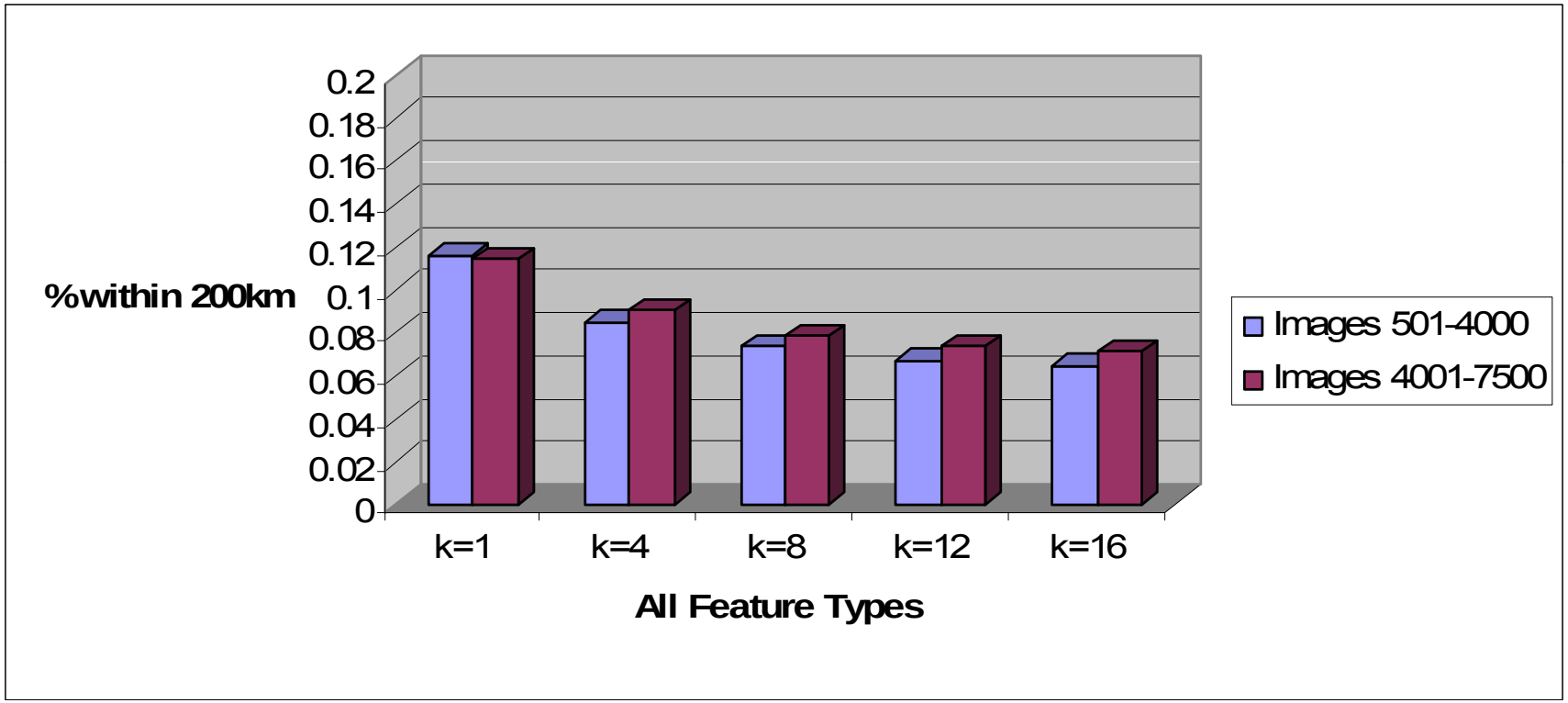
Results are much better with more test images.

Dataset, Accuracy, M2



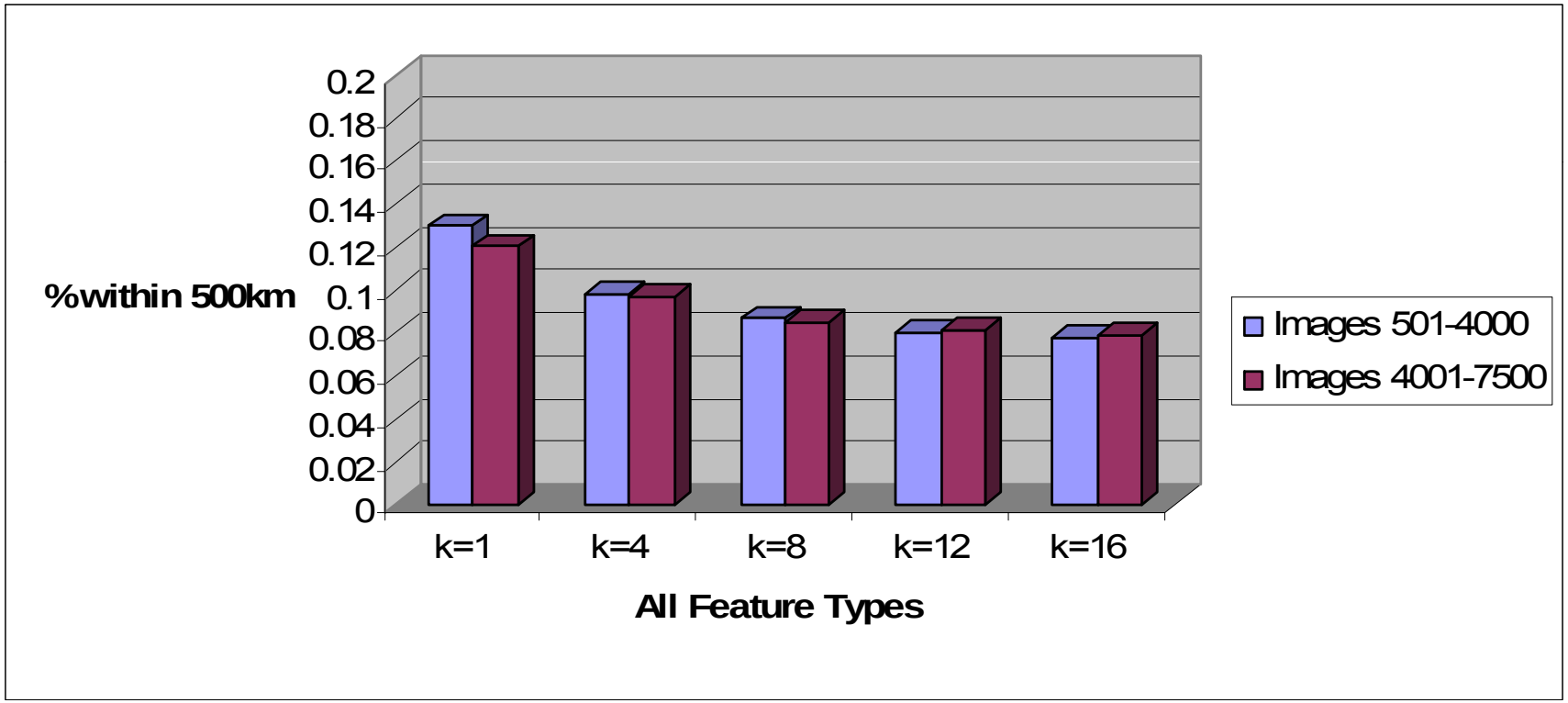
M2 performs worse than M1.

Dataset, % w/in 200km, M1



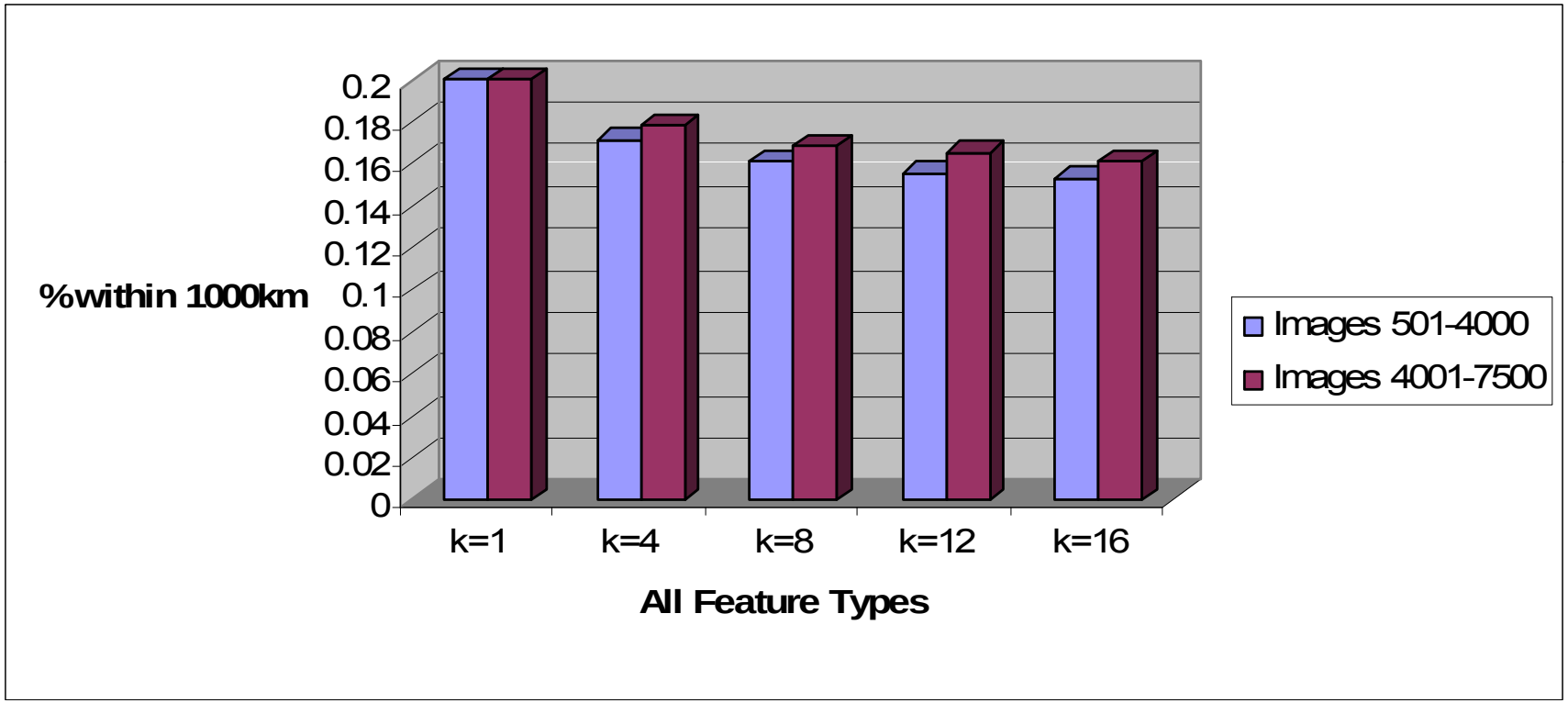
Again, with more test images, results are more similar to the authors'.

Dataset, % w/in 500km, M1



As expected, results improve when larger distance allowed.

Dataset, % w/in 1000km, M1



As expected, results improve when larger distance allowed.

Query Image (Argentina/Paraguay/Brazil)

Features: Tiny Images

Query Image (Barcelona)

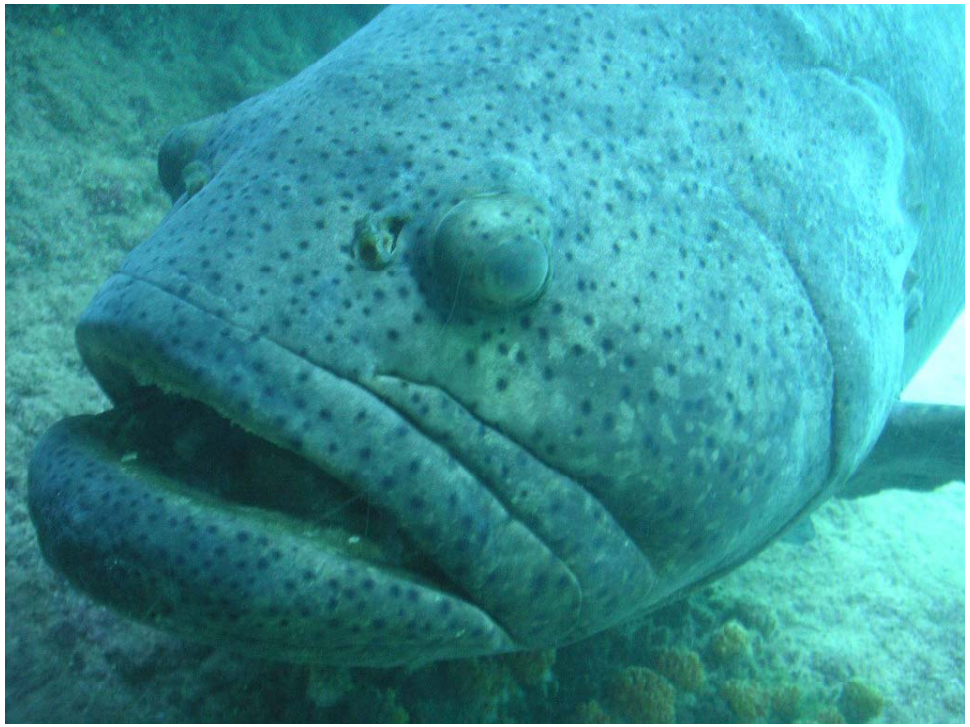
Chicago

Toronto

Features: Tiny Images

Query Image (Barcelona)

Features: Tiny Images

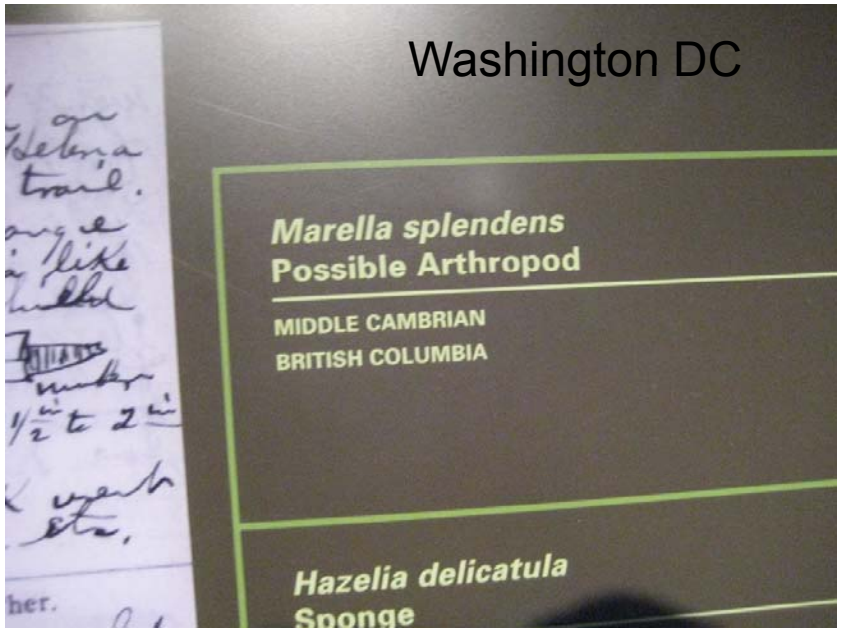


Query Image (Nassau, near Havana)

Features: Tiny Images

Query Image (Hyderabad)

Features: Tiny Images



Boston

Query Image (Athens)

Dallas

Rome

Features: Gist

Query Image (Guatemala)

Features: Gist

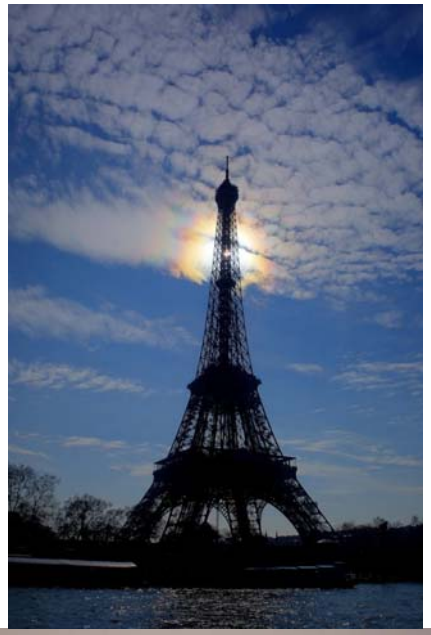
Query Image (Barcelona)

Features: Gist

Query Image (Aruba)

Features: Gist

Query Image (Florida)

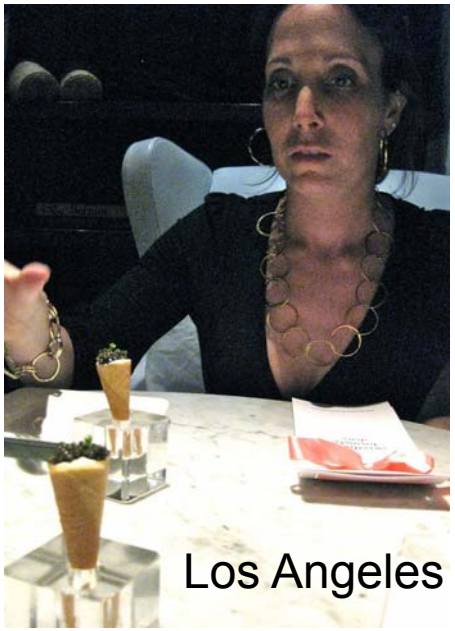


Paris

Moscow

Features: Gist

Query Image (Iceland)



Los Angeles

Melbourne

Features: Gist

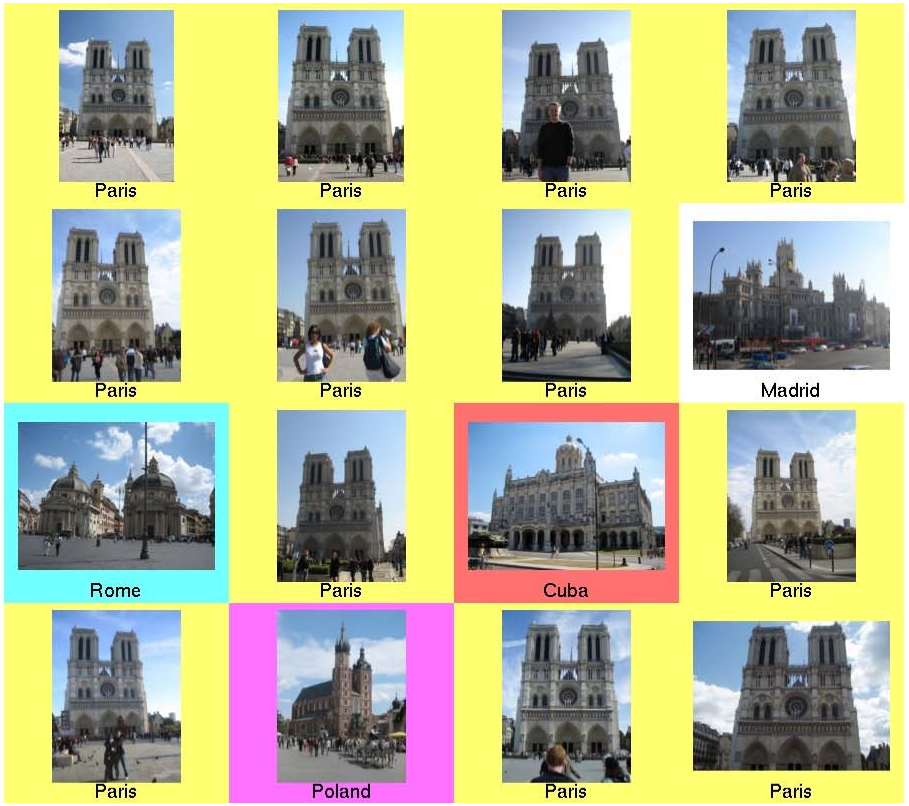
Query Image (Germany)

Toronto

Toronto

Features: Color

Hays 2008


Paris
Paris
Paris
Paris
Paris
Paris
Paris
Madrid
Rome
Paris
Cuba
Paris
Paris
Poland
Paris
Paris


Japan


Netherlands


Germany


Berlin


London


Moscow


London


Iran


California


London


Poland


Hungary

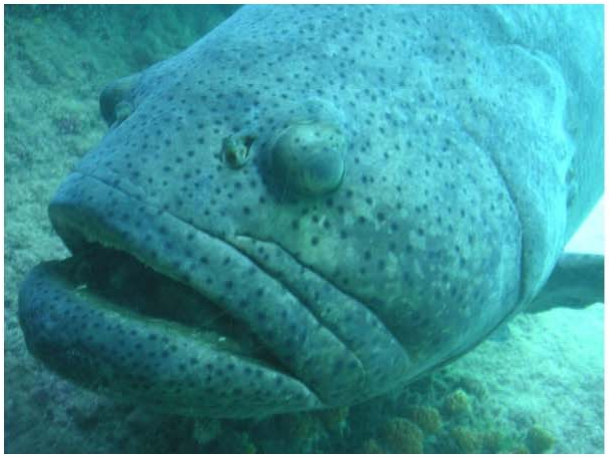

Rome

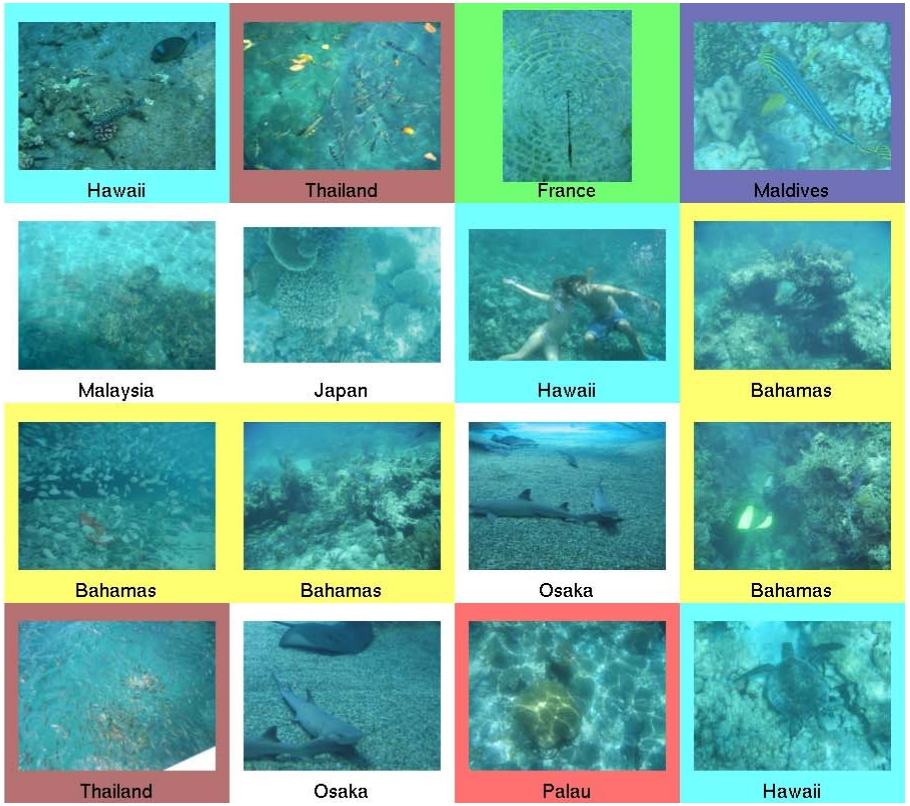

London


London


London

Hays 2008




Hawaii
Thailand
France
Maldives
Malaysia
Japan
Hawaii
Bahamas
Bahamas
Bahamas
Osaka
Bahamas
Thailand
Osaka
Palau
Hawaii

Hays 2008


Japan


wales


Japan


Japan


Malaysia


england


SouthAfrica


Florida


scotland


Japan


Japan


Mexico2


england


NewYork


Vermont


Austria

Hays 2008


Berlin


California


Monaco


Taipei


Greece


London


Barcelona


France


Colorado


Paris


India


Washington


Massachusetts


LosAngeles


Oregon


Mumbai

Hays 2008

Observations

- The image set is rather difficult
- Some suggestions are useful in various ways, some are very bad
- Scaling might improve results with a differently set σ
- This approach requires an enormous dataset to work well!

Discussion

- In what ways are the returned suggestions useful?
- Can we say the dataset is “noisy”?
- How can this method be improved?

References and Links

- J. Hays and A. Efros. IM2GPS: Estimating Geographic Information from a Single Image. CVPR 2008. http://graphics.cs.cmu.edu/projects/im2gps/
- A. Torralba, R. Fergus, and W. Freeman. 80 Million Tiny Images: a Large Dataset for Non-Parametric Object and Scene Recognition. PAMI 2008. http://people.csail.mit.edu/torralba/tinyimages/
- A. Oliva and A. Torralba. Modeling the Shape of the Scene: a Holistic Representation of the Spatial Envelope. IJCV 2001. http://people.csail.mit.edu/torralba/code/spatialenvelope/

References and Links (cont'd)

- P. Getreuer. Color Space Converter. Matlab Central. http://www.mathworks.com/matlabcentral/fileexchange/7744
- Distance Calculation. Meridian World Data. http://www.meridianworlddata.com/Distance-Calculation.asp
- Online Conversion – Unix time conversion. http://www.onlineconversion.com/unix_time.htm
- A. Mehrtash. demo links. http://users.ece.utexas.edu/~mehrtash/SceneRecognitionDemo/
- A. Kovashka. IM2GPS (Hays and Efros) Demo. http://www.cs.utexas.edu/~adriana/im2gps_demo.html