D10

# World Seer: A Real-time Geo-Tweet Photo Mapping System

The University of Electro-Communications, Tokyo, JAPAN   Keiji Yanai

Overview

- A system showing geotagged Twitter messages with photos on the streetviews as well as online maps
  - Photo Clustering & Search by date, words, userid and places
- "By-product" of the geo-tweet recording system.
  - "Main-product" is a recording system which monitors the Twitter stream. It stored 30 million geo-photo tweets.
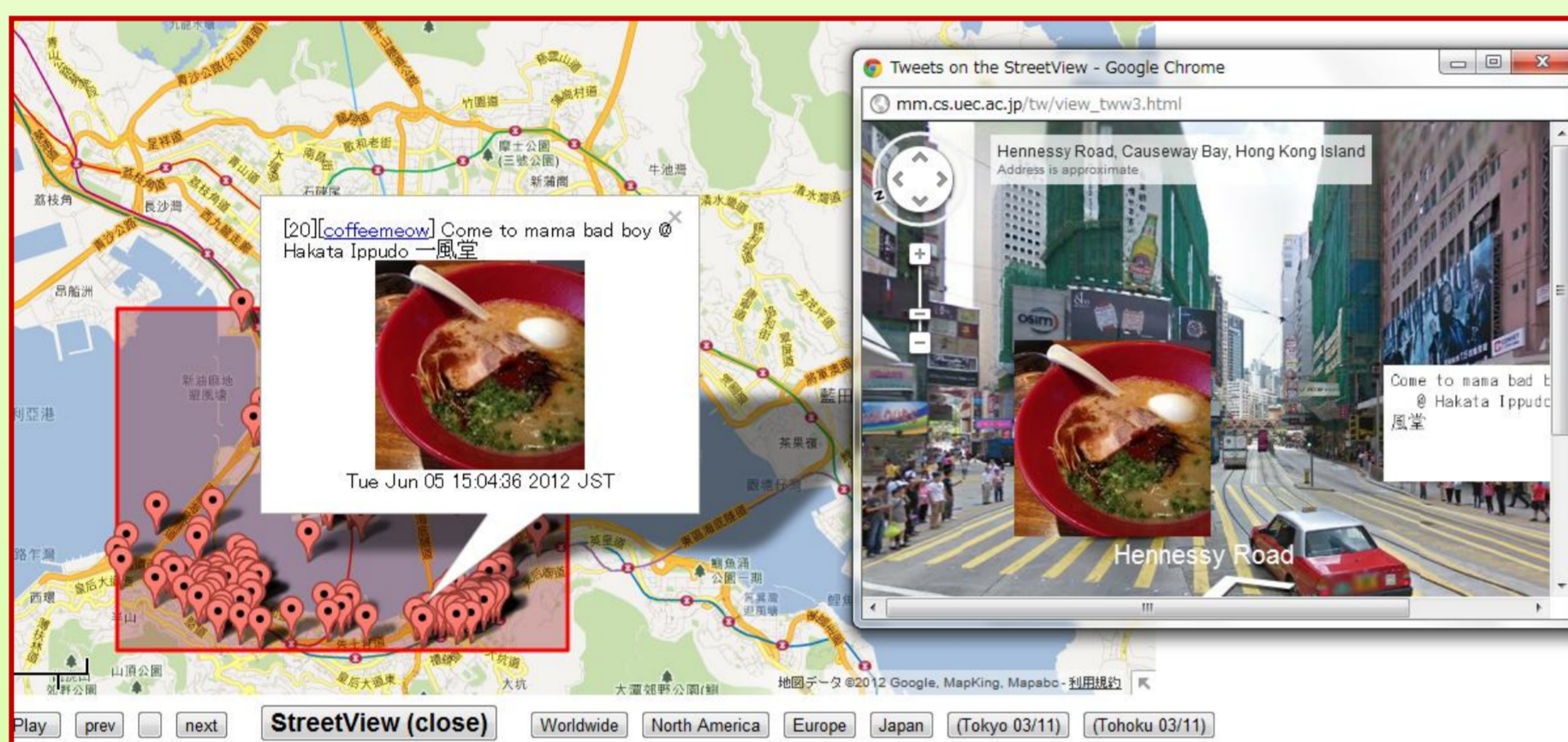
"Ramen noodle" on the streetview

"Clustering geo-tweet photos on March 11th 2011"

Backgrounds

◆ Some Twitter messages (Tweets) contain "geotags" and "photos" (we call them as "Geo-Photo Tweets")
◆ Twitter can be regareded as being one of other sources of geotagged photos than Flickr.
◆ Twitter provides "Twitter search API". However, the search term is limited for the latest 10 days.
◆ Twitter also provides "Twitter streaming API", via which we can receive Twitter message stream continuously.

◆To do researches with Twitter geotagged photos, we have to monitor the Twitter stream, pick up geo-photo tweets and store them into the own database for several months.

Objective: Collecting Geo-TwPhotos

◆ [Data Collection System] Store geo-photo tweets picked up from the Twitter stream into a DB continuously.
  • Now It has 28,000,000 geo-photo tweets for 16 months.
◆ [Visualizing System (Demo)] Map geo-photo tweets on the maps and the street-views. This system is a "by-product" system of the data collection system.

Potential Future Work

◆ Event detection with photos
  • Geo-tweet photos clustering and mining with tag analysis
◆ Cultural difference detection
  • Visual analysis of differences of something depending on places (food, way of life(culture), damage of disasters)

Twitter vs Flickr on GeoPhotos

◆ Everyday life (e.g. foods) <-> Travel or special event tweet messages <-> Keywords
  instant (online) uploading <-> offline uploading
  the same user uploads a few photos at once
                        <-> many photos at onces
  200,000 geo-photos/day <-> 150,000 photos/day
  (over 6 million nongeo-photos/day <-> 3million photos/day)

Characteristics of the photos are different from each other. We should use them considering their characteristics.

System Architecture

◆ [Backend] Geo-tweet collection system
  • Data collection via Twitter streaming API
  • MySQL + Perl
  • Can search the DB via Search API (not public)
  • The DB is being updated every 5 minutes
  • Download and extract visual features from some parts of the photos (Japanese photos)
◆ [Frontend] Geo-tweet photo mapping systems
  (1) A real-time mapping system
  • JavaScript code which runs on browsers
  • Feature extraction & K-means by Javascript (carried out on the client side)
  (2) Event detection system system
  • Offline feature extraction & clustering

Yusuke Nakaji and Keiji Yanai: **Visualization of Real World Events with Geotagged Tweet Photos**, IEEE ICME Workshop on Social Media Computing (SMC), (2012).