

Aspects of Auditory Perception and Cognition for Usable Display Resolution in Data Sonification

Johan Kildal
Nokia Research Center
P.O. Box 407, FI-00045 Nokia Group
johan.kildal@nokia.com

Abstract. Sonification of data via the mapping of values to frequency of sound is an auditory data analysis technique commonly used to display graph information. The goal for any form of graph is to display numerical information with accuracy and neutrality while exploiting perceptual and cognitive processes. Conveying information in frequency of sound is subject to aspects of pitch perception, largely overlooked to date, that can influence these properties of auditory graphing. This paper identifies some of these aspects and describes potential design limitations and opportunities derived from the musical nature of auditory data representations.

Keywords: Frequency categorization, auditory display, auditory graph, design.

1 Introduction

Parameter-mapping data sonification (PMDS) is an established data analysis technique in which numerical information is mapped to parameters of sound in order to perceptualise relations in the data [1]. As with any method to display quantified information, maximising the accuracy with which the information is made available to the senses is a desirable property. The accuracy that can be attained depends on the perceivable resolution of the information display, a function of the resolution of the display and the resolution of the senses involved. In data sonification, given the high pitch and temporal resolutions of the human auditory system (and the relatively large usable range for any of these parameters of sound), information mapped to frequency of sound and time can be conveyed with high accuracy. Research and experience have shown that optimum implementations of PMDS techniques are obtained by mapping relative numerical data to frequency of sound and laying them out temporally. This pairing of frequency and time has become the default design for auditory graphs [2].

An additional requirement is the uniformity of the psychophysical scaling used in the display. Focusing on the frequency-of-sound axis of the auditory displays considered here, the logarithm of the frequency offers a good approximation to a linear scale in the middle band of audible frequencies, within which equal increases of the logarithm of frequency are perceived as equal increases of pitch. This is particularly true for frequencies between 80 and 800Hz, where toneness is highest [3].

These characteristics of the frequency-of-sound scale are comparable – when not superior – to the most common visual quantitative data display scales, like distance on

an axis in a typical visual graph. In spite of this, this paper argues that there are aspects of human auditory perception and cognition which do not always permit representing numerical relations on a pitch scale in as neutral a way as with analogous visual scales. Furthermore, this paper argues that the way in which we interpret sounds in the frequency space can pose a limitation on the use of our fine frequency discrimination skills for the display of numerical information. These theses are grounded both on user experience observed with the use of frequency-and-time PMDS techniques (section 2), and on the current understanding of psychology of hearing (section 3).

2 Musical Nature of Sonified Data

In the many studies conducted in [4], users explored large numerical data sets interactively using frequency-and-time PMDS techniques, estimating magnitudes based on perceived pitch of piano notes and chords. The value-to-frequency mapping strategies employed in the sonifications made use of either the chromatic 12 tone equally tempered (12-TET) scale or the whole continuum of the frequency space. The participants that evaluated these techniques reached consistently high performance levels in overview and detail information-retrieval data-exploration tasks. During the interviews conducted to evaluate these techniques qualitatively (their usability and the exploratory strategies devised for different types of data sets), users often volunteered reports of their subjective experiences derived from listening to sonified data. A frequently-recorded comment was that upon extended use of these techniques the sonifications became more tiring to listen to. Many of their comments included aesthetic judgements regarding the audio representations of the data, with ratings ranging from appealing to (more frequently) ugly and even harsh. Both frequency spaces used (12-TET and whole continuum of frequency) received positive and negative aesthetic ratings.

The interesting aspect about receiving this kind of feedback was not the actual aesthetic tastes of the participants, but the fact that they volunteered them in interviews that dealt with pragmatic questions regarding interface usability and effective data mining. The frequent recording of such comments, together with multiple anecdotes about users attempting to improvise musical compositions during the more distended parts of the experiments (by searching for data that would render in particular frequency ranges, generating sonifications in rhythmic patterns, and even identifying elements of specific musical compositions and styles) provided enough anecdotal evidence to realise that the outcome of frequency-and-time PMDSs convey associated musical and aesthetic qualities, even when the tasks in hand are strictly analytical belonging to the domain of data mining. Indeed, research has already suggested that musical training can be an important factor in the use of PMDS techniques [5]. However, musical expertise is not required for users to be implicitly aware of the categories present in the different musical dimensions, and ultimately to be implicitly aware of the rules forming a particular musical context [6] that may arise from the sonification of data.

3 Cognitive limitations on usable display resolution

As stated above, we possess a very fine capacity to discriminate audible frequencies. The majority of the population can detect frequencies far smaller than a semitone, often of only a few cents, particularly across the middle range of audible frequencies (see [7] for a summary of related studies). This would suggest that the highest display resolution usable by most people is obtained by mapping values to the frequencies in the 12-TET scale, and that even the majority of users would benefit from the further-increased resolution obtained when values are mapped to the exact corresponding frequencies in the frequency continuum. In contrast, literature in musicology shows that, despite our fine frequency discrimination capabilities, no musical tradition makes structural use of more than 12 equal divisions in the octave¹. Burns ([8] p.218), offers the following explanation for this: "*The most likely reason for the adoption of a relatively small number of discrete intervals as the tonal material for music is that discretization or categorization is a typical, if not universal, strategy used by animals in order to reduce information overload and facilitate processing when subjected to the high-information-content signals and/or high information rates from a highly developed sensory system*". Due to the categorical perception of pitch (for which frequency shifts of any size are reliably labelled as one of the discrete intervals in familiar musical scales, like the 12-TET scale [9]), making use of the whole frequency continuum for the mapping of values may not provide the fine display resolution that could be expected. Instead, and unless required for local detailed analysis of a few data points, it could often be beneficial to select a relatively small subset of intervals per octave, and represent values rounded to those discrete frequencies. Such mapping strategy would be useful for overview explorations of larger data sets, where sonifications could represent data with sufficient accuracy, while being 'easier' to listen to.

When considering subsets of intervals for value-to-frequency mappings, it is necessary to investigate whether the obvious division in 12 equal intervals is still too large. For instance, in music (at least in the western tradition), although with exceptions, diatonic scales are used to create melodies, and not the full chromatic scale [8]. If the selection of such small collections of intervals for musical expression are part of a strategy to cope with complexity of information, similar criteria could also be relevant for the implementation of PMDS techniques intended to extract knowledge from complex data. However, as smaller sets of intervals are used, the chances of obtaining musical phrases with a strong tonal identity would be higher. This gives rise to the question of whether musical context (which can significantly improve pitch judgements [10]) is a desirable factor in a display of numerical information or if, on the contrary, the sense of expectation of musical events that might arise [11] could interfere with the objective exploration of numerical information.

¹ While there are traditions that include quarter tones or other intervallic divisions that are smaller than the western semitone, as well as embellishments like glissandi, none of them seem to make use of intervals smaller than the western semitone to construct structural elements in music, such as intervals in melodic lines. See [8] for a comprehensive discussion on this topic.

4 Conclusions

This paper has identified aspects of psychology of music and auditory perception and cognition that can be relevant for the design of value-to-frequency mapping strategies in auditory displays. Building on qualitative data collected from users of interactive sonification interfaces who explored large data sets, it is highlighted that auditory representations of data possess a musical nature that influences the subjective experience of data mining, even to an aesthetic level, and that they are generally rated as unpleasant if generated with relatively large sets of intervals per octave. The paper then discusses the implications of the categorical perception of the pitch space, arguing that, except for the cases in which detailed examination of data is required locally, most stages of the exploration cannot benefit from the complexity of information generated with very accurate mappings. These aspects of human auditory cognition should be included in the research agenda about PMDS techniques, and taken into account in their implementations. In particular, the effects of low-resolution displays in the performance and subjective experience with large auditory graphs should be investigated, including the implications of stronger tonal contexts that might arise as a result.

References

1. Hermann, T.: Taxonomy and Definitions for Sonification and Auditory Display. In: 14th International Conference on Auditory Display. Paris, France (2008)
2. Flowers, J.H.: Thirteen Years of Reflection on Auditory Graphing: Promises, Pitfalls, and Potential New Directions. In: First International Symposium on Auditory Graphs (AGS2005), 11th International Conference on Auditory Display. Limerick, Ireland (2005)
3. Huron, D.: Tone and Voice: A Derivation of the Rules of Voice-Leading from Perceptual Principles. *Music Perception* 19(1), pp. 1--64 (2001)
4. Kildal, J.: Developing an Interactive Overview for Non-Visual Exploration of Tabular Numerical Information. PhD Thesis, Department of Computing Science, University of Glasgow, UK (2008)
5. Neuhoff, J., Wayward, J.: Pitch Change, Sonification, and Musical Expertise: Which Way Is Up? In: 8th International Conference on Auditory Display. Kyoto, Japan (2002)
6. Bigand, E.: Contributions of Music to Research on Human Auditory Cognition. In: McAdams, S., Bigand, E. (eds.) *Thinking in Sound. The Cognitive Psychology of Human Audition*, pp. 231--277. Oxford University Press. (1993)
7. Moore, B.C.J.: Pitch Perception. In: *An Introduction to the Psychology of Hearing*, pp. 195--232. Academic Press (2004)
8. Burns, E.M.: Intervals, Scales, and Tuning. In: Deutsch, D. (ed.) *The Psychology of Music*, pp. 215--264. Academic Press (1999)
9. Burns, E.M., Campbell, S.L.: Frequency and Frequency-Ratio Resolution by Possessors of Absolute and Relative Pitch: Examples of Categorical Perception? *J. Acoust. Soc. Am.* 96(5), pp. 2704--2719 (1994)
10. Warrier, C.M., Zatorre, R.J.: Influence of Tonal Context and Timbral Variation on Perception of Pitch. *Perception & Psychophysics*, 64(2), pp. 198--207 (2002)
11. Huron, D.: *Sweet Anticipation: Music and the Psychology of Expectation*. MIT Press (2006)