

Spatio-Qualitative Data Visualization: SoftGIS and Weighted Average Visualization

Kamyar Hasanzadeh*

Department of Planning and Geoinformatics, Aalto University, Espoo, Finland

*Corresponding author: kamyar.hasanzadeh@aalto.fi

Received February 07, 2014; Revised March 20, 2014; Accepted March 26, 2014

Abstract In recent years there has been considerable breakthrough in acquisition of georeferenced qualitative data. These types of data have characteristics that distinguish them from quantitative datasets and therefore it is typically more challenging to discover knowledge from such spatio-qualitative data. SoftGIS is one of the most prominent attempts in collecting such spatio-qualitative data that is capable of providing informative data with applications in different disciplines. This paper uses this dataset in a case of urban experience in Helsinki in order to propound a visual technique that can help with knowledge discovery process. The visualization method proposed in this study, namely weighted average visualization (WAV), is tailored to meet specific characteristics of the aforesaid dataset and is capable of discovering patterns that are not visible through current approaches.

Keywords: *SoftGIS, visualization, spatio-qualitative data analysis, weighted average visualization*

Cite This Article: Kamyar Hasanzadeh, "Spatio-Qualitative Data Visualization: SoftGIS and Weighted Average Visualization." *Journal of Geosciences and Geomatics*, vol. 2, no. 1 (2014): 38-41. doi: 10.12691/jgg-2-1-6.

1. Introduction

Studying the qualitative data in their geographical context has the potential to reveal useful information in different studies, such as human geography, geology, urban studies and land use planning [13]. According to this, there has recently been a growing interest in applications of spatial technologies, and more specifically GIS, in studying the qualitative and subjective data [12].

A geographic information system (GIS) is an integration of hardware, software, and data designed to capture, store, manipulate, analyze, manage, and present all types of geographically referenced data [11]. In other words, GIS provides us with a wide range of tools and techniques that can provide a better understanding of different phenomena in their geographical context.

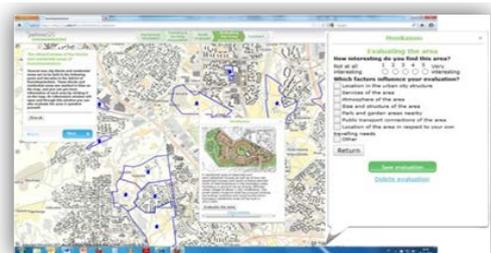


Figure 1. SoftGIS survey panel preview: <http://demo.asiatkartalle.fi/> (accessed on 5.2.2014)

Recent advances in spatial sciences and computer technologies have granted qualitative GIS with the possibility to be incorporated in the latest versions of

computer-aided data analysis [12]. This has resulted in a considerable breakthrough in acquisition of geocoded qualitative data. The innovation of 'SoftGIS', developed by Marketta Kytta and her team at Aalto University, refers to a collection of internet-based surveys (Figure 1) that allow the locality-based study of human experiences [6]. 'SoftGIS' provides a combination of 'soft' subjective data (qualitative) with 'hard' objective spatial data. SoftGIS has been capable of collecting large datasets for the use of urban planners interested in the development of more user-friendly physical settings [6].

With the considerable increase in professionals' access to soft geocoded data, the need for a solid analysis methodology has become more significant; one that suits the specific characteristics of the spatio-qualitative data and has the potential to extract the desired information from the huge mass of data. There typically two approach in studying and analysis of spatio-qualitative data; one that uses narration as a means of conveying the desired information [14] and one that uses a quantitative basis to analyze the data [13].

This research is based on the later approach and attempts to reach knowledge discovery through visual analysis of quantified SoftGIS dataset. Accordingly, this paper proposes a new visualization approach which has the potential to reveal useful information from SoftGIS or any other similar spatio-qualitative datasets.

2. Spatial Data Visualization

Visualization is a major component of information perception. One can take advantage of his or her visual

perception abilities in order to signify his or her cognition of the abstract data [10]. Moreover, visualization facilitates the understanding process of the users [7]. Therefore, it is important to benefit from the visualization techniques which best suit the data in order to amplify its most important underlying information.

The most common visualization methods used in geoinformatics are of maps and most people are familiar with them [9]. Moreover, in geoinformatics many types of thematic maps are also used. A thematic map can be used to emphasize the spatial pattern of one or more geographic attributes [11].

The potential users of SoftGIS data are urban authorities that are not necessarily experts in spatial sciences. Thus, the proposed methods not only must be informative, but also should be as simple as possible.

3. SoftGIS Data

The SoftGIS data was provided by Mapita Ltd. and it includes the recorded people's experiences in Espoo and Helsinki region in Finland. The dataset originally consisted of six categories of records, three categories of positive and three categories of negative records (atmosphere, appeal, social). For the simplicity of the study, the categories were generalized into two major classes of positive and negative records.

4. Qualitative Data Quantification

Studying qualitative data is normally more challenging than quantitative datasets. Accordingly, in disciplines involved with qualitative values (e.g. psychological sciences), it is often a good practice to quantify the values in order to improve the computational and comparison capacity of the qualitative datasets [1]. Since qualitative datasets contain nominal values, the computational techniques widely used for quantitative dataset are typically inapplicable to them. Computationally-enabling the qualitative data allows the application of a diversity of visual and analytical techniques that can help to make the analysis more feasible. Therefore, in this study, in order to overcome the aforesaid computational limitations, a simple quantification is used. Accordingly, the values +1 and -1 are assigned to the nominal *Positive* and *Negative* quantities respectively.

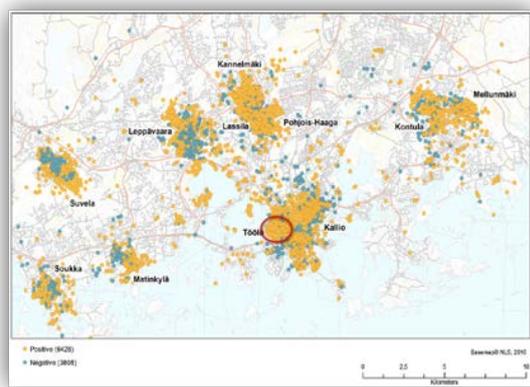


Figure 2. Point representation of Positive and Negative experiences. From [8]

5. Current Visualization Approaches

Kyttä *et al.* have proposed several methods for visual representation of SoftGIS data [8]. The simplest method used in the mentioned literature is a point map (Figure 2) in which the recorded experiences are divided into two classes of Positive and Negative and are demonstrated accordingly.

As it can be speculated in Figure 2, using a point representation of the recorded experiences has considerable limitations in conveying the desired information to the user. The first and most significant limitation is that in such a dense dataset it is not feasible to use point representation as there are too many overlapping features in the area. For instance one might assume that the marked region on the map dominantly contains positive records. However, as it can be observed in Figure 3, there are quite many negative records in that very same region as well, which are not visible in the first representation.

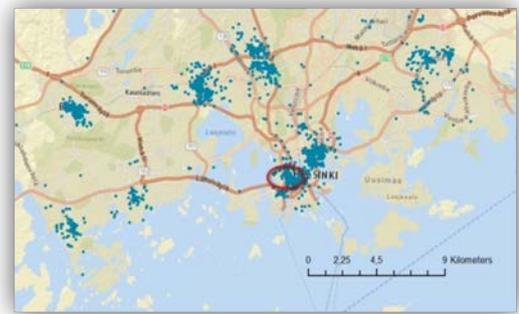


Figure 3. Negative records in Helsinki metropolitan area

In addition, observing large masses of point clouds does not provide any clear image on the overall recorded experiences of the neighborhoods and its variation throughout the region.

Another visualization used in the same literature is a ratio map which represents the proportion of positive markings to all markings within a cell of a certain size (Figure 4). Although in this visualization the data is managed in a more proper way, it is yet not highly practical to use as it requires continual reference to the legend. Moreover, it does not provide any information about the actual amount of negative and positive markings in each cell and this can result in confusion.

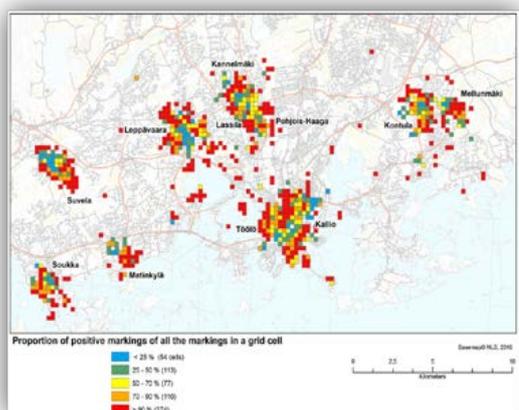


Figure 4. The share of positive and negative place markings. From [8]

Furthermore, another visualization method has been used in [8] that utilizes a Natural Neighborhood interpolation technique in order to create a continuous surface of the discrete markings (Figure 5).

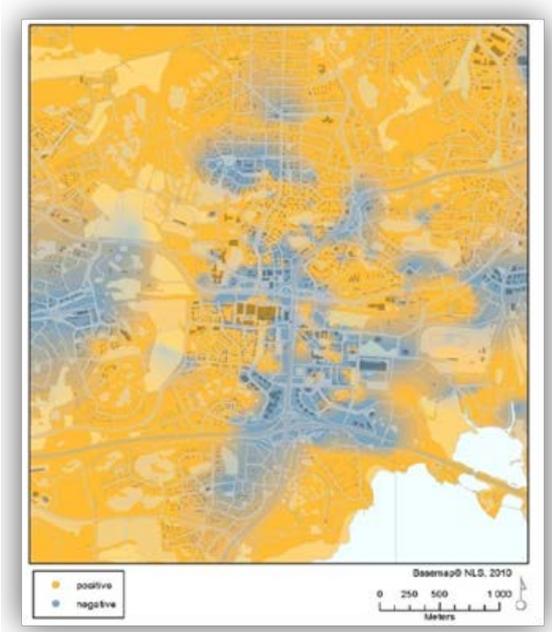


Figure 5. The natural neighborhood visualization of residents' positive (yellow) and negative (blue) experiences in a part of Espoo [8]

Although this visualization looks more attractive, it has several limitations. In order to explore its limitations, the visualization was recreated in this study (Figure 6) [13].

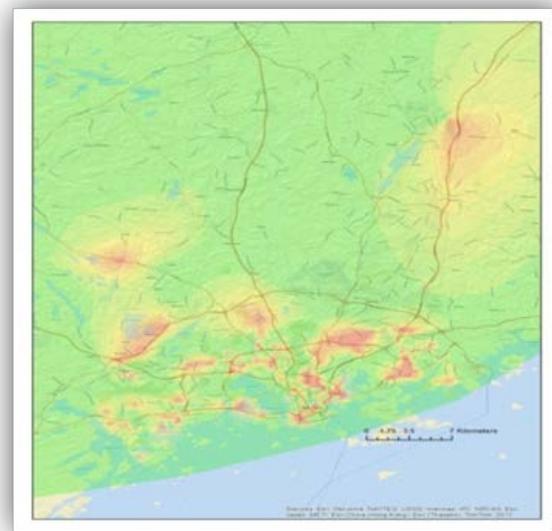


Figure 6. The NN visualization of residents' positive (green) and negative (red) impressions in Helsinki metropolitan area (Recreated)

The first and most significant limitation of this visualization is that it can be misleading. That is because in some areas the projections may be based on only a few observations (Figure 7). In other words, as a result of interpolation there might be values for locations for which there are not enough observations indeed. Moreover, as a result of natural neighborhood interpolation a contradicting record (in comparison to the dominant experience in its surrounding) can significantly influence the overall result in an area.

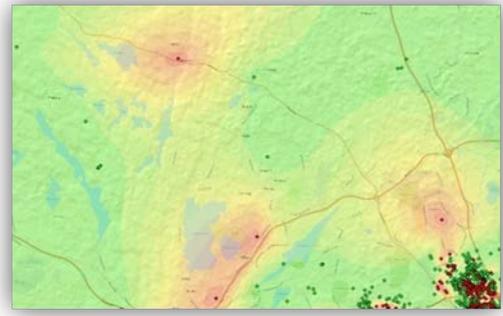


Figure 7. Limitations of using NN interpolation in SoftGIS data visualization

6. Weighted Average Visualization

The methods discussed to this point, were incapable of depicting a clear transition between overall experiences within areas. Moreover, for some purposes, such as urban planning, one may require a more precise understanding of the recorded locations. Thus we need to use a more location-based approach that provides a fair representation of all existing ideas.

The approach proposed and implemented in this study attempts to tackle the above discussed limitations by calculating a weighted average of markings within a predefined cell. Mathematically speaking, we are interested in calculating E (Experience) for every given cell as below:

$$E = \frac{n_n \times negative + n_p \times positive}{n_n + n_p}$$

Where n_n and n_p are the number of respectively negative and positive markings within the cell and negative and positive values are defined as the -1 and +1 respectively. Obviously, E is a continuous value in the range of [-1, 1] and the higher its value the more positive the overall experience would be.

The size and shape of the cells can considerably affect the visualization result. Obviously the larger the size of the cell, the less locationally precise it would be, though the visualization would become smoother. The optimum cell size in this case was considered as 300 meters. Finally the calculated E 's were visualized on a gradual scale as in Figure 8.

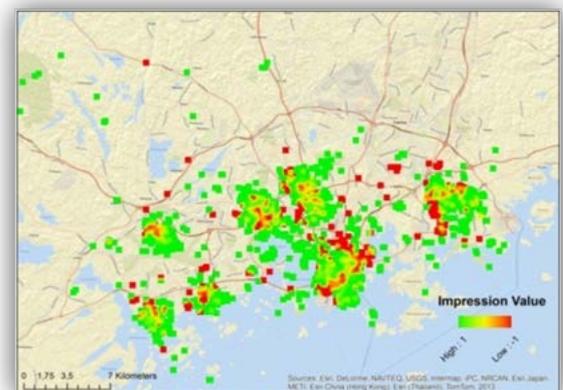


Figure 8. Weighted average impression map

In [Figure 8](#) we can see patterns which were not visible in the formerly discussed visualizations. For example in Helsinki city center, south of Kamppi (marked with a black oval in [Figure 9](#)), a highly negative experience can be observed. According to participants' views this is mostly associated with the city center traffic congestions and dirtiness.

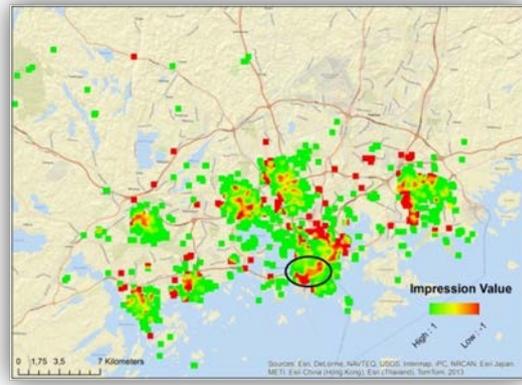


Figure 9. Negative impression pattern in Helsinki city center

In addition, this visualization preserves the locational quality of the dataset. Therefore, for each location the overall satisfaction of the residents can be extracted. This overall satisfaction is based on the average of negative and positive experiences within the area. Moreover, WAV is capable of capturing the gradual transition of experience states between different areas. This contributes to a smoother and more realistic image of the phenomenon.

7. Conclusion

Acquisition of spatio-qualitative datasets, such as SoftGIS data, is a rather recent achievement. Therefore studying methodology of such datasets still requires further research and thought. The current visual approaches discussed in the existing literature have considerable limitations that signify the need for a more reliable and informative approach which was the main purpose of this research. By considering the significant role of visualization in the perception of the abstract information, a weighted average visualization was proposed in this study that revealed interesting patterns. The aforesaid visualization preserves the locational quality of the dataset and provides a smoother visualization by extracting the gradual transition of the impression states throughout the area of study. Moreover, WAV represented a fair insight of the overall experiences by using an average of different recorded opinions.

In summary, if analyzed properly, SoftGIS or any similar geocoded subjective data can be considered as a great source of qualitative information with applications in various disciplines. One promising approach towards extracting knowledge from the SoftGIS data is to use

appropriate visualization. One that facilitates the perception of the qualitative information while maintaining its locational quality.

Acknowledgement

This research is a part of the master's thesis conducted at Department of Planning and Geoinformatics in Aalto University, Finland. My deepest appreciations go to Professor Kirsi Virrantaus for giving me the opportunity to work in the department of Geoinformatics lab. I would also like to thank Dr. Jussi Nikander and Dr. Paula Ahonen-Rainio for their brilliant comments and suggestions on my work.

And the last but not the least, I would like to thank the Mapita Ltd. group for providing the SoftGIS dataset, without which this research could not be accomplished.

References

- [1] Guttman, L. 1944. A basis for scaling qualitative data. *American sociological review*, pp. 139-150.
- [2] Kahila, M. 2008. Possibilities of Web-based softGIS Method in Revealing Residents Evaluation Knowledge of the Living Environment. In *FUTURE-Future Urban Research in Europe, The Electronic City Conference*, Bratislava.
- [3] Kahila, M. and Kyttä, M. 2011. Web-based SoftGIS method in research and urban planning practices. *The Electronic City*, 1 p. 199.
- [4] Kahila, M. n.p. Possibilities of Webbased softGIS Method in Revealing Residents Evaluation Knowledge of the Living Environment. [report].
- [5] Kahila, M., & Kyttä, M. 2006. The Use of Web-based SoftGIS-method in the Urban Planning Practices. In *Proceedings of the Conference on Urban Conditions and Life Changes*.
- [6] Kahila, M., & Kyttä, M. 2009. SoftGIS as a bridge-builder in collaborative urban planning. In *Planning support systems best practice and new methods*, Springer Netherlands, pp. 389-411.
- [7] Keim, D. A. 2001. Visual exploration of large data sets. *Communications of the ACM*, 44 (8), pp. 38-44.
- [8] Kyttä, M., Broberg, A., Tzoulas, T., & Snabb, K. 2013. Towards contextually sensitive urban densification: Location-based SoftGIS knowledge revealing perceived residential environmental quality. *Landscape and Urban Planning*, 113, pp. 30-46.
- [9] Nikander, J. 2012. Interaction and visualization methods in teaching spatial algorithms and analyzing spatial data. Ph.D. Aalto University.
- [10] Shneiderman, B. and Plaisant, C. 2004. *Designing the user interface*. Boston: Pearson/Addison Wesley.
- [11] Slocum, T. A., 2005. *Thematic cartography and geographic visualization*. Upper Saddle River, NJ: Pearson/Prentice Hall.
- [12] Verd, J. M. and Porcel, S. 2012. An Application of Qualitative Geographic Information Systems (GIS) in the Field of Urban Sociology Using ATLAS. ti: Uses and Reflections. 13 (2).
- [13] Hasanazadeh, K. 2014. *SoftGIS Data Mining and Analysis: A Case Study of Urban Impression in Helsinki*. Master's. Aalto University.
- [14] Kwan, M. P., & Ding, G. 2008. Geo-Narrative: Extending Geographic Information Systems for Narrative Analysis in Qualitative and Mixed-Method Research*. *The Professional Geographer*, 60 (4), 443-465.