# GROWING NEURAL GAS SONIFICATION MODEL FOR INTERACTIVE SURFACES

*Lukas Kolbe, René Tünnermann, Thomas Hermann*

Ambient Intelligence Group
Cognitive Interaction Technology - Center of Excellence (CITEC)
Bielefeld, Germany
{lkolbe,rtuenner,thermann}@techfak.uni-bielefeld.de

## ABSTRACT

In this paper we present our reimplementation of the Growing Neural Gas Sonification [1] for interactive surfaces such as our t-Desk [4] or touch-capable tablet PCs. Growing Neural Gas (GNG) [3] is an undirected learning algorithm that incrementally 'grows' a network graph into data distributions, revealing the data distributions' intrinsic dimensionality and aspects of its structure. The GNG Sonification (GNGS) [1] provides a method to interactively explore the GNG during the growing process, utilizing a Model-Based Sonification (MBS) [2] to convey audible information about the data distribution in addition to the visualization. The goal of our reimplementation was to be able to rapidly grasp the structure of the sonified and visualized data, to give the user the ability to conduct direct A/B comparisons between different (or similar) clusters within a data distribution. The direct bi-manual interaction as well as a simplified full-screen touchable user interface helps to focus on the exploration of the GNG rather than the interaction itself. We present and discuss different interaction metaphors for the excitation of the model setup in this MBS.

## 1. INTRODUCTION

The Growing Neural Gas is an undirected graph of vertices called neurons and connecting edges. The graph is adapted during training with a given dataset so that it represents the topological structure of the data distribution. For example, for two-dimensional data distributions the graph grows into a triangle mesh whereas three-dimensional regions will lead to interconnected tetrahedrons. While the neural gas grows and until the maximum number of neurons is reached, new neurons are inserted regularly at the place of maximum error. Connections between neurons age or get reinforced and neurons are moved in data space to minimize the quantization error with respect to the data. Fig. 1 shows an example of a GNG that grows into a two-dimensional dataset, learning its structure along the way. The small bright points show the underlying data distribution, the bright rings represent the neurons of the GNG and the lines between them are their connections. The line's thickness represents the age of the connection. The structure and intrinsic dimensionality become visible, until later in the learning process overfitting occurs and the structure diffuses again. Visually, this works very well only for two- and three-dimensional data. But what about higher dimensional data distributions? Other than projecting high-dimensional data into two- or three-dimensional space or selectively showing only a few of the data's dimensions, we have no means to visually grasp dimensionality properties.

In the GNGS, the user can 'pick' a neuron similar to picking a guitar string to induce an energy flow within the whole network. This energy flow is sonified and simultaneously visualized in the selected dimensions. The resulting sound is influenced by the amount of energy within each neuron and the number of connections it has to other neurons. This Model-Based Sonification,

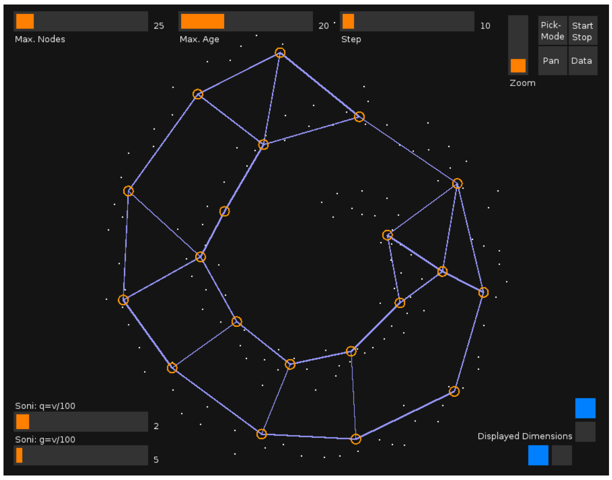


Figure 1: A GNG that grows into a tutorial data distribution. Using two-dimensional input data, the user quickly learns how the sonification correlates to the structure of the GNG. This knowledge can then easily be applied on higher-dimensional data.

specified in detail in the next section, can help the user to better understand higher-dimensional structures within the GNG than visualization alone because it conveys information from all available dimensions. Using our naturally well trained sense of listening, we are able to differentiate dimensionality structures.

## 2. GROWING NEURAL GAS SONIFICATION MODEL

A sonification model according to MBS [2] can be described by the following categories:

*Setup:* In this model, the connections in the GNG network serve as energy transducers between neurons. Each neuron emits a tone, whose frequency is determined by the number of connections emanating from it: for each connection, its frequency is increased by $4/3$[1].

*Dynamics:* Using the energy flow equation (1), the energy for each neuron is calculated. It decays over time, depending on parameters $g$ and $q$ (adaptable by the user) and the current state of the GNG network. The energy of each neuron determines the amplitude of its tone.

$$\frac{dE_i}{dt} = -gE_i(t) - \sum_{j \in I_N(i)} q \cdot (E_i(t) - E_j(t)) \qquad (1)$$

The parameter $g$ ascertains the exponential energy decay, $q$ determines the amount of energy that flows to every neigh-

[1] Corresponding to a quart in musical terms

Copyright http://creativecommons.org/licenses/by/3.0/

boring neuron each step. $E_i(t)$ describes the energy of neuron $i$, $I_N(i)$ is the set of neurons that are connected to neuron $i$. The default values of $g = 0.05$ and $q = 0.02$ lead to a slow dispersion of the induced energy throughout the network, yielding a clearly audible 'path' through it. Before the equilibrium state of silence is reached, the resulting tone becomes a signature for the sonified part of the network, as the energy is distributed near-uniformly among connected subgraphs.

*Link Variables:* The sonification is the superimposed sound signal of all existing neurons. It consists of one tone per neuron, with the pitch determined by the number of connections to other neurons and its amplitude determined by the current energy level of the neuron.

*Excitation:* There are three main modes of operation. In the *pick*-mode, the user induces energy into a neuron by picking it like a string on the guitar. The amount of energy induced is proportional to the distance the neuron is picked. The energy then propagates through the GNG, exciting other neurons along the way, until equilibrium is reached eventually. In the *continuous excitation*-mode, the energy level of the nearest neuron to the current position of the touch is set to constant 1 while the energy flow behavior is unchanged. As this mode allows for continuous excitation, moving the finger around a set of neurons induces a high energy level in each of them. This allows exciting a part of the network quickly to hear its signature, or to quickly compare different parts of the GNG. In the third mode, the GNG's learning process is sonified: every neuron has a constant energy level of $0.1$ and thereby no energy flows. The stationary sound is only influenced by sudden changes in the number of neurons and their connections.

*Listener:* The resulting sonification for all neurons is presented to the user as well as the synchronized visual feedback. For two-dimensional data, the sonification directly matches what the user can see on the screen. For higher dimensional data, the sonification often reveals more than the user is able to see at a time.

### 2.1. GNGS for interactive surfaces

Our application presents a 2D scatterplot of the data distribution to the user. He or she can select scatterplot variables with controls on the bottom-right for the $x$-axis and right-hand side for the $y$-axis. There are three controls in the top row of the screen to adapt the GNG algorithm:

- the maximum number of neurons for the GNG
- the maximum age of connections between neurons
- the learning rate parameter, e.g. speed of the learning process

Sliders to control aspects of the sonification are in the lower right corner:

- the energy flow rate parameter $q$ in eq. (1)
- the energy dissipation rate parameter $g$ in eq. (1)

The four buttons in the upper right corner start or pause the learning process, reset to its initial state, allow panning of the viewport and cycling through the different modes for the sonification. Additionally, a slider allows to zoom the viewport to also support single-touch devices.

Each neuron in the GNG is represented by a bright ring. If the neuron has an energy level greater than zero, a filled circle appears within it with it's diameter proportional to the energy level. The connections between the neurons are visualized by lines whose thickness represents the edges age.

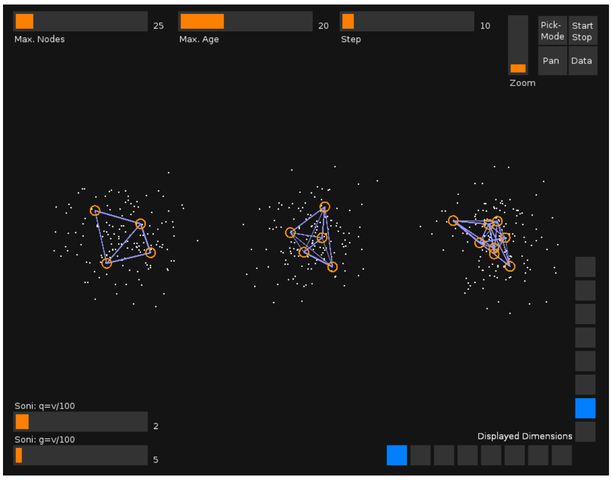


Figure 2: The user interface of the GNG Sonification application, showing the first two dimensions of the *three cluster* data distribution superimposed with a GNG during its adaption process.

### 2.2. Data distributions used for evaluation

To evaluate the GNGS, we devised several data distributions:

*Three cluster* distribution: containing 449 points in 8 dimensions, organized in three clusters of differing intrinsic dimensionality: 2, 4 and 8 each.

*Twisted snake* distribution: containing 800 points in 10 dimensions, a snake-like distribution when looked at in the first two dimensions, but with higher intrinsic dimensionality for each few hundred points resulting in a snake twisted into high dimensional space.

*Quiz* distribution: containing 800 points in 10 dimensions similar to the twisted snake, but the regions of different intrinsic dimensionality are spatially separated. There are two regions of the same dimensionality that are not visually distinguishable. By exploring the GNGS, the user is able to hear which regions are similar, hence the name of this data distribution.

### 2.3. Implementation Details

The GNG sonification is implemented in Python. Computation is handled by the *Modular toolkit for Data Processing* [5]. The user interface is implemented with *PyMT - A Multi-touch UI Toolkit for Pyglet* [6]. The sonification has been implemented in SuperCollider [7], utilizing Stinson's *OSC interface for Python* [8] for the interprocess communication.

### 2.4. Interaction

In this paper we mainly discuss exploring the *three cluster* dataset with the GNGS, but audio and video interaction examples for the other two datasets are provided on our website[2].

While growing into the *three cluster* data distribution, the GNG forms three separate networks. Fig. 3 shows snapshots of the learning process. This process is sonified, giving each existing neuron an energy level of $0.1$. In the first picture, two of the five neurons have only one connection, resulting in a low frequency for these two neurons. Their combined energy level is $0.2$, so the amplitude for this frequency is low as well. Three neurons have two connections, assigning them a higher frequency one quart above

[2]http://www.techfak.uni-bielefeld.de/ags/ami/publications/KTH2010-GNG

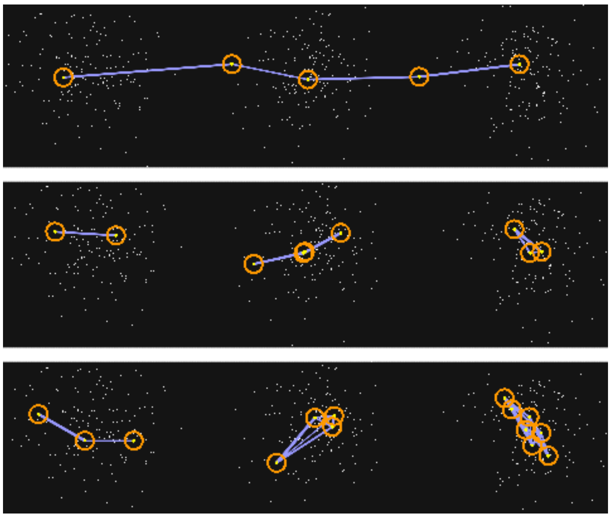

Figure 3: Learning-mode: every neuron in the GNG on top of the *three cluster* dataset has an energy level of $0.1$. The sound changes instantly when a neuron is removed or added, or the number of connections between them changes.

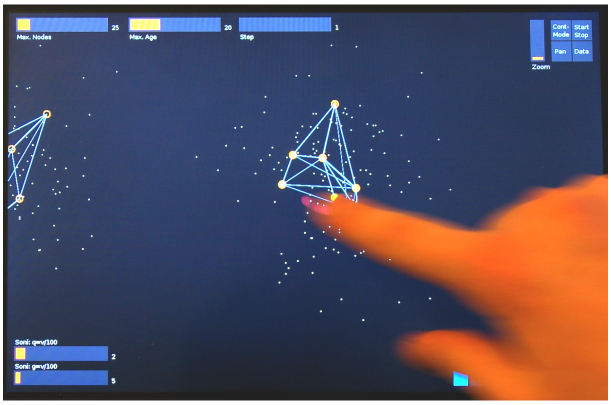

Figure 4: Continuous-mode: the right cluster was circled, inducing an energy level of 1 in each neuron within. Enables quick listening to the signature sound of the cluster.

the other. Their energy level is $0.3$, thus the amplitude is slightly higher. The resulting sound thus has a low pitch and amplitude. As more neurons appear with a higher number of connections during the learning process (the following two pictures in fig. 3), the sound becomes brighter and louder. The addition or removal of neurons as well as connections are clearly audible through sudden changes in pitch and/or amplitude of certain frequencies[3].

In the continuous excitation mode, the GNG can be explored after the learning process was stopped by the user. For as long as his or her finger is on the surface, the nearest neuron has a constant energy level of 1, inducing a steady flow of energy into the network. With swirling motions around a region of interest, a signature sound for this region can be produced. Fig. 4 shows the result of swirling with a finger around the right network in the *three cluster* distribution: all neurons within that network receive energy, and the resulting sound becomes the signature sound for this network. When the finger is lifted from the surface, the sonification immediately stops[4]. Using fast swirling or scribbling motions

[3]This can be heard in video example 1, available from our website

[4]Video example 2

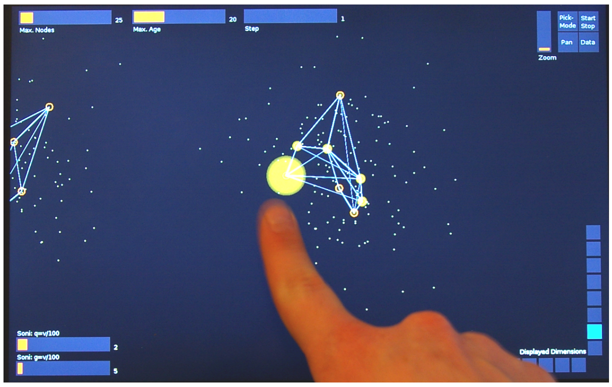

Figure 5: Pick-mode: Picking the leftmost neuron of the right cluster of the *three cluster* dataset. The resulting sound is very bright, as most neurons have many connections to other nodes, suggesting a high intrinsic dimensionality. It changes its pitch while the energy propagates through the neurons according to their number of connections, slowly fading until equilibrium is reached.

subsequently in different parts of the GNG allows for rapid A/B comparisons of the respective signature sounds, revealing structural differences or similarities.

Fig. 5 shows the picking mode. The user has picked the leftmost neuron of the right network, with the picking motion shown as a white arrow. The resulting sound starts much brighter than one would expect from looking at the two-dimensional scatter plot[5], as there are only four visible connections emanating from that neuron. But this neuron is folded into other dimensions as well and has connections to other neurons on nearly all eight dimensions, resulting in a very bright sound. As the energy propagates, the brightness gets a little lower while the volume slowly fades. As it seems, the other neurons in that network are not as well connected in the eight-dimensional space as the picked one.

Through changing the $g$ and $q$ parameters of the sonification, the user is able to alter how the energy flows within a network: $g$ determines how fast the energy decays in neurons, $q$ influences how fast the energy is transported along the edges. If enough energy is induced to spread throughout the whole network (e.g. the picked distance is far enough or the $g$ and $q$ parameters are set accordingly), the signature sound for this network becomes audible, similar to the swirling or scribbling motions in continuous mode.

Altering the maximum number of nodes or maximum age parameters, the GNG can be optimized to better learn a given data distribution.

Growing Neural Gas is an undirected learning algorithm, but there is no established decision criteria as to when it has fully grown into its data distribution. Overfitting occurs after a while and the learned structure becomes diffused again. The user has to make an informed decision as to when to end the learning process. The GNG Sonification Model provides a multi-modal and highly interactive tool to do just that.

Since the GNGS is totally invariant upon the choice of coordinate systems, the sonification allows an estimation of topology even if structure can not be visually guessed, e.g. if the data distribution is a two-dimensional data sheet twisted into a higher-dimensional subspace. Furthermore, GNGS provides information by sound that is complementary to the visually salient information, namely the connectivity of the graph in a dynamic form. GNGS provides thus interactive insight into relevant topological structures of complex distributions.

[5]Video example 3

## 3. DISCUSSION AND CONCLUSION

In this paper we have presented a reimplementation of the Growing Neural Gas sonification model, exploring new interaction possibilities with current interactive surfaces such as our tDesk, tabletop PCs or convertible touch-screen notebooks. The application enables different interaction modes: $a$) continuous excitation through motions on the surface and $b$) picking, analogue to picking a string on a guitar. There's also a $c$) non-interactive monitoring mode where the user just watches and listens to the GNG as it adapts its structure to the data distribution.

The main advantage of our approach is that a very natural and direct contact between the user and the explored data can be established, as the user intuitively interacts and explores the GNG, using the surface almost like a real, albeit two-dimensional, physical tool. The synchronization between the different components is crucial to enable closed-loop interaction. From the moment the user picks a neuron to the state of equilibrium silence, the visualization and sonification have to be in sync. Even slight differences among the representations can cause irritation and a loss of focus for the user. For the future, we plan to extend the GNGS application and our other Model-Based Sonifications with as-of-yet *untouched* aspects of continuous interaction with data distributions. For example, enabling the user to continuously deform data representations to perceive the resulting stress as informative sound.

In summary, the presented Growing Neural Gas Sonification enriches the available modes to interact with complex data and to perceive structure-related features as sound via MBS that can otherwise not easily be perceived. The tight coupling of visualization, sonification and the interactive surface in one interface contributes to a multi-modal experience and shows the potential to an increased level of understanding of structures in the data. In our ongoing research we plan to explore and evaluate how multimodal interaction modes as introduced here support the understanding of complex data.

## 4. REFERENCES

[1] T. Hermann and H. J. Ritter. Neural Gas Sonification - Growing Adaptive Interfaces for Interacting with Data. In E. Banissi and K. Börner, editors, *IV '04: Proceedings of the Information Visualisation, Eighth International Conference on (IV'04)*, pages 871–878, Washington, DC, USA, 07 2004. IEEE Computer Society.

[2] T. Hermann. Sonification for Exploratory Data Analysis. PhD thesis, Faculty of Technology, Bielefeld University, Febr. 2002 URL: http://sonification.de/publications/media/Hermann2002-SFE.pdf

[3] B. Fritzke. A growing neural gas network learns topologies. In G. Tesauro, D. Touretzky, and T. Leen, editors, Advances in Neural Information Processing Systems, volume 7, pages 625–632. The MIT Press, 1995.

[4] R. Tünnermann and T. Hermann. Multi-touch interactions for model-based sonification. In *ICAD '09: Proceedings of the 15th International Conference on Auditory Display*, Copenhagen, Denmark.

[5] T. Zito, N. Wilbert, L. Wiskott, P. Berkes. Modular toolkit for Data Processing (MDP): a Python data processing frame work. Frontiers in Neuroinformics (2008). URL: http://mdp-toolkit.sourceforge.net

[6] NUI Group. Python Multitouch. URL: http://pymt.txzone-.net/ (Jan. 2010)

[7] McCartney et al. SuperCollider: realtime audio synthesis and algorithmic composition. URL: http://supercollider.sourceforge.net/ (Jan. 2010)

[8] P. K. Stinson. scosc: SuperCollider OSC interface for python. URL: http://trac2.assembla-.com/pkaudio/wiki/SuperCollider (Oct. 2009)