Mode Explorer: Using Model-based Sonification to Investigate Basins of Attraction

Jiajun Yang
Ambient Intelligence Group,
CITEC, Bielefeld University
Bielefeld, Germany
jyang@techfak.uni-bielefeld.de

Thomas Hermann
Ambient Intelligence Group,
CITEC, Bielefeld University
Bielefeld, Germany
thermann@techfak.uni-bielefeld.de

ABSTRACT

This paper presents a novel interactive auditory data exploration method to investigate features of high-dimensional data distributions. The Mode Explorer couples a scratching-interaction on a 2D scatter plot of high-dimensional data to real-time dynamical processes, excited in data space at the nearest mode in the probability density function (pdf) obtained by kernel-density estimation. Specifically, the sign-inverted pdf is used as a potential function in which test particles perform oscillations at low friction, yielding signals that can directly be played back as sound. This Model-based sonification approach is used to interactively search the distribution for different modes, learn about their details, i.e. the Hessian matrix at the mode, and thus enable a non-parametric parameter selection for appropriate bandwidth.

CCS CONCEPTS

•Human-centered computing → Auditory feedback; Sound-based input / output; **•Mathematics of computing** → *Multivariate statistics;*

KEYWORDS

Model-based Sonification, Exploratory Data Analysis, Kernel-density estimation

ACM Reference format:
Jiajun Yang and Thomas Hermann. 2017. Mode Explorer: Using Model-based Sonification to Investigate Basins of Attraction. In *Proceedings of AM '17, London, United Kingdom, August 23–26, 2017,* 4 pages.
DOI: 10.1145/3123514.3123525

DOI: 10.1145/3123514.3123525

1 INTRODUCTION

Exploratory Data Analysis (EDA) refers the process to gain knowledge and detect patterns of data when an explicit knowledge of the data is absent. EDA relies heavily on data visualization, yet visualizing high-dimensional data is still challenging. When analyzing multivariate data, it is often necessary to apply dimensionality reduction in order to 'fit' data features into the visually observable domain [4]. This neglects our highly developed listening skills, which we use to extract meaningful information in many aspects. Likewise for EDA, sound can best be exploited by connecting any human exploratory activity with informative 'side effects' that reflect properties of the data.

In this paper this principle is applied by augmenting 2D scatter plots of high-dimensional data by information about the underlying high-dimensional density function. This is particularly relevant as properties of such high-dimensional scalar fields are difficult to visualize at all: best is to plot sectional views, yet as the data space is high-dimensional, a 2D plane is only very limited and potentially misleading information about the real distribution. Applying dimensionality reduction, e.g. by projecting the data on 3-5 principal components, alleviates the problem to some degree, but fails to be useful in situations where the underlying data model exhibits a higher intrinsic dimensionality.

On the other hand, knowing the mode structure of a data distribution is not only meaningful as it guides subsequent modelling (e.g. whether to apply clustering methods, and how many clusters etc.), it also may be instrumental to find a suitable sweet spot between under- and overfitting, by providing information about the appropriate kernel density estimation (KDE) bandwidth.

Combining auditory and visual exploration might enrich the users' intuition into the data, yet it can also be instrumental to accelerate the completion of certain tasks, such as outlier detection, data segmentation, preparation of clustering runs. At this time, we start to tap into these areas and just report the models and interactions in our work in progress system. As to the auditory part, *Model-Based Sonification* [3] is applied, a sonification technique that starts from defining a dynamic model that involves the given data, offering interaction types (such as shaking, hitting, scratching the data) to excite the dynamic model which in turn leads to a time series that is used as audio signal, i.e. as the sonification. The coherence of the model definition enables the application of the model to all sorts of data and bypasses any subjective application/data-specific parameter mapping, rendering the sonification useful even without reference. As we constantly interpret signals effortlessly, we can trust that meaningful patterns are automatically learned and exploited by users, same as we can gain information about the floor just through the sound of footsteps. As the sonification models connect directly to physical systems we can hope that the inferences are less culturally biased than auditory icons, they furthermore explore an analogue correspondence between structure and sound, without any user-biased reduction of information.

As previous work we set on and extend the Particle Trajectory Sonification model [1], originally proposed for cluster analysis, and the Markov-chain Monte-Carlo simulation-based Sonification model [2], where the authors sonified stochastic processes to tap to transition frequencies between modes of high-dimensional distributions.

2 FROM KERNEL-DENSITY ESTIMATION TO SONIFICATION MODELS

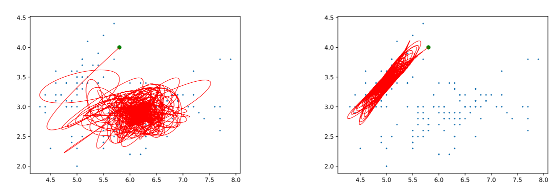
Given a dataset X with N observations, which has an underlying probability density $p(\vec{x})$ we seek to estimate. In statistics, a standard way is KDE [5], a non-parametric method to estimate the underlying distribution by a weighted overlap of data distributions using localized kernel functions $K_\sigma(s)$, where s is the distance between a data point and the vector $\vec{x}$ of interest, stating

$$\hat{p}(\vec{x}) = \frac{1}{N}\sum_{i=1}^{N} K_\sigma(||\vec{x}-\vec{x}_i||) = \frac{1}{\sigma N}\sum_{i=1}^{N} K\left(\frac{||\vec{x}-\vec{x}_i||}{\sigma}\right) \quad (1)$$

If $K \sim \mathcal{N}(x;\mu = 0,\sigma)$, and σ takes an appropriate value, i.e. between the smallest and largest point pair distance, the resulting density function exhibits one mode for each separated cluster of data points. With smaller σ, more and more density modes appear until ultimately each data point leads to a separate mode. Likewise with increasing σ at some point all data points contribute to a single mode around the sample mean.

The starting point for our sonification model is to interpret the sign inverse pdf as a potential function, i.e. $V(\vec{x}) = -\hat{p}_\sigma(\vec{x})$, which provides the basis for a conservative acoustic system when being probed by a point mass m. Integrating the equation of motion $m\ddot{\vec{x}}(t) + F_r(\dot{\vec{x}}) = -\nabla V(\vec{x})$ with a friction force $F_r(\vec{v}) = -(1-r)\vec{v}$ proportional to the velocity $\vec{v}(t) = \dot{\vec{x}}(t)$ delivers trajectories that converge towards a mode, in other words a potential trough or local minimum.

How would a particle probing V sound? First we need to define what variable constitutes the audio signal. We define the audio signal to be the squared velocity as it is a scalar that oscillates as particles move. Intuition tells us that as the particle's total energy E dissipates the particle will be more and more confined to $\{\vec{x}|V(\vec{x}) < E\}$ and ultimately, i.e. at low energy be trapped within a potential function that can well be approximated by a multivariate harmonic potential, thus yielding a mixture of sine waves whose frequencies are directly related to the eigenvalues of the Hessian matrix at the mode. We can expect more chaotic trajectories and thus noisy signals at higher energies. Since the absolute attractive force $|\nabla V(\vec{x})|$ decreases weaker than linearly for high distances from the mode, we can expect a pitch increase as a particle converges to the mode. Interaction video V1.mov (see Link section) illustrates a number of particle sounds starting from different locations and for kernels of different σ. Apparently, the intuition holds true, yet we cannot anticipate all details such as subtle changes of amplitude modulations and mode splits which become apparent particularly with very low friction (i.e. r very close to 1).



(a) $\sigma = 1\mathbf{std}(X)$ **(b)** $\sigma = 0.5\mathbf{std}(X)$

Figure 1: Examples of the trajectories of a particle injected in the same coordinate but with different σ.

A demonstration of the particle moving across the data space can be seen in Fig. 1a & 1b. In the figures, a 150×4 data is given and the particle was inserted at the same position in both cases. However, σ differed with Fig. 1a having a value equals the standard deviation resulting only one cluster and thus the particle converged to the dataset mean. On Fig. 1b, on the other hand, the particle converged to the local minimum of the left cluster.

The spectrogram in Fig. 2 illustrates an example of the trajectory. At the beginning the particle experiences attractive forces from multiple modes thus it has a more chaotic movement, as it moved closer to the nearest mode of the density function, the nearest trough become the predominant force and others contributed to less and less of the gravitation. As a result, the movement becomes more harmonic (i.e. less spectral complexity). Ideally, if the particle converged to the minimum, the signal will become a mixture of sinusoids.

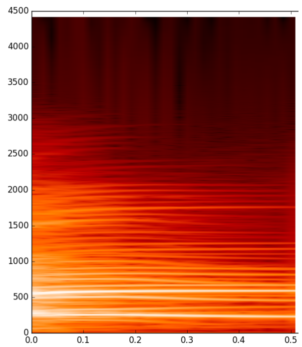


Figure 2: Spectrogram of a squared velocity vector.

This sonification is known as Particle Trajectory Sonification and discussed in more detail in [1]. As new contribution, here we derive it from KDE and relate it to finding the appropriate bandwidth. We provide an Python implementation (using Cython) that renders signals in real-time and thus allows continuous exploration and interactive change of control parameters as supplementary material.

3 MODEL CONTROL PARAMETERS

The sonification model depends on few intuitively clear parameters: the bandwidth σ, the particle mass m, the friction coefficient r, an integration time constant Δt, and the utilized audio sampling rate f_s. We discuss their interrelation and how meaningful controls are isolated for interacton. The numeric integration of the equation of motion is obtained by

$$\vec{v}[n+1] = r\vec{v}[n] + \Delta t \cdot \frac{-\nabla V(\vec{x}[n])}{m} \quad (2)$$

$$\vec{x}[n+1] = \vec{x}[n] + \Delta t \cdot \vec{v}[n+1] \quad (3)$$

It takes $M = -\ln(2)/\ln r$ integration steps for the half time (i.e. -6 dB) decay, thus the half time is $\tau_{1/2} = -\frac{\ln(2)}{f_s \ln r}$. We take $\tau_{1/2}$ rather than r as it is more intuitive. We use $f_s = 11025$ as this delivers sufficient audio quality at controlled computational load. Furthermore Δt controls the integration step lengths, yielding faster movement per integration step, whereas the particle mass m controls inertia and thus slower acceleration at equal force. Both cancel each other and it suffices to set $m = 1$ and control the integration via Δt alone. The most critical parameter is σ as its change affects the Jacobian at any point. Acoustics tells us that the oscillation near to a local mode has angular frequency $\omega = \sqrt{k/m}$ with k being the curvature at the mode. However, from the definition of V with a Gaussian kernel K_σ we see that curvature in the mode is $2/\sigma^2$, so that $\omega = \sqrt{2/\sigma^2 m}$, thus it makes sense to use an effective mass $m_{eff} \propto m_0/\sigma^2$ to compensate pitch changes as σ is varied. Note that the effective pitch will nonetheless change, which is due to the fact that with larger σ more data points collectively contribute with attractive force to pull at a particle, so that pitch gives as more an information about cluster size, and remains independent of σ. Fig. 3 shows how σ affects the clusters’ resolution.

4 SYSTEM FEATURES: MODE EXPLORATION

Based on the previously described method, we can detect how close a particle is towards the mode of a cluster. However, it is hard to truly listen to the sound when a particle has converged to a mode as it is usually towards the end of the sound clip and tends to be quiet (see V1.mov). Thus, we create a new scanning method which bypasses the main part of the convergence by a gradient descent for hearing the sound of a particle’s harmonic movement specifically around a mode (i.e. the maximum value of the pdf).

Fig. 4 is a screenshot of the video file V2.mov, which contains a full demonstration of interacting with the Mode Explorer program to sonify the modes of a $N = 400$, 5-dimensional data. A Wacom tablet is used with a pen stylus, which is modified with a 3-axis accelerometer[1] attached to detect the pen’s tilting angle. The stylus serves two functions: 1. Touch or scratch on a 2D projection of the data to initiate the mode explorer and the sonification; 2. Tilt the pen from lower to straight-up to control σ between zero to twice the data standard deviation, to interactively adjust the smoothness of the KDE.

When a particle is inserted at $\vec{x}$, the particle starts the motion based on the potential function and motion law defined in Eq. 1,2,3 through certain iterations with maximal friction ($r = 0$) until it converges to a local minimum of V. No sound is produced during this mode searching phase. Once converged, the particle is given an impulse (i.e. small kinetic energy by setting v) allowing it to vibrate around the mode with small friction ($r \approx 0.999$), thus causing decaying sound. Due to the differences in cluster sizes, a larger cluster will allow particles to spin with a higher frequency than a small one, resulting in different pitches between clusters. From the video (see at 00:50”), we can see that three clusters were stably formed with different tones separating them apart. Aligning with Fig. 3, σ dictates the resolution of clustering. When it increases, more neighbouring data points will join the cluster, resulting a higher pitch.

5 DISCUSSION

In real life, we often scratch on different surfaces and hear their acoustical response, which quickly depicts information about the material and texture of a physical object. In the

[1] The device uses the rapid prototyping tool for physical computing named Brix2.

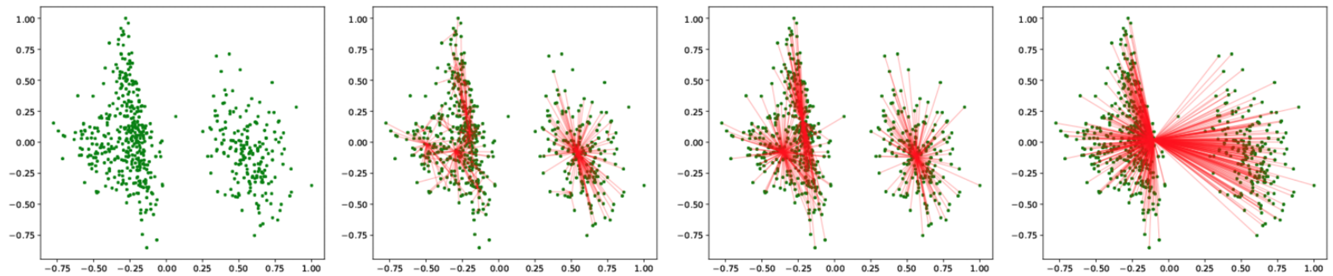


Figure 3: Demonstration of how σ affects the mode detection of a 600×5 data. From left to right: $\sigma = 0.01std(data)$; $\sigma = 0.3std(data)$; $\sigma = 1std(data)$; $\sigma = 2std(data)$.

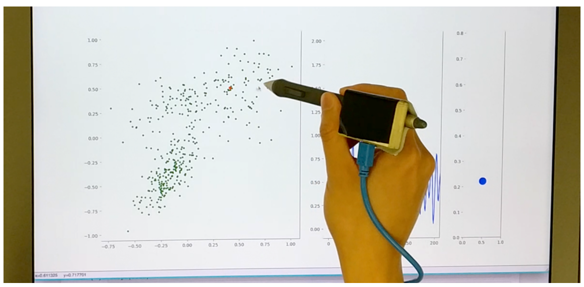


Figure 4: Screenshot from example file ModeExplorer.mov.

Mode Explorer Sonification, we can likewise 'scratch the data surface' of a high dimensional data set and hear its clustering structure. The proposed system provides a physical model-based solution to sonifying complex data with a highly responsive interaction means. Using the potential function to sonify the particle's kinetic energy not only avoids introducing subjective decisions as required in parameter mapping sonification, but more importantly, it allows studying the clustering structure without any dimension reduction – we reach into and experience the truly high-dimensional pdf.

The algorithm is suitable for both visualization and sonification as currently presented. The visual part is the projected trajectory as the particles moves around the data. However, due to the gradient descent this trajectory becomes confined to a small blob around the mode and information is better audible than visible. We argue that the sonification approach can bring in a few extra benefits: Firstly, the trajectory of the particle unfolds in a high-dimensional space, while the visualization of the trajectory discards all but 2 dimensions. Using sound is quick and responsive and can work very naturally in this context (as seen in V2). However, as a work-in-progress, we have not tested our hypothesis in user studies.

Using a continuous interaction we can continuously explore those areas of data where attraction basins touch, and gain immediate experience of the structure of classification borders (if we assume clusters to constitute different classes).

6 SUMMARY

The paper presents a dynamic system to investigate basins of attraction of high-dimensional data via Model-Based Sonification. We introduced a new algorithm based on the Particle Trajectory Sonification Model which focuses on sonifying local maxima in the pdf of the data, corresponding to separated clusters when using an appropriate bandwidth. Our aim is to establish sonification as unobtrusive yet informative side product of continuous interaction with visualizations, allowing us to carry over the richness of real-world sonic explorations to enhance our understanding of complex data. Our current research points to the importance of responsiveness (i.e. low-latency) and natural physical interactions to be critical factors to facilitate the multimodal integration of visual and auditory exploration.

ACKNOWLEDGMENTS

This research was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG). We wish to thank Akhil Jain for programming support.

LINK

Supplementary materials at DOI: doi.org/10.4119/unibi/2912254

REFERENCES

[1] Anonymous. 1999. Listen to Your Data: Model-based Sonification for Data Analysis. *Advances in intelligent computing and multimedia systems* 8 (1999), 189–194.

[2] Anonymous. 2001. Sonification of Markov Chain Monte Carlo Simulations. Georgia Institute of Technology.

[3] T. Hermann. 2011. *The Sonifcation Handbook*. Logos Verlag, Chapter Model-Based Sonification, 399–425.

[4] Samuel Kaski and Jaakko Peltonen. 2011. Dimensionality Reduction for Data Visualization [appllications corner]. *IEEE Signal Processing Magazine* 28, 2 (2011), 100–104.

[5] Bernard W. Silverman. 1986. *Density Estimation for Statistics and Data Analysis*. Vol. 26. CRC press.