ONTOLOGICAL APPROACHES TO SENSOR DATA ANALYSIS

Sensing Urban Space: from Street View Recognition, Event Inference to Understand Urban Behavior

Fan Zhang\textsuperscript{12}, Lezhi Li\textsuperscript{23}, Hui Lin\textsuperscript{1}

May 30, 2016

\textsuperscript{1}Institute of Space and Earth Information Science, The Chinese University of Hong Kong
\textsuperscript{2}Senseable City Lab, Massachusetts Institute of Technology
\textsuperscript{3}Graduate School of Design, Harvard University

http://www.iseis.cuhk.edu.hk/
Dataset – Street Level Images

Covering:
More than 5 million miles of roads, 3,000 cities, 50 countries,
Data Collection

100,000+ images collected from google street view API, within a 6 km radius. Images are extracted on a grid, 4 directions for each point.
Contents

- Objective
- Way to discover **Visual Patterns** in Urban Area

- Hand-Design Features
  - Greenery
  - Building Area
  - Entropy
  - Colors (RGB)

- Deep Learning Features
  - Learning Features **Intelligently**
Hand-Design Visual Features – Building Color

Fig. Building facade color map
Hand-Design Features – *Entropy*

Fig. Image entropy: visual noise

Low entropy – industrial districts
Hand-Design Features

Image Entropy

Greenery

Openness

Building Color
Demo

http://lezhili.me/cv_streetview/
Recognizing Cities/Neighborhoods from Street View images

- **Deep Learned Visual Features**

Training
- Image + Label

Testing
- New Image

Label
- "Beacon Hill"

Predicted Label
- "Beacon Hill"
Deep Learned Visual Features

Recognizing Cities/Neighborhoods from Street View

Boston (1), score 0.996
San Francisco (8), score 0.002
Chicago (2), score 0.002

New York (5), score 0.963
San Francisco (8), score 0.036
Chicago (2), score 0.000

Singapore (9), score 0.976
Tokyo (10), score 0.023
Hong Kong (3), score 0.000
Deep Learned Visual Features

Feature Visualization – Discovery the Hierarchical Feature Structure of Image sets
Deep Learned Visual Features

- In order to evaluate visual similarity in a more comprehensive, hierarchical way, and an automated process is required for feature extraction.

- **Deep Learning vs. Traditional Machine Learning**
  - Deep Artificial Neural Networks (3 layers to N layers..)
  - Automatic **Feature Extraction** Process
  - Convolutional Neural Nets (CNN)
Deep Learned Visual Features

Different Task to Discover Visual Features

- What’s the visual feature of
  - “Cities Name”
  - “Neighborhood Name”

- What’s the visual feature with a high/low
  - “Populations”
  - “Crime”
  - “Income”

- What’s the visual feature with a high/low
  - “Human Perception”
  - “Human Emotion”
The diagram illustrates the process of visual perception assessment inference. It begins with a street view image stream, which is then analyzed for various attributes:

- **Color/Light**: Tint, Brightness
- **Geometric**: Spatial Boundary
- **Texture**: Surface Properties

These attributes lead to scene content, which includes:

- Building, Road, Vehicle, Pedestrian, Tree
- Area, Length, Width, Quantity

The scene context further divides into:

- Residential Area
- Central Business District
- Park
- Natural Landscape

This context, along with cognitive descriptions such as Beautiful, Quiet, Safety, Greenery, Ugly, Noisy, etc., leads to a rating score, which is the final output of the visual perception assessment inference process.
Deep Learned Visual Features

Different Task to Discover Visual Features

What's the visual feature of "Cities Name"?
What's the visual feature of "Neighborhood Name"?
What's the visual feature with a high/low "Populations"?
What's the visual feature with a high/low "Crime"?
What's the visual feature with a high/low "Income"?
What's the visual feature with a high/low "Human Perception"?
What's the visual feature with a high/low "Human Emotion"?

Image Spatial Features

Images Temporal Features
From **Image** Semantic Recognition to **Event** Inference

*Earth Quake*

- July 2008
- August 2011
- May 2015

Destruction in Onagawa, Japan after the 2011 earthquake
➢ From **Event** Inference to Urban Development **Behavior**
• Conclusion

• Street View Images have Infinite Dimension

• Deep Learning Helps Spatial knowledge discovery

Thank you!