Applied Geography 76 (2016) 147–153

Contents lists available at ScienceDirect

Applied Geography

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

Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery

Li Yin [a, *], Zhenxin Wang [b]

[a] Department of Urban and Regional Planning, University at Buffalo, The State University of New York, Buffalo, NY 14214, USA
[b] Center for Human-Engaged Computing, Kochi University of Technology, 185 Miyanokuchi, Tosayamada-Cho, Kami-Shi, Kochi 782-8502, Japan

A R T I C L E I N F O

Article history:
Received 31 May 2016
Received in revised form
13 September 2016
Accepted 17 September 2016
Available online 25 September 2016

Keywords:
Street design features
Enclosure
Walkability
Machine learning

A B S T R A C T

One major limitation currently with studying street level urban design qualities for walkability is the often inconsistent and unreliable measures of streetscape features across different field surveyors even with costly training due to lack of more objective processes, which also make large scale study difficult. The recent advances in sensor technologies and digitization have produced a wealth of data to help research activities by facilitating improved measurements and conducting large scale analysis. This paper explores the potential of big data and big data analytics in the light of current approaches to measuring streetscape features. By applying machine learning algorithms on Google Street View imagery, we generated objectively three measures on visual enclosure. The results showed that sky areas were identified fairly well for the calculation of proportion of sky. The three visual enclosure measures were found to be correlated with pedestrian volume and Walk Score. This method allows large scale and consistent objective measures of visual enclosure that can be done reproducibly and universally applicable with readily available Google Street View imagery in many countries around the world to help test their association with walking behaviors.

1. Introduction

The recent advances in sensor technologies and digitization have produced a large amount of data at fine resolution. These new data can help research activities by facilitating improved measurements or new variables of interest (Ruppert, 2013), and conducting large scale study at micro level. Big data such as Google Street View imagery that documents most North American cities in high resolution and is available and accessible online to everyone provides new opportunities to measure characteristics of physical cities at fine spatial scales. These data and big data analytics make it increasingly possible to better measure the physical city and study behaviors that may be influenced by the urban built environment (Yin, Cheng, Wang, & Shao, 2015). This paper aims to explore the potential of big data and big data analytics in the light of current approaches to measuring street level urban design qualities for walkability. Emphasis is placed here on how Google Street View imagery and machine learning techniques can help us approach conventional questions on evaluating streetscape features with objective measurement of visual enclosure or possibility of better identification.

How people use the built environment depends on what the spatial structure that designers and planners created offers. Multidisciplinary researchers and practitioners have been exploring ways to improve physical activity outcomes through modifying the built environment. "Urban design may influence people's choices and behavior in the use of the built environment. This influence, however, remains assumption and unclear until urban design qualities can be defined, quantified, measured, and tested empirically" (Yin, 2014, p. 273). Ewing and Handy (2009) developed quantitative but subjective measurement protocols of five categories of urban design qualities on their contribution to street walkability through field survey, including visual enclosure, imageability, human scale, transparency, and complexity. They demonstrated in detail how to quantify and rate street level urban design features by surveys to develop street design metrics for walkability. As Ewing and Clemente (2013) suggested, the challenge is to move from subjective definitions and measures to operational objective ones that can be done reliably and consistently across raters to capture the essence of each category of street level urban design quality. Purciel, Neckerman, and et al (2009) translated some of the variables in these five categories using GIS.

* Corresponding author.
E-mail address: liyin@buffalo.edu (L. Yin).

http://dx.doi.org/10.1016/j.apgeog.2016.09.024
0143-6228/

However, there are still two major limitations with developing such street design metrics: 1) the unfeasible or inefficient large scale field observation in terms of time and cost; and 2) the often inconsistent and unreliable measures of street design qualities across different raters even with costly training due to lack of more objective processes. Moreover, there is no established standard procedure to handle the error margin involved to ensure the comparability and reliability from different raters.

In this paper, we explore and discuss how applying machine learning algorithms to Google Street View imagery can help study the morphology of the built environment, in particular, by generating objectively some street level urban design measures related to walkability. We use visual recognition techniques and visual images of the streetscape to create variables on visual enclosure for over 300 sampled street blocks across Buffalo, New York. Big data sources such as Google Street View imagery and big data analytics like machine learning technology together can help fill the gaps on current approaches to measuring street level urban design qualities that have been traditionally subjective and limited based on small samples. The results can provide new insight into the pedestrian friendly design and urban planning.

2. Assessing street level urban design qualities for walkability

Recent studies have shown that obesity and obesity-related diseases are associated with lack of physical activity and unhealthy diet (Papas et al., 2007). A promising approach to addressing this problem is through various interventions on the modification of the built environment to create streetscapes and neighborhoods that are walkable and livable. Planners and urban designers who are concerned with designing and planning the built and landscape elements in space and how people interact with these elements play an important role in shaping the built environment to encourage walking (Ewing & Handy, 2009).

Many studies have characterized and measured the built environment using the D variables – *density, design, diversity, destination accessibility*, and *distance to transit*, and explained travel mode choice and walking frequency in terms of these variables (Saelens, Sallis, & Frank, 2003; Ewing & Cervero, 2010; Frank et al., 2006; 2007; Sallis and Glanz, 2006; 2009; Smith et al., 2008; Marshall & Garrick, 2010; Boer, Zheng, Overton, Ridgeway, & Cohen, 2007). One of the Ds, *design*, has usually been represented by street network characteristics of the studied area and measured by street network density or block size. However, pedestrian activities can vary significantly from street block to street block even within a small geographic area with similar street network density (Desyllas & Duxbury, 2001). The currently widely used *design* variables are not sufficient to reflect the micro and street level environment and its impact on pedestrian experience, as discussed in classic urban design readings like Lynch (1960) and Hedman (1984). Hajrasouliha and Yin (2015) argued that the current street network connectivity measures only captured physical connectivity and suggested the important role of visual connectivity on its impact of pedestrian volumes. Other *design* variables have been occasionally used including street width, sidewalk coverage, number of trees, building setbacks, etc. The findings are mixed, however, with regard to their impact on physical activity or pedestrian activity (Desyllas & Duxbury, 2001).

A set of indicators and matrixes can help systematically quantify the impact of components of the street environment on the level of walkability to compare street and neighborhood design. Clemente et al. (2005) and Purciel and Marrone (2006) developed a field manual with detailed measurement protocols for coding the street physical features contributing to the five urban design qualities and provided detailed guidance for the field work and survey that are focusing on the street level experience.

One of the five qualities is visual enclosure (Ewing & Handy, 2009; Purciel et al., 2009). Two variables used to measure visual enclosure are proportion of sky ahead the street and across the street. These two variables were used to measure the amount of sky visible from a standing point on a street, with trees, buildings, street lights, and other street furniture and man-made objects as visual obstruction. Designers and literature suggested that this information on enclosure measure how the built environment encapsulated the pedestrian and related it to people's perceived confinement of space and sense of intimacy, and a place's livability and sense of security (Ewing & Handy, 2009; Porta and Renne, 2005; Zhang et al., 2012). These two sky visibility variables can be influenced by many factors that have been used to characterize the built environment linked to physical activity in health-related research such as street width, building height, and presence of trees. Section height to width ratio has been used to measure the degree of sense of visual enclosure (Carmona & Tiesdell, 2007). Tree canopy can help to reach some level of intimacy on streets. Streets that are wide, with lower buildings, small building setback, and without trees have relatively high level of sky exposure and low level of visual enclosure.

To estimate the proportion of sky for measuring visual enclosure, Purciel et al. (2009; p15) suggested to form a frame of vision box using the thumb and pointer fingers of both hands “that is visible when you look ahead with your line of sight parallel to the ground” and “hold it up to your face”, then “move it away until you can see all four sides” (See Fig. 1). This method is subjective to field surveyors' heights and individual definitions and interpretations of how far away is the right distance to see all four sides. Different field surveyors may get dissimilar proportion of sky numbers for the same location even with systematical training. This raises the problem of reliability. The current literature suggested moving to operational objective street design measures for street walkability studies. However, little has been done on objectively measured sky related visual enclosure variables because of the limitation on both data and method.

3. Big data and big data analytics for planning and designing walkable streets

Big data is defined by high volume, velocity, variety, and variability of information. Big data analytics can analyze information with high volume, velocity, variety, and variability better than conventional tools and helps to uncover patterns, correlations, and other useful information. The internet and a wide range of programs online are important sources of big data. Google Street View provides street level imagery along most streets in the U.S. and many countries around the world, and is available over the Internet and through Google Earth software. The total road length covered is more than 5 million kilometers. It allows users to view streetscape and experience walking down streets in the virtual street environment.

Photographs and sections have been used by designers and planners for many years for graphic representations of spaces for analysis and design. Google Street View imagery provides more information and flexibility than these traditional representations for analyzing street morphology and its impact on pedestrian volume or movement. It also covers much larger area than traditional small sampled areas limited by time and budget. Google Street View has been used in several studies for streetscape audits. Rundle, Bader, Richards, Neckerman, and Teitler (2011) suggested Google Street View effective for auditing walkable street environments. They found high levels of agreement between measurements based on audits from field work and from Google Street