Applied Geography 93 (2018) 103–111

Contents lists available at ScienceDirect

Applied Geography

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

From image descriptions to perceived sounds and sources in landscape: Analyzing aural experience through text

Olga Chesnokova*, Ross S. Purves

Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

ARTICLE INFO

Keywords:
Sound experiences
User generated content
Natural language processing
Landscape

ABSTRACT

The importance of perception through all the senses has been recognized in previous studies on landscape preference, but data on aural perception, as opposed to the visual, remains rare. We seek to bridge this gap by analyzing texts that describe more than 3.5 million georeferenced images, created by more than 12000 volunteers in the Geograph project. Our analysis commences by extracting and automatically disambiguating descriptions that potentially contain verbs and nouns of sound (e.g. rustle, bellow, echo, noise) and adjectives of sound intensity (e.g. deafening, quiet, vociferous). Using random forests we classify more than 8000 descriptions based on the type of sound emitter into geophony (e.g. rustling wind, bubbling waterfall), biophony (e.g. gulls calling, bellowing stag), anthrophony (e.g. roaring jets, rumbling traffic) and perceived absence of sound (e.g. not a sound can be heard) with a precision of 0.81. Further, we additionally classify these descriptions as negative, neutral and positive using an Opinion Lexicon and GloVe word embeddings. Our results show that sentiment classification gives an additional level of understanding of descriptions classified into different types of sound emitters. We see that geophony, biophony and anthrophony cannot be uniquely classified as positive or negative. Our results demonstrate how text can provide a valuable, complementary to field-based studies, source of spatially-referenced information about aural landscape perception.

1. Introduction and background

What is the contribution of sounds to the way people perceive landscapes? And how can we gather information about such perceptions over large spatial scales? User Generated Content (UGC) has proven to be a suitable source for research questions dealing with such phenomena as people's perception of sense of place (Jenkins, Croitoru, Crooks, & Stefanidis, 2016), conceptualizations of natural features (Derungs & Purves, 2016), olfactory perception (Quercia & Schifanella, 2015), visual perception of landscapes (van Zanten et al. 2016) and assessment of the collective value of protected areas (Levin, Mark, & Brown, 2017). In this study we investigate another subjective phenomenon, namely aural perception of landscapes in UGC, with the underlying future aim of integrating sound information in landscape preference models.

Aural perception is an important constituent in landscape preference assessment (Brown & Brabyn, 2012; Sherrouse, Clement, & Semmens, 2011; Tudor, 2014) and is typically integrated using field surveys (Pilcher, Newman, & Manning, 2009) or laboratory sessions (Benfield, Bell, Troup, & Soderstrom, 2010; Manyoky, Wissen Hayek, Heutschi, Pieren, & Grêt-Regamey, 2014). However, these methods do not allow large regions to be characterized and are time consuming. We assume that aural perception of landscape is present in some written descriptions associated with photographs uploaded by individuals in UGC since photographs have been argued to be a good source of information related to shared experiences of places (Fisher & Unwin, 2005), and sound is one important element of such experiences. The following example vividly illustrates such use of language at an individual level: "If you press your nose to the computer screen, you might just catch the scent of the wild garlic, and if you listen carefully you should hear the song of willow warbler and blackcap.[1]" However, if we wish to analyze such descriptions, then important questions remain with respect to how they can be extracted, how common they are, and what properties they have.

1.1. Sound experiences

Although our sensory experience of nature is by definition multisensory, the visual is often privileged in both research and policy. Thus, despite the introduction of 'soundscape', 'acoustic ecology' and

* Corresponding author.
E-mail addresses: olga.chesnokova@geo.uzh.ch (O. Chesnokova), ross.purves@geo.uzh.ch (R.S. Purves).

[1] http://www.geograph.org.uk/photo/824881.

https://doi.org/10.1016/j.apgeog.2018.02.014
Received 22 December 2017; Received in revised form 20 February 2018; Accepted 21 February 2018

'soundscape ecology' (Southworth, 1969; Schafer, 1993; Pijanowski, Farina, Gage, Dumyahn, and Krause, 2011), aural perception is often of secondary importance in modelling landscape preferences. To relate sound to landscape preference it is important to consider the influence of perceived sound emitters as natural or unnatural (Fisher, 1999), rather than simply decibel values, since we do not hear abstract sounds, but "we hear the way *things sound*" (p. 40 Morton, 2009). Krause (2008), in collaboration with Gage, developed a useful taxonomy for sound emitters in landscape, identifying geophony (non-biological natural sounds), biophony (sounds produced by animals) and anthrophony (human-generated sounds).

Fisher (1999) claims that as soon as we perceive a sound as natural it has a positive aesthetic quality. Thus, similar sounds when perceived as being emitted by a jet engine or a waterfall would be considered unpleasant or "majestically powerful," respectively (p. 28–29 Fisher, 1999). Carles, Barrio, and De Lucio (1999) in their study of sound influence on landscape value note that similar to findings in visual perception, water sounds are typically positively connoted. Furthermore, discordant scenes, for example with positive visual (e.g. a water body) and negative aural cues (e.g. the sound of a busy road) were considered to be especially disturbing. In a series of soundwalks reported on by Pérez-Martínez, Torija, and Ruiz (2018), visitors characterized the sounds of certain emitters as being unpleasant, with, for instance, bird calls dominating, and thus detracting from landscape aesthetics. The negative effects of anthrophony are reported by Pilcher et al. (2009) to be especially important in wild areas, natural parks and other protected areas, where the intrusion of anthropogenic sounds is more disturbing. All of these studies provide us with useful clues as to how aural perception influences landscape perception, but none of them are easily applied across large regions.

1.2. User generated content and extraction of subjective phenomena from language

Our starting point is the hypothesis, based on an initial exploration of content, that UGC can be used to estimate aural perception of landscapes in the British Isles. This hypothesis is supported by previous work which has shown that, for example, tags associated with Flickr images or Tweets content have strong associations with place (Jenkins et al. 2016; Rattenbury, Good, & Naaman, 2007) or that olfactory perception of urban landscapes can be explored through UGC (Quercia & Schifanella, 2015). The same team of researchers also generated maps of urban noises using tags (Aiello, Schifanella, Quercia, & Aletta, 2016) by relating particular terms (e.g. church, car, dog) to particular sounds. However, their study implicitly links sounds to terms without clear evidence of the actual perception of sounds at a location. Similarly, analysis of spectrograms recorded by acoustic sensors (e.g. Pijanowski, Villanueva-Rivera, et al. 2011) does not allow a direct link between the presence of sounds and their perception by humans.

In this paper we build on previous work in two key ways. Firstly, the methods currently used in estimation of aural perception are time consuming and are not suitable for large regions. Using UGC provides an opportunity to explore the link between aural perception and landscapes across the British Isles. Secondly, in the case of recorded sounds presented in laboratory sessions the nature of a sound is abstracted from its context in the landscape. Therefore, we here set out to explore the efficacy of a range of methods for extracting and classifying textual descriptions related to aural perception of sounds, and apply sentiment analysis methods to explore the extent to which landscape descriptions related to different sound emitters can be characterized as positive, neutral or negative. We then explore, quantitatively and qualitatively how aural perception is characterized in our corpus, zooming in to explore local patterns in the description of sound experiences and zooming out to characterize the prominence and distribution of different sound experiences.

2. Data and methods

2.1. Data and study region

As a corpus we used descriptions associated with georeferenced pictures collated through the crowdsourced project Geograph British Isles. Geograph was launched in 2005 with the aim of documenting landscapes through the combination of representative pictures of a location and associated textual descriptions referring to individual grid squares at a granularity of 1 km in Great Britain and Ireland. Geograph contains simple game play elements, with the first contribution to a grid square being awarded more points, and has an active community of more than 12000 users. Similar to most UGC, contributions are biased, with a small number of users[2] contributing the majority of the data, but in previous work it has been shown that descriptions are not strongly biased by individual users, perhaps because of the clear aims and moderation of the uploaded photographs. Furthermore, in a survey carried out by the projects' initiators, users stated that it was important to be sure that the photographs and descriptions are archived for generations to come, and that they be used for educational purposes and promotion of local history. Since no mobile version of Geograph exists we assume that descriptions are written when photographs are uploaded from the desktop computer, though we found evidence that some users take notes in the field.[3] The data used in this paper were downloaded in June 2016, and consisted of more than 5 million photographs, of which more than 3.5 million also had a textual description, and are available under a Creative Commons Attribution-ShareAlike 2.5 License.

2.2. Method overview

Our approach to extracting, classifying and evaluating aural descriptions from the corpus involved three distinct methodological steps:

1. Extraction of descriptions referring to either experienced sounds or perceived absence of sound
2. Classification of the extracted descriptions according to a taxonomy of sound emitters
3. Allocation of sentiment values to each classified description of sound

Fundamental to our work in the first two tasks was the development of an annotated corpus, which was used to evaluate the quality of our extraction rules, and to serve as training and test data for our classifier. Fig. 1 gives an overview of the key steps carried out and described below.

2.2.1. Rules of annotation

As is typical in work on natural language, we created an annotated dataset to, firstly, better understand the properties and use of language in our corpus, secondly, to provide training data for our classifier, and thirdly to evaluate the efficacy of our methods. The annotated dataset contained examples of either descriptions referring to perceived sounds (and thus, not *per se* all detectable sounds) or their perceived absence and we classified these examples according to the type of referenced sound emitter (Table 1).

Descriptions of the following cases were all annotated as related to sound experience:

- aural perception at the moment the photograph was taken, for

[2] Detailed demographic data about users are not available, but based on a survey carried out by the project initiators it appears that users are in general more likely to be over 50 and male.

[3] I made a note on the map that whilst photographing this, the larks were almost deafening! Source: http://www.geograph.org.uk/photo/902702.