CRYSTALLIZATION SONIFICATION OF HIGH-DIMENSIONAL DATASETS

T. Hermann

Faculty of Technology
Bielefeld University, Germany
`thermann@techfak.uni-bielefeld.de`

H. Ritter

Faculty of Technology
Bielefeld University, Germany
`helge@techfak.uni-bielefeld.de`

ABSTRACT

This paper introduces *Crystallization Sonification*, a sonification model for exploratory analysis of high-dimensional datasets. The model is designed to provide information about the intrinsic data dimensionality (which is a local feature) and the global data dimensionality, as well as the transitions between a local and global view on a dataset. Furthermore the sound allows to display the clustering in high-dimensional datasets. The model defines a crystal growth process in the high-dimensional data-space which starts at a user selected "condensation nucleus" and incrementally includes neighboring data according to some growth criterion. The sound summarizes the temporal evolution of this crystal growth process. For introducing the model, a simple growth law is used. Other growth laws which are used in the context of hierarchical clustering are also suited and their application in crystallization sonification offers new ways to inspect the results of data clustering as an alternative to dendrogram plots.

In this paper, the sonification model is described and example sonifications are presented for some synthetic high-dimensional datasets.

1. INTRODUCTION

Understanding the structure of high-dimensional data is one of the great challenges in data analysis. The research field of data mining [1] is concerned with the development of techniques to detect patterns or hidden regularities in such data. For the selection of a data mining technique, however, the human researcher first needs some evidence about what pattern or structure there might be in the data at hand [2]. Such evidence can be gained by exploratory data analysis, which therefore is an important element for solving data mining tasks. Whereas during the last decades mainly visualization techniques were considered to present high-dimensional data[3], recently more and more techniques for displaying data by sound are being considered[4].

In contrast to visualizations which are often static, sound is an intrinsically dynamic medium. In our environment sound allows us to follow the evolution of dynamical processes, and in many situations we are furthermore able to intervene into our environment and to manipulate the sounding object or even cause sounds by exciting an object. Obviously our auditory system has evolved to extract useful information about a sounding object from this kind of indirect connection between object properties and sound. This motivates to look for processes within the context of data exploration which can be used for an acoustic presentation in the same manner. Model-Based Sonification [5] provides a framework for the definition of dynamical data-driven models which may generate sound in this way. A sonification model defines (i) a data-driven setup of elements, (ii) its dynamical evolution and (iii) how a user can interact with the model.

The Crystallization Sonification Model provides a spatial setup of unconnected data points, a condensation law which determines how data points are integrated into a crystal, and an interaction type, which determines how a user can initiate crystal growth by supplying a condensation nucleus. Integration of data into the crystal set changes the shape of the crystal and adds acoustic energy to the crystal solid. The shape determines the crystal's acoustic properties (its timbre) while energy is related to sound level. The sound provides information about the intrinsic (local) dimensionality of data [6], determined by an incremental update of the crystal dataset covariance matrix.

The paper is structured as follows: Section 2 motivates Crystallization Sonification. In Section 3, the model is introduced in detail and the sound rendering is explained. Section 4 then gives sound examples for some synthetic datasets. The paper closes with a conclusion.

2. CRYSTALLIZATION SONIFICATION

The Data Crystallization Model defines a technique to browse high-dimensional data for exploring both their neighborhood structure and the intrinsic dimensionality of the data.

Regarding instrument sounds in our real environment, we find that sounds can often be divided into two parts: a transient phase and a stable state corresponding to quasi-stationary oscillations. For identification of instrument classes, the transient phase is particularly important. In the dynamical evolution of the sound, detailed information is given about the instrument and there are specialists which are even able to distinguish violins made from different violin builders by listening to them.

The aim of the following sonification model is to define a dynamical complex sound pattern in such a way that the global data distribution determines the "instrument characteristics" whereas local properties determine the initial phase of the sonification. For this purpose a spatial model is used, in which each data point is represented by a point in a vector space. Starting at a user selected vector which is determined by browsing a scatter plot of the given dataset, a crystallization process is initiated by including neighboring points successively into the growing "crystal solid". With each point inclusion, the acoustic properties of the solid changes. The dynamics is defined so that the sound reflects the intrinsic dimensionality of the data. The sonification stops when all data points have been included into the solid. As this process resembles the process of crystallization starting at the condensation nucleus, this metaphorical name was chosen for the model.

For sound generation, modal synthesis is applied [8]. The

modes are taken to be integer multiples of the fundamental frequency (harmonics). The energy within the modes is determined by the eigenvalue spectrum of the data covariance matrix.

The Crystallization Model provides information about the intrinsic dimensionality of the data. Furthermore, although this is a merely side effect, clustering can be perceived.

3. MODEL DESCRIPTION AND SONIFICATION DESIGN

Setup

The Crystallization Model is based on a spatial model. The model space is an Euclidian vector space $\mathbb{R}^d$ in which data points are located at $\mathbf{x}_i$, $i = 1, \ldots, N$.

The setup does not include any dynamical elements. The model of a crystal is only used as a metaphor to simplify the interpretation of the sonifications.

Dynamics

Starting from a user selected coordinate vector $\mathbf{x}_c$ called the condensation nucleus, a crystal growth starts. This means that data points are included one by another into the set of points $I_c(t)$ which form the crystal at time t. For data point inclusion, there are various possibilities. One method is to order all points according to their distance to $\mathbf{x}_c$ and include them in this order. Alternatively, one could include at each step the point which has the smallest distance to the crystal. Generally, the different strategies for agglomerative clustering [7] can be applied here.

To start with, the following simple inclusion rule is used: inclusion of a point with coordinate vector $\mathbf{x}$ is done at sonification time

$$t = \mathrm{map}(\|\mathbf{x} - \mathbf{x}_c\|, [0, \max_i \|\mathbf{x}_i - \mathbf{x}_c\|], [0, T]) \qquad (1)$$

where T is the total duration of the sonification. $\mathrm{map}(x, [a, b], [c, d])$ maps a value x linearly from a source interval $[a, b]$ to a destination interval $[c, d]$.

The crystal timbre is given by a superposition of its different acoustic modes and modal synthesis is applied for sound rendering. How timbre depends on the crystal set is described below.

Initial State and Excitation

Let $I_c(t)$ denote the subset of data points in the crystal at time t. The initial conditions are given by the empty set $I_c(0) = \{\}$. As there is no crystal existing at the beginning, the initial sound of the model is silence. Excitation means in this model to initiate a crystallization process. This is done by selecting a starting point for the crystal growth.

Interaction Types

The user interacts with the system by using the mouse pointer. Clicking the mouse into a 2D scatter plot of the dataset triggers the sonification. The coordinates of the nearest neighbor to the mouse coordinates in the scatter plot are taken for the condensation nucleus vector $\mathbf{x}_c$. Alternatively, the mouse pointer coordinates can be used to determine the position within the visualized subspace and column means are used for the other vector components.

Data-Sound Linkage

The sonification is computed from the temporal evolution of the crystal. As no explicit dynamics is defined, the sound pressure is not bound directly to any dynamic element. Details about the linkage between data properties and modal structure of the crystal are given below.

Listener

The listener is not localized with respect to the model space and thus a monaural sonification is rendered. A second (stereo) audio channel thus remains free for other purposes like e.g. comparison of two sonifications of the same dataset started from different points or comparison of sonifications for different datasets.

Sound Synthesis

For sound computation, modal synthesis [8] is applied. This means that the sonification is computed by using additive synthesis

$$s(t) = \sum_{i=1}^{d} a_i(t) \sin(\omega_i(t)t + \phi_i) \qquad (2)$$

where the mode frequencies ω_i and the amplitude of these oscillations a_i are time-variant functions.

For the implementation, the time series $a_i(t)$ and $\omega_i(t)$ are discretized using control points on a 5 msec grid. The synthesis is done by superimposing the modes into an output vector $s(t)$. $a_i(t)$ and $\omega_i(t)$ are interpolated between control points. The basic building block of the sound thus is a time-variant sine oscillator with amplitude and frequency control. All initial phase values ϕ_i are set to 0. They are not relevant for the auditory perception since all modes are in the audible frequency range and multiples of the fundamental frequency so that rhythmical elements like beats cannot occur.

Data-Model Assignment

The model determines the inclusion order of data points into a growing crystal. This does not suffice to determine the sonification. Additionally it has to be defined how the crystal's acoustic properties depend on an included element at each time t.

The aim of the model is to encode information about the crystallization process into the temporal and spectral evolution of sound. For this a closer look to the temporal evolution of timbre and envelope of physical acoustic objects is taken now. Many musical instruments produce almost periodic oscillations of air pressure. Periodic oscillations can be characterized by a set of Fourier series coefficients. The modes are harmonic oscillations whose frequencies $f_n = nf_1$ are integer multiples of a fundamental frequency f_1. A simple description of timbre can be given by the vector of partial tone amplitudes $(a_1, a_2, \ldots)$, also called timbre vector. Figure 1 shows a timbre vector for a violin and a metal flute.

However, synthesized periodic signals that use the appropriate timbre vectors still sound different from the real instrument. That is because the temporal evolution of timbre is missing which is also an essential part for the characterization of musical instruments. Especially the attack phase (usually the first 50-80 msec) is important for humans to classify and distinguish instrument sounds. Another important aspect for the identification of an instrument is the temporal evolution of sound level, the amplitude envelope
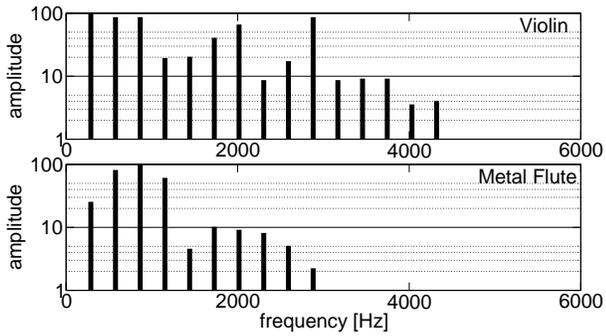
Figure 1: Timbre of two musical instruments, characterized by their timbre vectors. The plot shows the amplitudes of the partial tones, which are called harmonics. Data from [9].

of the signal. These two parts are connected with each other. To give an example, with increasing amplitude of a mass attached to a real spring, the oscillation deviates from harmonic motion due to nonlinearities of the force laws. So with progressing decay of energy, the timbre vector converges to a vector with all energy in the fundamental while the amplitude decays. Figure 2 shows the transient phenomenon for two musical instruments. In many instruments the most important variations of timbre occurs in the transient phenomenon.
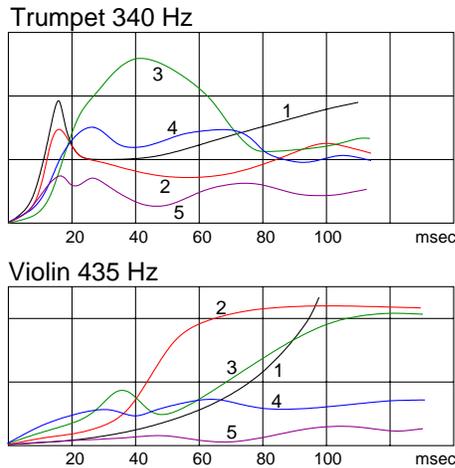

Figure 2: Transient Phenomenon for tones from two musical instruments, given by the temporal evolution of the partial tones, here shown for the first 5 partial tones. Data from [9].

While an instrument can be characterized by its timbre vector, in our model the crystal $I_c(t)$ at time t can be characterized by its dataset covariance matrix

$$\mathbf{C}_{xx}(t) = \frac{1}{|I_c(t)|} \sum_{\mathbf{x} \in I_c(t)} (\mathbf{x} - \bar{\mathbf{x}}(t))^T (\mathbf{x} - \bar{\mathbf{x}}(t)), \qquad (3)$$

which describes the variance structure of the crystal subset. The crystal mean is here given by

$$\bar{\mathbf{x}}(t) = \frac{1}{|I_c(t)|} \sum_{\mathbf{x} \in I_c(t)} \mathbf{x} \; . \qquad (4)$$

The idea of the crystallization model is to perceive the temporal evolution of this matrix during the crystallization process. The covariance matrix is known to include information about the intrinsic dimensionality of the dataset. The eigenvalues λ_i, $(i = 1, \ldots, d)$ of $\mathbf{C}_{xx}$ are the variances along the principal axes of the crystal dataset.

Using the ordered elements of the eigenvalues as strength of the harmonics provides the assignment from data to acoustic properties. For each time t the eigenvalues $\lambda_1(t) \geq \lambda_2(t) \geq \ldots \geq \lambda_d(t)$ of $\mathbf{C}_{xx}(t)$ are computed. The partial tone amplitudes $a_i(t)$ are then determined from $\lambda_i(t)$ by using

$$a_i(t) = \left(\frac{\lambda_i(t)}{\sum_i \lambda_i(t)} \right)^p . \qquad (5)$$

To smooth the timbre evolution, values between successive control points are again interpolated. Using $p = 0.5$, the square root in eq. (5) forces the acoustic energy $E = \sum_i a_i^2$ to be constant over time. Then the timbre vector moves on the surface of a d-dimensional sphere. However, perceptually, the timbre change can be amplified by setting $p > 1$.

Unfortunately, due to the ordering of the eigenvalues, the energy in the fundamental always dominates the sound. The main effect of this assignment for the timbre vector is that data distributions with a high intrinsic dimensionality have more energy in higher harmonics of the sound and thus have a brighter or more brilliant sound, whereas data distributed along a single or only few dimensions will sound like a pure sine tone.

The mode frequencies $\omega_i(t)$ are multiples of the fundamental frequency ω_0. As a generalization, within this model

$$\omega_i(t) = (1 + (i - 1)h_\omega)\omega_1(t) \qquad (6)$$

is used, where h_ω is the harmonics factor. Setting $h_\omega = 1$ yields the harmonics series of periodic oscillations. To enlarge the spectral range, larger values for h_ω can be assigned. The acoustic effect of h_ω is to stretch the spectrum and to amplify perception of timbre changes.

The metaphor of crystallization is furthermore used to derive the amplitude envelope of the sound: assume that the inclusion of each data point into the crystal sets free a constant amount of "crystallization energy". This energy dissipates in acoustic radiation. Thus, the energy stored in the crystal decays exponentially and increases on each point inclusion. The energy $E(t)$ is given by

$$\frac{dE}{dt} = -\gamma E + \sum_i g_i \delta(t - t_i) , \qquad (7)$$

where t_i is the inclusion time for data point $\mathbf{x}_i$ and g_i is the amount of energy used for sound radiation. The solution is an exponential energy decay with discontinuous steps of size g_i at t_i. As we are mainly interested in points that change the crystal's variance structure we use g_i proportional to the change of the first eigenvector of $\mathbf{C}_{xx}(t)$ in inclusion of $\mathbf{x}_i$.

Thus the more a data point inclusion changes the variance ellipsoid, the louder it gets audible. As a side effect, the stabilization of the main variance axis with growing crystal size causes a smooth decay of sound level.

So far, the fundamental frequency remains constant during the sonification. As a modification of the basic model, the perceptually salient attribute pitch can be used to further enrich the sound.

In musical instruments, the fundamental frequency often scales inversely to the length of an acoustic tube or string. Similarly, in this model, pitch is used to inform the listener about the "size" of the crystal. With increasing crystal size, the variance $V_t = \mathrm{tr}(\mathbf{C}_{xx}(t))$ stabilizes to $V_X = \mathrm{tr}(\mathbf{C}_{xx}) = \mathrm{tr}(\mathbf{X}^T\mathbf{X})$. Using

$$\omega_1(t) = \hat{\omega}/\mathrm{map}(V_t, [0, V_X], [1, 2]), \tag{8}$$

with the map function as introduced before, the pitch decreases by an octave with increasing variance.

A compromise must be found as for very small pitch variations, the variance growth cannot be resolved and for very large pitch variations, the timbre variations are perceptually masked. According to the authors' experiences, a pitch range of about 3 halftones realizes a suited balance.

4. EXAMPLES

To demonstrate the sound space of the Crystallization Sonification Model and to introduce the connection between the sound and structures in the data, a series of synthetic datasets are used.

The first sound examples (s. Table 1) are crystallization sonifications for a Gaussian cluster in a 5-dimensional data-space, started at (a) the center, (b) the tail and (c) far outside. Figure 3 shows the sound signal and its spectrogram for (a). The sound exam-
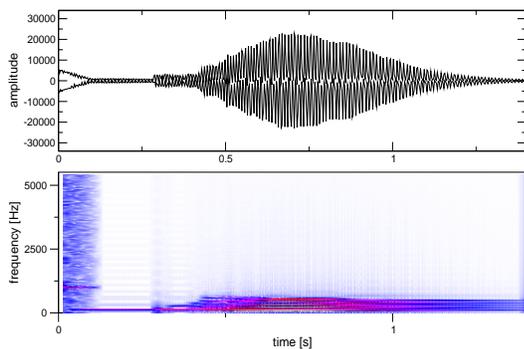

Figure 3: Signal Plot and Spectrum for a Crystallization Sonification started from the center of a 5d Gaussian distribution.

Sound:	E1a (from center) E1b (from tail)
	E1c (from outside), sounds at [10]
Duration:	1.4 secs

Table 1: Sound Examples for Crystallization Sonifications for a 5d Gaussian cluster dataset.

ples illustrate basic aspects of the sonification. At the beginning, a high pitched tick sound is audible which marks the beginning of the sonification. During the sonifications, the progress in crystal growth can be heard from the pitch decay. Obviously, the growth is slower with excitation at the center of the cluster. The brightness is constant during the sonification. This implies that the crystal covariance ellipsoid does not change significantly in shape. The initial sounds distinguish however due to the different location of the condensation nucleus. As expected, all three sonifications sound the same at the end of the crystallization process.

The next sound examples (s. Table 2) use a dataset with 300 records drawn from a mixture of two Gaussian clusters in a 6-dimensional data-space. Data points are drawn from

$$0.3\mathcal{N}(-0.75\hat{x}_0, 0.1\mathbf{I}_6) + 0.7\mathcal{N}(0.75\hat{x}_0, 0.25\mathbf{I}_6), \tag{9}$$

where $\mathcal{N}(\mu, \mathbf{\Sigma})$ is the normal multivariate distribution. Figure 4 shows a scatter plot of the dataset and the condensation nucleus used for the examples. The sonifications depend on the excita-
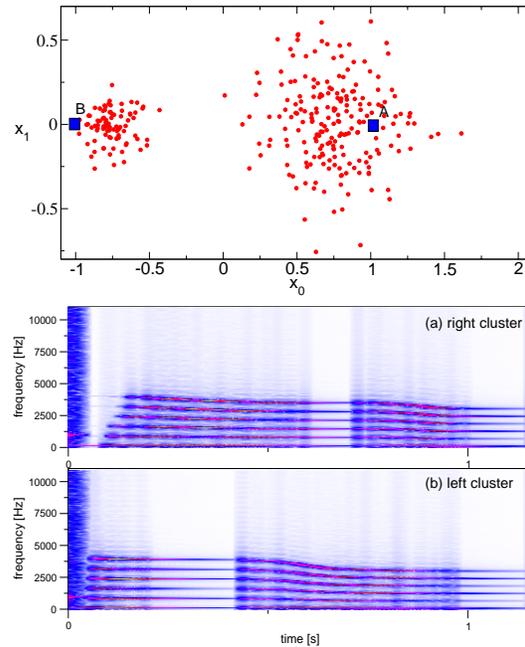

Figure 4: Scatter plot of the mixture of two Gaussians in 6d and the Spectrograms of a Crystallization Sonification. (a) is the spectrogram started within the smaller cluster (b) is the spectrogram for a condensation nucleus within the right cluster shown above.

Sound:	E2a (started in A), E2b (started in B),
	sounds available at [10]
Duration:	1.4 secs

Table 2: Sound Examples for Crystallization Sonifications for the mixture of two Gaussian clusters in 6d.

tion point: it can be heard that the frequency decay and thus the variance increase is faster in example E2a, where the condensation nucleus lies within the bigger cluster. In the second example, the pitch remains constant on a plateau which indicates the separation of the clusters. On the begin of inclusion of data points from the larger cluster, the pitch decays rapidly.

The following sound examples (s. Table 3) demonstrate the effect of the energy decay constant γ in eq. (7) on the crystallization sonifications. As the amplitudes of the harmonics decay exponentially with time and each point inclusion increases the energy, a large decay time has the effect of smoothing the envelope whereas with small decay times the individual point inclusions can be resolved.

Sounds:	E3a, E3b, E3c, E3d, E3e, E3f $\tau_{1/2} = \{0.001, 0.005, 0.01, 0.05, 0.1, 0.2\}$ sounds at [10]
Duration:	1.4 sec

Table 3: Sound Examples for Crystallization Sonification for the mixture of two Gaussians dataset in 6d. The energy decay time constant $\tau_{1/2} = (\ln 2)/\gamma$ is varied between the examples.

Depending on the interest in the data, listening to the sonifications on different time scales is necessary. Changing the total duration for the whole crystallization process allows to choose the temporal resolution. The next sound examples (s. Table 4) differ in the total duration T of the crystallization process. With increasing duration, the granularity of the crystal growth more and more gets audible. However, to facilitate the comparison of sonifications, and to accelerate browsing the data, durations less than 2 seconds proved useful, as the whole sound then fits well into short-term auditory memory.

Sound:	E4a … E4f using $T = \{0.2, 0.5, 1, 2, 4, 8\}$, sounds at [10]
Duration:	between 0.2 sec and 8 sec

Table 4: Sound Examples for Crystallization Sonifications for the dataset with two Gaussians in 6d data-space, varying the total duration T of the crystallization process.

The following sound examples are computed for a dataset with 3 clusters (C_0, C_1, C_2) in $\mathbb{R}^{10}$ with intrinsic dimensions $d(C_1) = 2$, $d(C_2) = 4$, $d(C_3) = 8$. The scatter plot used for selecting the condensation nucleus is shown in Figure 5. In the sound exam-

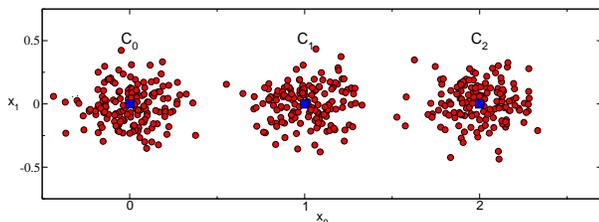Figure 5: Scatter plot of dataset ($d = 10, N = 450$) drawn from mixture of three Gaussians with covariance matrix of different rank and thus different intrinsic dimension.

ples (s. Table 5) it can be heard that the initial brilliance depends on the starting position. However, with proceeding crystal growth, the sound gets more and more mellow since the whole distribution becomes more and more similar to a linear data distribution. The sonification uses $T = 1.5$ sec, $\tau_{1/2} = 0.05$ sec, $h_\omega = 2$, $p = 2$ and a pitch range of 0.14 octaves.

The following example presents Crystallization Sonifications for a 5-dimensional dataset sampled from a mixture of a uniform distribution in the (x_0, x_1)-plane with $(-0.5 < x_0 < 0.5, -5 < x_1 < 5)$ and a $\mathcal{N}\{0, 0.04\mathbf{I}_5\}$ distribution shown in Figure 6. The structure of the dataset is that two manifolds of different dimensionality intersect. Crystallization was started from different points in model space. Sound examples are compiled in Table 6.

Sound:	starting point: C_0 (E5a), C_1 (E5b), C_2 (E5c) sounds at [10]
Duration:	1.9 sec

Table 5: Sonifications for a mixture of three clusters with intrinsic dimensionality 2, 4, and 8 in a 10d data-space.
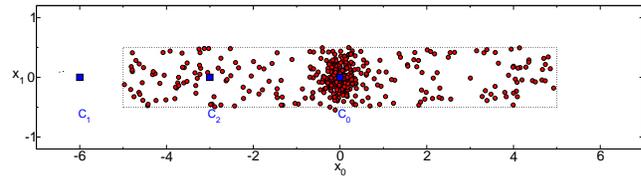

Figure 6: Mixture of a 2d uniform distribution and a 5d Gaussian.

Sound:	condensation nucleus in (x_0, x_1)-plane at C1=(-6,0) E6a, C2=(-3,0) E6b, C3=(0,0) E6c
Duration:	2.16 s

Table 6: Sound examples for DCS for a mixture of a 2d distribution and a 5d cluster.

Starting from the edge (C_1, C_2) the initial timbre consists of two harmonics, as the distribution is two-dimensional. With an increasing number of points, the crystal becomes more and more one-dimensional until the data cloud at the origin is included. Inclusion can be perceived from an increase of harmonics and an increased volume. Inclusion of the rest of the data points then again decreases the sound brightness. This sound is very different from the sonification where the condensation nucleus is set to the origin: here the timbre begins complex, with five harmonics and timbre changes, developing to a rather pure sine sound at the end.

5. CONCLUSION

Crystallization Sonification provides a technique for browsing high-dimensional data. Mouse interactions in a 2d scatter plot visualization both supply the location of interest and start the rendering of the sonification. The metaphor of crystallization helps the listener to interpret the sound. The sonification aims at making explicit use of the temporal evolution of timbre for encoding information about the data. Crystallization sonification can be easily added to existing visual displays and can be used for identifying intrinsic dimensionality and for comparing the effective local and global data dimensionality. As a side effect, clusters can be perceived from amplitude waves in the sound. This basic model is representative for a class of models which encode dynamical system evolutions into sound evolutions. In machine learning there are many similar dynamical processes whose temporal evolution provides interesting information about the dataset, for instance the adaptation of self-organizing maps or artificial neural networks. We think that suited sonification models involving such processes are able to increase the researcher's understanding of structures in the data as well as the adaptation process itself.

In the context of data clustering, hierarchical clustering algorithms are frequently used. The results of such clustering techniques are often represented visually in form of dendrogram plots. Agglomerative clustering [7] involves growing clusters, which is very similar to the presented condensation process, and thus the

sonification model may be used to inspect the clustering progress and its results as an extension to dendrogram plots. Such extensions and applications of the model are subject to ongoing research. Finally, we think that the use of process analysis from physics, such as "crystallization" in the present example, may provide helpful in the construction of sonifications that can be edited at higher semantic levels.

6. REFERENCES

[1] U. M. Fayyad et al., Ed., *Advances in Knowledge Discovery and Data Mining*, MIT Press, 1996.

[2] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, 1977.

[3] S. H. C. du Toit, A. G. W. Steyn, and R. H. Stumpf, *Graphical Exploratory Data Analysis*, Springer-Verlag, New York, 1986.

[4] G. Kramer, Ed., *Auditory Display - Sonification, Audification, and Auditory Interfaces*. Addison-Wesley, 1994.

[5] Thomas Hermann and Helge Ritter, "Listen to your data: Model-based sonification for data analysis," in *Advances in intelligent computing and multimedia systems, Baden-Baden, Germany*, G. E. Lasker, Ed. 1999, pp. 189–194, Int. Inst. for Advanced Studies in System research and cybernetics.

[6] J. Bruske and G. Sommer, "Intrinsic dimensionality estimation with optimally topology preserving maps," *IEEE Trans. of Pattern Analysis and Machine Intelligence*, vol. 20, no. 5, pp. 572–575, 1998.

[7] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*, Prentice-Hall, Englewood Cliffs, NJ, 1988.

[8] J. M. Adrien, "The missing link: modal synthesis," in *Representations of Musical Signals*, G. De Poli, A. Piccialli, and C. Roads, Eds. MIT Press, 1991.

[9] Bergmann and Schaefer, *Lehrbuch der Experimentalphysik*, Walter de Gruyter, Berlin, 1974.

[10] Thomas Hermann, "Sonification for exploratory data analysis – demonstrations and sound examples," `http://www.techfak.uni-bielefeld.de/ ~thermann/projects/index.html`, 2002.