ENTROPY SONIFICATION

Chandrasekhar Ramakrishnan

Native Systems Group
ETH Zürich
Zürich, Switzerland
cramakrishnan@acm.org

Steven Greenwood

Institut für Design- und Kunstforschung
Hochschule für Gestaltung und Kunst
Basel, Switzerland
steven@telesthetic.org

ABSTRACT

We present *entropy sonification*, a technique to bring interesting data to the foreground of a sonification and push uninteresting data into the background. To achieve this, the data is modeled as an information source, and the underlying sonification is converted into sound grains. Information-theoretic attributes of the data are used to determine the amplitude envelope and duration of the grains. The information source model adds an additional avenue for control. By altering the information source model, one can focus on different aspects of the data via *entropy zooming*.

[Keywords: Sonification Design, Interaction Design, Information Theory, Granular Synthesis]

1. INTRODUCTION

How should a sonification command the user's attention? It is this question that motivates the present research. Think of an automobile engine. When driving a car, even one we are not familiar with, we quickly become acquainted with the sound of its properly functioning engine. Once we have reached that point, we do not have to actively pay attention to the engine sound anymore because it provides us little information. But when something is wrong, the engine is revving too high, for example, we immediately become aware of the sound again and intuitively know how to react.

It is probably fair to say that creating a sonification that enables a similarly intuitive interaction between the user and the data is the goal of many, if not all, sonification techniques. That is, the sonic feedback provided by a sonification should be rich enough that a user who is well acquainted with it immediately understands its meaning and automatically knows how to react.

Some sonification designs focus on creating richness in the translation of data to sound and trust that this richness carries over and supports an intuitive interaction with the data once the user has learned to interpret the sonification. This is a reasonable approach: after all, in the example above, the sound of the automobile engine has not been designed to optimize communication with the user. The communication is a byproduct of the users familiarity with the mode of interaction.

Along these lines, de Campo proposes a general framework for designing sonifications with the aim that "structures/patterns in the data (which are not known beforehand) emerge as perceptual entities in the sound which jump to the foreground..." [2]. In this framework, the way data communicates and emerges as a perceptual entity is a byproduct of considering the dimensionality and sampling rate of the data with the desired number of streams of output and picking an appropriate sonification technique on that basis.

This is a perfectly legitimate approach. Other sonification designs, however, directly confront the issue of interaction and communication. In the development of auditory icons, Gaver ([4], Section 3) argues for using everyday sounds that are already familiar to the user and argues against mapping data parameters to musical parameters precisely because everyday sounds can be intuitively interpreted more easily.

As with Gaver's work, the present research is concerned with designing a sonification that can be understood by an audience with little or no training. This work was developed to support a sound installation presenting a sonification in an art museum for an audience that is not necessarily familiar with the auditory display of data and does not have the time to thoroughly acquaint themselves with the sonification. One possible response is to simplify the conversion of data to sound in the hope that this makes the sonification more accessible. This, however, runs the risk that the sonic material then becomes too meager and is both boring to a lay audience and insufficient for conveying the data to an expert one.

We have chosen to retain the full multi-dimensional complexity of the sonification and focus instead on drawing the listeners' attention to the interesting elements of the data set. So even if the listener does not necessarily know what the sound is telling him or her about the data, s/he at least knows what to pay attention to. The technique for doing this is entropy sonification. It is a way of guiding the listener to the significant information in the data set using rhythmic variation and accentuation, a standard tool from music composition and therefore very likely familiar to many listeners.

We present entropy sonification by first providing motivation (Section 2), and then describing the necessary information theory background (Section 3) before giving a precise definition (Section 4) and two illustrative examples (Sections 5 and 6). Entropy sonification adds an additional means of control: the description of the data as an information source. Section 7 considers the selection of different information source descriptions for one data set and introduces entropy zooming.

2. ENTROPY SONIFICATION

Entropy sonification is based on the following assumptions about the way sound should provide feedback about a system: a working system always makes sound; this sound normally disappears into the background and becomes insistent only when something is abnormal or amiss. These assumptions are restated and summarized in Table 1.

The automobile example presented above conforms to these assumptions. In this case, though, the design of the engine sound itself does nothing to support this type of feedback: the human brain must do all the work to push the engine sound into the background or foreground. We, on the other hand, would like to design a sonification that does its best to provide clues as to what is important and what can be ignored.

In order to realize this, we need a measure that determines, given some data, which elements are typical and which ones are

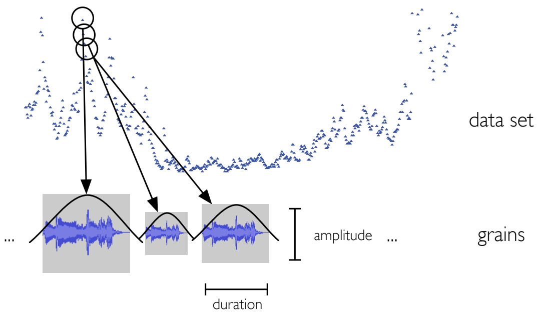

Figure 1: *Schematic Depiction of Entropy Sonification.* Amplitude and duration are functions of the ratio of information content to information entropy.

1. Elements (of a data set or event stream) that are typical should be heard, but unobtrusive
2. Elements that are atypical should be highlighted, but within the constraints of the sonification

Table 1: *Entropy Sonification Goals*

unexpected. Information theory is one possible framework for quantifying the amount of "typicalness" or "unexpectedness" of a data point.

Information theory treats data as the product of an information source, which is simply modeled as a stochastic process ([14], section on discrete noiseless systems). This can be justifiably criticized as a gross oversimplification. Nonetheless, information theory has been successfully applied to a variety of problems in data transmission and compression and has proven to be a powerful framework for creating and analyzing algorithms for communicating information.

The information source model forms the basis of entropy sonification. This model is used to "tell" the sonification which data points to highlight and which ones to push into the background. The highlighting/backgrounding of data points is accomplished though rhythmic variation: important points are accented and have a longer duration than others. Rhythmic variation was chosen since it is known that rhythm is a more salient aspect of music cognition than melody or timbre. Dowling and Harwood summarize research on the importance of rhythm vs. melody as follows: "It is noteworthy that the rhythmic aspects of these stimuli overrode their pitch-pattern aspects in determining listeners' responses. This clearly demonstrates the importance of rhythm in music cognition." ([3] p. 193)

To control rhythmic variation, we need to vary the loudness, duration, and attack of the sound. Granular synthesis is a natural technique for realizing this. The urparameters of a grain are duration and amplitude envelope or window (Roads' *Microsound* [13], discussion in Chapter 3), which map perfectly to the variables we need to control. Figure 1 presents a graphical depiction of this scheme.

One final issue must be settled – the order in which the grains are played. In some cases, as in the examples we present, there is a natural order to the data which can be used by the sonification as well. In situations where no natural order exists, an ordering must be artificially produced, for example, using the distance from a user selected point.

3. INFORMATION THEORY BACKGROUND

We provide the minimum fundamentals of information theory necessary for the presentation of the sonification technique. For a more extensive treatment, we refer the reader to the source, Shannon and Weaver's *The Mathematical Theory of Communication* [14], or David MacKay's *Information Theory, Inference, and Learning Algorithms* [8] for a broader and more modern presentation.

An **information source** X is determined by two sets of N elements: an alphabet $A_X = \{a_1, a_2, ..., a_N\}$, which defines the set of possible values, and a probability distribution $P_X = \{p_1, p_2, ..., p_N\}$, which defines a probability for each element of A_X. We use the notation $P(a)$ to denote the probability of an element of the alphabet. In particular, $P(a_i) = p_i$.

The **Shannon information content** of an element a of the alphabet A_x is defined as

$$h(a) = log_2 \frac{1}{P(a)}. \tag{1}$$

We will just use "information content" or "information" to mean Shannon information content.

The **information entropy** of an information source X is defined as

$$H(X) = \sum_{a \in A_X} P(a) log_2 \frac{1}{P(a)} = \sum_{a \in A_X} P(a)h(a) \tag{2}$$

In other words, the information entropy is the expected value, or average, of information over the information source.

4. SONIFICATION DESIGN

Given an information source, X, and a data set or event stream, E, which uses the alphabet A_X, we define the entropy sonification of E as a granular synthesis process that generates a grain for each element of E. Each grain has the property that its duration and volume are functions of the ratio of information to entropy. We do not make any other requirements on the content of the grain.

Thus, duration and volume are defined by functions f and g, respectively:

$$\text{duration of grain} = duration(x) = f(h(x)/H(X)) \tag{3}$$

$$\text{volume of grain} = volume(x) = g(h(x)/H(X)) \tag{4}$$

4.1. Sonification Duration

Let us consider the special case where the duration function is multiplication by a constant, d:

$$duration(x) = d \frac{h(x)}{H(X)} \tag{5}$$

In this case, the expected value of the duration of a grain is:

$$\mathcal{E}[duration] = \frac{d}{H(X)} \sum_{a \in A_X} P(a)h(a) = d \tag{6}$$

Thus, given N points of data, $e_1 ... e_N$, the expected length of the sonification is:

$$\mathcal{E}[length] = \sum_{i=1}^{N} \mathcal{E}[duration(e_i)] = \sum_{i=1}^{N} d = Nd \tag{7}$$

In other words, if grains are played one after another without overlap, the total duration of the sonification should be the duration per grain times the number of data points, assuming the information source accurately represents the data. This is a nice property which can be used to gauge the accuracy of the information source description.

4.2. Uniform Distribution

Another special case to consider is an information source based on the uniform distribution; thus all elements of the alphabet have equal probability. Using the notation $|A_X|$ to represent the number of elements in A_X, this gives the following expressions for information content of an element of the alphabet and entropy of the information source:

$$h(a) = log_2 \frac{1}{P(a)} = log_2 |A_X|, a \in A_X \tag{8}$$

$$H(X) = \sum_{a \in A_X} P(a)h(a) = log_2 |A_X| \tag{9}$$

Since every symbol has the same information, which is equal to the entropy of the information source, this results in a sonification in which every grain has the same duration and envelope. This information source model, the uniform information source, is useful as a basis for comparison.

4.3. Examples and Limitations

In the following two sections, we present examples of entropy sonification. For the purposes of exposition, we restrict ourselves to the simple case of discrete information sources and two dimensional data sets. This is not a limitation of the theory, but a desire to keep the exposition simple. Entropy sonification can be applied to any dimensionality of data; only a description of the data as an information source is necessary.

Our examples use simple histograms as the basis of information-source models. In general, any stochastic process can serve as an information source. And though we only present and discuss discrete information sources, the theory extends to continuous information sources [14], and entropy sonification extends naturally to this situation as well.

5. EXAMPLE 1: TEXT

Taking our lead from Claude Shannon, we first turn our attention to a sonification of English-language text. Text has two attributes that make it a good candidate for entropy sonification: it is a discrete information source, as opposed to a continuous information source such as one defined by a Gaussian distribution; and letter frequency tables to define the model are widely available.

The source code and audio for this example may be downloaded from our web site [12].

5.1. Data

The "data," in this case text, is taken from the beginning of Thomas Pynchon's *Gravity's Rainbow* [11]:

> A screaming comes across the sky. It has happened before, but there is nothing to compare it to now.
>
> It is too late. The Evacuation still proceeds, but it's all theatre. There are no lights inside the cars. No light anywhere. Above him lift girders old as an iron queen, and glass somewhere far above that would let the light of day through. But it's night. He's afraid of the way the glass will fall – soon – it will be a spectacle: the fall of a crystal palace. But coming down in total blackout, without one glint of light, only great invisible crashing.

5.2. Grain Waveform Design

The text sonification architecture is implemented in version 3 of James McCartney's programming language, SuperCollider [9]. Though this is a sonification, not a text-to-speech system, it does take a few cues from speech synthesis. Letters are converted to sound differently based on whether the letter is a vowel or consonant. Vowels are synthesized as a combination of a fundamental frequency and two formants with frequencies specified by Peterson and Barney's table [10]. Consonants are synthesized as filtered noise.

The uniform information source applied to this grain design yields the sonification displayed in Figure 2.

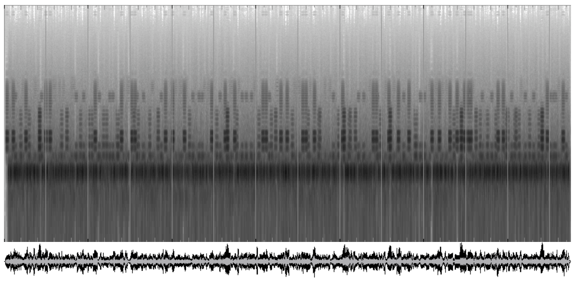

Figure 2: *Text Sonification Using the Uniform Distribution in Frequency and Time Domains; Duration: 0:34*

5.3. Model

Using a letter frequency table [15] that provides frequencies for 27 characters, the 26 letters of the English alphabet plus space, we can compute the information content for each letter, plus the information entropy for the alphabet as a whole, as displayed on the next page in Figure 4. In our model, we ignore punctuation marks such as commas, periods, etc.

This model applied to the text produces the sonification shown here in Figure 3.

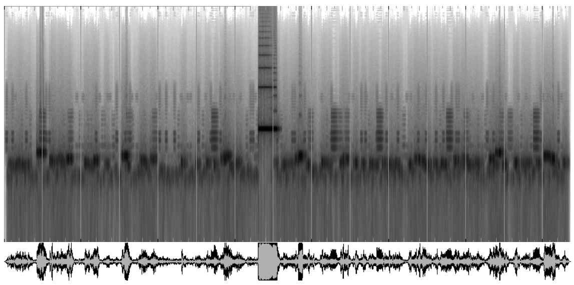

Figure 3: *Text Sonification in Frequency and Time Domains; Duration: 0:37*. Letters with lower probability stand out, in particular the "q" near the middle of the sonification.

6. EXAMPLE 2: TED SPREAD

This example is presents a sonification of the TED spread. The TED spread is a measure of the stress or risk in the banking system [17]. One way to measure stress in the banking system is to look at

Character	Probability	Information
[ ]	0.190	2.393
E	0.102	3.298
T	0.074	3.762
A	0.066	3.919
O	0.061	4.035
I	0.056	4.153
N	0.056	4.166
H	0.054	4.206
S	0.051	4.299
R	0.046	4.449
D	0.037	4.760
L	0.033	4.943
U	0.023	5.455
M	0.021	5.608
C	0.019	5.703
W	0.019	5.718
F	0.018	5.837
Y	0.017	5.921
G	0.016	5.957
P	0.013	6.254
B	0.012	6.442
V	0.009	6.828
K	0.007	7.243
X	0.001	9.480
J	0.001	10.288
Q	0.001	10.288
Z	0.001	10.966
Entropy		4.078

Figure 4: *English Letter Probabilities and Information*

the difference between the interest rates banks charge one another when lending money to each other and the interest rate provided by a "low-risk" investment. If banks are safe and trust one another, this difference should be small.

As with the previous example, the source code and audio may be downloaded from our web site [12].

6.1. Data

The TED compares the returns provided by U.S. T-Bills, the low-risk investment, with the LIBOR (London Interbank Offered Rate) rate banks charge one another (this was formerly referred to as a EuroDollar contract; the names "T-Bill" and "EuroDollar" are the source of the name "TED"). Current and historical data necessary to compute the TED spread is available from the U.S. Federal Reserve [1]. This data is graphed in Figure 5.

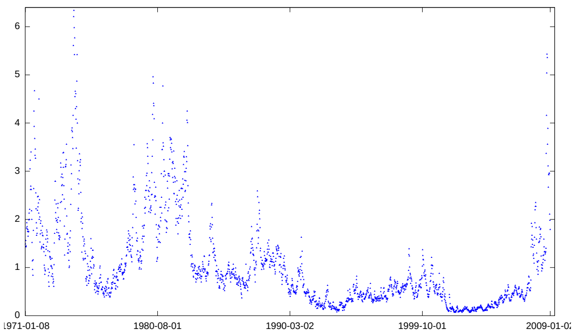

Figure 5: *The Weekly TED Spread, 1971 – 2009*

6.2. Grain Waveform Design

Grains are noise filtered by two bandpass filters, one for each component of the TED Spread. The center frequencies of the filters are determined by the interest rate on T-Bills and LIBOR contracts, respectively. The quality, Q, parameter of the filter and stereo pan are determined by the value of the TED Spread itself. The smaller the TED Spread, the smaller the Q value and the narrower the stereo separation.

The sonification in Figure 6 is the result of this grain design and the uniform information source.

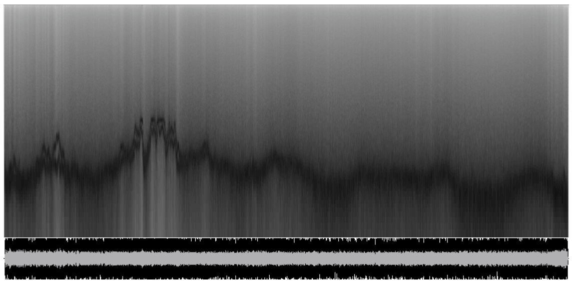

Figure 6: *TED Sonification Using the Uniform Distribution in Frequency and Time Domains; Duration: 2:04*

6.3. Model

To apply entropy sonification, we need a model of the TED spread as an information source, which we construct with the help of a histogram. One could construct a more precise model, but even this simple one yields a sonification that brings out many salient features of the data set.

Figure 7 shows the distribution of values in the TED spread series. For the purposes of computing information content, we concern ourselves solely with which band in the histogram each value falls. This model is summarized in Figure 8.

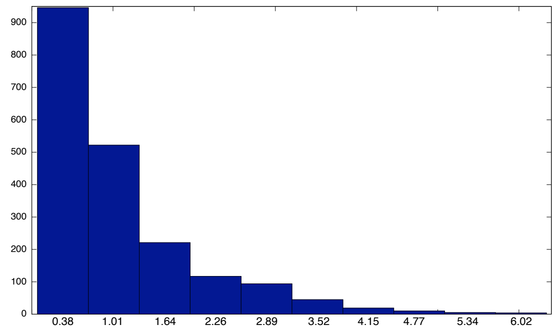

Figure 7: *A Histogram of the TED Spread*. The x-axis is labeled with the center of each bin; the y-axis is the number of values that fall within the bin.

The resulting sonification is shown in Figure 9.

7. CONTRIVED MODELS

In the examples above, we have constructed information sources based on empirical analysis of the subject of the sonification. This is not the only way to produce an information source. In some situations, there might be an *a priori* model derived from other considerations. Another possibility is the construction of a model

Bin	Bin Max	Probability	Information
1	0.63	0.48	1.05
2	1.25	0.26	1.94
3	1.88	0.11	3.18
4	2.51	0.06	4.10
5	3.14	0.05	4.41
6	3.76	0.02	5.47
7	4.39	0.01	6.72
8	5.02	0.005	7.64
9	5.64	0.002	8.64
10	6.27	0.002	8.97
Entropy			2.08

Figure 8: *TED Bin Probabilities and Information*

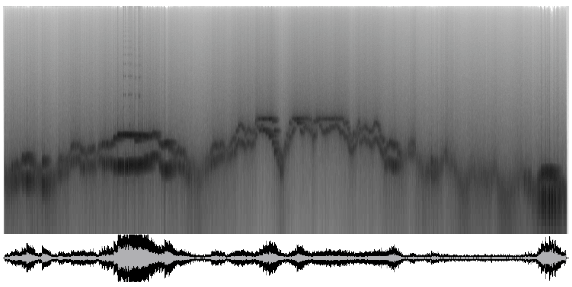

Figure 9: *TED Sonification in Frequency and Time Domains; Duration: 1:57*

that is intentionally false for the purpose of focusing on a particular aspect of the data set.

7.1. Entropy Zooming

The goal of entropy sonification is to make interesting elements of a data set stand out. An artifact of using Shannon information as the basis for entropy sonification is the implicit assumption that interesting elements are rare ones. This may not actually be the case. The interesting elements could actually be the common ones, and the rare ones might be uninteresting.

This problem is solved with entropy zooming. Entropy zooming involves constructing a model of a data set as an information source with the goal of bringing out particular characteristics of the data. We can do this by starting with an existing model of the data (e.g., one derived from analysis) and *reducing* the probability of the data with the characteristics we want to focus on and increasing the probability of the other data.

To illustrate this, we take the TED spread example and define a new information source model to focus on the smaller values of the TED spread by making them less likely according to our new, "zoomed" model. This gives us a new model, summarized in Figure 10, which results in the sonification in Figure 11.

8. ENTROPY SONIFICATION AND MODEL-BASED SONIFICATION

Hermann and Ritter provide a taxonomy of sonification techniques [6], breaking them down into the following categories:

1. Auditory Icons
2. Earcons
3. Audification
4. Parameter Mapping
5. Model-Based Sonification

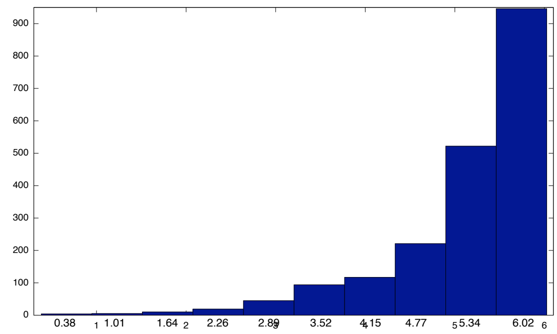

Figure 10: *Alternative Histogram of the TED Spread for Zooming*

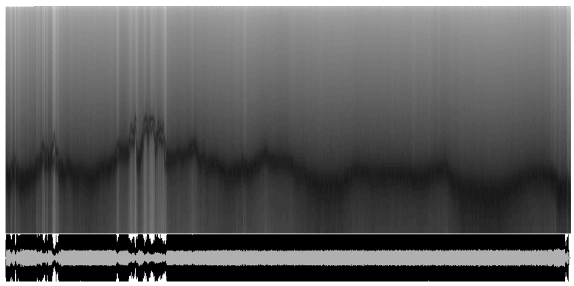

Figure 11: *Zoomed TED Sonification in Frequency and Time Domains; Duration: 8:34*

Entropy sonification is most similar to parameter mapping, but it is different in two significant ways. First, entropy sonification does not map data points to sound-synthesis parameters; rather, it maps *information-theoretic properties* of data points to sound-synthesis parameters. Second, whereas choosing from the techniques listed above is a mutually exclusive choice, entropy sonification may be used in combination with any other sonification technique.

In this capacity, the combination of entropy sonification and model-based sonification presents a particularly interesting case study. The goal of model-based sonification is to interpret data such that it becomes a kind of "sounding instrument" [6]. There are a variety of ways to achieve this, as demonstrated by T.Hermann in his dissertation [7]. In fact, one of the proposed models, the Markov chain Monte Carlo (McMC) simulation sonification model, actually incorporates a kind of entropy sonification: "The amplitude is used to communicate the 'interestingness' of the mode by using loud grains for modes which are rarely visited" (ibid. p. 105).

But whereas the use of amplitude in the McMC simulation sonification model is an artifact of that particular model design, the technique described in this paper may be applied to any model-based sonification. For example, crystallization sonification [5] is a particularly good candidate to augment with entropy sonification. Crystallization sonification defines a process for generating a timbre, given a data set in $\mathbb{R}^n$ and a point in $\mathbb{R}^n$. That is to say, crystallization sonification is a technique for creating an instrument from a data set. Entropy sonification, on the other hand, generates a rhythm from a data set, or, in other words, a score independent of the instrument used to play it. This appraisal applies generally to the two techniques: model-based sonification defines an instrument; entropy sonification defines a score. This is precisely what makes their combined use potentially so interesting.

9. CONCLUSIONS AND FUTURE WORK

The research described in this paper, entropy sonification, presents a technique for focusing the user's attention on particularly interesting portions of a sonification. It accomplishes this systematically, using well-understood ideas from information theory. The mechanics of the technique make it possible to define and apply different measures of what is interesting, allowing entropy zooming.

Though the research is presented here in an idealized form, it has been developed to support *eMotion*, a multi-disciplinary research project at the Institut für Design- und Kunstforschung in the Hochschule für Gestaltung und Kunst Basel FHNW [16]. In this project, we are investigating how museumgoers interact with artworks by collecting realtime data from visitors to the Kunstmuseum in St. Gallen, Switzerland as they wander through an exhibition conceived specifically for this purpose in June and July 2009. In addition to being stored for later analysis, the data will be presented to the visitors in the form of a visual projection and sound installation. We expect to have more to report in the future about our experience using entropy sonification in the context of this sound installation.

Furthermore, we think information theory is a theoretical framework with unexplored potential for not just designing sonifications, as we have done here, but also analyzing and comparing sonification techniques. Information theory enjoyed a period of popularity in music theory, with prominent exponents such as Leonard Meyer. It thus seems plausible that an undertaking which applies information theory to the realm of auditory display could also yield illuminating results.

10. ACKNOWLEDGEMENTS

The authors would like to thank Prof. Curtis Roads, Dr. Oliver D. King, and Prof. Alberto de Campo as well as the eMotion team for their helpful criticisms and comments. This work was generously funded by the Schweizerischer Nationalfonds and Ubisense.

11. REFERENCES

[1] Board of Governors of the Federal Reserve System, "Federal Reserve Statistical Release," <http://www.federalreserve.gov/releases/H15/data.htm> [accessed 9 Jan 2009]

[2] A. de Campo, "Toward a Data Sonification Design Space Map," in *Proc. of the 2007 International Conference on Auditory Display*, Montréal, Canada, June 2007.

[3] W.J. Dowling and D.L. Harwood, *Music Cognition*, Academic Press, San Diego, California, 1986.

[4] W. W. Gaver, "The SonicFinder, a prototype interface that uses auditory icons," *Human Computer Interaction*, vol. 4, no. 1, pp 67–94. 1989.

[5] T. Hermann, H. Ritter "Crystallization Sonification of High-Dimensional Datasets," in *Proc. of the 2002 International Conference on Auditory Display*, Kyoto, Japan, July 2002.

[6] T. Hermann, H. Ritter "Listen to your Data: Model-Based Sonification for Data Analysis," in *Int. Inst. for Advanced Studies in System Research and Cybernetics*, 1999, pp 189–194.

[7] T. Hermann, "Sonication for Exploratory Data Analysis," Ph.D. thesis, Bielefeld University, Bielefeld, Germany, Feb 2002.

[8] D.J.C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, Cambridge, United Kingdom, 2003.

[9] J. McCartney, "SuperCollider: a New Real Time Synthesis Language," in *Proc. of the 1996 International Computer Music Conference*, Hong Kong, 1996

[10] G.E. Peterson, and H.L. Barney, "Control Methods Used in a Study of the Vowels," *Journal of the Acoustical Society of America*, vol. 24, no. 2, pp. 175–184, Mar. 1952.

[11] T.R. Pynchon, *Gravity's Rainbow*, (1987 ed.), Penguin Putnam Inc, New York, New York, U.S.A., 1973.

[12] C. Ramakrishnan. "Entropy Sonification Implementation in SuperCollider." <http://www.illposed.com/research/entropy_sonification/>

[13] C. Roads, *Microsound*, MIT Press, Cambridge, Massachussetts, 2004.

[14] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Illinois, United States, 1949.

[15] B. von Sydow, "Mono-, Bi and Trigram Frequency for English," 17 Oct 2008, <http://www.cs.chalmers.se/Cs/Grundutb/Kurser/krypto/en_stat.html> [Accessed 22 Dec 2008].

[16] M. Tröndle, S. Greenwood, et. al. *eMotion*, <http://www.mapping-museum-experience.com/> [accessed 20 Mar 2009].

[17] Wikipedia contributors, "TED Spread," *Wikipedia, The Free Encyclopedia*, 13 Dec 2008, <http://en.wikipedia.org/w/index.php?title=TED_spread&oldid=257712580> [accessed 8 Jan 2009].