Assessing Neighborhood Conditions using Geographic Object-Based Image Analysis and Spatial Analysis

Chi-Feng Yen
Center for Human Dynamics in the Mobile Age, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.
cyen@sdsu.edu

Ming-Hsiang Tsou
Center for Human Dynamics in the Mobile Age, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.
mtsou@sdsu.edu

Chris Allen
Center for Human Dynamics in the Mobile Age, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, USA.
ccchris.allen@gmail.com

Abstract

Traditionally, understanding urban neighborhood conditions heavily relies on time-consuming and labor-intensive surveying. As the growing development of computer vision and GIScience technology, neighborhood conditions assessment can be more cost-effective and time-efficient. This study utilized Google Earth Engine (GEE) to acquire 1m aerial imagery from the National Agriculture Image Program (NAIP). The features within two main categories: (i) aesthetics and (ii) street morphology that have been selected to reflect neighborhood socio-economic (SE) and demographic (DG) conditions were subsequently extracted through geographic object-based image analysis (GEOBIA) routine. Finally, coefficient analysis was performed to validate the relationship between selected SE indicators, generated via spatial analysis, as well as actual SE and DG data within region of interests (ROIs). We hope this pilot study can be leveraged to perform cost- and time- effective neighborhood conditions assessment in support of community data assessment on both demographics and health issues.

2012 ACM Subject Classification Computing methodologies → Computer vision

Keywords and phrases neighborhood conditions assessment, geographic object-based image analysis, spatial analysis

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.70

Category Short Paper

1 Introduction

Socio-economic and demographic data are fundamental components of understanding neighborhood makeup and health condition [13]. Conventional approach to investigate the socio-economic and demographic condition within urban areas relies heavily on time-consuming and labor-intensive surveying that usually causes the lag of socio-economic and demographic changes (e.g. American Census Survey) [3]. With the growing development of computer vision, remote sensing and GIS technology, the lag of socio-economic and demographic data in urban areas can be effectively compromised to some degree. This pilot study aims to

10th International Conference on Geographic Information Science (GIScience 2018).
Editors: Stephan Winter, Amy Griffin, and Monika Sester; Article No. 70; pp. 70:1–70:7
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

examine the utility of RS imagery in urban neighborhood conditions assessment at census block within given zip codes, specifically using a popular image classification paradigm: GEOBIA. The research focus is to find SE indicators, generated from the extracted features using spatial analysis, that can represent certain SE or DG conditions in given census-block neighborhoods.

2 Remote Sensing on Neighborhood Condition Assessment

Considerable researches have demonstrated the feasibility of RS satellite imagery in estimation of DG data. [2] exploited both high- and medium- resolution satellite imagery to estimate population distribution for areas lacking census data in support of disaster resilience in Haiti. [1] filled DG data gap within developing countries using per-pixel population estimates generated by a classification and regression trees (CART) and multi-resolution satellite imagery. More recently, [3] identified the significant associations between the presence of specific vehicle models and voter preferences across 200 cities in the US. using convolutional neural network (CNNs) and the Google Street View image dataset.

In terms of SE condition, the contextual features, such as the size of building structures, the abundance of vegetation cover and etc., can reflect the different socio-economic status (SES) in urban areas, particular in residential areas [8, 9]. Specifically, lower SES usually accompanies less vegetation cover and swimming pools but high density of residential buildings [7, 11, 12]. In addition, SES information plays an important role to understand neighborhood health conditions.[4] indicated that neighborhood characteristics (e.g. SE and built environments) impact cancer incidence or outcomes. [6] found that women in high-SES neighborhoods have higher breast cancer-specific survival than in low-SES neighborhoods. [13] illustrated that lower SES is highly related to poorer health condition in developing countries because of the close proximity of people living and insanitary settlement. These characteristics could directly or indirectly ascend disease spread within the neighborhoods.

Although RS has been widely-applied for neighborhood condition assessment, providing more detailed perspectives of neighborhood characteristics is still on demand. Here, we take advantages of ortho high-resolution aerial imagery and GEOBIA to provide detailed and accurate neighborhood conditions assessment. The methodology will be elaborated in the following paragraphs.

3 Methodology

3.1 ROIs

Three ROIs: (i) 92130 Carmel Valley, (ii) 92120 Del Cerro and (iii) 92113 Logan Heights were selected to represent high, medium and low SES, respectively. The selection of ROIs was based on household income data, derived from city-data.com.

3.2 GEOBIA

The term of GEOBIA is specifically for GIScience because of requiring the knowledge in geographic information (GI) to segment and classify RS imagery. Moreover, the objects of GEOBIA are usually associated with natural features (e.g., grassland) or artificial features (e.g., building) [10]. These unique emphases set apart GEOBIA from object-based image analysis (OBIA), which is more used in other disciplinesdisciplines (e.g., computer vision and biomedical imaging) [5]. This pilot study applied GEOBIA to extract selected features from NAIP imagery. The features are within two main categories, aesthetics and street morphology, that have potential to reflect SE and DG conditions within neighborhoods.

3.3 Feature Extraction and Building SE Indicators

Vegetation cover and swimming pool were two primary features extracted from NAIP imagery via GEOBIA. The following SE indicators were built from these two features via spatial analysis:

(i) Aesthetics:
- Percentage of vegetation cover (veg_percent)
- Swimming pool density(sp_den)
- Percentage of swimming pool area (sp_percent)
- Number of swimming pool (sp_num)

where: percentage of vegetation cover = the area of vegetation within each census block / each area of census block; swimming pool density = the number of swimming pool within each census block / each area of census block; percentage of swimming pool area = the area of swimming pool within each census block / each area of census block; number of swimming pool = total number of swimming pools within each census block.

(ii) Street Morphology:
- Road density (rd_den)
- Road junction density (rd_junction_den)

where: road density = total road length within each census block / each area of census block; road junction density = total number of road junctions within each census block / each area of census block.

3.4 Selection of SE and DG variables

To validate whether the given SE indicators can represent certain SES or DG makeup, we selected the following SE and DG variables from 2015 American Census Survey data.

- T1115_INCOM: Median income; total household income
- F1115_MHV: Median housing value
- T1115_PROF: Total professional, scientific, management, administrative employed civilians age 16 and older
- P1115_I75: Percent individuals with income below / over $75,000

Coefficient analysis was subsequently performed to assess the association of the SE indicators with surveyed SE and DG data.

4 Results and Discussion

Due to the limitation of pages, this paper will only show the results of high and medium SES in this section.

4.1 Feature Extractions based on GEOBIA

Figure 1a-e shows a subset of NAIP image, swimming pool and vegetation cover in high SES and medium SES, respectively.

Figure 1 The NAIP image with the detection of vegetation cover and swimming pools in high SES and medium SES in San Diego County.

4.2 The correlation of SE Indicators as well as surveyed SE and DG variables

The coefficient outcomes of high SES and medium SES were demonstrated in Table 1 and Table 2, respectively. The highest positive / lowest negative correlation between each SE indicator and individual surveyed variable was highlighted by red / blue.

Table 1 shows that swimming pool density and percentage of swimming pool area have the highest positive correlation with total household income. Number of swimming pool yields the highest positive correlation with total professional, scientific, management, administrative employed civilians age 16 and older. Percentage of vegetation cover yield the highest positive correlation with percent individuals with income below / over $75,000, while road density has the highest negative correlation with this SE variable.

Table 2 demonstrates that three swimming pool-related indicators have the highest positive correlation with median housing value. Percentage of vegetation cover yield the highest positive correlation with total professional, scientific, management, administrative employed civilians age 16 and older. Two road-related indicators have the greatest negative correlation with percent individuals with income below / over $75,000.

Table 1 The coefficient analysis of high SES neighborhood (92130 Carmel Valley).

Features / SE or DG variables	T1115_INCOME	T1115_PROF	T1115_MHV	P1115_I75
sp_den	0.60	-0.18	0.42	0.49
sp_percent	0.62	-0.18	0.44	0.51
sp_num	0.45	0.53	0.45	0.45
veg_percent	0.25	0.28	-0.05	0.37
rd_den	-0.35	-0.27	-0.43	-0.62
rd_junction_den	-0.13	-0.21	0.13	-0.11

Table 2 The coefficient analysis of medium SES neighborhood (92120 Del Cerro).

Features / SE or DG variables	T1115_INCOME	T1115_PROF	T1115_MHV	P1115_I75
sp_den	0.31	-0.16	0.44	0.42
sp_percent	0.29	-0.09	0.47	0.40
sp_num	0.10	-0.01	0.49	0.42
veg_percent	0.29	-0.20	0.26	0.44
rd_den	-0.18	0.08	0.13	-0.22
rd_junction_den	-0.39	0.29	-0.12	-0.57

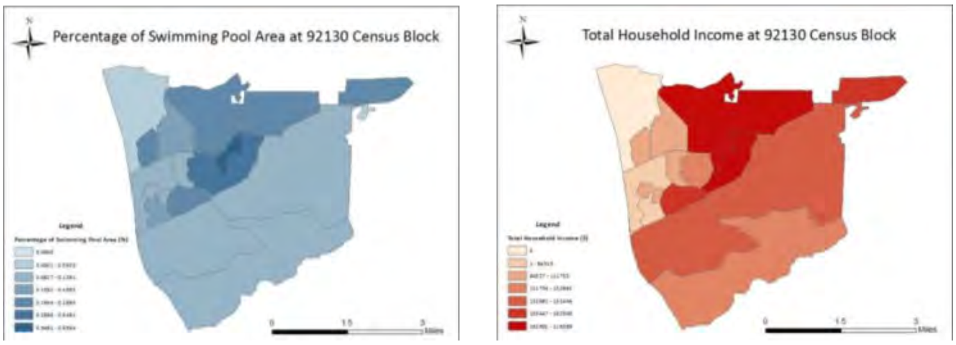

Figure 2 Coefficient=0.62

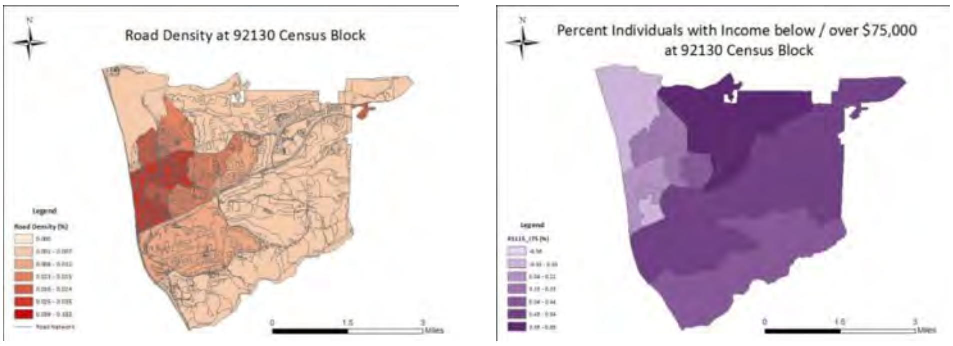

Figure 3 Coefficient=0.-62

Here we highlighted few pairs of SE indicators and the surveyed variables via Geovisualization (Figure 2–4).

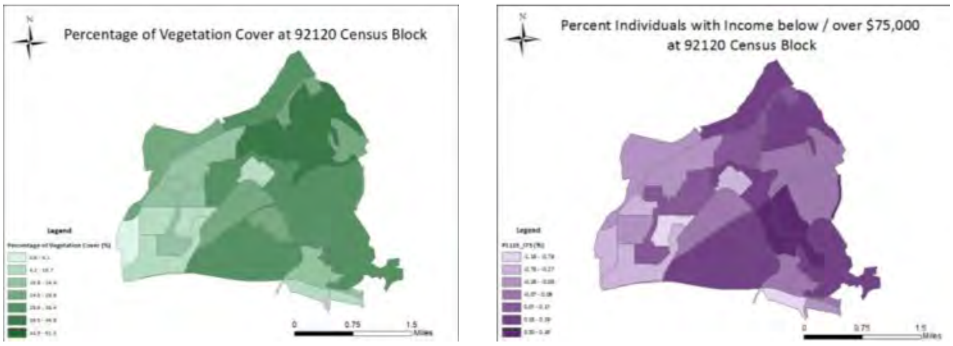

Figure 4 Coefficient=0.44

5 Conclusion

Although all highest positive correlation between SE indicators and surveyed variables are not significant, some certain SE indicators show the potential to assess specific SE or DG conditions within ROIs. Specifically, swimming pool-associated indicators have the greatest correlation with total household income at high SES and with median housing value at medium SES. Vegetation indicator yields the highest correlation with percent individuals with income below / over $75,000 at both high and medium SES. In terms of negative correlation, road density has the greatest negative correlation with percent individuals with income below / over $75,000 at high SES, while road junction density meets the greatest negative correlation with this DG variable. In the near future, we plan to incorporate Google Street View into our framework to provide different angle of features that have potential to represent neighborhood conditions.

References

1. Derek Azar, Ryan Engstrom, Jordan Graesser, and Joshua Comenetz. Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data. *Remote Sensing of Environment*, 130:219–232, 2013.
2. Derek Azar, Jordan Graesser, Ryan Engstrom, Joshua Comenetz, Robert M Leddy Jr, Nancy G Schechtman, and Theresa Andrews. Spatial refinement of census population distribution using remotely sensed estimates of impervious surfaces in haiti. *International Journal of Remote Sensing*, 31(21):5635–5655, 2010.
3. Timnit Gebru, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei. Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, 114(50):13108, 2017. URL: `http://www.pnas.org/content/114/50/13108.abstract`.
4. Scarlett Lin Gomez, Sally L Glaser, Laura A McClure, Sarah J Shema, Melissa Kealey, Theresa HM Keegan, and William A Satariano. The california neighborhoods data system: a new resource for examining the impact of neighborhood characteristics on cancer incidence and outcomes in populations. *Cancer Causes & Control*, 22(4):631–647, 2011.
5. Geoffrey J. Hay and G. Castilla. *Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline*, pages 75–89. Springer, 2008.
6. Theresa HM Keegan, Salma Shariff-Marco, Meera Sangaramoorthy, Jocelyn Koo, Andrew Hertz, Clayton W Schupp, Juan Yang, Esther M John, and Scarlett L Gomez. Neighbor-

hood influences on recreational physical activity and survival after breast cancer. *Cancer Causes & Control*, 25(10):1295–1308, 2014.

7 Xiaojiang Li, Chuanrong Zhang, Weidong Li, Yulia A. Kuzovkina, and Daniel Weiner. Who lives in greener neighborhoods? the distribution of street greenery and its association with residents' socioeconomic conditions in hartford, connecticut, usa. *Urban Forestry & Urban Greening*, 14(4):751–759, 2015. doi:10.1016/j.ufug.2015.07.006.

8 D Stow, A Lopez, C Lippitt, S Hinton, and J Weeks. Object-based classification of residential land use within accra, ghana based on quickbird satellite data. *International journal of remote sensing*, 28(22):5167–5173, 2007.

9 Douglas Stow. *Geographic object-based image change analysis*, pages 565–582. Springer, 2010.

10 Douglas A Stow, Christopher D Lippitt, and John R Weeks. Geographic object-based delineation of neighborhoods of accra, ghana using quickbird satellite imagery. *Photogrammetric Engineering & Remote Sensing*, 76(8):907–914, 2010.

11 Douglas A Stow, John R Weeks, Sory Toure, Lloyd L Coulter, Christopher D Lippitt, and Eric Ashcroft. Urban vegetation cover and vegetation change in accra, ghana: Connection to housing quality. *The Professional Geographer*, 65(3):451–465, 2013.

12 Francisco J Tapiador, Silvania Avelar, Carlos Tavares-Corrêa, and Rainer Zah. Deriving fine-scale socioeconomic information of urban areas using very high-resolution satellite imagery. *International journal of remote sensing*, 32(21):6437–6456, 2011.

13 John R Weeks, Arthur Getis, Douglas A Stow, Allan G Hill, David Rain, Ryan Engstrom, Justin Stoler, Christopher Lippitt, Marta Jankowska, and Anna Carla Lopez-Carr. Connecting the dots between health, poverty and place in accra, ghana. *Annals of the Association of American Geographers*, 102(5):932–941, 2012.