Computers, Environment and Urban Systems 54 (2015) 240–254

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Extracting and understanding urban areas of interest using geotagged photos

Yingjie Hu [a,*], Song Gao [a], Krzysztof Janowicz [a], Bailang Yu [b], Wenwen Li [c], Sathya Prasad [d]

[a] STKO Lab, Department of Geography, University of California Santa Barbara, Santa Barbara, CA 93106, USA
[b] Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, 500 Dongchuan Rd., Shanghai 200241, PR China
[c] GeoDa Center for Geospatial Analysis and Computation, School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ 85287, USA
[d] Applications Prototype Lab, Esri Inc., 380 New York Street, Redlands, CA 92373, USA

ARTICLE INFO

Article history:
Received 19 November 2014
Received in revised form 1 September 2015
Accepted 1 September 2015
Available online xxxx

Keywords:
Areas of interest
AOI
Social media
Flickr
DBSCAN
Chi-shape
Tag extraction
Photo analysis
Data mining

ABSTRACT

Urban areas of interest (AOI) refer to the regions within an urban environment that attract people's attention. Such areas often have high exposure to the general public, and receive a large number of visits. As a result, urban AOI can reveal useful information for city planners, transportation analysts, and location-based service providers to plan new business, extend existing infrastructure, and so forth. Urban AOI exist in people's perception and are defined by behaviors. However, such perception was rarely captured until the Social Web information technology revolution. Social media data record the interactions between users and their surrounding environment, and thus have the potential to uncover interesting urban areas and their underlying spatiotemporal dynamics. This paper presents a coherent framework for extracting and understanding urban AOI based on geotagged photos. Six different cities from six different countries have been selected for this study, and Flickr photo data covering these cities in the past ten years (2004–2014) have been retrieved. We identify AOI using DBSCAN clustering algorithm, understand AOI by extracting distinctive textual tags and preferable photos, and discuss the spatiotemporal dynamics as well as some insights derived from the AOI. An interactive prototype has also been implemented as a proof-of-concept. While Flickr data have been used in this study, the presented framework can also be applied to other geotagged photos.
© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Urban areas of interest (AOI) refer to the areas within an urban environment which attract people's attention. Such areas may contain city landmarks, commercial centers, and recreational zones, or may simply provide a scenic view of the city. The concept of urban AOI is different from *urbanized area*, as the former puts additional emphasis on people's interests. Consequently, an urbanized area (e.g., a regular residential neighborhood) may not necessarily also be an AOI. Unlike the well-defined administrative districts, the boundaries of urban AOI are vague. This is because urban AOI are subjective: given a familiar city, most people will have a list of interesting areas in mind; yet, their lists may differ due to different ages, cultures, education backgrounds, personal interests, and so forth. Similarly, agreeing on certain AOI does not imply agreeing on their spatial extents and delineations. In this respect, AOI are related to the concept of *vague place* (Cohn & Gotts, 1996; Jones et al., 2008; Liu et al., 2010; Montello et al., 2003).

Urban AOI have great meanings in multiple application domains. For tourists, AOI highlight the interesting zones within a city, and can therefore be used to support trip planning of travelers. For city planners, AOI reveal the regions which receive high exposure among the general public. Accordingly, these regions could be assigned higher priorities when there are only limited resources for urban planning projects, such as city beautification (Espuche et al., 1991; Gandy, 2006). Since AOI are often visited by a large number of people, transportation analysts can examine these regions to understand traffic flows and human mobility patterns (Batty, 2007; Yuan & Raubal, 2012). In addition, information service providers can display targeted information based on AOI (e.g., highlighting the hotels within the AOI of a city).

AOI exist in the perception of people, and as a result, it is difficult to capture AOI using traditional data collection methods. In urban studies, remote sensing data have been often used to monitor the status of a city (Hsu, Elvidge, & Matsuno, 2013; Shi et al., 2014; Yu et al., 2010). While good at detecting physical phenomena (e.g., land use types), remote sensing data are unable to observe interests of people. Alternatively, human participant survey, such as the one employed by Montello et al. (2003), could be used to uncover AOI. While such surveys provide valuable insights, they are labor-intensive, time-consuming, and do not scale well.

Social media, such as Twitter and Flickr, record the interaction between people and their surrounding environment (Mckenzie, Adams, & Janowicz, 2013). Compared to remote sensing images, social media data contain valuable information about the behavior of people in

* Corresponding author.
E-mail addresses: yingjiehu@geog.ucsb.edu (Y. Hu), sgao@geog.ucsb.edu (S. Gao), jano@geog.ucsb.edu (K. Janowicz), blyu@geo.ecnu.edu.cn (B. Yu), wenwen@asu.edu (W. Li), sprasad@esri.com (S. Prasad).

http://dx.doi.org/10.1016/j.compenvurbsys.2015.09.001
0198-9715/© 2015 Elsevier Ltd. All rights reserved.

geographic space. In most cases, data from social media platforms can be retrieved through their public APIs at low costs. While social media data have often been criticized for the representative issue, i.e., the users may not constitute a representative sample of the entire population (Chou et al., 2009), they are nevertheless generated by millions of people from different countries throughout the world.

Among the many types of social media data, geotagged Flickr photos possess a high suitability for exploring urban AOI. One major advantage of Flickr data is that they reflect the interest of people towards locations (Crandall et al., 2009; Kennedy et al., 2007; Li & Goodchild, 2012). This can be distinguished from some other social media data, such as geotagged tweets, which are not necessarily related to the locations they originate from. For example, a user may tweet about a humanitarian crisis, such as a particular drought, in East Africa from her office in Santa Barbara, California. While the location of her office is attached to that tweet, her real interest is in the events in Africa. In addition, the openness of the Flickr API allows the retrieval of publicly available data throughout the world and since the year of 2004. This adds an interesting temporal component to the data. Many other social media data are either constrained by user permissions (e.g., Facebook) or limited API accessibility for long-term data (e.g., Twitter and Foursquare). Besides, existing research shows that a major proportion of Flickr photos were taken in urban areas (Crandall et al., 2009; Hollenstein & Purves, 2010), and this gives Flickr data one more advantage for studying urban AOI. Back in March 2013, Flickr already had 87 million registered members and more than 8 billion photos (Jeffries, 2013).

This paper aims at extracting urban AOI and understanding them from spatial, temporal, and thematic perspectives. Geotagged Flickr data from six different cities in six different countries have been retrieved for this study. From a spatial perspective, this research examines the locations and extents of AOI in multiple cities and countries. From a temporal perspective, this research investigates the evolution of AOI in the past ten years (2004–2014), and compares the evolution patterns in developed and developing countries. From a thematic perspective, this research extracts the semantics of AOI from the textual tags and user contributed photos, and uncovers the thematic topics underpinning these AOI. The contributions are listed as follows:

- We develop a framework for extracting and understanding urban AOI from geotagged photo data. This framework is not restricted to Flickr data, but can also be applied to other geotagged photos.
- To generate proper polygon representations for AOI from point clusters, we design an experiment to identify an optimal parameter for the chi-shape algorithm, which balances the emptiness and the complexity of the generated polygons.
- We examine the extracted AOI from spatial, temporal, and thematic perspectives. Additional insights, such as the changes of landmark exterior and untypical AOI, are also discovered and discussed.
- We design and implement an online prototype that allows readers to explore the extracted AOI. This prototype can be accessed at: http://stko-exp.geog.ucsb.edu/urbanAOIs/index.html.

The remainder of this paper is organized as follows. Section 2 discusses related work on points of interest, vague places, volunteered geographic information, as well as extracting hot zones and landmarks from spatial footprints. Section 3 describes the dataset used in this work, and Section 4 presents the framework for extracting and understanding urban AOI from geotagged photos. Section 5 provides a discussion on the spatiotemporal dynamics of AOI and the insights acquired from the extracted AOI. Section 6 describes the interactive prototype implemented based on the proposed framework. Finally, Section 7 summarizes this work and discusses future directions.

2. Related work

2.1. AOI, POI, and vague place

The concept of AOI is closely related to two notions in literature, namely *point of interest* (*POI*) and *vague place*. POI represents individual locations (e.g., a restaurant or a landmark) which are of interest to people (Mckenzie et al., 2014; Yoshida, Song, & Raghavan, 2010). In contrast, an AOI may contain multiple co-located geographic features, such as the restaurants on a pedestrian street or several nearby landmarks (Elias, 2003; Raubal & Winter, 2002). AOI may also include the areas that do not have prominent landmarks but simply provide scenic views (e.g., the areas in Paris that provide a good view of the Eiffel Tower). Thus, AOI can incorporate the *reversed viewshed*: unlike the traditional *viewshed* which represents the visible areas from a fixed point, *reversed viewshed* identifies the areas from which a given target can be seen (Fisher, 1996; Senaratne, Bröring, & Schreck, 2013). In terms of geometric representations, AOI are typically represented as polygons instead of points. Such a polygon representation has several advantages. First, it enables new functionalities that require spatial relations. For example, applications of geographic information retrieval often need to query information within a specific region (Jones et al., 2008), and AOI, which represent the boundaries of these regions, can be used for more accurate information search. Second, from a computational perspective, it is generally more efficient to perform operations on a polygon than on a set of points (Akdag, Eick, & Chen, 2014). In addition, polygons convey information about the shape of a region, and can be employed for shape-based analysis (Li et al., 2014). Finally, from a cognitive perspective, polygons can provide simple and accessible representations for areas compared with clustered points.

The concept of vague boundary is important in the context of AOI. While disagreement may exist on the extent of AOI in a city, Montello et al. (2003) showed that there is often a general consensus on the core area of a place (e.g., *downtown Santa Barbara*). Similarly, in this study, we are interested in identifying the AOI which are agreed upon by a significant number of people. While vague places can often be referred to by vernacular names (Hollenstein & Purves, 2010), AOI may not have a particular name due to the co-existence of multiple attractions. In addition to the work from Montello et al., there are also other studies on vague places. For example, Jones et al. (2008) used the place names harvested from Web search to identify the boundaries of vague places based on kernel density estimation. Davies et al. (2009) discussed the user needs and implications for modeling vague places in GIS.

2.2. Volunteered geographic information and spatial footprints

Volunteered geographic information (VGI) provides an alternative data source for many geospatial applications (Elwood, 2008; Goodchild, 2007; Hu et al., 2013; Sui, Elwood, & Goodchild, 2012). Although lacking standard quality control, some VGI has been shown to have equally good quality as authoritative data (Haklay, 2010; Li & Goodchild, 2012). *Spatial footprint* is a special type of VGI generated by people on social media platforms (Cheng et al., 2011). When posting stories or sharing information on social media, people can geo-reference the posts by attaching their locations. Compared with traditional VGI platforms (e.g., OpenStreetMap) on which people directly contribute geographic information (Haklay & Weber, 2008), *spatial footprints* can be considered as indirectly contributed VGI, i.e., the major goal of users is usually to share information rather than to contribute data.

The geotagged Flickr photos used in this research are one type of spatial footprints. A number of studies have been conducted based on Flickr data. For example, Hollenstein and Purves (2010), Intagorn and Lerman (2011), as well as Li and Goodchild (2012) have used Flickr data to investigate the boundaries of vague place concepts. Keßler et al. (2009) and Gao et al. (2014) employed Flickr data to enrich

Download English Version:

https://daneshyari.com/en/article/6921941

Download Persian Version:

https://daneshyari.com/article/6921941

Daneshyari.com