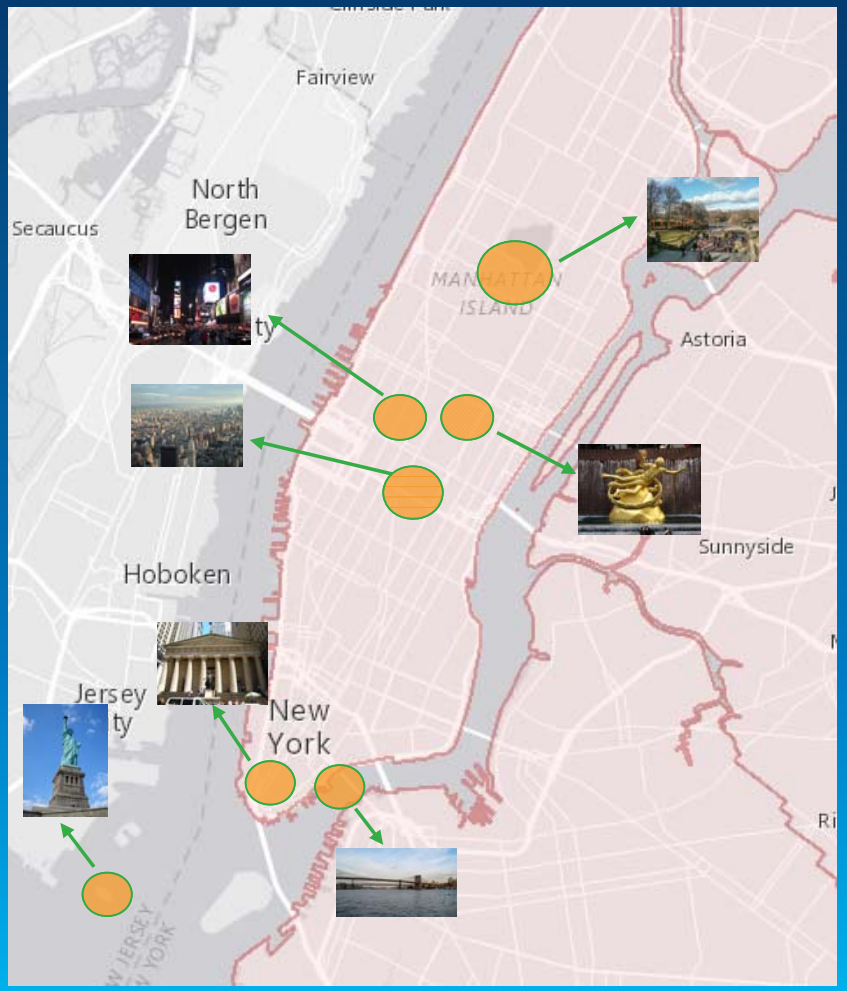
Extracting and Understanding Urban Areas of Interest Using Geotagged Photos

Yingjie Hu, Song Gao, Krzysztof Janowicz, Bailang Yu, Wenwen Li, Sathya Prasad

Outline

- Introduction: what is urban areas of interest (AOI)?
- Potential data sources to extract AOI
- A framework for extracting and understanding AOI
- Knowledge and insights derived from the extracted AOI
- Conclusions and future work

1. Introduction: what is urban areas of interest (AOI)?



1. Introduction: what is urban areas of interest (AOI)?

- Urban AOI are defined by people's perception
- Different people may have different opinions
  - The list of AOI
  - The boundaries of AOI

Can we identify urban AOI agreed by many people? How?

2. Potential data sources to extract AOI

- **Remote sensing images:**
  - **Commonly used in urban studies (e.g., detecting land cover types)**
  - **Unfortunately, remote sensing data don’t record the perception of people**

A RS image of Shanghai

A photo of Shanghai

2. Potential data sources to extract AOI

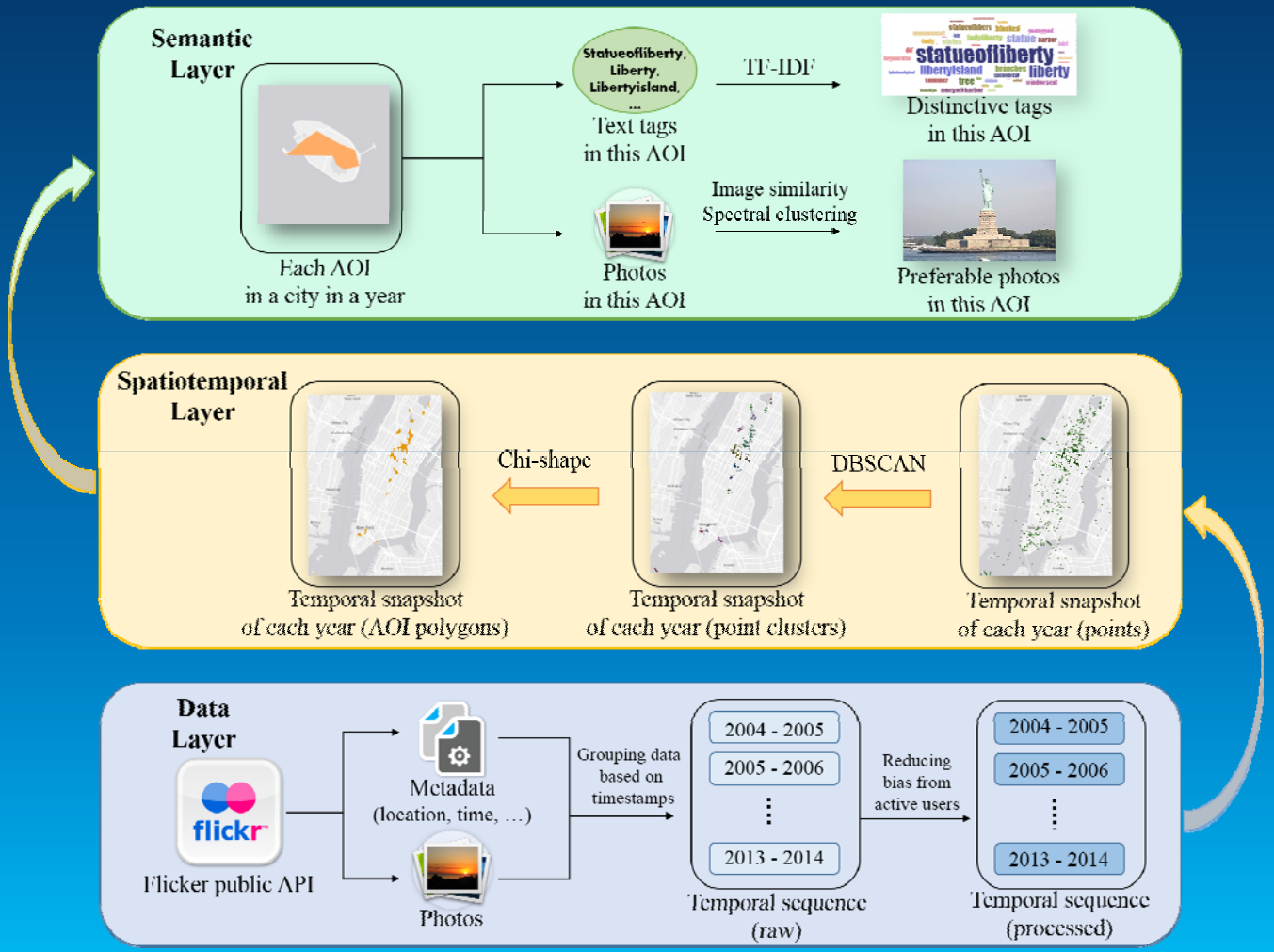
- **Human participant survey**
  - **Example: Dan Montello (2003): Where is Downtown Santa Barbara?**
  - **Requires a lot of time and human efforts**

Please tell us the areas you consider interesting in the city, and draw them on the map.

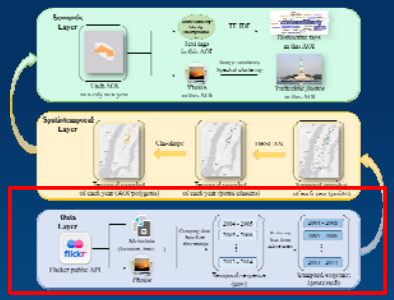
2. Potential data sources to extract AOI

- **Social media data**
  - **Provide records for people's interactions with the urban environment**
  - **Many social media data contain location information**
    - **Geotagged Tweets**
    - **Geotagged Flickr photos**
    - **Foursquare checkins**
    - **…**
  - **Can be retrieved from public APIs**

3. A framework for extracting and understanding AOI

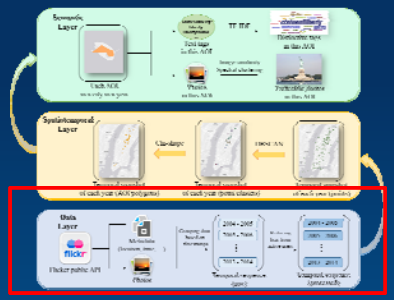


3.1 Layer 1: Data preprocessing



- **Why Flickr data?**
  - **Reflect locations people consider interesting**
  - **Cover a timespan of the past 10 years**
  - **Publicly available through APIs**
  - **Large number of users (around 100 million users)**

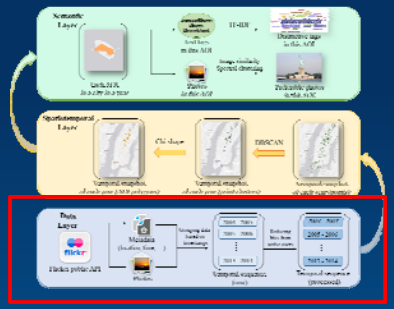
3.1 Layer 1: Data preprocessing



- **Cities: New York, London, Paris, Shanghai, Mumbai, Dubai**
- **Timespan: 2004 - 2014**
- **Method: Flickr public API**

City	# User	# photo
New York City	2,751	2,761,542
London	2,357	2,876,013
Paris	3,019	1,456,298
Shanghai	1,775	254,123
Mumbai	1,901	55,532
Dubai	2,176	89,457

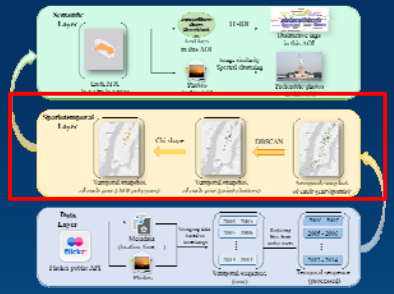
3.1 Layer 1: Data preprocessing



- Constructing temporal snapshots
  - One-year time window (10 snapshots in total)
  - Each snapshot contains Flickr data in one year (e.g., 2008 – 2009)

- Removing dominance effect from active users
  - A small number of active users contribute a huge number of data
  - Most users contribute a few data records
  - Keep only one record for each user within a neighborhood radius

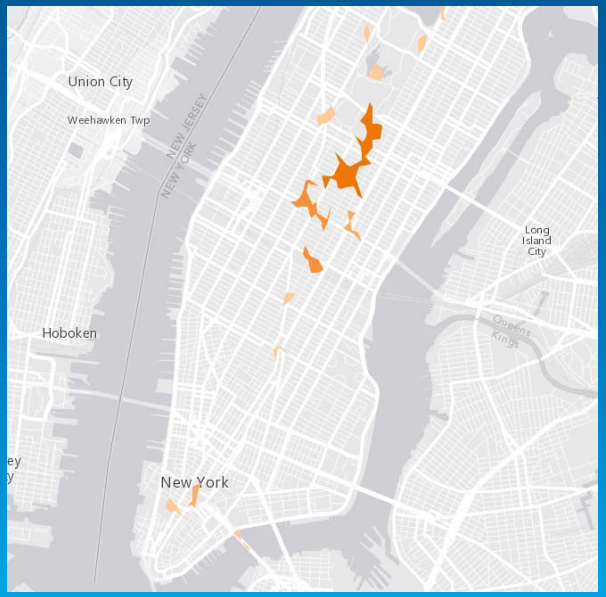
3.2 Layer 2: clustering and area construction



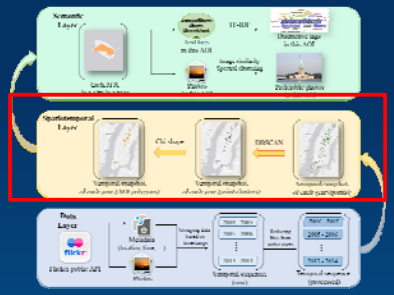
- To identify the areas which are interesting to many people
- We abstract it into a clustering problem

Clustering

Region construction



3.2 Layer 2: clustering and area construction

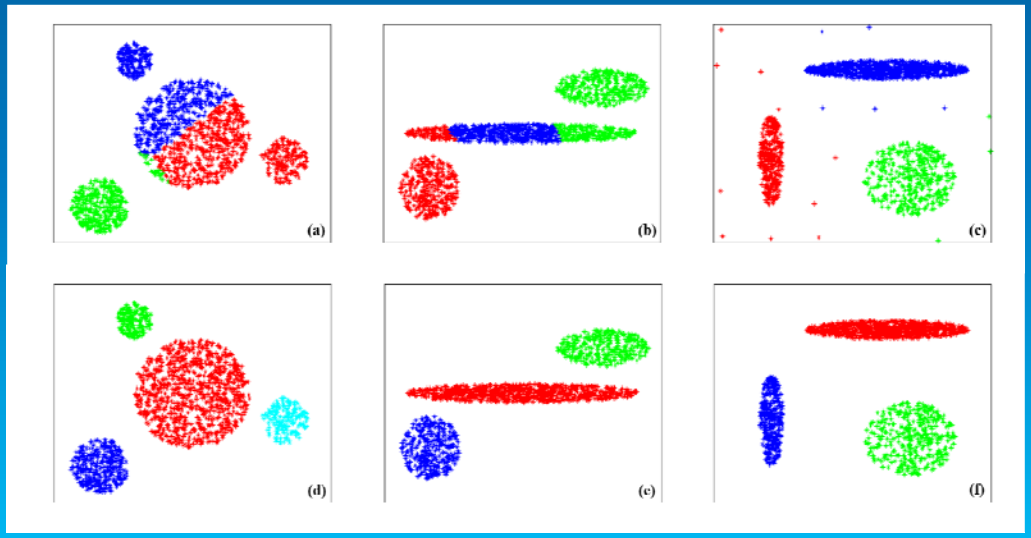


- Clustering method: DBSCAN (Density Based Spatial Clustering of Applications with Noise)

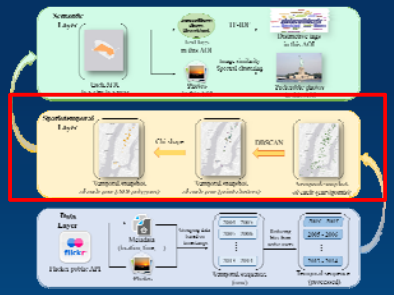
- Advantages of DBSCAN in extracting AOI
  - Doesn't require a pre-determined number of clusters
  - Clusters can be any arbitrary shape
  - Robust to noise

- K-means

- DBSCAN

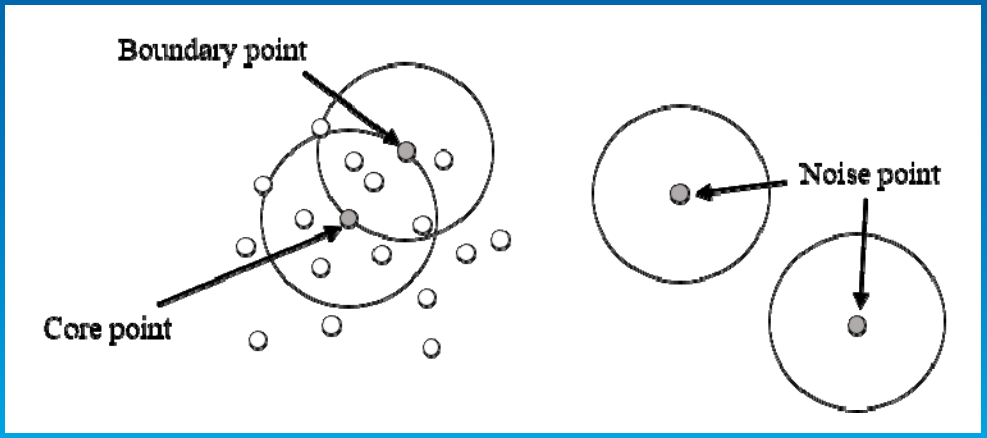


3.2 Layer 2: clustering and area construction

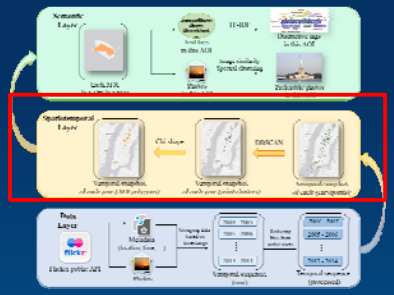


- DBSCAN requires two parameters: search radius (Eps) and minimum number of points (MinPts)

- Example: Eps is the circle radius, and MinPts is the minimum number of points in this radius (in this case: 8)



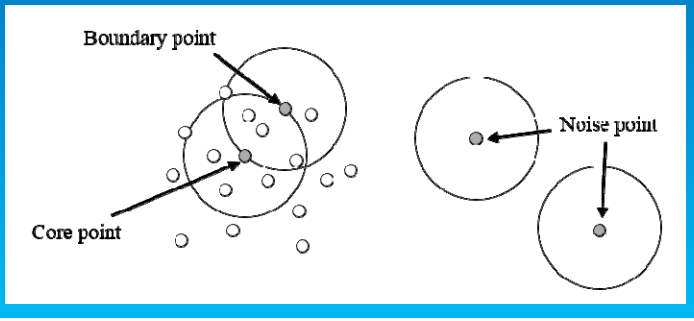
3.2 Layer 2: clustering and area construction



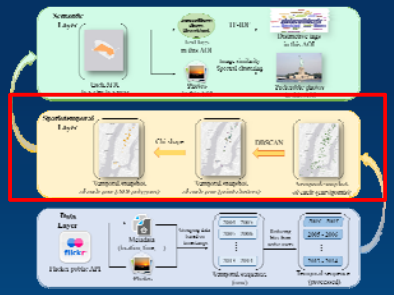
- Eps and MinPts together define the minimum density threshold, and clusters are formed at locations where the density is larger than this threshold

- Larger Eps will produce clusters in large geographic scale, and small Eps will produce clusters in small scale
  - E.g., 200 meters could detect neighborhood-level clusters
  - E.g., 5 kilometers could detect city-level clusters

- MinPts determines the significance of the derived clusters
  - High MinPts requires more people to agree
  - Low MinPts generate more but less significant clusters

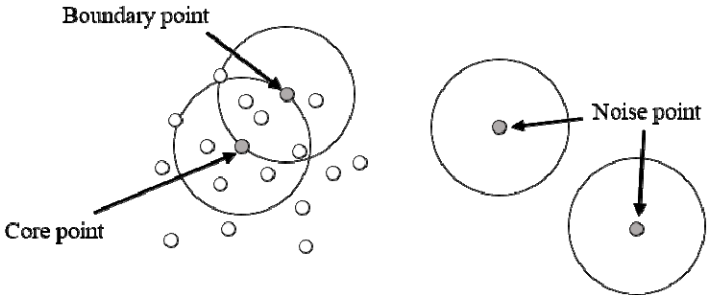


3.2 Layer 2: clustering and area construction

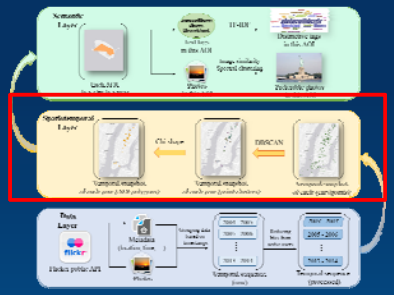


- Selecting Eps and MinPts by iterative experiments
  - Two cities: NYC and Shanghai
  - Iterate Eps from 100 to 500 meters
  - Iterate MinPts from 1% to 5%
  - 7 human participants to evaluate the result
  - Select 200 meters for Eps and 2% for MinPts

- Two parameters together determine the meaning of AOI
  - In this study, AOI are city regions which have been visited by at least 2% of different people given a radius of 200 meters.
  - In other studies, AOI can be defined differently using a different set of parameters



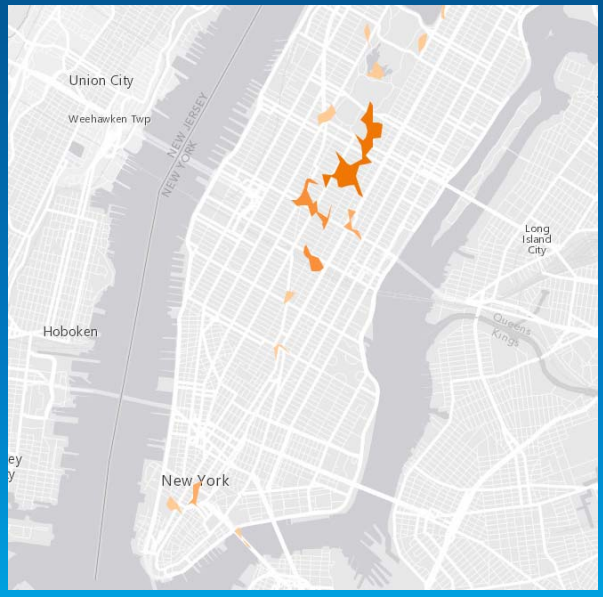
3.2 Layer 2: clustering and area construction



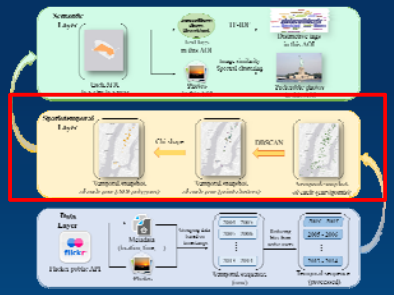
- **Constructing polygons from point clusters**

Clustering

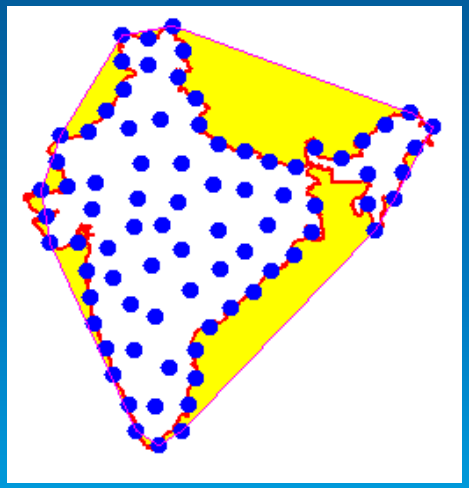
Region construction



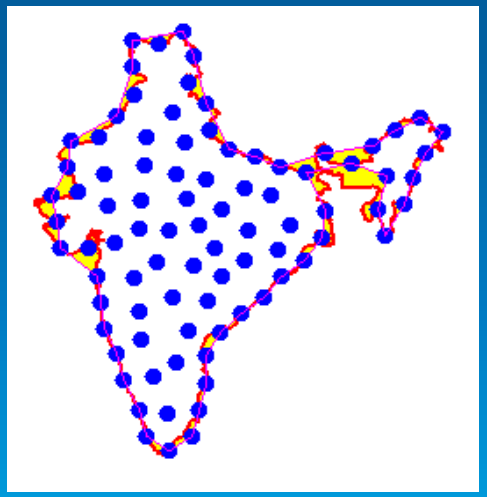
3.2 Layer 2: clustering and area construction



- **Constructing polygons from point clusters**
  - **Convex hull: provides the smallest convex polygon, but often contains empty regions**
  - **Chi-shape (concave hull): more accurate delineation of the shape; the generated polygon is not convex**

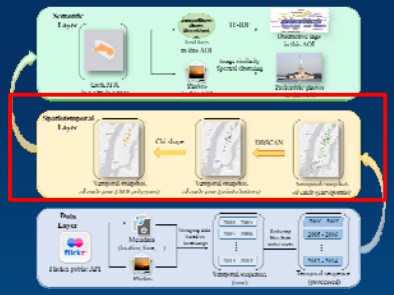


Convex hull



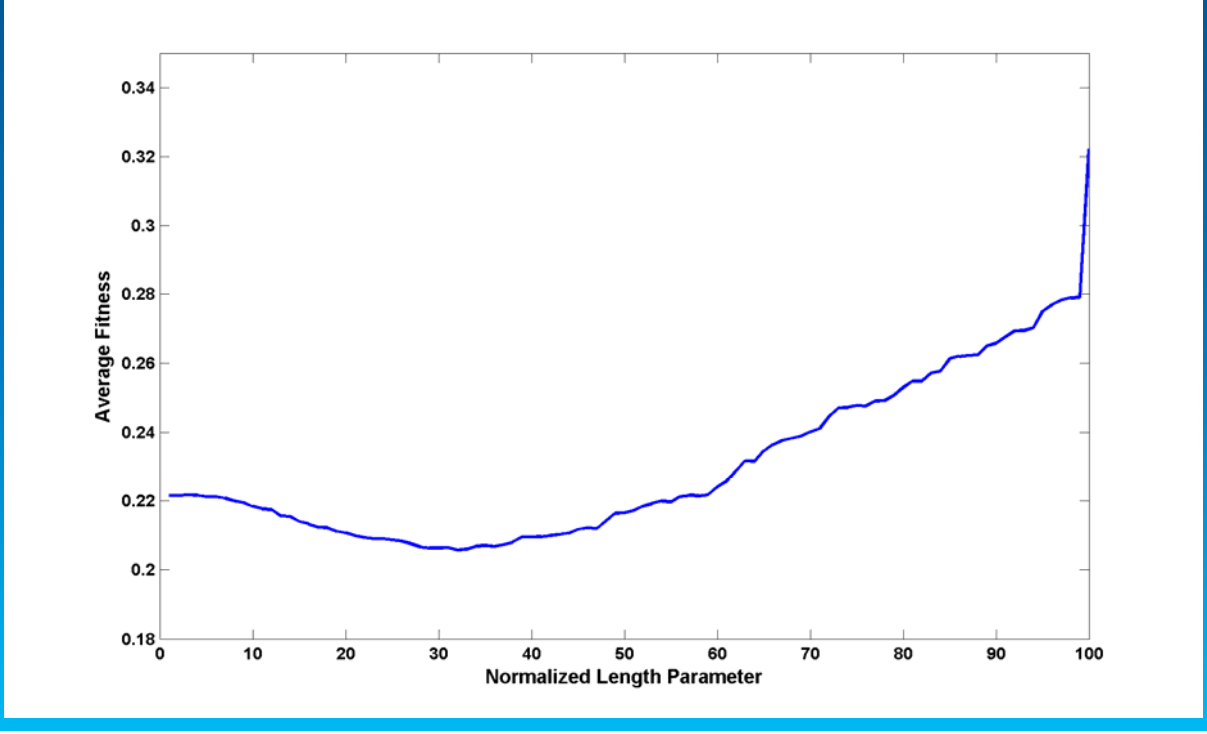
Concave hull

3.2 Layer 2: clustering and area construction



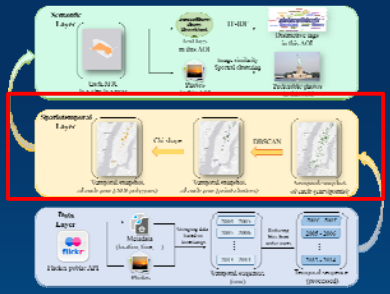
- **Chi-shape algorithm requires a parameter of lambda in [1, 100]**
  - **Using fitness function to determine lambda**

$$\Phi(P, D) = Emptiness(P, D) + C * Complexity(P)$$

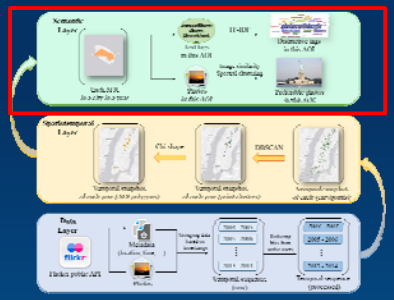


3.2 Layer 2: clustering and area construction

- **Polygons generated using different lambdas**

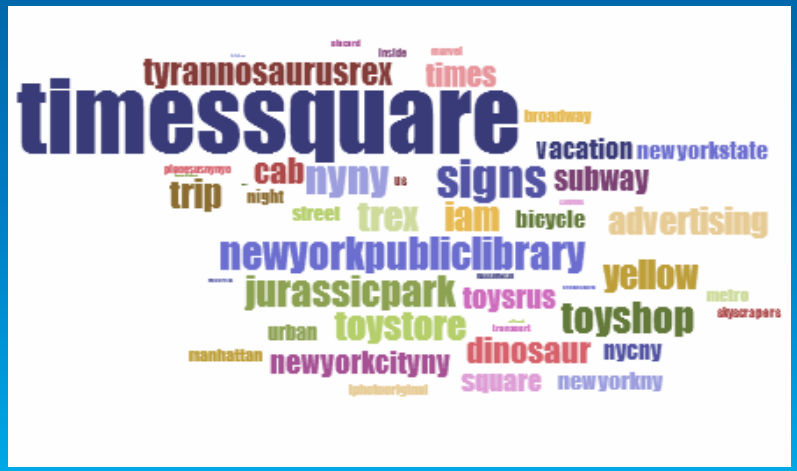


3.3 Layer 3: understanding AOI



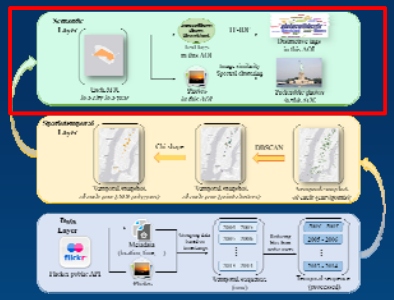
- What are the major topics that help form the AOI
- Data: textual tags attached to photos
- Challenge: some tags are common to many AOIs
  - E.g., “Paris” and “France” are very common to AOIs in the city of Paris
- Method: term-frequency and inverse document frequency (TF-IDF)

Eiffel tower area, Paris, 2014

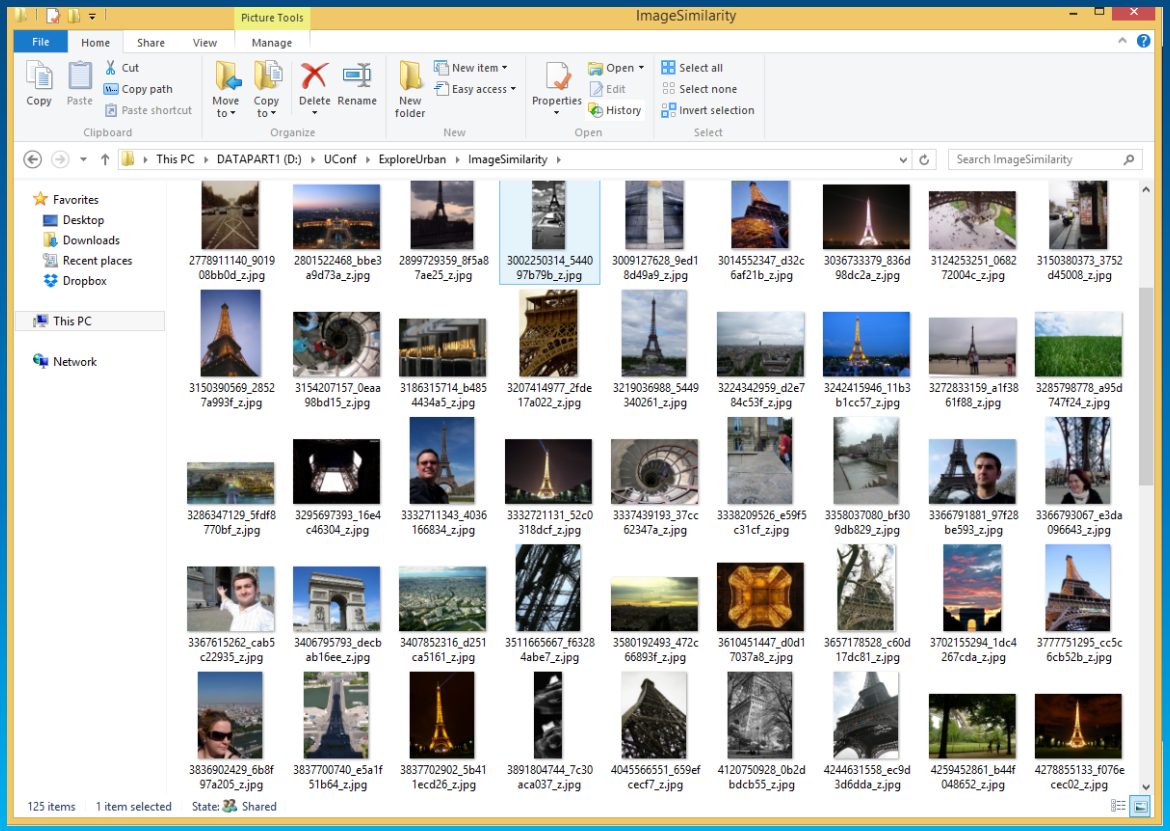


Time Square area, New York, 2005

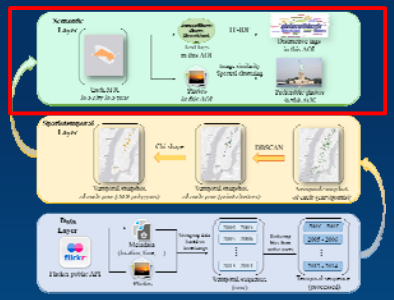
3.3 Layer 3: understanding AOI



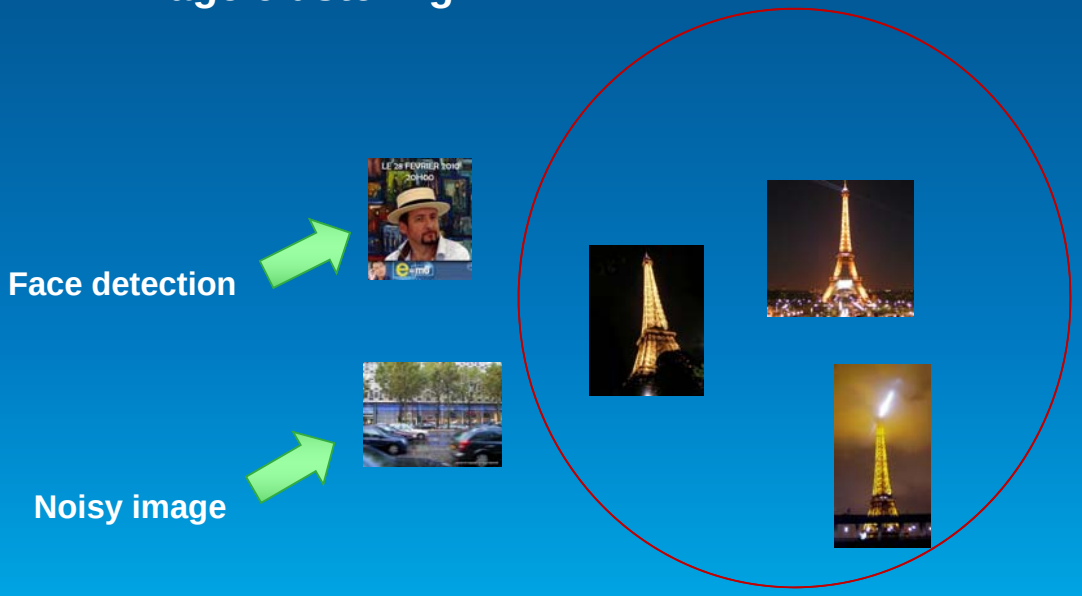
- What are people looking at in these AOIs?
- Data: Flickr photos
- Challenge:
  - Photos taken in an area have random qualities
  - A huge number of photos to process
    - 1,456,298 in Paris
    - 2,761,542 in New York
    - ...



3.3 Layer 3: understanding AOI

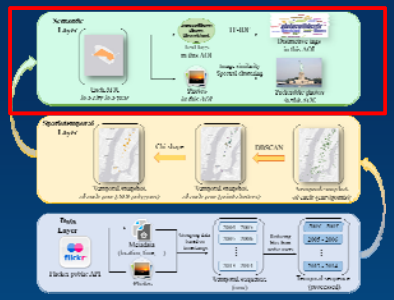


- An automatic workflow combining multiple methods:
  - Human face detection using OpenCV library
  - Image similarity comparison
  - Image clustering



Representative image

3.3 Layer 3: understanding AOI



- **Method:**
  - **An automatic workflow combining multiple methods:**

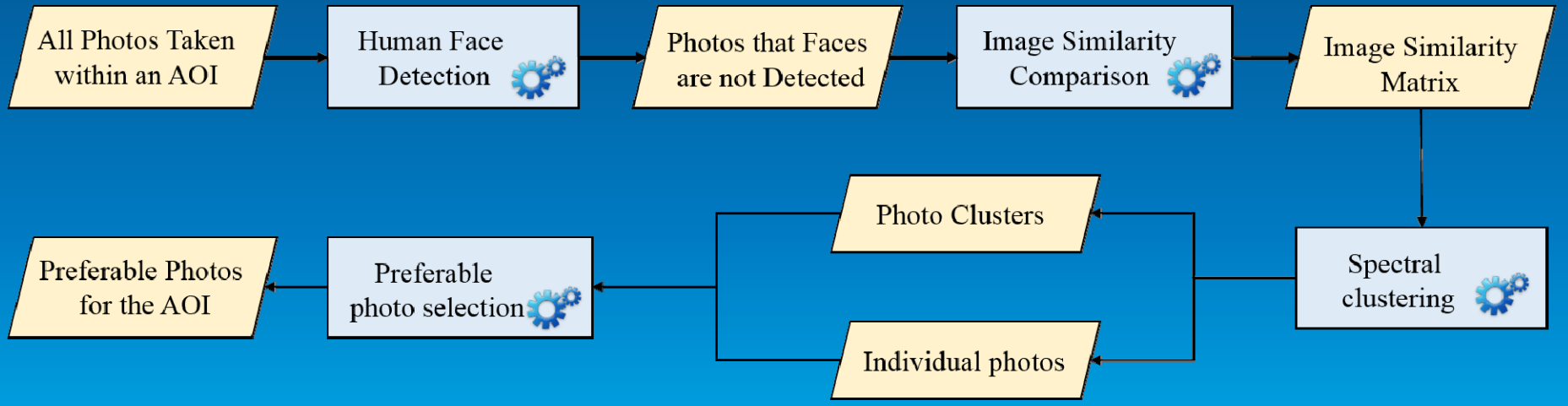
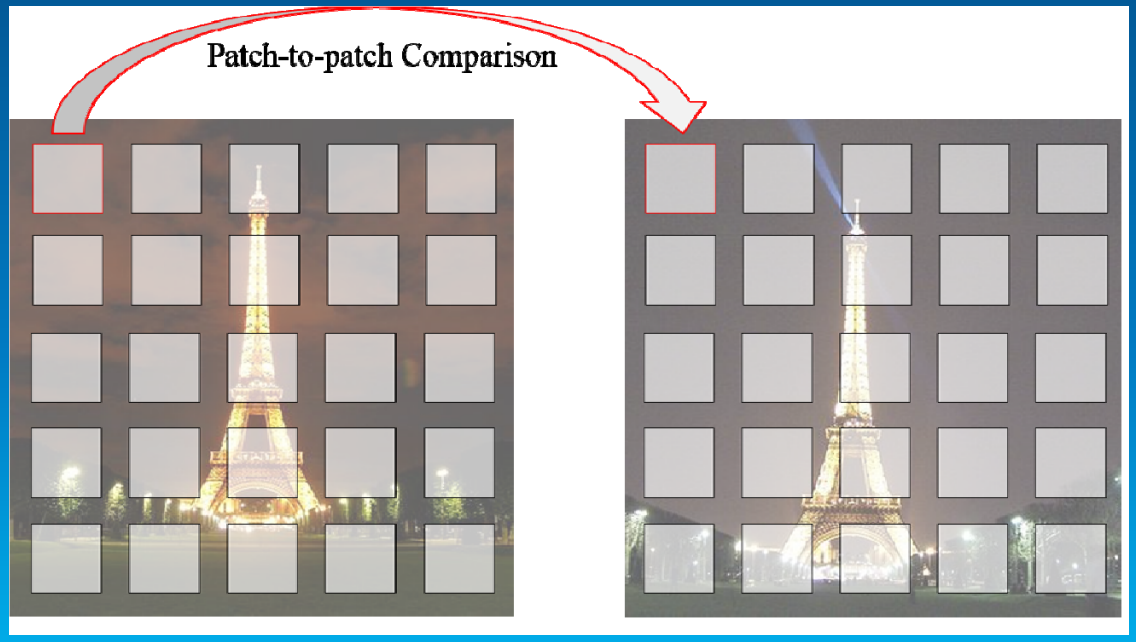


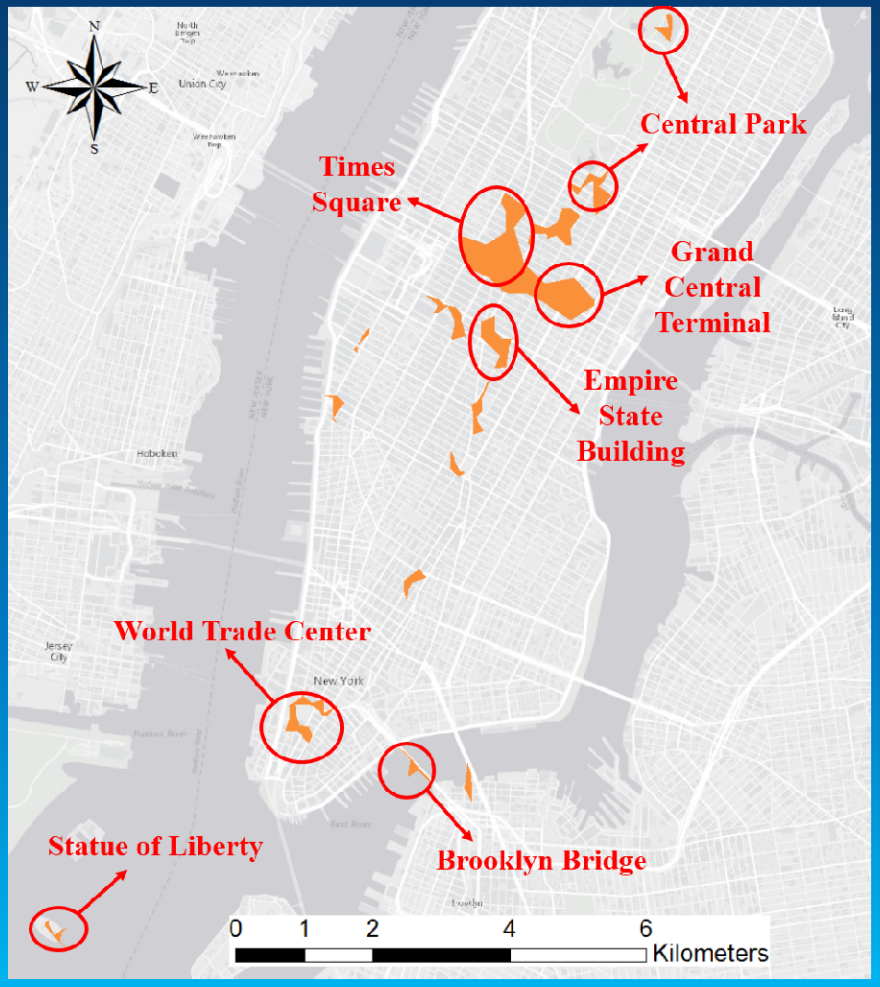
Image comparison

Patch-based comparison: adaptive to slightly distorted images

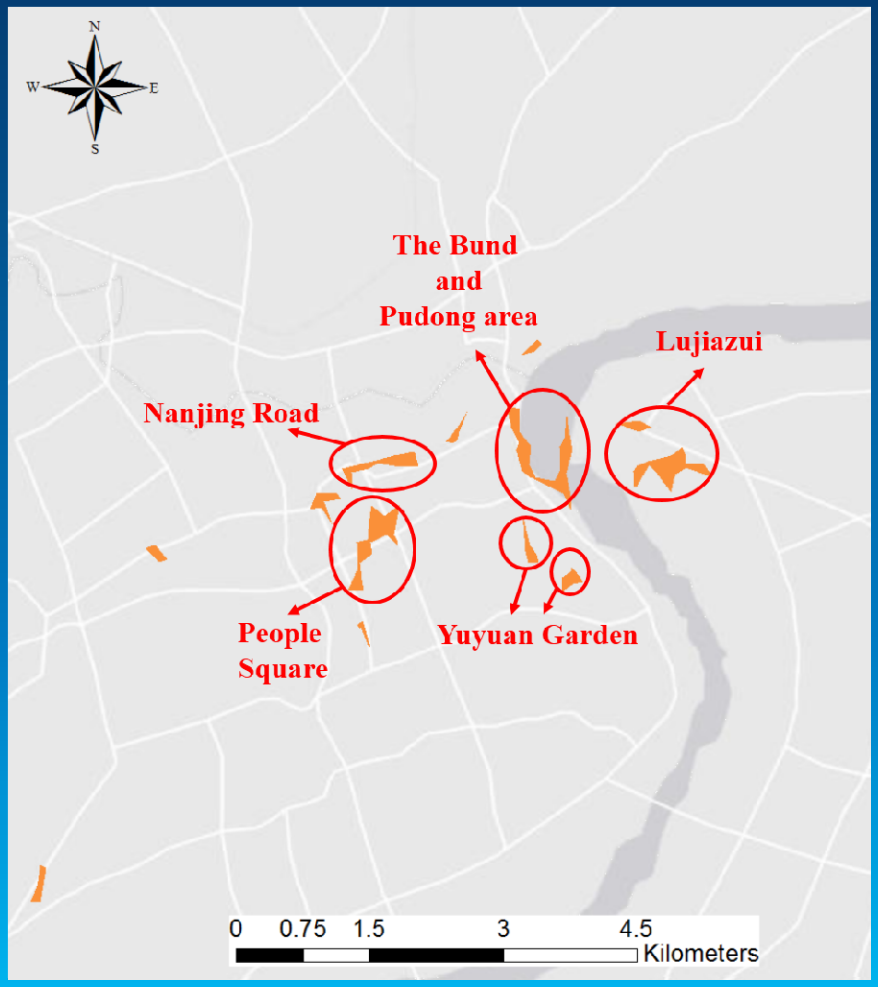


4. Knowledge and insights derived from AOI

4.1 Spatial distribution of the AOI



(a)

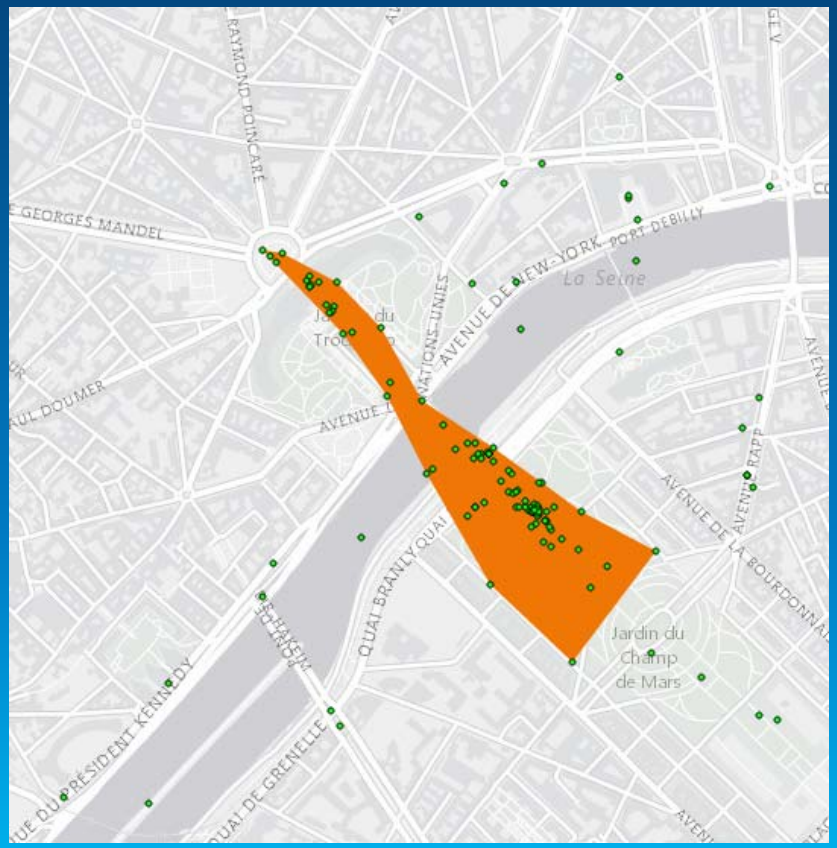


(b)

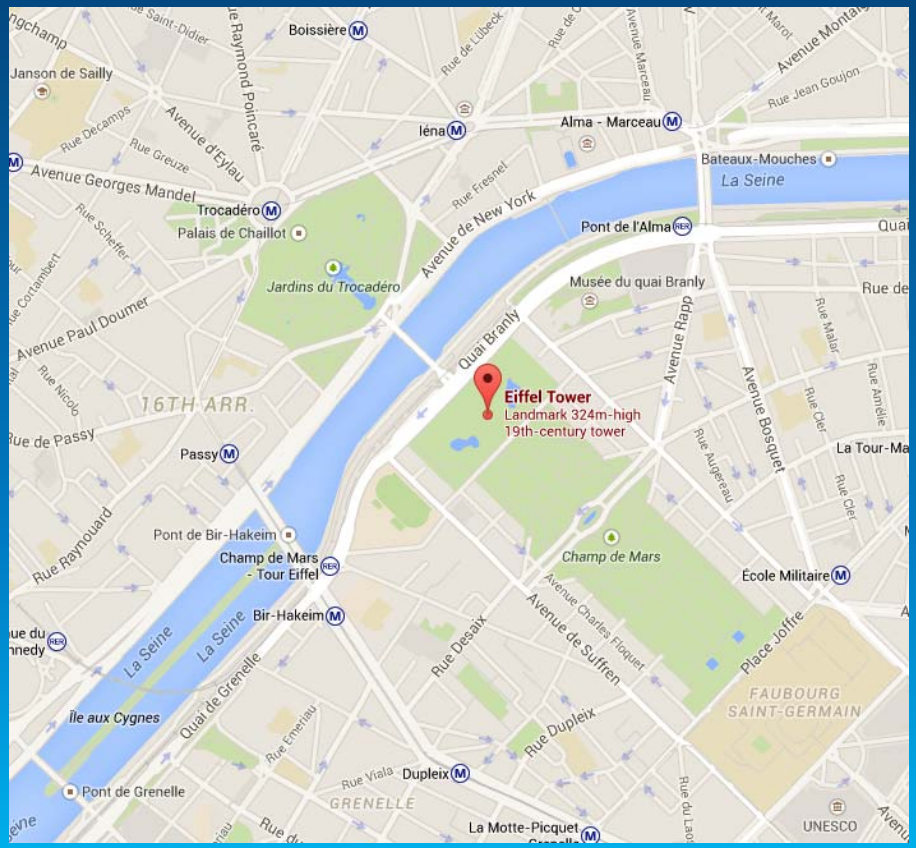
4. Knowledge and insights derived from AOI

4.1 Spatial distribution of the AOI

- **Eiffel Tower AOI**

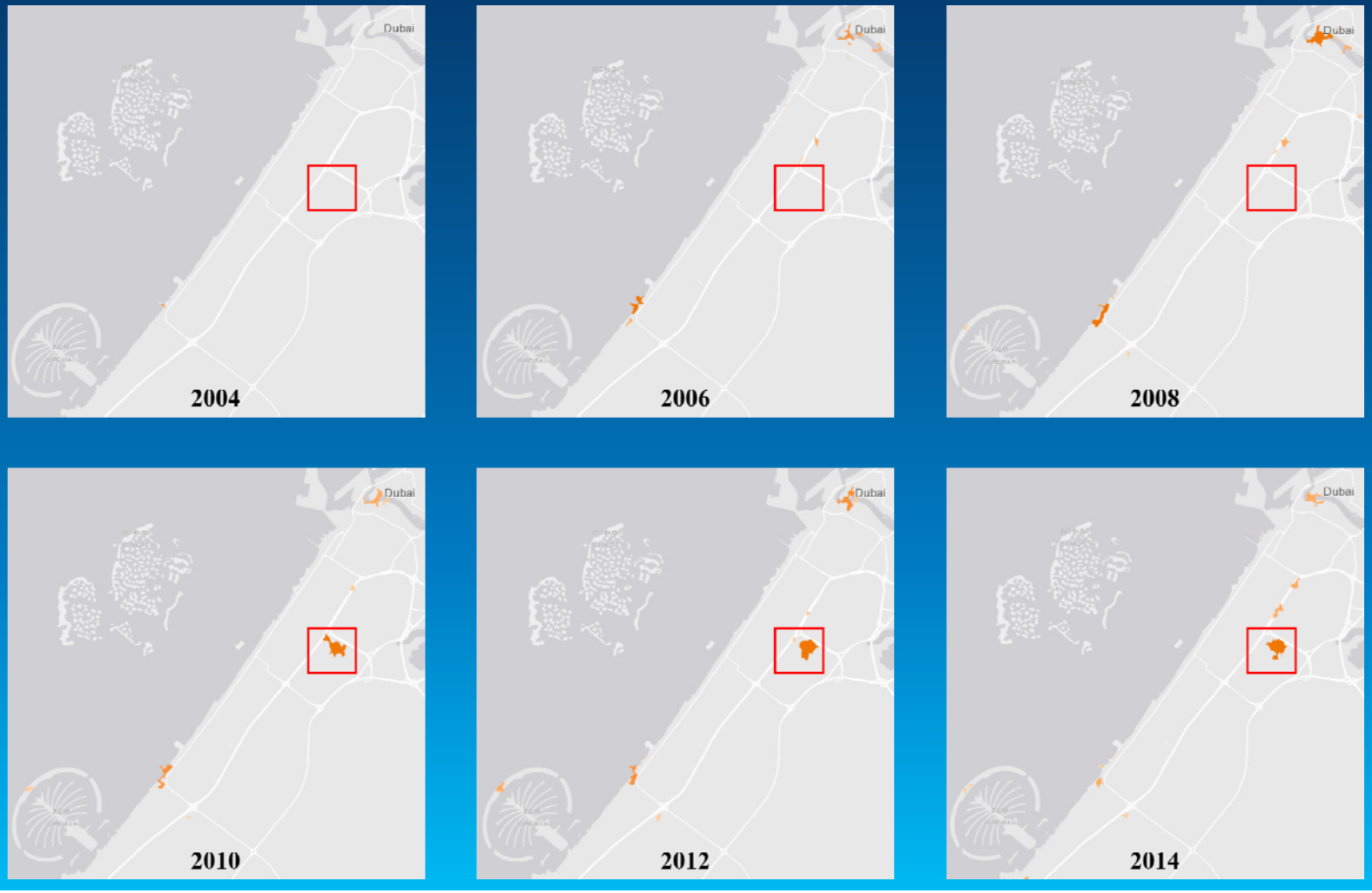


- **Eiffel Tower from Google Map**



4. Knowledge and insights derived from AOI

4.2 Spatiotemporal dynamics of AOI (an example of Dubai)

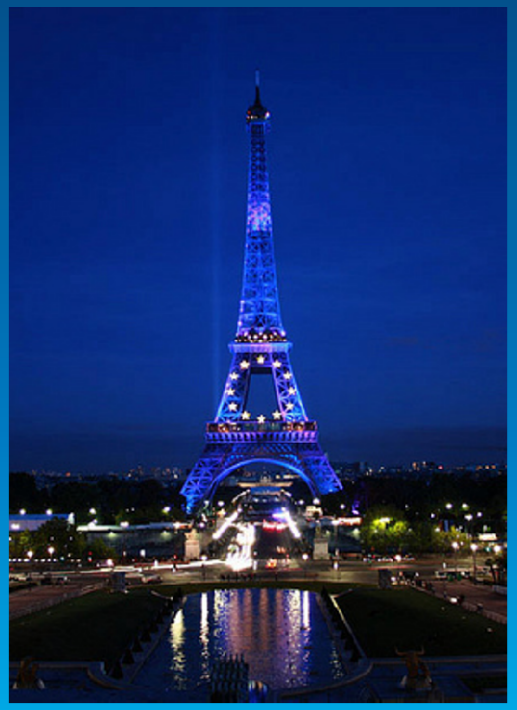


4. Knowledge and insights derived from AOI

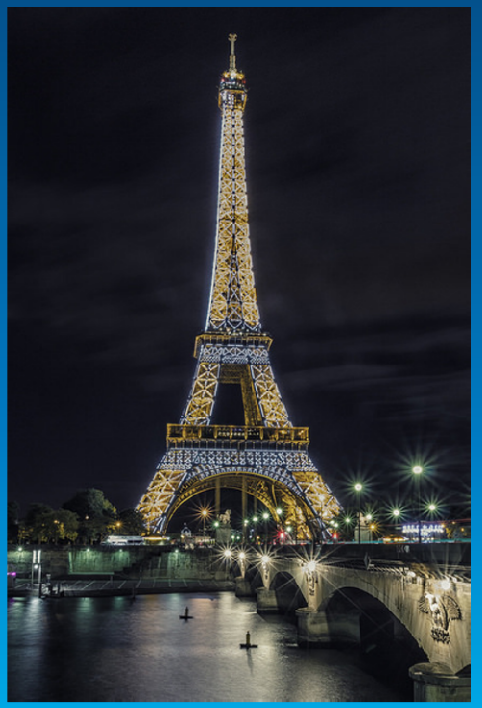
4.3 Historical slideshow

- Reveal the changes of landmarks through the lens of general people
- Valuable documentary for museums

2005



2008



2014

4. Knowledge and insights derived from AOI

4.4 Untypical AOI

- AOI are not necessarily the regions which provide aesthetical views

An online prototype for the extracted AOI

- **http://stko-exp.geog.ucsb.edu/urbanAOIs**

5. Conclusions and future work

- Urban AOI are areas within an urban environment that attract people's attention
- We develop a framework for extracting and understanding urban AOI from geotagged photo data
- We derive spatiotemporal knowledge and other insights from the extracted AOI
- We develop an interactive online demo to visualize the extracted AOI
- Seasonal variability of AOI could be an interesting future direction

Thank you!

Yingjie Hu
yingjiehu@umail.ucsb.edu
http://geog.ucsb.edu/~hu

Urban AOI: http://stko-exp.geog.ucsb.edu/urbanAOIs