Analyzing Relationship between User-Generated Content and Local Visual Information with Augmented Reality-based Location-based Social Networks

Abstract: Location-based Social Networks (LBSN) have become an important source of spatial data for geographers and GIScientists to acquire knowledge of human-place interactions. A number of studies have used geotagged data from LBSN to investigate how user-generated content (UGC) can be affected or correlated to the external environment. However, local visual information at micro-level, such as brightness, colorfulness, or particular objects/events in the surrounding environment, is usually not captured and thus becomes a missing component in LBSN analysis. To provide a solution to this issue, we argue in this study that the integration of Augmented Reality (AR) and LBSN proves to be a promising avenue. In this first empirical study on AR-based LBSN, we propose a methodological framework to extract and analyze data from AR-based LBSN and demonstrate the framework via a case study with Wallame. Our findings bolster existing psychological findings on color-mood relationship and display intriguing geographic patterns of the influence of local visual information on UGC in social media.

Keywords: augmented reality, location-based social networks, geographically weighted regression, visual data mining, social media

1 Introduction

Online social networks have grown from niche curiosities to a global phenomenon that is responsible for a significant fraction of overall Internet usage (Kumar, Novak, & Tomkins, 2010). People use online social networks for numerous purposes including expressing opinions, connecting with friends, sharing personal experiences, providing professional or educational assistance, marketing, or entertaining customer and clients. If shared publicly, user-generated content (UGC) which can be based on various media such as text, audio, photos, video or 3D objects, becomes easily accessible for data analysis. Online social networks have become an important source of valuable social media data for many domains in today's “big data” era (Ellison, Steinfield, & Lampe, 2007; Mika & Web, 2007; Palen, 2008).

Location-based social networks (LBSN), a subcategory of online social networks, can be regarded as the convergence between location-based services and online social networks (Fusco, Michael, Michael, & Abbas, 2010). Through the prevalence of smartphones equipped with positioning modules, LBSN enable users to easily generate, view, and interact with spatial information. All LBSN data is tightly coupled with location information acquired from multiple sources including GPS sensors embedded in smartphones, place information expressed by user-generated messages or pictures, estimated distance to cell towers or Wi-Fi hotspots, etc. As a result, LBSN provide an important basis to analyze human behavior in society and even predict spatiotemporal patterns of social events.

Traditionally, LBSN acquire locational information as a string of addresses or a set of coordinates. Researchers can then associate such information with relevant UGC to implement further analyses, such as discovering people's behavioral patterns and delineating place boundaries. A number of studies have used geotagged data from LBSN to investigate how user sentiment can be affected or correlated to the external environment. For instance, Hannak et al. (2012) noticed in a study that changes in humidity and temperature can have significant influence on user sentiment on Twitter. Similarly, Li, Wang, and Hovy (2014) compared weather data from NOAA and tweets in same spatiotemporal settings and found out that people's moods on Twitter correspond to a number of weather factors including temperature change, precipitation, and hail.

In spite of works focusing on macro-level external factors like weather, local visual information at micro-level, such as brightness, colorfulness, or particular objects/events in the surrounding environment, is usually not captured and thus becomes a missing component in LBSN analysis. For example, a Twitter user may post a statement "Life is good!" while enjoying a cup of coffee in a café. There may be multiple factors at the scene which contribute to such positive statement: cozy lighting of the environment, a friendly and charming waitress serving him, or the aroma from a cup of coffee in front of him. These important elements, especially color, of the surrounding environment cannot be captured by a mere string of address, a pair of coordinates, or even the name of the café, which are most common types of spatial information in LBSN. Numerous studies in the field of cognitive psychology have shown various kinds of significant color-mood relations across cultures and people (NAz & Epps, 2004; Ou, Luo, Woodcock, Wright, & Application, 2004; Palmer & Schloss, 2010; Wilms & Oberfeld, 2018). While most LBSN allow users to capture and post visual content such as images and videos, such content often fails to reflect the immediate background environment when people post their message. This is largely because that localness cannot always be assumed when users create content on LBSN. People may choose to create content in one spatiotemporal setting and publish it in another. According to Johnson, Sengupta, Schöning, and Hecht (2016), the localness assumption holds in about 75% of cases in three large LBSN (Twitter, Flickr, and Swarm), resulting in potentially misrepresentative conclusions in pertinent studies.

In regard of the need to acquire and associate local visual information with UGC, we argue in this study that the integration of Augmented Reality (AR) and LBSN proves to be a promising approach. AR overlays real world views or scenes with virtual objects or information generated by computers that appear to visually coexist in the same space. We discussed in a previous paper how AR-based LBSN allow users to create, share, store, and modify AR content and highlighted the research potential it would bring to the GIScience community (Liu & Fuhrmann, 2018). Thanks to the spatial nature of AR, AR-based LBSN can provide us with UGC that is organically interwoven with surrounding visual environment.

In this paper, we first discuss the status quo of AR-based LBSN and briefly review relevant practices on geotagged photo analysis. After that, we propose a methodological framework to extract and analyze data from AR-based LBSN. We demonstrate this framework via a case study of data from Wallame, which is currently the most popular AR-based LBSN on market. Findings from the case study are then discussed to explain the link between UGC and local visual information. Being the first study using data from AR-based LBSN, we confirm in our findings the existence of color-mood correlations from an empirical standpoint and contribute to a broader discussion in LBSN analysis of the importance of local visual information. We conclude the paper with discussion on limitation of our study and directions of future work.

2 Related Work on AR-based LBSN

As a crucial technology in computer visualization and mixed reality, AR has been studied and applied in industry for quite a long time, including the field of GIScience, and especially in areas related to geovisualization. As early as of 1997, a prototype system that combines the overlaid 3D graphics of augmented reality with mobile computing was proposed (Feiner, MacIntyre, Höllerer, & Webster, 1997). Sun, Li, Zhang, Wang, and Wu (2007) brought forward the concept of ARGIS, which digitally describes, stores and controls the objective geographic world, meanwhile, integrates such descriptions into the real world, offers the space information of a designated object and supplies the outdoor mobile information interaction. In recent years, as the computing capacity of mobile devices becomes more and more powerful, developments and research on mobile AR have been rapidly emerging. In Paelke, Sester, and Sensing (2010), researchers explored the design space of augmented paper maps in which maps are augmented with additional functionality through a mobile device to achieve a meaningful integration between device and map that combines their respective strengths. In Zhang, Han, Hao, and Lv (2016), the authors designed a mobile AR based underground pipeline information system, which respectively realized a computer vision based version (CV-version) and a sensor based version (Sensor-version).

Despite the fact that the notion of AR has a long-standing history in GIScience, AR-based LBSN are still at a nascent phase and has gained limited attention from academia. The concept of AR-based LBSN is rooted in the notion of AR 2.0, proposed by Schmalstieg, Langlotz, and Billinghurst (2011). Similar to Web 2.0, the concept of AR 2.0 opens the door to crowdsourced AR information by shifting the authority over AR content creation and dissemination from enterprises and governments to the general public. Thankfully, with an increasing exposure to various types of AR-based LBSN (e.g., Sekai Camera, Layar, Wikitude, and Libre Geo Social) on market over the last decade, researchers have started to realize its significance from the GIScience perspective. In MacIntyre, Hill, Rouzati, Gandy, and Davidson (2011), researchers presented the design and implementation of the Argon AR Web Browser and combined it to the popular KML language (the spatial markup language for Google Earth and Maps), allowing users to easily develop and share 2D and 3D content using existing web technologies and Liao and Humphreys (2015) have identified emerging uses of mobile AR-based LBSN through qualitative interviews with Layar users. Scholz and Smith (2016) presented a framework to apply AR-based LBSN in marketing places to engage consumers more effectively. Trojan, Chudáček, and Chrastina (2017) implemented a web-based platform to dynamize static data for popular AR-based LBSN such as Layar and Wikitude.

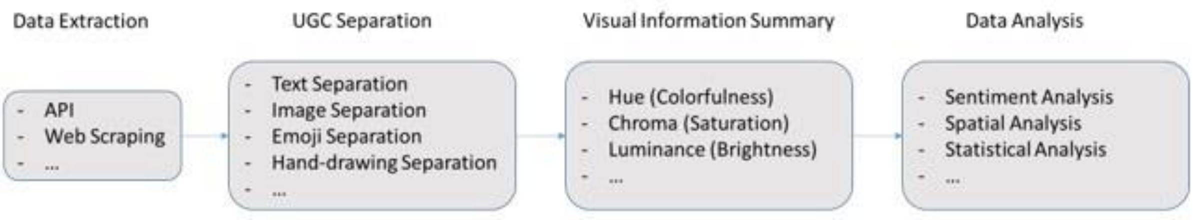
As AR-based LBSN is still a novel concept to both the public and the GIScience community, we have found virtually no study that implement analysis over data extracted from existing AR-based LBSN. However, leaving the AR concept aside, there has been continuous efforts discovering the connection between characteristics of UGC and local visual information, especially through studies of geotagged photos and videos in LBSN. For example, Hu et al. (2015) have matched textual tags of a particular area of interest with preferred photos generated from spectral clustering. In Kaneko and Yanai (2016), the authors proposed a similar system which makes use of "geo-photo tweets" which are tweets including both geotags and photos in order to visually and geographically cluster various events together. Both representative photos and geotags were selected to describe events at places. In Antoniou et al. (2016), researchers extracted and classified geo-tagged photos and their metadata from Flickr, Panoramio and Geograph to delineate a land cover map. A similar study is conducted by Tracewski, Bastin, and Fonte (2017), where the authors used a deep learning network to filter and classify geo-tagged photos for land use characterization. Dunkel (2015) evaluated crowdsourced geo-data photos from Flickr for visualizing landscape perception within the context of environmental perception analysis. He used density of photos, spatial distribution of tags, and label priority ranking altogether to demonstrate that analyzing crowdsourced data may contribute to a more balanced assessment of the perceived landscape.

In light of existing research on geotagged photos in LBSN, we present this study in purpose of proposing AR-based LBSN as a valuable avenue to harvest both UGC and local visual information simultaneously. In recent years, AR as a unique feature of visualization has been widely advertised and applied in video chatting or photo editing in common social media, including Facebook, Snapchat, and iMessage. Besides offering a tool that creates 3D virtual avatars for people, AR also provides a unique opportunity for people to leave virtual traces in the physical reality, which is the focus of our present study on AR-based LBSN. As people virtually interact with the physical reality through AR, we may be able to closely observe how local visual information affect online content generated by social media users.

3 Methodology

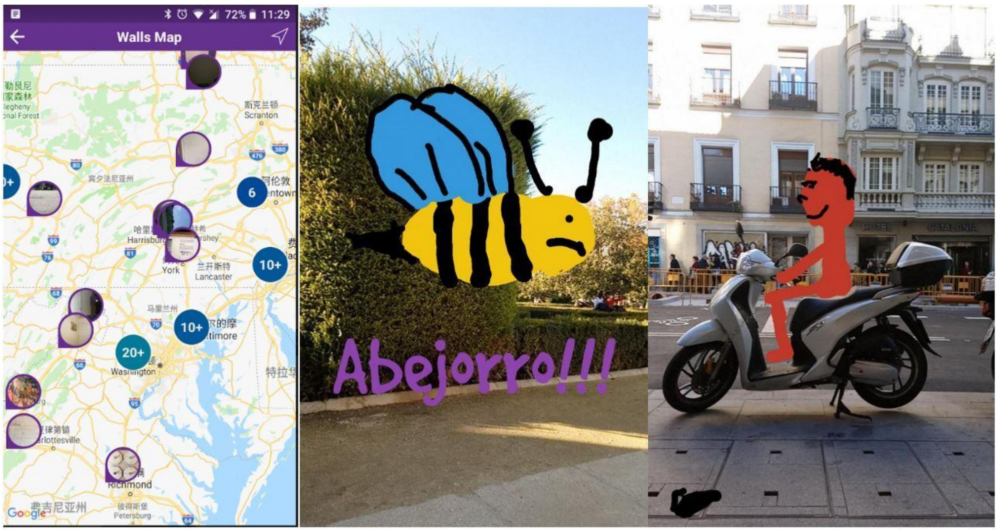
Before implementing the study, we first propose here a methodological framework to describe our general workflow (Figure 1). The framework primarily derives from our work on Wallame, while it is certainly expandable and applicable to other AR-based LBSN.

Figure 1. Framework of AR-based LBSN analysis



Wallame is currently the most popular AR-based LBSN platform with over 200,000 users worldwide (estimation based on the number of Google Play Store downloads). This smartphone application allows users to take a picture of the surrounding environment and overlay this picture with creative content (text, freehand drawing, and image), generating a so-called “wall”. WallaMe then uses the device’s positioning function to geotag this “wall” and aligns it in 3D space with reference to visual patterns of the background image. Once processed, an icon of this “wall” will be indicated on an online map and its UGC will become visible in reality to other WallaMe users who are visiting the same location (Figure 2). Specifically, people can hold their smartphone and use its camera to replicate this “wall” in Wallame, in order to read the virtual message created by its author at this particular location. Users can then perform generic actions on LBSN such as sharing, liking, commenting on each other’s “walls” on Wallame, and adjusting privacy settings so that only selected users or groups have access to the posted content. In summary, WallaMe allows users to create, view and share virtual doodles in reality.

Figure 2. Screen capture of WallaMe “walls” map and two “wall” examples.



As shown in the framework, our study on Wallame is divided into four parts: data extraction, UGC separation, visual information summary, and data analysis. Each part is elaborated below:

3.1 Data Extraction:

AR-based LBSN can host various types of data, including text, image, audio, video, and even 3D objects. To extract these data, commercialized ones, such as Layar and Wikitude, provide publicly accessible application programming interface (API) to the public at a price, which is similar to most well-established LBSN (e.g., Twitter, Facebook, Flickr, and Foursquare). However, smaller AR-based LBSN, such as Wallame and World Brush, often do not provide APIs to the public. Therefore, web scraping approaches are sometimes needed to extract their data, which is the case in our present study with Wallame. As data from Wallame ("walls") are hosted on Google map and are freely accessible in JSON format by the general public, we were able to browse and scrape them in a Microsoft Edge web browser. It should be noted that to avoid legal disputes, we contacted Wallame in the beginning of our study and acquired their permission on our scraping activity.

3.2 UGC Separation:

An intrinsic nature of AR-based LBSN is that virtual and real elements are displayed to users in a mixed form, whether in 2D or 3D. As a result, it is necessary to recognize virtual information created by users and separate it from real information of the visual background in order to implement further analysis. Wallame allows users to generate "walls" by overlaying text, image, emoji and freehand drawing over a background photo. In our study, we focused on recognizing and extracting textual content from the scraped "walls". We used Google's optical character recognition (OCR) tool through its Cloud Vision API to complete this task, largely because Wallame is used all over the world in multiple languages (and English is not the dominant language) and Google's OCR tool supports the recognition of most major languages in the world well (Walker, Fujii, & Popat, 2018). Google's Cloud Vision service has been reliably utilized in numerous GIS-related fields including ecosystem services (Richards & Tunçer, 2017), indoor navigation (Serrão et al., 2015), location-based advertising (Vignesh Kandasamy, Madhu, Gupta, Niveditha, & Bordoloi, 2018), etc. Before we applied the OCR tool to all collected records, we validated its performance using 585 manually labeled records to determine its accuracy rate.

In addition to text recognition, Google's OCR tool also provides the function to delineate box boundary of detected text. Using this information, we were able to crop out textual content from the scraped "walls" and fill missing pixels with surrounding pixels using inward interpolation, so that visual features of the background can be mostly preserved. This step was implemented in MATLAB 2019a.

3.3 Visual Information Summary:

Once UGC was separated from the scraped “walls”, we computed various measures to summarize characteristics of visual information in the background. Our first step was to convert RGB color scheme into HCL (Hue-Chroma-Luminance) color scheme for all images. HCL color space was originally proposed by Sarifuddin and Missaoui (2005) and has been widely adopted by researchers in various visualization scenarios and computer-vision studies including meteorological research (Stauffer, Mayr, Dabernig, & Zeileis, 2015), visual surveillance (Maddalena & Petrosino, 2008), and text extraction (Kim, Lee, & Kim, 2009). HCL is a polar transformation of the uniform CIELUV color space. Compared with common color spaces such as RGB, HSV, or HSL which are more suitable and understandable for machines, HCL is designed to be perceptually based and more human friendly (Zeileis, Hornik, & Murrell, 2009). With HCL color space, we can directly acquire measures on color (hue), saturation (chroma) and brightness (luminance) as people visually perceive the local environment, and make reasonable implications of how these measures might influence people’s mood and expression. The transformation was first implemented in MATLAB 2019a to convert RGB into CIELab color space (color represented in three dimensions of L, a, and b). Then, while the lightness measure L remained unchanged, the conversion of a and b to c (chroma) and h (hue) was implemented using the following formulas (Zeileis et al., 2009):

$$c = \sqrt{a^2 + b^2}$$

$$h = \mathrm{atan2}(b, a) \ (h = h + 2\pi \ \text{ if } h < 0)$$

As RGB values were transformed into HCL values for each pixel in a background image, we computed mean and standard deviation values of all pixels' HCL values in the image to summarize visual characteristics of the background image. In addition, we adopted a useful, human-tested metric of colorfulness suggested by Hasler and Suesstrunk (2003). In their study, 20 non-expert participants were asked to rate 84 images on a 1-7 scale of colorfulness. Then, the authors derived a simple metric that correlated with the results of the participants via a series of experimental calculations. They found through these experiments that a simple opponent color space representation along with the mean and standard deviations of these values correlates to 95.3% of the survey data. The colorfulness metric M was calculated as below (in RGB color space):

$$rg = r - g$$

$$yb = 0.5 * (r + g) - b$$

$$\sigma_{rgyb} = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2}$$

$$\mu_{rgyb} = \sqrt{\mu_{rg}^2 + \mu_{yb}^2}$$

$$M = \sigma_{rgyb} + 0.3 * \mu_{rgyb}$$

It should be noted that there exists more complicated and robust descriptors of image such as SIFT (Lowe, 2004), SURF (Bay, Tuytelaars, & Van Gool, 2006), and WLD (Chen et al., 2009). However, most of them emphasize on feature/pattern recognition in order to enhance machine learning performances in computer vision. In our study, we primarily focused on human perception of color in the local environment as our goal is to understand how local visual information that is immediately perceivable to human eyes relates to UGC on social media. In the future, we would like to incorporate and examine other dimensions of human visual perception such as texture (i.e., complexity and regularity of the background image) in our framework.

3.4 Data Analysis

To summarize characteristics of textual content created by Wallame users, we first employed sentiment analysis on the extracted text. Sentiment analysis is an important component of social media analysis in the geospatial domain and a number of relevant studies have been applied. Ballatore and Adams (2015) built a multi-dimensional model of place emotion based on the vocabulary in the WordNet-Affect lexicon and applied the model to about 100,000 travel blog posts in order to explore the emotional structure of places. Babac and Podobnik (2016) through their sentiment analysis on a harvested Facebook textual dataset found that men and women similarly express hard emotions such as anger or fear, while there is a significant difference in expressing soft emotions such as joy or sadness regarding to soccer games. Wang et al. (2016) performed text mining on geotagged tweets to detect people's emotions in the response to a wildfire. Wang and Zhou (2016) performed sentiment analysis on citywide TripAdvisor hotel reviews and found that spatial dependence exists in customer satisfaction. Sentiment analysis can also be implemented beyond textual information. For example, Kang et al. (2019) found out the relationship between human emotions and environmental factors based on a happiness ranking list of places generated by human emotions calculated over 2 million faces detected from over 6 million photos. Raja et al. (2018) presented a taxonomy on affect recognition from a variety of physical sensors including phone sensors, body sensors, vehicle sensors, etc.

In this study, because of Google Cloud's wide coverage of languages and relatively reliable performance, we again adopted its sentiment analysis service provided via the Cloud Natural Language API which has been used extensively in commercial applications as well as in academia (Tran, Nguyen, Nguyen, & Golen, 2018). Google's sentiment analysis generates two measures: score and magnitude. The score of a document's sentiment ranges from -1 to 1 and indicates the overall emotion of a document (-1 being most negative and 1 being most positive). The magnitude of a document's sentiment ranges from 1 to infinity and indicates how much emotional content is present within the document, which is often proportional to the length of the document. The analysis was implemented in Python 3.4. In addition to these two measures, we computed text length (i.e., number of characters) of each Wallame record and normalized the magnitude measure with it. We used sentiment score, normalized sentiment magnitude, and text length as dependent variables of UGC in our study. In terms of independent variables, we used all six calculated visual measures (mean and standard deviation values of hue, chroma, and luminance respectively) plus a colorfulness measure computed according to Hasler and Suesstrunk (2003). Each Wallame "wall" contains a pair of latitude/longitude coordinates which enables us to perform spatial analysis.

As all variables are collected, we first employed an ordinary least squares (OLS) regression model on each dependent variable and all independent variables to examine the general association between them. Each OLS model was carried out with a spatial weights matrix, which applied a method of adaptive bandwidth optimized by AICc (corrected Akaike Information Criterion) values on a Gaussian kernel, to diagnose the model's spatial dependence. Such parameters were chosen due to the study's exploratory nature, as we lack prior knowledge on data from AR-based LBSN. With the spatial weights matrix, we applied Moran's I analysis, which measures how autocorrelated features are spatially, on each model's residuals and found significant signs of spatial autocorrelation. This suggested that a spatial model, such as geographically weighted regression model (GWR), was needed to better describe our data than OLS. GWR is a variation of the OLS model by adding a level of modeling sophistication that allows the relationships between the independent and dependent variables to vary by locality (Fotheringham, Brunsdon, & Charlton, 2003). It is widely used as an exploratory technique to analyze features that are spatially heterogeneous. Briefly speaking, GWR constructs an OLS equation for every feature in the dataset, which incorporates the dependent and independent variables of features falling within the aforementioned bandwidth of each target feature. The OLS and GWR models were implemented in GeoDa (Anselin, Ibnu, & Youngihn, 2006) and ArcGIS Pro respectively.

4 Findings

4.1 Dataset

We scraped data directly from the Google map embedded in Wallame for a week in January 2019. Each Wallame record ("wall") is consisted of a picture with UGC overlaid on the background, anonymous user ID, as well as associated spatiotemporal information (i.e., geographic coordinates and time stamp). As shown in Table 1, in total we acquired 46,591 unique "walls" globally, generated by 15,434 unique WallaMe users. We detected 31,911 "walls" containing textual information, of which 28,801 "walls" had valid sentiment score and magnitude information, and 16,228 of them were non-zero (meaning that the sentiment is not neutral). From "walls" containing text, we randomly chose 585 and manually labeled them to examine the accuracy of Google's OCR tool. We acquired a correct rate of 95.8% which was considered to be sufficient to proceed with further analysis.

(Table 1 inserted here)

Figure 3.1-3.3 displays heat maps of all collected "walls" across the world as well as in different regions, from which we can observe that Wallame is mostly used in central Europe, metropolitan areas in North America, and parts of Latin America. Countries with high Wallame usage include Poland, Israel, USA, Spain, Hungary, Mexico, and Denmark.

Figure 3.1. Global heat map of public “walls” on Wallame

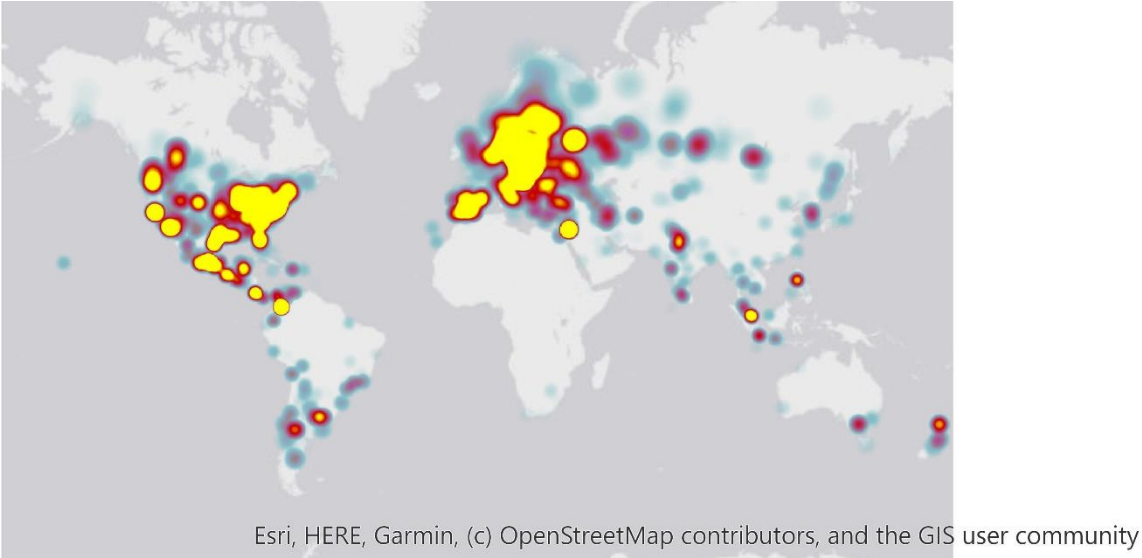


Figure 3.2. Heat map of public “walls” on Wallame in North and Central America

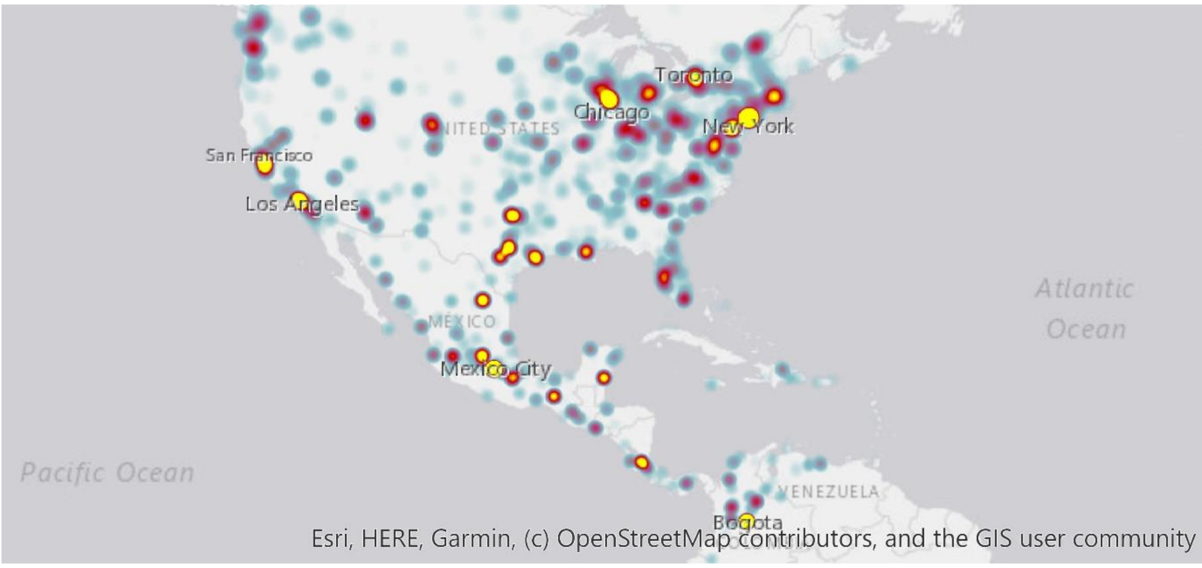
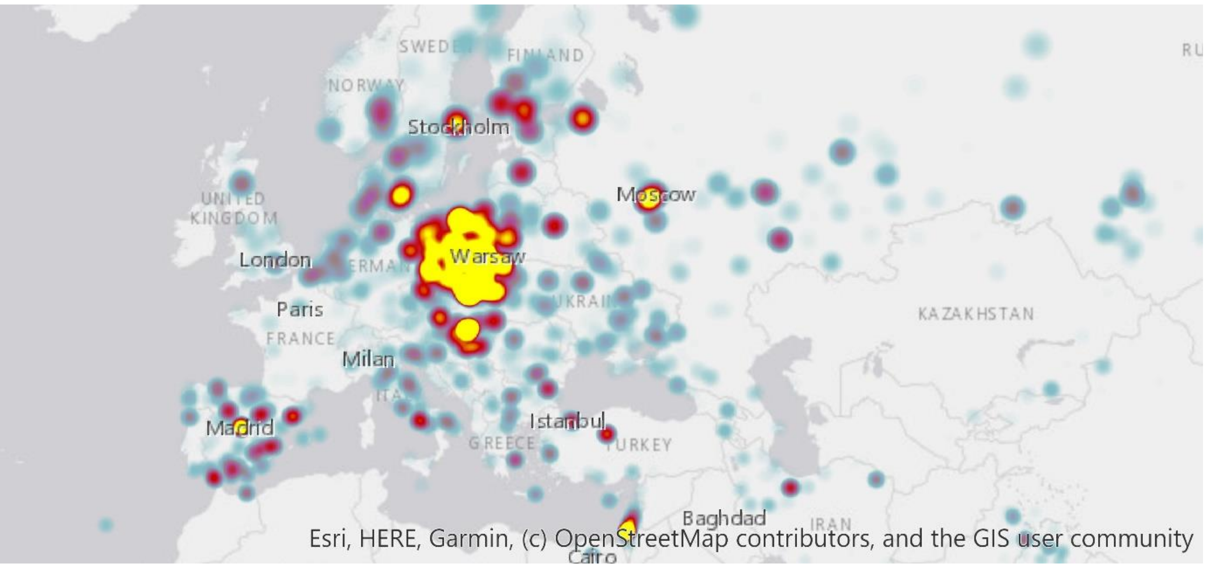


Figure 3.3. Heat map of public “walls” on Wallame in Europe



4.2 Model Results

Table 2 displays statistically significant ($p<0.05$) coefficients in all three OLS models. We can see that the colorfulness metric as computed in Hasler and Suesstrunk (2003) is the only factor that is significantly correlated to sentiment score, indicating a positive correlation between a colorful environment as perceived by human eyes and people's mood expressed online. In addition, most visual factors are negatively correlated to normalized sentiment magnitude, suggesting that people are likely to be more emotional online in dark places with monochromatic color background. In terms of text length, it appears that people tend to write more content online in a bright environment with hard lighting.

(Table 2 inserted here)

Significantly large results from Moran's I analysis on the standard residuals of all three OLS models indicate that a spatial model was needed to better fit our data.

Therefore, we employed GWR models on the dependent variables to observe how local visual information correlates to UGC at local scales. Table 3 shows a comparison of model diagnostics between our OLS and GWR models, including adjusted R^2 (the larger the better), AICc (the smaller the better), and Moran's I on standard residuals (the closer to 0 the better). We can notice that for all three dependent variables, the GWR model outperforms the OLS model by obtaining larger adjusted R^2, smaller (or very similar) AICc, and Moran's I of standard residuals that are remarkably closer to 0 (or statistically insignificant in the case of normalized sentiment magnitude).

(Table 3 inserted here)

The GWR coefficient maps display some intriguing trends. Evident spatial patterns exist across all independent variables for nearly all coefficients, and three instances are presented here. Figure 4.1 shows that for the sentiment score variable, influence of local colorfulness clearly increases from southeast to northwest in Europe. Figure 4.2 shows that for the normalized sentiment magnitude variable, influence of average chroma is particularly low in Hungary and Balkan area but becomes very high in Italy and Ukraine. Figure 4.3 shows that for the text length variable, Italy and Florida are two noticeable local clusters, high and low respectively, of the coefficient of luminance standard deviation. Such spatially clustering phenomenon, though varying in extent, can be observed for all significant coefficients in our GWR models.

Figure 4.1. Map of regression coefficient of colorfulness (dependent variable: sentiment score)

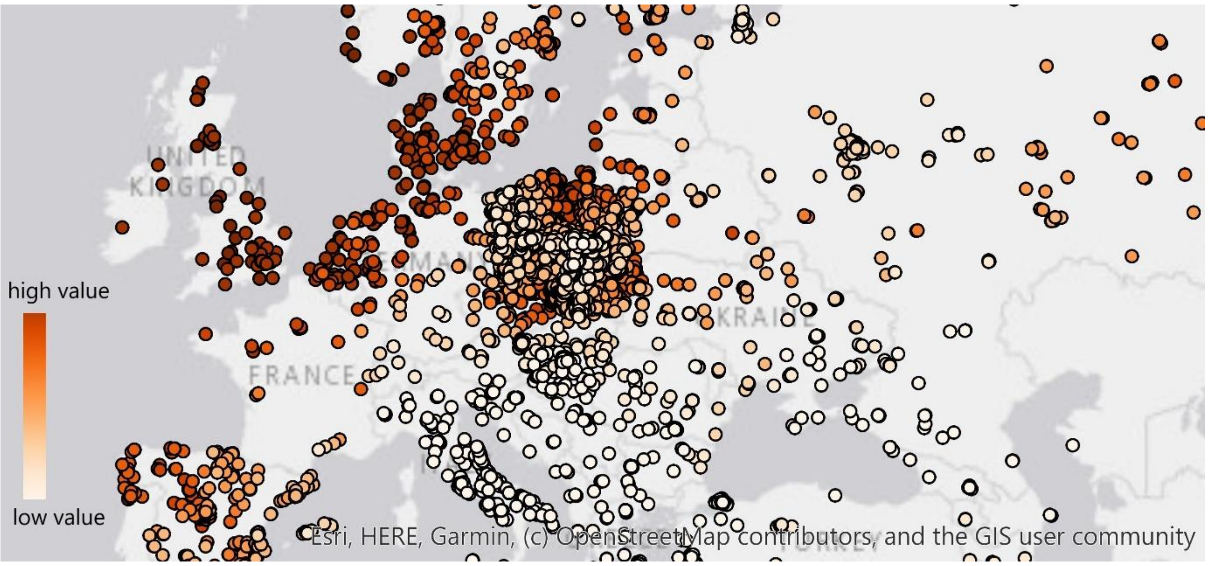


Figure 4.2. Map of regression coefficient of average chroma (dependent variable: normalized sentiment magnitude)

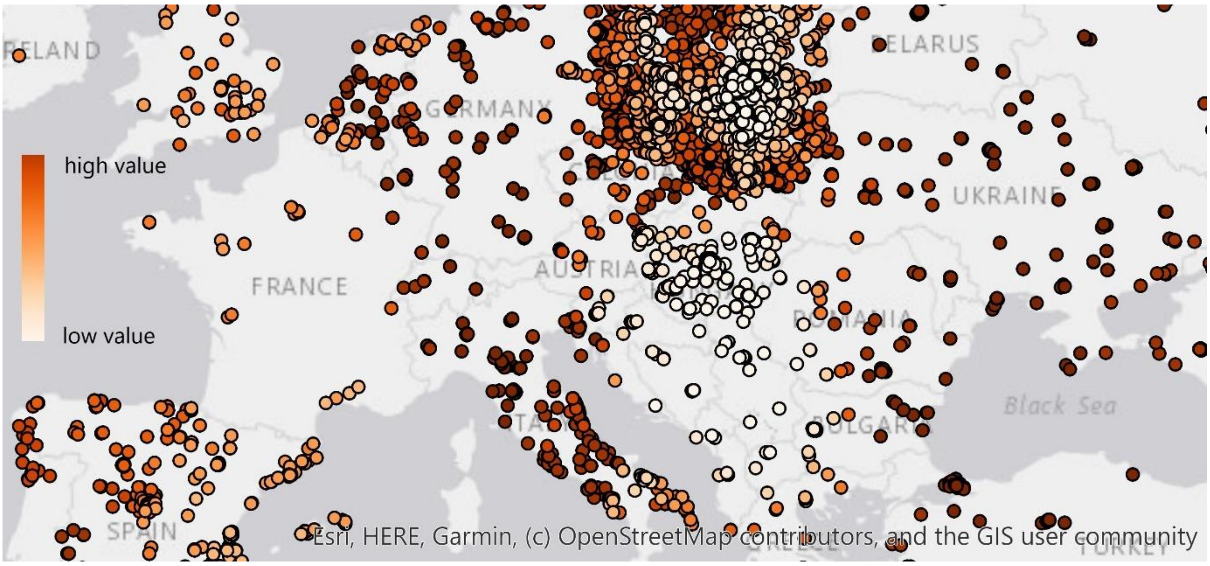
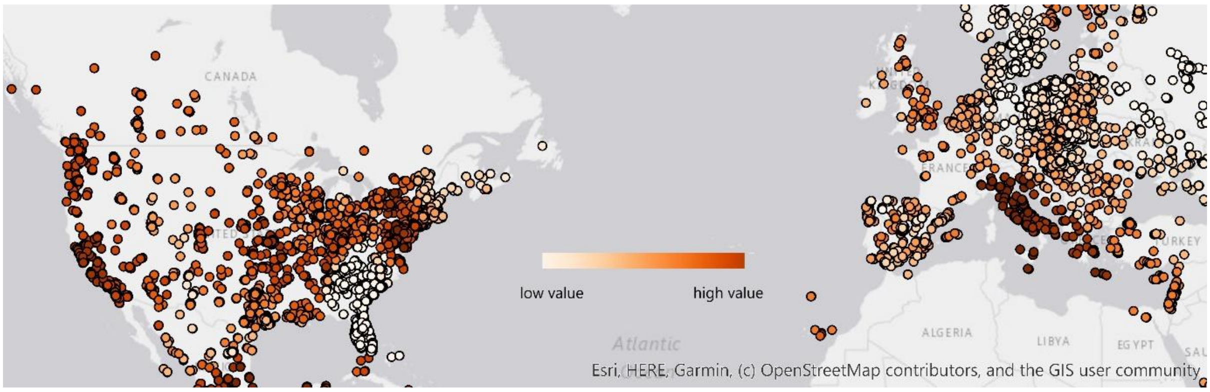


Figure 4.3. Map of regression coefficient of luminance standard deviation (dependent variable: text length)



5 Discussion

Is what we see related to what we say on social media? Our study affirms this. OLS results from our study suggest that the polarity and extent of people's emotion are significantly correlated to a number of visual environmental factors including colorfulness, brightness, and visual intensity. In addition, these visual factors also associate with the length of text people post online.

The effect of visual information, color in specific, on people's mood is a long-standing research topic and a number of theories have been established in the field of psychology. In Ou et al. (2004), four color-emotion models (warm–cool, heavy–light, active–passive, and hard–soft) were developed and people had a uniform response to them regardless of their culture. In NAz and Epps (2004), a survey among 98 college students revealed that principle hues comprised the highest number of positive emotional responses, followed by the intermediate hues and the achromatic colors. Palmer and Schloss (2010) in their research advanced our understanding on color-mood relationship by articulating an ecological valence theory in which color preferences arise from people's average affective responses to color-associated objects. More recently, Wilms and Oberfeld (2018) controlled all three perceptual color dimensions (hue, saturation, and brightness) in their study and found out that effects on human emotion are not only determined by the hue of a color, as is often assumed, but by all the three color dimensions as well as their interactions. As most of these psychological studies use surveys or psychophysical experiments among a small group of observers, our study on AR-based LBSN bolsters their conclusion by offering a novel approach to collect first-hand data online at large scale.

Some researchers have probed image-sharing social media such as Instagram to see if visual information contained in images can help us detect people's emotional status. In Reece and Danforth (2017), the authors extracted statistical features from Instagram photos to diagnose depression of users. Their models outperformed general practitioners' average unassisted diagnostic success rate. Manikonda and De Choudhury (2017) also extracted visual attributes in Instagram photos, compared them with linguistic attributes in accompanying texts, and found that images with a variety of distinct visual cues serve as a vehicle of expression of distress, helplessness and social isolation to certain individuals. Our present study complements these studies by analyzing visual information as an external contributing factor, instead of being an internal expression, of people's emotions. It should be noted that different from both Instagram-based and survey/experiment-based psychological studies, our study on AR-based LBSN collected local visual information as ambient data, that is, information neither purposefully presented to observers nor deliberately generated by social media users.

Therefore, we hope that our work would be meaningful to future studies on how people's emotions are associated with surrounding visual information.

Because residuals from our OLS models display some spatial autocorrelation, it is reasonable that our GWR models outperform them as the geographic factor is accounted for. However, it is somewhat surprising to us that clear geographical patterns exist for nearly all coefficients in all three models. Considering the wide usage of Wallame globally, it is plausible that people living in different environments, speaking different languages, and originating from different cultures hold different sensitivity levels towards local visual information. For example, in the sentiment score model, high clusters of the coefficient of colorfulness can be generally found at high-population areas, which may suggest that urban-living people's mood are more sensitive to the variety and vibrancy of colors in their daily life. Although it is difficult to give a satisfactory explanation to all geographical patterns, our study definitely discovered new possibilities and research directions on the topic of color-mood relationship.

Currently, AR-based LBSN is still a niche curiosity to limited population. However, with recent development in both AR-supporting hardware (e.g,. high-resolution screen and light, wearable device) and AR applications (e.g., gaming, advertising, education, etc.), a future of AR-based social media is prospective. In this regard, our research accentuates the value of AR-based LBSN and provides a novel framework towards analyzing UGC and local visual information through it. We believe that appropriate analytical methods of AR information, such as the one introduced in this paper, will prepare researchers in the GIScience domain to better understand micro-level human-place interactions.

6 Limitation and Future Works

Being the first empirical study on AR-based LBSN, our present work has a number of limitations. First, we only worked with Wallame and only extracted textual information from its records as UGC, while a large portion of AR content is comprised of emojis, images, videos, or virtual 3D models. In the future, we would like to expand our work to analyze non-textual information from AR-based LBSN. Second, as mentioned in section 3, the local visual environment is a complex subject that can only be partially represented by colors. For future studies, we would like to incorporate other dimensions of human visual perception, such as shape, texture, or even specific objects/events/people in the visual background, into our model to better describe the relationship between local visual environment and characteristics of UGC. Third, because currently Wallame does not provide an API to access its AR content, we had to use a scraping method and separate UGC from its visual background using Google's OCR tool. While the OCR tool has been tested and proven effective by several previous studies as well as our own validation dataset, there inevitably remains some inaccuracies and loss of data which can propagate into the analysis result. In the future, as Wallame (or other AR-based LBSN) becomes more popular and provides API to the public, we look forward to working with its data in better quality.

Reference

Antoniou, V., Fonte, C., See, L., Estima, J., Arsanjani, J., Lupia, F., . . . Fritz, S. J. I. I. J. o. G.-I. (2016). Investigating the feasibility of geo-tagged photographs as sources of land cover input data. *5*(5), 64.

Babac, M. B., & Podobnik, V. (2016). A sentiment analysis of who participates, how and why, at social media sport websites. *Online Information Review*.

Ballatore, A., & Adams, B. (2015). *Extracting place emotions from travel blogs.* Paper presented at the Proceedings of AGILE.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). *Surf: Speeded up robust features.* Paper presented at the European conference on computer vision.

Chen, J., Shan, S., He, C., Zhao, G., Pietikainen, M., Chen, X., & Gao, W. (2009). WLD: A robust local image descriptor. *IEEE transactions on pattern analysis and machine intelligence, 32*(9), 1705-1720.

Dunkel, A. (2015). Visualizing the perceived environment using crowdsourced photo geodata. *Landscape Urban Planning, 142*, 173-186.

Ellison, N. B., Steinfield, C., & Lampe, C. J. J. o. C. M. C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *12*(4), 1143-1168.

Feiner, S., MacIntyre, B., Höllerer, T., & Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal Technologies, 1*(4), 208-217.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*: John Wiley & Sons.

Fusco, S. J., Michael, K., Michael, M., & Abbas, R. (2010). *Exploring the social implications of location based social networking: an inquiry into the perceived positive and negative impacts of using LBSN between friends.* Paper presented at the Mobile Business and 2010 Ninth Global Mobility Roundtable (ICMB-

GMR), 2010 Ninth International Conference on.

Hannak, A., Anderson, E., Barrett, L. F., Lehmann, S., Mislove, A., & Riedewald, M. (2012). *Tweetin'in the rain: Exploring societal-scale effects of weather on mood.* Paper presented at the Sixth International AAAI Conference on Weblogs and Social Media.

Hasler, D., & Suesstrunk, S. E. (2003). *Measuring colorfulness in natural images.* Paper presented at the Human vision and electronic imaging VIII.

Hu, Y., Gao, S., Janowicz, K., Yu, B., Li, W., & Prasad, S. (2015). Extracting and understanding urban areas of interest using geotagged photos. *Computers, Environment and Urban Systems, 54*, 240-254.

Johnson, I. L., Sengupta, S., Schöning, J., & Hecht, B. (2016). *The geography and importance of localness in geotagged social media.* Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.

Kaneko, T., & Yanai, K. (2016). Event photo mining from twitter using keyword bursts and image clustering. *Neurocomputing, 172*, 143-158.

Kang, Y., Jia, Q., Gao, S., Zeng, X., Wang, Y., Angsuesser, S., . . . Fei, T. (2019). Extracting human emotions at different places based on facial expressions and spatial clustering analysis. *Transactions in GIS, 23*(3), 450-480.

Kim, E., Lee, S., & Kim, J. (2009). *Scene text extraction using focus of mobile camera.* Paper presented at the 2009 10th International Conference on Document Analysis and Recognition.

Kumar, R., Novak, J., & Tomkins, A. (2010). Structure and evolution of online social networks. In *Link mining: models, algorithms, and applications* (pp. 337-357): Springer.

Li, J., Wang, X., & Hovy, E. (2014). *What a nasty day: Exploring mood-weather relationship from twitter.* Paper presented at the proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management.

Liao, T., & Humphreys, L. (2015). Layar-ed places: Using mobile augmented reality to tactically reengage, reproduce, and reappropriate public space. *New Media & Society, 17*(9), 1418-1435.

Liu, C., & Fuhrmann, S. J. T. i. G. (2018). Enriching the GIScience research agenda: Fusing augmented reality and location-based social networks. *22*(3), 775-788.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision, 60*(2), 91-110.

MacIntyre, B., Hill, A., Rouzati, H., Gandy, M., & Davidson, B. (2011). *The Argon AR Web Browser and standards-based AR application environment.* Paper presented at the 2011 10th IEEE International Symposium on Mixed and Augmented Reality.

Maddalena, L., & Petrosino, A. (2008). A self-organizing approach to background subtraction for visual surveillance applications. *IEEE Transactions on Image Processing, 17*(7), 1168-1177.

Manikonda, L., & De Choudhury, M. (2017). *Modeling and understanding visual attributes of mental health disclosures in social media.* Paper presented at the Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.

Mika, P. J. W. s. s., services, & Web, a. o. t. W. W. (2007). Ontologies are us: A unified model of social networks and semantics. *5*(1), 5-15.

NAz, K., & Epps, H. J. C. S. J. (2004). Relationship between color and emotion: A study of college students. *38*(3), 396.

Ou, L. C., Luo, M. R., Woodcock, A., Wright, A. J. C. R., & Application. (2004). A study of colour emotion and colour preference. Part I: Colour emotions for single colours. *29*(3), 232-240.

Paelke, V., Sester, M. J. I. J. o. P., & Sensing, R. (2010). Augmented paper maps: Exploring the design space of a mixed reality system. *65*(3), 256-265.

Palen, L. J. E. Q. (2008). Online social media in crisis events. *31*(3), 76-78.

Palmer, S. E., & Schloss, K. B. J. P. o. t. N. A. o. S. (2010). An ecological valence theory of human color preference. 200906172.

Raja, M., Exler, A., Hemminki, S., Konomi, S. i., Sigg, S., & Inoue, S. (2018). Towards pervasive geospatial affect perception. *GeoInformatica, 22*(1), 143-169.

Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science, 6*(1), 15.

Richards, D. R., & Tunçer, B. (2017). Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services*.

Sarifuddin, M., & Missaoui, R. (2005). *A new perceptually uniform color space with associated color similarity measure for content-based image and video retrieval.* Paper presented at the Proc. of ACM SIGIR 2005 workshop on multimedia information retrieval (MMIR 2005).

Schmalstieg, D., Langlotz, T., & Billinghurst, M. (2011). Augmented Reality 2.0. In *Virtual realities* (pp. 13-37): Springer.

Scholz, J., & Smith, A. N. (2016). Augmented reality: Designing immersive experiences that maximize consumer engagement. *Business Horizons, 59*(2), 149-161.

Serrão, M., Shahrabadi, S., Moreno, M., José, J., Rodrigues, J. I., Rodrigues, J. M., & du Buf, J. H. (2015). Computer vision and GIS for the navigation of blind persons in buildings. *Universal Access in the Information Society, 14*(1), 67-80.

Stauffer, R., Mayr, G. J., Dabernig, M., & Zeileis, A. (2015). Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations. *Bulletin of the American Meteorological Society, 96*(2), 203-216.

Sun, M., Li, M., Zhang, F., Wang, Z., & Wu, D. (2007). *Hybrid tracking for augmented reality GIS registration.* Paper presented at the 2007 Japan-China Joint Workshop on Frontier of Computer Science and Technology (FCST 2007).

Tracewski, L., Bastin, L., & Fonte, C. C. J. G.-s. i. s. (2017). Repurposing a deep learning network to filter and classify volunteered photographs for land cover and land use characterization. *20*(3), 252-268.

Tran, T., Nguyen, D., Nguyen, A., & Golen, E. (2018). *Sentiment analysis of marijuana content via facebook emoji-based reactions.* Paper presented at the 2018 IEEE International Conference on Communications (ICC).

Trojan, J., Chudáček, S., & Chrastina, P. (2017). Augmented Reality As A New Way Of Exploring The City: Unified Platform For Data Providers. *WIT Transactions on Ecology and the Environment, 210*, 161-169.

Vignesh Kandasamy, S., Madhu, A., Gupta, P. K., Niveditha, A., & Bordoloi, K. (2018). Location based Advertising for Mass Marketing. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42*, 5.

Walker, J., Fujii, Y., & Popat, A. C. (2018). *A web-based ocr service for documents.* Paper presented at the Proceedings of the 13th IAPR International Workshop on Document Analysis Systems (DAS), Vienna, Austria.

Wilms, L., & Oberfeld, D. J. P. r. (2018). Color and emotion: effects of hue, saturation, and brightness. *82*(5), 896-914.

Zeileis, A., Hornik, K., & Murrell, P. (2009). Escaping RGBland: Selecting colors for statistical graphics. *Computational Statistics & Data Analysis, 53*(9), 3259-3270.

Zhang, X., Han, Y., Hao, D., & Lv, Z. (2016). ARGIS-based outdoor underground pipeline information system.

Journal of Visual Communication and Image Representation, 40, 779-790.

Table 1. Summary of scraped Wallame data

Total number of walls	46,591
Number of walls containing text	31,911
Number of walls with sentiment	28,801
Number of walls with non-neutral sentiment	16,228
OCR correct rate (based on N=585 sample)	95.8%
Number of unique users	15,434

Table 2. Significant ($p<0.05$) OLS[1] coefficients of three dependent variables

	Sentiment Score	Normalized Sentiment Magnitude	Text Length
Hue mean	p>0.05	p>0.05	-0.1158
Hue std.dev[2]	p>0.05	0.0000	0.3014
Chroma mean	p>0.05	-0.0001	p>0.05
Chroma std.dev	p>0.05	-0.0001	p>0.05
Luminance mean	p>0.05	-0.0001	1.1060
Luminance std.dev	p>0.05	-0.0001	1.2987
Colorfulness	0.0004	0.0000	-0.2896

[1] OLS: ordinary least squares

[2] Std.dev: standard deviation

Table 3. Comparison of model diagnostics between OLS and GWR[3] models

		Sentiment Score	Normalized Sentiment Magnitude	Text Length
OLS	Adjusted R^2	0.0003	0.0102	0.0321
	AIC[4]	2787.5100	-159807	353127
	Moran's I (std.res[5])	0.1351	0.1163	0.2631
GWR	Adjusted R^2	0.098	0.0154	0.0704
	AIC	-51811.9267	-159729.6736	352193
	Moran's I (std.res)	0.0016	p>0.05	0.0051

3 GWR: geographically weighted regression

4 AIC: Akaike information criterion

5 Std.res: standard residuals

Figure Legends

Figure 1. Framework of working with AR-based LBSN

Figure 2. Screen capture of WallaMe "walls" map and two "wall" examples.

Figure 3.1. Global heat map of public "walls" on Wallame

Figure 3.2. Heat map of public "walls" on Wallame in North and Central America

Figure 3.3. Heat map of public "walls" on Wallame in Europe

Figure 4.1. Map of regression coefficient of colorfulness (dependent variable: sentiment score)

Figure 4.2. Map of regression coefficient of average chroma (dependent variable: normalized sentiment magnitude)

Figure 4.3. Map of regression coefficient of luminance standard deviation (dependent variable: text length)