Representing Spatial Geographical Data via variations of Volume and Tempo in Sound.

Nick Bearman and Nicholas J. Tate

Department of Geography, University of Leicester, University Road, Leicester, LE1 7RH, UK
Tel: +44 (0)116 2523823, Fax: +44 (0)116 2523854
Email: nickbearman@yahoo.co.uk, n.tate@le.ac.uk

KEYWORDS: sonification, spatial data, sound, volume, tempo

1. Introduction

The majority of spatial data are represented visually which can lead to an overload of information if too many data are represented. This study explores and evaluates the methods of representing spatial data using sound, with the aim of increasing the amount of spatial data that can be communicated without information overload. This study was performed for an undergraduate BSc Geography dissertation, and concludes that using variations in sonic volume and tempo can be used to represent spatial data. A particular application of this is using sonic data to augment visually displayed spatial information, but further research is needed in order to develop an effective application.

2. Scientific Background

Currently sound is a widely used medium for transferring information, both verbally and non-verbally. Multimodal communication is when more than one mode of communication is used at once - for example, sound (talking) and visual (body language). Mehrabiam (2006) found that for a typical spoken message, 55% of the understanding of the message came from the visual aspect, and only 45% from the audio aspect. Currently, the typical human's auditory system is significantly underutilised – Brewster (1994) reported that humans can tell the difference between any 2 of 400,000 sounds, and can remember up to 49 sounds at any one time.

When information is represented visually, users must focus their attention on the specific output device in order to receive the presented information. However, sound is omni-directional so it can be heard without concentrating specifically on the output device. Sounds are usually used for monitoring or warning purposes, as sound is good at representing rapidly changing data, rather than data that needs to be referred back to. McRoberts and Sanders (1992) have suggested that skills in non-verbal auditory perception are liked to spatial awareness. Currently geographical data tend to be presented in a visual format, whether it is a static map or a fully-fledged Geographical Information System, and sound is rarely used to display this type of data.

It is possible to represent many types of geographical data sonically – most easily are the data sets that represent actual sounds in the real world, such as the noise generated by an airport or traffic on a road. Normally this could be presented as a map with noise contours, but it could also be represented sonically with the sound being the noise of the aircraft at a specific location. Other data, which is not sonic in nature, could also be represented this way using a generic sound. This may be a better way of representing both audio and non-audio based data sets, because it would provide more flexibility in representing two data parameters – for example aircraft noise and aircraft frequency. Sound has a restriction on the amount of data it can contain, because of the fact that the majority of people can only access a limited amount of data from it. Jacobson (2002) reported that there were only 3 factors in sound that could be varied – tone, volume and tempo. Other factors in a sound could be used (for example having different instruments) but it would be too difficult for people to extract one data set from this. Jeong and Jacobson (2002) concluded in their study that tone was not effective at demonstrating variation in a data set, so volume and tempo will be investigated in this experiment. Some or all of these factors could be used to augment visual information with sonic information. This has the potential to create a very powerful method of interacting with multiple data sets.

The actual sound that will be used is important to select, as it will provide the main interface for the data. A series of beeps at a constant interval will be the easiest to understand – the tempo or volume of the beeps could then be varied to represent the variation in the data set. Using a series of beeps would allow the same data representation system to be used for a large number of different data sets which, after the initial learning period, would reduce subsequent learning periods.

The aim of the research reported in this paper is to develop a method to represent spatial geographical data using different aspects of sound, and to evaluate how effective these methods are on a sample of university students and staff. The technique created should be independent of the data, so it can be used to represent any type of spatial data, and should primarily be used for explaining data rather than predicting data. The effectiveness of volume and tempo will be compared and gender, musical ability and geographical experience will be analysed to see if they have an effect. The future uses of this method of representing spatial data will also be considered.

3. Methodology

The sampling design was intended to have a total size of approximately 50 people, with an even split between musical ability/no musical ability, male/female and geographical experience (geography or non-geography degree). The methodology involved interviewing 61 people, and asking them to complete a number of computer based exercises. The questionnaire asked for the data used to categorise people into different groups (gender, musical ability and geographical experience). Answers to the computer based exercises were also written on the questionnaire, with preferences on volume or tempo.

A Visual Basic computer program was written to assess the participant's ability to use volume and tempo to differentiate between different sounds (Ordering Exercise), and to locate specific points within a variation of sounds (Population Density Exercise). Figure 1 shows the different parts of the program, and the order in which they are shown to the participant.

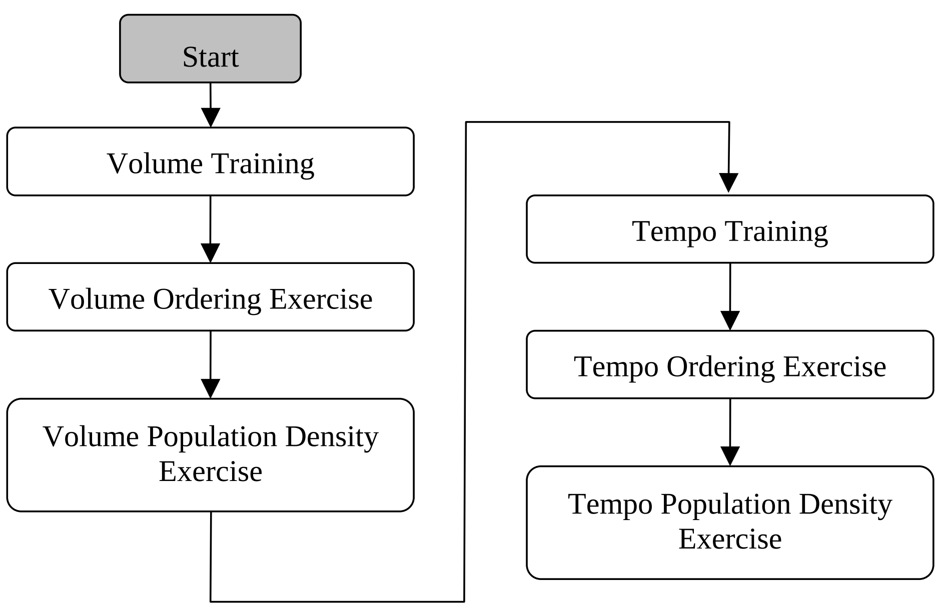

Figure 1. Flowchart of the stages the participant is taken though in the computer program.

The Training section allows the user to familiarise themselves with how data will be represented using sound. The ordering exercise allows the user to play six sounds, each with a different volume (or tempo) and asks them to write them down on the questionnaire, in order from lowest to highest.

The Population Density Exercise (see Figure 2) is the most demanding of the exercises, and should provide the most information about representing spatial variation with sound. The task was explained as the area in the rectangle representing a map, with the sound representing the population density, with higher values being represented by higher volume or tempos. The participants were asked to locate the three 'cities' and record the X Y coordinates. All activities were timed.

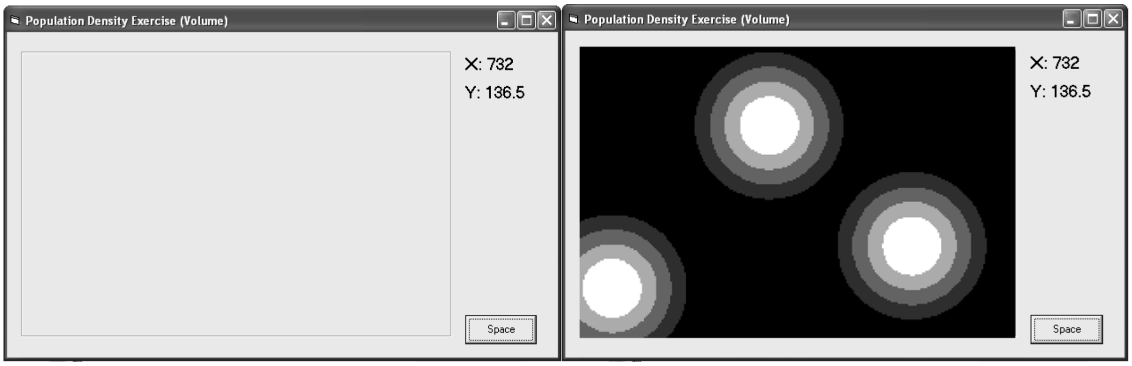

Figure 2. The Population Density Exercise shows the blank screen to the left shown to the user, and asks them to locate the 3 high volume (or tempo) locations. The screen to the right (not visible to the user) shows the variation in sounds played by the computer when the mouse is moved over that section, with a higher volume (or tempo) illustrated by the lighter colours.

4. Results & Analysis

There were a total of 61 participants, and the results are summarised in the tables below.

Attribute	***Number of people***	
Gender	Male = 31	Female = 30
Degree	Geography = 33	Non-Geography = 28
Musical Ability	Musical Ability = 33	No Musical Ability = 28

Table 1. Number of people involved. Total = 61.

	Average time for			
Category	**Volume Order**	**Volume City**	**Tempo Order**	**Tempo City**
All	1:01	1:32	0:55	1:15
Male	1:01	1:31	1:00	1:20
Female	1:01	1:33	0:49	1:10
Musical Ability	1:08	1:39	1:03	1:20
No Musical Ability	0:56	1:24	0:52	1:18
Student				
- Geog*	1:00	1:30	0:57	1:19
- Non-Geog*	1:03	1:39	0:54	1:04

Table 2. Average times (mins and seconds) for the exercises for different groups. Note: the ones marked * are sub-sets – i.e. the first one is people who are students and do a geography degree.

For the population density exercise, the average time for locating the three cities was 1:32 for volume and 1:15 for tempo. Overall, the majority of people preferred tempo over volume for both exercises, and this was reflected with better results for tempo.

	Musical Ability	Gender	Degree	Volume Cities	Tem Cities	City Result
Volume Order	0.657	0.51	0.43	0.304		
Tempo Order	0.667	0.613 (0.671327)	0.509 (0.653334)			
Musical Ability		0.251	0.011	0.381	0.119	
Degree				0.467	0.883	
Gender				0.61	0.782	
City (correct)	0.038	0.78	0.804			3.03
Order (time)	0.921	0.034	0.517 (0.590898)			
Order (result)	2.788	0.148	3.409			

Table 3. Chi-Square Results for the tests performed. Numbers in brackets are Fishers exact test results.

4.1 Analysis

The overall average times for tempo were lower than volume for both the ordering exercise (6 seconds lower) and the population density exercise (17 seconds shorter). This, combined with a larger number of correct tempo results for the ordering exercise (11 more) and a larger number of correctly identified cities (15 more) means that people generally found tempo easier to understand than volume.

A series of chi-square tests were performed on the data, and when the expected counts were less than 5, a Fishers exact test was performed (Rowntree 2000, Robson 1994). Of the relationships tested, only a few were statistically significant. These were the relationship between musical ability and correct identification of the cities (p-value = 0.038) and gender and time taken for the ordering exercise (p-value = 0.034). Participants with musical ability took longer on the city location exercise, but were more likely to correctly identify them. There was little or no difference between the genders for the volume exercises, but for the tempo exercises females were significantly faster than males.

5. Conclusions

One of the main objectives of the study was to understand how effective sound is at representing geographical data. Two ways of measuring effectiveness were used – the ordering exercise which asked people to compare six different sounds and put them in order, and the population density exercise, which asked people to locate the sound with a particular characteristic.

The results showed that people preferred tempo over volume, and were also more likely to get the correct results with tempo. It is possible that a learning factor made tempo appear easier than volume, but this is unlikely because out of 61 people, 44 preferred tempo. In a future study, half the participants could do volume then tempo, and half tempo then volume, but a larger sample would be needed for this. Many people likened the tempo sound to that of a Geiger counter, which probably made it easier to understand as it was a familiar sound. Gender seemed to have no significant effect on the volume results, but females were faster than males for the tempo exercises. Musical ability seemed to have a significant effect on the population density exercise, but not the ordering exercise. Participants with geographical experience were faster on the volume exercises than people without geographical experience, but the opposite was true on tempo exercises.

Due to the fact that little research has been done in this area, particularly in the discipline of geography, there are a number of aspects of the methodology that could be improved. The volume and tempo sounds in the population density exercise were calculated differently, resulting in a different resolution gradient for the cities. When asking whether people preferred volume or tempo, a closed, graded question would have provided more information for the analysis than the question used. The issues of novelty effects, which Geelan (2005) described as "the single greatest besetting sin in tech-related evaluations: ignoring the motivational effects of the cool new tech" should have been considered. It is possible that people put more effort into understanding this method of data representation because of the fact that it was new and something they hadn't seen before. If it was used on a much more day-to-day basis, people would potentially put less effort into understanding and using it, so may have lower levels of comprehension than demonstrated in this study.

This method of representing data could be used on an audio map, where a map (of any sort) would be augmented with an audio 'layer' of data, which could represent population density, height, the extent of radioactive contamination or potentially any spatial data. This also allows more information to be provided to people without causing information-overload, by using multi-modal communication. Depending on the data sets involved, augmenting spatial information with sonic information could be a very powerful, adaptable and flexible way of representing and interacting with spatial data. The sound gradient in the population density exercise (how quickly the volume or tempo of the sound changes with movement) could be altered and tested to see if interpretation could be made easier. Overall, tempo is a more effective way of representing geographical data than volume using sound. Further research is required to refine the process used to create the sound representing the data, but this study has shown that it definitely has potential as a technique to present and interact with data.

6. Acknowledgements

Nick Bearman would like to thank his dissertation supervisor Nick Tate, and all the participants of the study.

7. References

Brewster, S.A. (1994) Providing a Structured Method for Integrating Non-Speech Audio into Human-Computer Interfaces, PhD thesis, (downloaded from http://www.dcs.gla.ac.uk/~stephen/ThesisIntroAbs/thesisIntro.html on 04/07/2005).

Geelan, D. (2005) *Evaluating Scientific Visualisations (and their educational effects): Some issues with qualitative and multimethod approaches,* presentation at GRC Visualization, Oxford, slide 16.

Jacobson, R.D., (2002) Representing Spatial Information through Multimodal Interfaces, *Proceedings of the Sixth International Conference on Information Visualisation*, The Computer Society.

Jeong, W. and Jacobson, R.D. (2002) Exploratory user study of haptic and auditory display for multimodal information systems. In: McLaughlin, M. L., Hespanha, J.P., and Sukhatme, G.S. (eds) *Touch in virtual Environments: Haptics and the design of interactive systems, IMSC Series in Multimedia,* Prentice Hall: New York, pp. 194-204.

McRoberts, G. W. and Sanders, B. (1992) Sex differences in performance and hemispheric organization for a nonverbal auditory task, *Perception & Psychophysics,* **51,** pp. 118–122.

Mehrabiam, A (2006) Training the Trainers workbook p47, *NUS NSLP Train the Trainers residential training course,* National Union of Students: London.

Robson, C. (1994) *Experiment, Design and Statistics in Psychology,* Penguin Books, London, pp. 89-112.

Rowntree, D. (2000) *Statistics Without Tears*, Penguin Books: London, pp 150-154.

Biography

Nick Bearman completed his BSc(hons) in Geography at University of Leicester in 2006, and this paper is based on that dissertation. In summer 2006 he was involved in a project looking at how a Virtual Reality theatre could be used to assess people's spatial memory skills. He will be returning to Leicester in 2007 to study for an MSc in GIS.