Active Query Sensing for Mobile Location Search

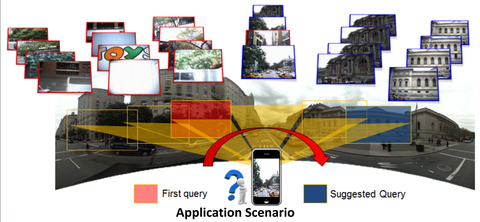
Felix X. Yu Rongrong Ji Shih-Fu Chang

Digital Video and Multimedia Laboratory, Columbia University, 10027, New York, United States

SCOTTSDALE ARIZONA ACM MULTIMEDIA 2011 USA

Abstract

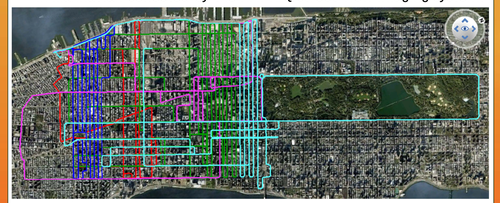
- When the first query fails to find the right target (up to 50% likelihood), how should the user form his/her search strategy in the subsequent interaction?
- We propose a novel *Active Query Sensing* system to suggest the best way for sensing the surrounding scenes while forming the **second** query for location search.



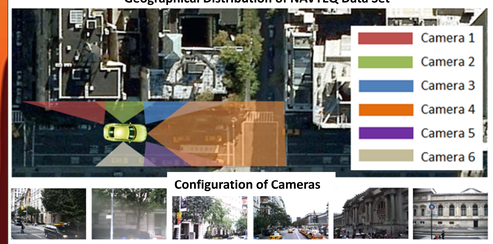
Application Scenario

NAVTEQ 0.3M NYC Data Set

- The data set consists of about 300,000 images of 50,000 locations in Manhattan collected by the NAVTEQ street view imaging system.



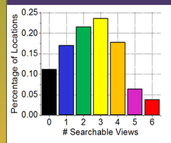
Geographical Distribution of NAVTEQ Data Set



Configuration of Cameras

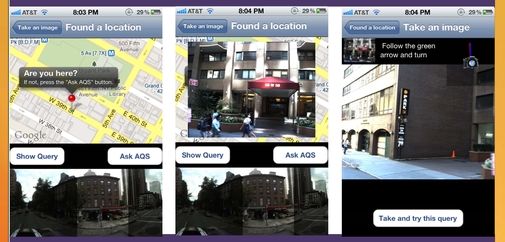
- To simulate the mobile location search scenarios, we manually cropped queries from Google Street View in 226 randomly chosen locations covered by the above mentioned routes in NYC.
- For each location, six query images are cropped from viewing angles similar to the view orientations used in the database. This results in 1,356 images with angles and ground truth locations tags.

Problem Justification



- Most locations are recognizable only with a subset of the views.
- Each location has a unique subset of preferred views for recognition.
- *Location Dependence*: different locations have different degrees of "difficulty".
- Locations of the same search difficulty do not significantly cluster together.
- There is no single dominant view that can successfully recognize all locations.

User Interface



Problem Formulation

Assumption: User is unlikely to take "junk image" with no hope for finding the true target. The first query, even unsuccessful, can be used as "probe" to narrow down the solution space.

Examples of "junk" images Examples of query images

- A subsequent query should maximize the discriminability (uncertainty reduction) over candidate locations, narrowed down by the query already been taken, by selecting the best query view.

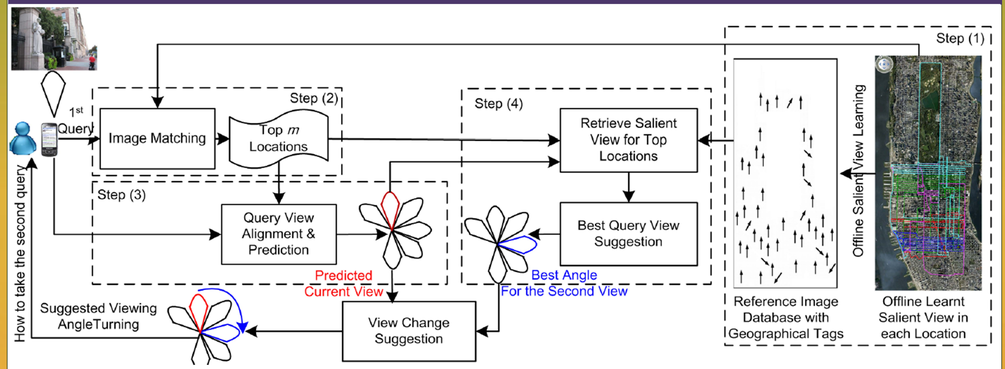
$$v_2 = \max_v (H(l|q_1) - H(l|q_1, v))$$

- Without showing mathematical details, we found the above process can be well approximated by

$$\max_v \sum_{l \in L_N} Saliency(l, v)$$

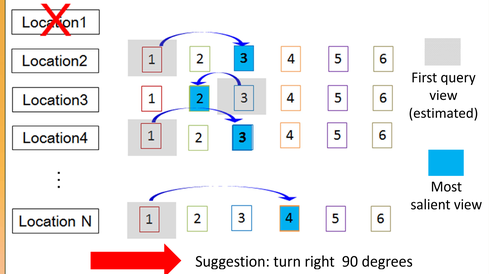
- L_N is the candidate locations narrowed down by the first query.

System Architecture



- Offline analysis: to discover the best query view(s) for each location indexed by the system.
- Online process: to estimate the most likely view of the first query (which has failed); to suggest optimal view change to the user.

Online Analysis



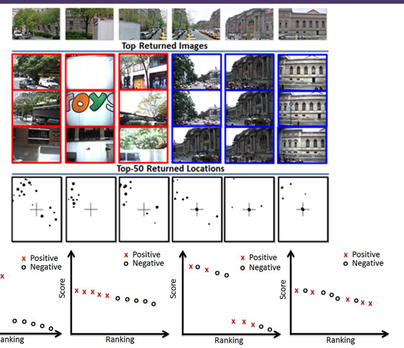
- There are two ways for estimating the first query view.

Train view classifiers offline

View alignment based on the image matching

- We use a majority voting mechanism to estimate the optimal view angle change and suggest the user turns to the most salient view.

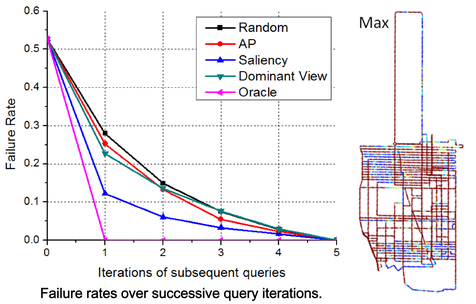
Offline Analysis



- An ideal score distribution is the one that has maximal separation between the scores of the positive results and those of the negative ones.
- Predict online search performance by using the reference image of each {location, view} to retrieve the same location:

$$Saliency = \frac{\sum_{i=1}^{N} \sum_{j=1}^{i} score(j) rel(j)/i}{\sum_{i=1}^{N} \sum_{j=1}^{i} score(j) \overline{rel}(j)/i}$$

Experiment



Failure rates over successive query iterations.

- We simulate the first query with a randomly chosen viewing angle.
- Only 47% of the random query images succeed, resulting in a 53% failure rate after the first query.
- We evaluate the performance of reducing the failure rates in subsequent queries by using different active query strategies.
- The performance gain achieved by the saliency based AQS scheme is quite impressive – 12% error rate after only one additional query.
- The offline measures of saliency can also be used to "grade" the searchability of each location, indicating the locations where the system performs more robustly.

Conclusions and Future Works

- We have developed a novel Active Query Sensing system that actively suggests the best query strategy if/when the first visual query fails.
- This the first effort in actively guiding users to achieve more satisfactory experience in using mobile visual search.
- We develop saliency measurement based on score distributions to predict the robustness of each query view and the search difficulty of each location.
- In the future, we will extend the idea to multi-view object search.

Which view will be the best query?

Acknowledgement

We thank NAVTEQ for providing the NYC image data set, Dr. Xin Chen and Dr. Jeff Bach for their generous help, and Tongtao Zhang for designing the mobile interface.