Exploring Sketch-based Sound Associations for Sonification

Lars Engeln[1], Mandy Keck[2]

[1]*Technische Universität Dresden, Chair of Media Design, Dresden, Germany*
[2]*University of Applied Sciences Upper Austria, Hagenberg, Austria*

Abstract

The interpretation of sounds can lead to different associations and mental models based on the person's prior knowledge and experiences. Thus, cross-modal mappings such as verbalization or visualization of sounds can vary from person to person. The sonification provides complements to visualization techniques or possible alternatives to more effectively support the interpretation of abstract data. Since sonifications usually map data attributes directly to auditory parameters, this may conflict with users' mental models.

In this paper, we analyze various sketch-based associations of sounds to better understand users' mental models in order to derive understandable cross-modal correlations for sonification. Based on the analysis of different sketches from a previously conducted user study, we propose three semantic-auditory channels that can be used to encode abstract data.

Keywords

sonification, mental model, sketching, auditory encoding, sound, visual encoding

1. Introduction

Cross-modal mappings have been well investigated over the past decades. Several studies have investigated how stimulations in one modality can influence the perception of information in another modality [1]. For instance, the *Bouba/Kiki* effect showed that subjects associated invented words such as *Bouba* with curvy shapes, whereas words such as *Kiki* were associated with jagged shapes.

Such findings have been identified repeatedly in recent years in similar research [2]. Thus, cross-modal associations between pitch and various visual features such as brightness, size, and position [1] and correlations between timbre and shapes [3] have been demonstrated. Moreover, there is evidence that such cross-modal correspondences are already active in infancy [4].

In recent years, the investigation of sonic expressions through mental representations by means of graphic sketches has emerged [5, 6]. Traditionally, sketching is part of creative thinking and a valuable thinking and communication tool in design [7, 8]. In the design disciplines, it has already been shown that sketching works in dialogue with the mental imagination and can be used as an external approximation for the inner image [7].

In the context of sound retrieval, visually sketching the desired sounds also appeared to be a viable and attractive workflow, as textual descriptions of sound material often do not reflect users' mental concepts [5]. Further preliminary work examined sketches as input for synthesizers [9]. In our previous work, we investigated hand-drawn sketches of spectral shapes in the context of a visual sound design [10], as well as for sound retrieval with machine learning approaches [11].

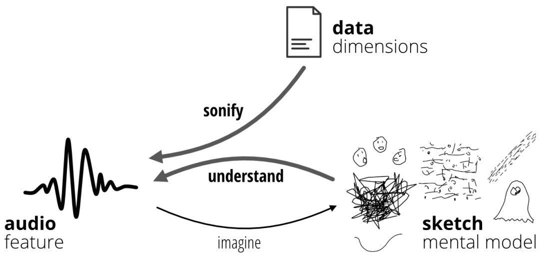

Figure 1: Supporting sense making from sketches as mental model for sonification.

In this paper, we examine sketches as mental models for sounds and extract properties from these sketches that may be relevant for sonification. A corpus of sketches was collected in our previous user study [11], in which subjects were asked to sketch their impressions of a range of sounds. The aim of the study was to learn more about the users' mental models when listening to different sounds. These sounds were usually short (up to 10 seconds) and ranged from natural and environmental sounds to instruments and synthetic soundscapes. The participants were free to sketch their associations when listening to each sound, but time was limited to capture intuitive associations. In the following, we will explore these associations and discuss how they can be used for the sonification of abstract data.

WAVA: AVI 2022 Workshop on Audio-Visual Analytics, June 07, 2022, Rome, Italy
lars.engeln@tu-dresden.de (L. Engeln);
mandy.keck@fh-hagenberg.at (M. Keck)

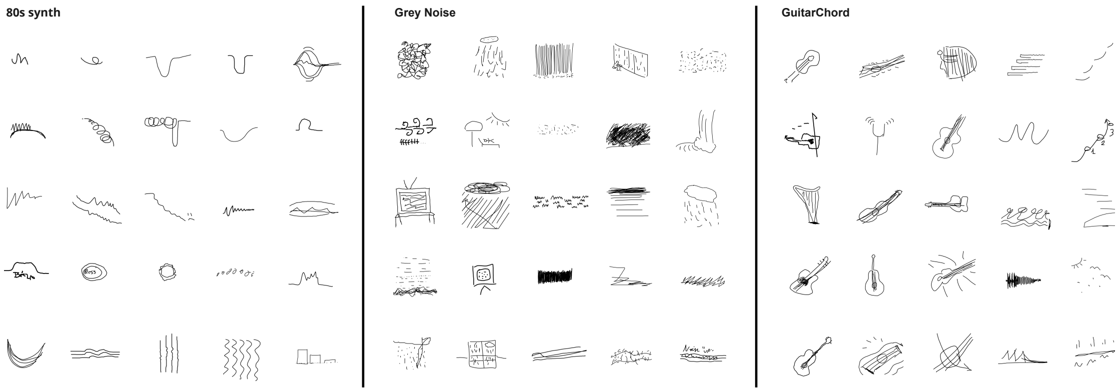

Figure 2: Exemplary overview of sketches derived from three different sounds with mainly abstract sketches (left), heterogeneous metaphorical associations (middle), and consistent, homogeneous metaphorical associations (right), (sound files for these three sketches: https://tinyurl.com/sketchedsound)

2. The Means of Mental Models for Sonification

Additional to common auditory channels such as loudness, pitch, and timbre (i.e. brightness, complexity), the perceived impressions of a sound should also be considered for sonification. By using the sketches from our previous user study [11], we can investigate how different sounds are perceived. When many people sketch the same or a similar image they associate with a sound, one can infer a general, homogeneous mental model. Figure 2 (right) shows various associations for the sound "GuitarChord" that leaves little room for interpretation and suggests a homogeneous mental model. On the other hand, a multitude of different sketches and thus different interpretations of a sound suggest a very heterogeneous mental model (see Figure 2, middle). When used for sonification, e.g., as a categorical mapping in the form of hearcons, it should be ensured that the sounds are clearly distinguishable from each other in their interpretation. Furthermore, it seems to be relevant for sonification, whether a sound is understood as mainly abstract (see Figure 2, left) or by possible metaphorical associations (see Figure 2, right).

By trying to understand how people formulate their mental model, information can be derived for the sonification of abstract data (see Figure 1). There are two ways of doing this: One way is to explore what graphic elements were used in the sketches to represent certain auditory features. Especially for audio-visual approaches, this can provide indications for the cross-modal mapping. On the other hand, the general structure and higher-level characteristics of the sketches can be utilized.

3. Interpretations of Sound

As mentioned in section 2, two main categories emerge by analyzing the sketches, representing different semantic interpretations of sounds. In Figure 3, these categories are distinguished into *abstract* and *metaphorical* associations. While the metaphorical category (see Figure 3, right) tends to associate the sound with the source that generates the sound (e.g., a car, a guitar, or hands clapping), the abstract category tends to describe the perceived properties. On the one hand, these properties can describe *temporal* aspects of sounds represented by waveforms, shapes, curves, and sinusoids (see Figure 3, left). The analysis of sounds and associated sketches revealed that participants often attempted to map tonality, loudness, or the combination of more complex sounds over time onto linear or radial visual structures. On the other hand, the abstract associations encompass the properties of a sound not only as a temporal structure but also as a *general impression* (see Figure 3, middle). These abstract images contain lines and shapes sketched while listening to the sound, such as round shapes, pointed edges, even spirals, or swirling, crisscrossing lines, and may indicate characteristics of a sound such as its sharpness, harmony or emotional impression.

4. Mapping of Data Attributes

In considering the different associations of sound, the question arises which of these properties can be used for the sonification of abstract data. As different visual channels can be used to visualize abstract data, auditory channels can also support the interpretation of abstract data. The table in Figure 4 shows suitable visual channels

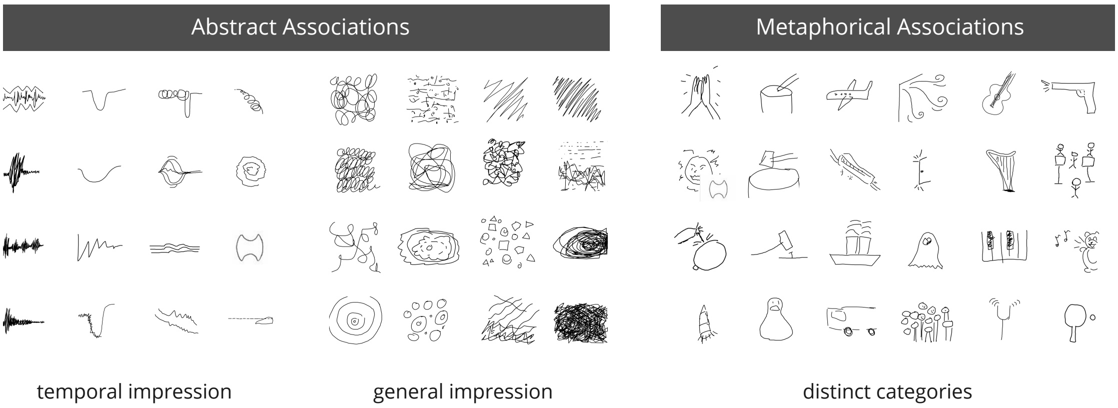

Figure 3: The sketches can be distinguished in abstract and metaphorical associations.

for different data attributes according to [12].

Nominal data attributes can be encoded by identity channels such as color or shape that are easily distinguishable from each other and do not associate any inherent order. These identity channels can be complemented by metaphorical sound associations: different categories of abstract data can be mapped to the sound sources, e.g., different instruments or associations, such as vehicles, animals, or tools (cp. Section 2). The analysis of sketches can support here in finding sounds with a homogeneous mental model that are easy to distinguish, such as a *pistol shot* and a *water drop* - instead of choosing sounds that are very similar such as different vehicles.

Ordered data attributes can be encoded by magnitude channels such as position, size, luminance and angle that suggest an implicit order of the abstract data. According to [12], different ordering directions can be distinguished: sequential, diverging and cyclic. Sequential ordering directions can be found in temporal structures, as one can distinguish e.g. different levels of loudness, pitches, and fluctuations.

Ordered data can also be sorted in a divergent ordering direction, such as data from temperatures or differences in elevation. In information visualization, magnitude channels and identity channels are often combined in this case, such as diverging color scales or movements in different directions. In this case, the general impressions associated with the sounds can be used to encode a spectrum of different properties. Examples are pairs of opposites like sharp to dull, aggressive & calm, or harmonic and dis-harmonic. These are derived from auditory-perceptual features or from the visual appearance of the mental model. The use and correct interpretation of these general impressions requires further research. For example, it would be interesting to evaluate to what extent the sharpness or dullness of a sound can be used for encoding uncertainty in data.

Data Attributes	Ordering Direction	Visual Channels	Semantic-auditory Channels
Nominal	-	identity channels: spatial region, color, shape, motion	source of the sound (generator)
Ordered	Sequential ⟶	magnitude channels: position, size, tilt, luminance	loudness, tonality, fluctuation
Ordered	Diverging ⟷	magnitude (+identity) channel	sharp < > dull aggressive < > calm, harmonic < > disharmonic

Figure 4: Mapping of data attributes to visual and semantic-auditory channels.

5. Conclusions

Audio-visual associations such as sketches can provide an extended knowledge about sounds that can be applied to sonification. By considering these associations, sounds can be described more meaningfully and the interpretation of abstract information can potentially become more effective. Based on the analysis of sketch-based associations, three different semantic-auditory channels have been proposed that can be combined with visual channels of information visualization to encode abstract data. The investigation to what extent these proposed channels can be used to expressively and effectively encode and discriminate abstract data is part of our future work.

References

[1] K. K. Evans, A. Treisman, Natural cross-modal mappings between visual and auditory features, Journal of Vision 10 (2010) 6–6. URL: https://doi.org/10.1167/10.1.6. doi:10.1167/10.1.6.

[2] C. P. Davis, H. M. Morrow, G. Lupyan, What does a horgous look like? nonsense words elicit meaningful drawings, Cognitive Science 43 (2019). doi:10.1111/cogs.12791.

[3] M. Adeli, J. Rouat, S. Molotchnikoff, Audiovisual correspondence between musical timbre and visual shapes, Frontiers in Human Neuroscience 8 (2014). doi:10.3389/fnhum.2014.00352.

[4] O. Ozturk, M. Krehm, A. Vouloumanos, Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds, Journal of Experimental Child Psychology 114 (2013) 173–186. doi:10.1016/j.jecp.2012.05.004.

[5] P. Knees, K. Andersen, Searching for audio by sketching mental images of sound, in: Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval - ICMR '16, 2016. doi:10.1145/2911996.2912021.

[6] S. Delle Monache, D. Rocchesso, Sketching sonic interactions, Foundations in Sound Design for Embedded Media, A Multidisciplinary Approach (2019) 79–101. doi:10.4324/9781315106359-4.

[7] G. Goldschmidt, Manual Sketching: Why Is It Still Relevant?, Springer International Publishing, Cham, 2017, pp. 77–97. URL: https://doi.org/10.1007/978-3-319-56466-1_4. doi:10.1007/978-3-319-56466-1_4.

[8] B. Buxton, Sketching User Experiences, Interactive Technologies, Morgan Kaufmann, Burlington, 2007. doi:https://doi.org/10.1016/B978-012374037-3/50045-6.

[9] S. Löbbers, M. Barthet, G. Fazekas, Sketching sounds: an exploratory study on sound-shape associations (2021). URL: https://arxiv.org/abs/2107.07360. doi:10.48550/ARXIV.2107.07360.

[10] L. Engeln, R. Groh, Cohearence of audible shapes—a qualitative user study for coherent visual audio design with resynthesized shapes, Personal and Ubiquitous Computing (2020) 1–11. doi:10.1007/s00779-020-01392-5.

[11] L. Engeln, N. L. Le, M. McGinity, R. Groh, Similarity analysis of visual sketch-based search for sounds, in: Audio Mostly 2021, AM '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 101–108. URL: https://doi.org/10.1145/3478384.3478423. doi:10.1145/3478384.3478423.

[12] T. Munzner, Visualization Analysis and Design, AK Peters Visualization Series, CRC Press, 2014. doi:https://doi.org/10.1201/b17511.