

Images and Perceptions of Neighbourhood Extents

Paul Clough
University of Sheffield
Dept. of Information Studies
Sheffield, S1 4DP, UK

+44 114 2222664

p.d.clough@sheffield.ac.uk

Robert Pasley
University of Sheffield
Dept. of Information Studies
Sheffield, S1 4DP, UK

+44 114 2222666

r.pasley@sheffield.ac.uk

ABSTRACT

In this paper, we describe an experiment in which we use an on-line questionnaire to elicit people's perception of the extents of smaller vague regions, such as neighbourhoods. Our approach uses images of street scenes rather than landmarks or placenames.

Categories and Subject Descriptors

H.1.2 [Models and Principles]: User/Machine Systems. - *Human factors*

General Terms

Experimentation, Human Factors.

Keywords

Vague regions, user study, questionnaire

1. INTRODUCTION

Vernacular regions are commonly referred to by people in everyday life, but often do not exist in digital gazetteers (DGs), such as the ADL gazetteer [3]. This may reduce the effectiveness of GIR systems, which rely on a link between textual and spatial definitions in order to index spatially, as well as textually. Efforts to automatically create definitions of these regions have been made [1, 4]. These efforts usually assume that vernacular terms will co-occur with known placenames [4], or landmarks [5]. However as the size of region reduces, fewer landmarks (or placename references) may exist to derive the region boundary. We therefore proposed a simple alternative using images: we assume people will recognise street scenes if shown an image of them, have some idea of where these scenes are, and be able to say whether the scene is inside a region, outside of it or on the boundary.

2. PREVIOUS WORK

Previous work has focused on defining vernacular regions for human geography [6]. This has been done via ethno-telephony (using company names that reflect region names in 'phone directories). Later with the advent of the Web, "a vast storehouse

of knowledge" [2], web mining has been used to gather suitable documents. Geo-references can be extracted from these documents, which can then be grounded to spatial references and these can be used to form a representation of the region [1]. Such efforts can rely on assumptions about the mining process and the distribution of placenames on the web and typically deal with larger regions.

Mansbridge [5] used a street survey methodology to elicit definitions for "Sheffield City Centre" (a city with population 500,000 in the UK), and "The Midlands" (a region of about 120km x 150km in the UK). This was done by stopping people and asking them about various points. Two questions were asked, of the form "Which of the following places are in Sheffield city centre?" For Sheffield this was followed by a list of 38 landmark names and placenames for "The Midlands". The allowed responses for each landmark/placename were 'Yes', 'No', 'On the Boundary' and 'Don't know where it is.' These responses were processed manually to allow definition of region boundary.

3. METHODOLOGY

An on-line survey similar to that of [5] was carried out. Additionally two new regions, The Peak District and Hunters Bar, were added. The Peak District is a predominantly rural area near Sheffield. Hunters Bar is a residential neighborhood in Sheffield, and is the focus of this research. Hunters Bar is thought to be about 0.5km² and contains a roundabout of the same name. We felt, however, that the roundabout may not be at the centre of the region, and that we could not expect to define the ground truth for Hunters Bar using such an approximation. The main problem, however, was that Hunters Bar does not contain geo-references as found in existing publicly-accessible spatial resources and, although it contains streets, some of those streets are long and only a small part of them would be in Hunters Bar.

We looked for images of Hunters Bar using popular web photo resources, but found insufficient geo-tagged images. Therefore, we created our own set of images by taking a set of pictures evenly spread around the roundabout itself, going out to beyond where we felt Hunters Bar *might* extend to. We were limited by non-descript areas, and the street network. Because we require a decision for each of the points, and because we had created a questionnaire for 4 regions, there was pressure to keep the number of points to a minimum for practical reasons. The points were derived from the position the photograph was taken; any other geo-coding would have required further judgments.

The format of the photo-based question was similar to the other three, "For each of these places, would you say it is in Hunters Bar?" Following this were 32 numbered images, shown in a size

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GIR'10, 18-19th Feb. 2010, Zurich, Switzerland

Copyright © 2010 ACM ISBN 978-1-60558-826-1/10/02... \$10.00

that allowed them to be part of a multi-row question, but still clearly visible. The idea was for people to be able to take a quick decision on each image. We emailed invitations to participate to a University of Sheffield email list of around 25,000 staff and students. It was assumed that since Hunters Bar is an area associated with the University, that members of this list would be familiar with it and able to participate. The questionnaire began with some general questions to collect information such as age, location and length of association with the regions. We also collected comments on the questionnaire design itself.

4. RESULTS

We received 590 valid responses (617 started but did not complete it). This is a small percentage of the potential respondents, perhaps because many emails are received from this list. Of the 590 responses, 260 of these left comments. The age of respondents is biased toward being young: about 75% were under 26, 15% were 26-35 the other age ranges being below 5% each. The number of responses in the “Do not know where it is” column ranged from 32% up to 85% for each image. The higher counts are due to backstreet scenes, poor quality images and too much focus on generic street furniture, as well as people’s lack of familiarity. For people with up to 6 years of association with the region an average of 17/33 images were unknown; for those with 6 years or more 15/33 images were unknown. For those with no association of the region, 24/33 images were unknown.

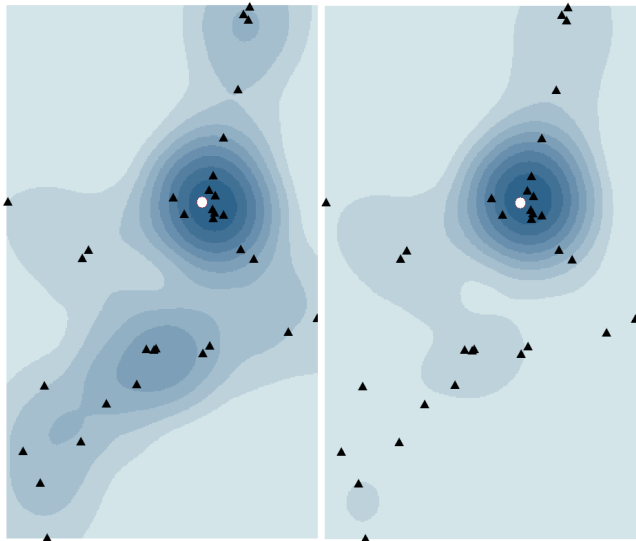


Figure 1: Un-weighted points (left), frequency weighted (right)

We created two Kernel Density Estimates (KDEs): one for all the points un-weighted, the other using the frequency of the yes (y) answers plus half the frequency of the on the edge (e) answers (weight = $y + 0.5e$). Figure 1 shows the resulting KDEs with black triangles indicating the location of images in the questionnaire and the dot indicating the Hunters Bar roundabout.

The centre of each KDE is very close to the roundabout, suggesting that it is hard to escape the underlying point density. However it may be that our original selection of points was sufficient to match the perceptions we are testing and that the un-weighted points coincide with the weighted ones. We notice that, as well as a slight shift northwards and a larger hotspot, the effects of outlying points is reduced. It appears that the roundabout (white dot) is quite central to the region as defined by

KDE. However, it is also clear that the area of this vague region does extend beyond just this one landmark (commonly used to locate Hunters Bar) to a wider region which is likely to be more representative of Hunters Bar in practice. This is useful for certain vague regions, especially smaller ones, where clear landmarks may not exist or be difficult to locate with existing spatial data sources. We cannot say anything about the size of the region because criteria for thresholding the KDE are hard to define (it is easier to find a boundary based on area assumption).

5. QUESTIONNAIRE COMMENTS

We received various comments on the questionnaire, some of which highlight aspects that could be improved. The most common comment was that the images used were too small, too dark and sometimes unclear. Second to this, people drew a complete blank on some of the images which seemed to discourage them and affect their motivation, more than originally expected. Some international students commented that Sheffield’s architecture is all very generic and difficult to distinguish.

6. CONCLUSIONS

Gathering people’s perception on the extent of smaller regions, such as neighbourhoods, using named landmarks is often difficult because such landmarks just do not exist. In this paper we have explored whether people could judge the extent of a small region in Sheffield, Hunters Bar, using the membership of visualized street scenes presented as images. The coordinates of these street scenes and a weighting derived from perceptions of membership can be used to derive a KDE. Some methodological aspects require more attention, such as the quality and selection of images which seem to affect people’s motivation to participate fully. The method of display used might also benefit from an improved design. It may be that surveys based on street scenes are capable of higher resolution, since they do not rely on recognising named landmarks to identify places. Future work might include testing text based surveys using business names, and exploring different ways of finding suitable images.

7. ACKNOWLEDGMENTS

Work partially supported by the EPSRC and Ordnance Survey (CASE/CAN/06/67).

8. REFERENCES

- [1] Blessing, A. & Schütze, H. (2008), Automatic acquisition of vernacular places, In Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services, 662-665.
- [2] Chakrabarti, S. (2003), *Mining the Web: Discovering Knowledge from Hypertext Data*, Morgan Kaufmann.
- [3] Hill, L. L. (2000), 'Core Elements of Digital Gazetteers: Placenames, Categories, and Footprints', *Lecture Notes in Computer Science* **1923**, 280-290.
- [4] Jones, C., Purves, R.S., Clough, P., and Joho, H. (2008) Modelling Vague Regions with Knowledge from the Web, *International Journal Geographic Information Systems (IJGIS)*, **22(10)**, 1045-1065.
- [5] Mansbridge, L. (2005), 'Perceptions of imprecise regions in relation to geographical information retrieval,' Master's thesis, Information Studies. University of Sheffield.
- [6] Zelinsky, W. (1980), 'North America's Vernacular Regions', *Annals of the Association of American* **70(1)**