DOI: 10.61782/fa.2025.0708

FORUM ACUSTICUM
EURONOISE 2025

PSYCHOACOUSTIC ANALYSIS OF VISCOELLASTICALLY DAMPED MEMBRANES USING PHYSICS-INFORMED MACHINE LEARNING

Martínez Orellanos, Cristhiam Fidel[1*] **Bader, Rolf**[1]
[1] Institute for Systematic Musicology, University of Hamburg, Germany
Neue Rabenstr. 13, 20354 Hamburg, Germany

ABSTRACT

The perceived sound of percussion instruments is significantly influenced by Viscoelastic Damping (VD), a mechanism whereby damped energy is temporarily stored and released back into the system, modifying the frequency spectrum. Performers harness this influence by choosing different materials and damping additives to shape the sound of their instruments. However, this process is still not fully understood. To address this, we apply physics-informed machine learning within a psychoacoustically grounded framework to investigate how changes in VD parameters influence the perceived sound of musical membranes. By analyzing a physics-informed dataset, we track spectral centroid differences and spectral flux across time and frequency ranges. Results show that the perceptual influence of VD strongly depends on where along the frequency spectrum the damping occurs, particularly in relation to the fundamental frequency. A temporal drift of the affected spectral region toward lower frequencies was also observed. Self-organizing maps (SOMs), initialized with physically informed features, were used to explore these complex, high-dimensional relationships. This method offers new insights into how specific damping conditions shape perceived sound and provides a foundation for more detailed analyses using alternative time–frequency representations in future research.

Keywords: *viscoelastic damping, machine learning, psychoacoustics*

**Corresponding author*: *crismartinez@t-online.de.*

1. INTRODUCTION

The sound of musical instruments is highly influenced by the physical properties of the materials with which they are constructed. For percussion instruments, membranes are typically made from materials like leather and polymers. The internal damping of these materials is largely determined by their viscoelastic properties, leading to what is known as Viscoelastic Damping (VD). VD is characterized by the interrelation between stress, strain and time; the present stress is the result of a finite history of previous strains [6]. Given that in viscoelastic materials there is a phase difference between stress and strain, the present stress is damped by the accumulated strain. If by the time the present stress is completely damped there is still energy stored in the history of strains, this energy is released back into the system, starting the cycle again. This is known as the memory effect and leads to a highly nonlinear relation between the parameters of VD and the resulting frequency spectrum [3].

Despite its complex behavior, musicians actively make use of VD either by exploring the diverse sonic possibilities that different materials offer (e.g., when selecting a membrane material), or by intentionally using VD to alter the instrument's original sound [2, 17]. Although this practice is widespread, little is known about how VD influences the perceived characteristics of sound. Specifically, how do modifications in VD parameters lead to particular changes in the perceived sound characteristics.

To investigate the relationship between perception and musical acoustics, different music information retrieval algorithms have been developed to analyze the frequency and temporal characteristics of musical instrument signals during the onset of the sound [4, 15]. Moreover, the development of such features, as well as the advent of data-driven methods have propelled the use of

machine learning and physical modeling in psychoacoustics [5, 14, 16]. Still, many of these approaches primarily target genre or instrument classification and transcription. With regard to damping in drums, an analysis of the influence of drumhead damping patterns in the perception of the sound using convolutional neural networks and physical modeling is presented in [1]. Nonetheless, mapping perceptible changes in the sound of a musical membrane to variations within the VD parameter space remains challenging. This is due to a lack of data reflecting consistent parameter changes in VD, as well as the need for tools capable of navigating a broad parameter space while flexibly and accurately analyzing transient dynamics from a psychoacoustic perspective. To address this, the authors introduced a method in [9], employing physics-informed machine learning to investigate the relationship between systematic variations in VD and the resulting frequency spectra, thereby revealing key characteristics of VD's physical behavior.

Still, a psychoacoustic analysis implies additional challenges. On the one hand, it requires flexible tools for analyzing the transient phase of percussive sounds, which is characterized by its high intricacy [11]—this being the time window during which VD primarily operates, as will be discussed in the next section. On the other hand, it demands a better understanding of which time and frequency features the auditory system extracts and integrates into perceptual categories such as timbre or pitch. Indeed, the fact that VD is embraced by musicians suggests that the auditory system is capable of detecting and extracting patterns from its complex relationships and behaviors. Still, this latter question remains under active investigation. Therefore, the present study is based on an analysis framework developed in previous work [10], which is based on findings from the existing literature related to important temporal cues for the perception of musical sounds—also outlined in the following section. However, a detailed analysis of the auditory cues underlying sound perception, lies beyond the scope of this work.

Accordingly, the aim of this study is twofold: (a) to employ physics-informed machine learning, grounded in a psychoacoustically relevant analysis framework, to identify underlying patterns and relationships between VD and measured features, in order to (b) investigate which changes in the frequency spectrum caused by VD modify the perceived characteristics of the sound—specifically, when, where, and how VD produces these modifications. Correspondingly, this article is structured as follows: Section 2 describes the methods, i.e. the analysis framework, physics-informed dataset data pipeline and experimental design. Sections 3 and 4 present the results and their interpretation, while Section 5 summarizes the conclusions of the study.

2. METHOD

2.1 Physics- and Psychoacoustics-Based Analysis Framework

Despite the highly nonlinear relation between the parametric modifications and the resulting frequency spectra, the musical application of VD points towards the ability of the brain to extract patterns from among these complex physical components. To understand this process, both the physical dynamics of VD and the brain's temporal cues for perceiving musical sounds must be considered. To this end, an analysis framework addressing this has been presented in previous research [10] and is summarized below.

Key physical characteristics of VD in musical membranes, identified in [3, 9], include: (1) the memory effect causes viscoelastically damped frequencies to decay non-exponentially, with a nonlinear amplitude behavior starting around 40-50 ms after an initial constant decrease; (2) modifications to the model parameters lead to nonlinear amplitude fluctuations in neighboring frequencies, with the width of the affected region yet to be determined; (3) VD influences sound most significantly during the first 120 ms, although this temporal interval may vary with frequency; (4) a noticeable increase in spectral fluctuation occurs around 25 ms [10]; (5) the FDTD model requires a minimum number of periods of the damped frequency for the damping to develop completely.

Concerning the perception of musical sounds, previous studies have highlighted relevant integration times for non-percussive and percussive-harmonic sounds [12, 15]. A time window of approximately 10 ms reflects the auditory system's state of inertia, as it requires around 10 ms to establish frequency filter widths on the basilar membrane. Auditory changes occurring above this threshold affect the envelope of the sound, while changes below it influence timbre perception [12]. The interval between 10 ms and 50 ms is closely associated with pitch perception; once the frequency filters are established, their sensitivity plays and important role [12]. For frequencies below 500 Hz, differences as small as 1.8 Hz can be perceived, whereas above this threshold, the detection limit increases with frequency [18]. Changes during this time window also shape timbral characteristics. A broader window of 100–150 ms

represents the upper bound beyond which sensitivity to frequency fluctuations remains relatively constant [12]. If such fluctuations persist up to 250 ms, the pitch percept tends to weaken, and the sound is perceived as more noise-like.

Although the influence of VD unfolds continuously over time, defining specific time points that are both physically and perceptually relevant serves as a starting point for understanding how VD affects the acoustic features upon which the auditory system forms perceived characteristics of a sound. These temporal cues therefore constitute the analytical framework of this study and are summarized in Table 1.

Table 1: Analysis framework.

Time (ms)	Description
10	Filter formation in the basilar membrane. Timbre and envelope perception
25	Increase in spectral flux
50	Non-linearities in amplitude of viscoelastically damped frequencies begin. Pitch and timbre perception. Sensitivity to frequency variations
100	Perception of harmonicity - noisiness
120	Influence of VD decreases

2.2 Physics-Informed Dataset

The behavior of viscoelastic damping can be systematically controlled and modified through four parameters defined in a Finite-Difference Time-Domain (FDTD) model of a viscoelastically damped membrane developed by the second author in [3] (see Table 2). By simulating a circular membrane with a tension of 2284 Pa, a density of 300 kg/m^3, a thickness of 3 mm, a radius of 10 cm, and fixed boundary conditions ($u_{\text{rim}} = 0$), and varying only the parameter space from Table 2, a physics-informed dataset was generated in [9]. For the present study, a subset of 220 time series was extracted from this dataset (see Table 3). Each sample has a duration of 500 ms and a sampling rate of 96 kHz. The selection of viscoelastically damped frequencies—hereafter referred to as VD$_{freq}$—is constrained by the FDTD model, which requires integer multiples of periodicity to maintain numerical stability. Furthermore, when $\Re\{E(s)\} = 0$, no viscoelastic damping is applied; this case is referred to as the *no-damping* condition.

Table 2: Parameters of the FDTD.

Parameter	Description
VD$_{freq}$	Viscoelastically damped frequency
T	Integration time. Length of memory effect
$\Re\{E(s)\}$	Damping coefficient
γ	Rate of decay

Table 3: Parameter space of the dataset.

Parameter	Fixed Value	Min. Value	Max. Value	Increase
VD$_{freq}$ (Hz)	700, 6400			
T (ms)	10			
$\Re\{E(s)\}$		0*	0.003	0.0003
γ		1/100	1/10	1/10

* $\Re\{E(s)\} = 0$ implies no VD (no-damping case).

2.3 Psychoacosutic Features

2.3.1 *Spectral Centroid Differences*

The spectral centroid has been shown to be strongly related to the sensation of brightness in a sound [13]. In accordance with the analysis framework described in Section 2.1, changes in spectral centroid caused by VD at discrete time points are investigated by calculating the difference between the centroid values of a data sample and those of the no-damping case at the same time points. Previous studies have found that the frequency range along which the spectral centroid varies can change considerably across different time points [10]. Therefore, the calculated differences are normalized to facilitate comparison across time. As a result, the centroid differences are calculated as:

$$CD_i(t) = \frac{C_i(t) - C_{\text{ref}}(t)}{CD_{max}(t) - CD_{min}(t)}, \qquad (1)$$

where $C_i(t)$ and $C_{\text{ref}}(t)$ denote the spectral centroids of sample i and the no-damping case, respectively, calculated at the discrete time point t. If $CD_{\text{max}}(t) = CD_{\text{min}}(t)$, then $CD_i(t) = 0$. The results are visualized as a heat map following the approach proposed in [10], with damping coefficients along the x-axis, decay rates along the y-axis, and centroid values mapped to color. This visualization enables a comprehensive overview of how changes in the parameter space of VD affect the spectral centroid at discrete time points.

2.3.2 *Spectral Flux per Mel Band*

Building on the approaches proposed in [11] and [4] to evaluate spectral dissimilarities in a logarithmically filtered Mel spectrogram between discrete time frames t and $t - \mu$, a measure of spectral dissimilarity based on the decibel scale and preserves values per Mel band was introduced in [10] and is calculated as:

$$SF(t, m) = X_{\text{dB,filt}}(t, m) - X_{\text{dB,filt}}(t - \mu, m), \quad (2)$$

where $X_{\text{dB,filt}}$ denotes the decibel-filtered Mel spectrogram, t the discrete time point, m the Mel band index, and $\mu = 1$. Quarter-tone Mel bands were employed in the present study. The objective of this approach is to investigate how and where VD affects the frequency spectrum, enabling the localization of spectral regions where VD-induced changes occur. This, in turn, provides insight into how the spectral centroid is affected and which frequency ranges are most impacted by the energy introduced through the memory effect.

2.4 Self-Organizing Maps (SOM)

Self-Organizing Maps (SOM) transform non-linear statistical relationships in high-dimensional data into geometric relations on a 2D $n \times n$ grid [8]. The dimensionality of each neuron in the SOM matches that of the feature vectors extracted from the dataset. Both neurons and data samples are normalized before training. During training, the Euclidean distance between each sample and neuron is computed to determine the Best Matching Unit (BMU). The BMU's weights and those of its neighboring neurons are updated iteratively using a Gaussian neighborhood function. This process continues for P iterations, refining the SOM structure. Cluster assessment is performed using a U-Matrix, with values ranging from 0 (complete similarity) to 1 (max. distance).

2.5 Data Pipeline and Experiments

Figure 1 depicts the data pipeline of the proposed method. For the frequency-domain transformation, a Short-Time Fourier Transform (STFT) was applied using a Bartlett window with a window size of 85.3 ms and a hop length of 1 ms. Since the highest simulated frequencies in the FDTD model reach approximately 15 kHz, only frequency bins up to this value were considered for the calculation of psychoacoustic features under each VD_{freq} condition.

At this stage, the pipeline branches into two processes. On one hand, the spectral centroid is calculated from the magnitude coefficients of the STFT frequency bins. Centroid differences at the discrete time points defined in the analysis framework are then computed and analyzed with respect to changes in the VD parameter space (Experiment 1). On the other hand, a decibel-scaled Mel spectrogram with quarter-tone Mel bands is computed, and spectral flux per Mel band is calculated, also at the time points t defined in the analysis framework.

Subsequently, the dissimilarity coefficients obtained from the spectral flux are segmented into three frequency ranges: (1) one octave below, (2) one octave around, and (3) one octave above VD_{freq}. For each range, a peak-picking algorithm is applied using minimum width and distance constraints to retain relevant information. The spectral dissimilarities corresponding to the extracted peaks are then isolated, normalized, and incorporated into the feature vector. Principal Component Analysis (PCA) is used to initialize the SOM's weights, and one SOM is trained per frequency range at each time point. Note that each SOM involves three stages of dimensionality reduction: frequency range selection, peak-picking, and PCA for initialization. The results are then analyzed with respect to both the VD parameters and the centroid differences (Experiment 2).

Each SOM is configured with $n = 20$ neurons, a Gaussian neighborhood radius $o = 1.5$, an initial learning rate $\eta(\ell) = 0.5$, and $P = 1000$ iterations. For visualization, the centroid difference coefficients and the parameter-space annotations of the FDTD dataset are used as metadata: the inner color of each data point represents the damping coefficient $\Re\{E(s)\}$ (darker = stronger damping), the outer color indicates the centroid difference (darker = higher difference), and the shape size encodes the rate of decay γ (larger = slower decay). The U-Matrix employs a black-to-white scale, where black denotes complete similarity and white indicates maximum dissimilarity. In this context, the term *physics-informed* refers to the principle of observational bias in physics-informed learning, where observational data inherently reflect the physical laws governing the system and thus serve to embed these laws into the machine learning model [7].

Both the analysis of centroid differences and the segmentation of the frequency spectrum into distinct ranges are key steps in the proposed method (see Fig. 1, rounded purple rectangles), as they directly address the central research question posed in the introduction—namely, where, when, and how VD influences the perceived char-

acteristics of a sound. Results are presented in the following section.

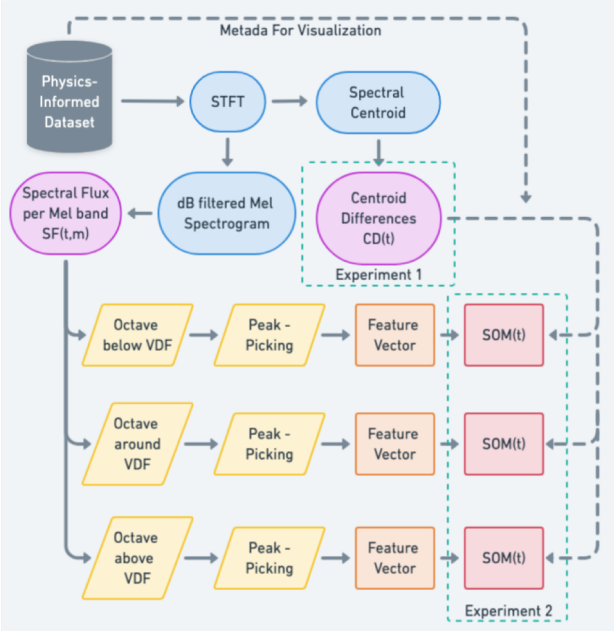

Figure 1: Overview of the data pipeline. An STFT with a Bartlett window (85.3 ms size, 1 ms hop) is applied. The pipeline branches into two processes: (1) spectral centroid, leading to centroid difference analysis (Experiment 1) and (2) computation of spectral flux per quarter-tone Mel band. Flux dissimilarities are segmented into frequency ranges relative to VD_{freq}, followed by peak picking, normalization, PCA initialization, and SOM training. Analysis is conducted in relation to VD parameters and spectral centroid differences (Experiment 2)

3. RESULTS

3.1 Experiment 1 - Centroid Differences

Figure 2 shows the analyses for the centroid differences for each VD_{freq} across the four time points defined in the analysis framework. For the 700 Hz case, at t =10 and 25 ms the centroid difference increases linearly with higher damping and slower decay, while at t = 50 and 100 ms the difference continues to grow with increasing damping but focuses towards $1/20 < \gamma < 1/70$. Moreover, at t = 50 ms some configurations of the parameter space of VD lead to a lower spectral centroid than that of the no-damping case (see Fig. 2 top row third column left to right, three first columns within the heat map).

In the 6400 Hz case, results show that at t = 10 and 25 ms the stronger the damping and the faster the decay the lower the spectral centroid. This trend persists at t = 50 ms, but the decrease is no longer linear throughout——for instance, the centroid decreases less at medium damping ($0.09 < \Re\{E(s)\} < 0.18$) than at higher damping values (≥ 0.21). At t = 100 ms, changes in the centroid become highly nonlinear; while medium damping values produce the greatest differences, this index varies with almost every decay rate. Since in both damped frequencies the 10 ms and 25 ms cases behave similarly, the 10 ms time point is discarded for the SOM analysis presented below.

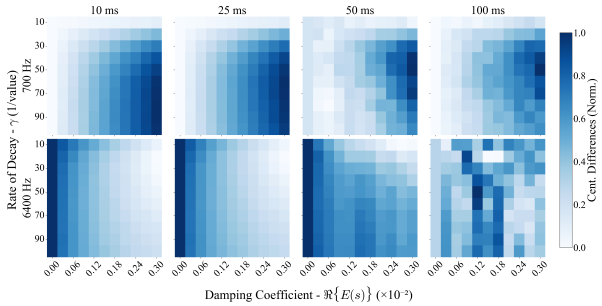

Figure 2: Centroid difference analysis for VD_{freq} = 700 and 6400 Hz. Each column, from left to right, represents one of the four time points defined in the analysis framework, increasing sequentially. Within each heat map, the damping coefficient increases left to right and the rate of decay decreases top to bottom. The darker the color, the higher the centroid difference. For the 700 Hz case, at t = 10 and 25 ms, centroid differences increase with higher damping and slower decay. At t = 50 and 100 ms, differences continue to rise but focus on damping values between $1/20 < \gamma < 1/70$. Notably, some configurations at t = 50 ms result in a lower spectral centroid than the no-damping case. For the 6400 Hz case, at t = 10 and 25 ms, stronger damping and faster decay lead to lower spectral centroids. This trend becomes nonlinear at t = 50 ms. At t = 100 ms, centroid changes become highly nonlinear. Based on these observations, the 10 ms time point is excluded from the SOM analysis.

3.2 Experiment 2 - Physics-Informed Self-Organizing Maps

Figure 3 shows the results for VD_{freq} = 700 Hz. In the octave below the VD_{freq} (first row top to bottom), at t = 25 ms data is distributed in three different regions following an increase in the damping coefficient. This distribution shows that an increase in both the damping coefficient and, in a lesser degree, the rate of decay lead to a higher centroid difference. At t = 50 ms, centroid values for low damping coefficients ($\Re\{E(s)\} < 0.12$) fall below those of the no-damping case represented by a lighter outer color, with the degree of reduction varying based on the decay rate. Furthermore, the gradual transition between clusters based on damping coefficients becomes

more diffuse, particularly separating data with a fast rate of decay. At t = 100 ms the gradual transition becomes even more fuzzy, still, centroid difference increases with increasing damping coefficient. In all three time points the clustering patterns of the SOM strongly match the centroid difference patterns found in Experiment 1 (see Fig. 2 top row). In the octave around the VD$_{freq}$ (Fig. 3 second row), at t = 25 ms the linear relation between the damping coefficients and the centroid differences remains. Moreover, the regions created by the U-Matrix group data according to their centroid difference values, although this distribution breaks the gradual transition found in the octave below the VD$_{freq}$. Small clusters based on damping coefficient are visible, but not clearly formed. At t = 50 ms, the overall data distribution aligns with the corresponding centroid differences heat map (see Fig. 2 center map). The U-Matrix groups data according to similar centroid difference values, clearly separating the two data samples with the highest difference. At t = 100 ms no pattern is found. In the octave above VD$_{freq}$, the previously observed clustering pattern related to the damping coefficient does not emerge at any of the three time points. However, small regions associated with centroid difference values are formed by the U-Matrix, particularly at t = 50 ms, where areas corresponding to low centroid differences (top left) and high centroid differences (top right) become noticeable. No distinct pattern is observed at t = 100 ms.

For the VD$_{freq}$ = 6400 Hz results are shown in Figure 4. At t = 25 ms, a gradual transition from low to high damping coefficients is displayed by the clustering pattern across all frequency ranges, which matches the trends found in Experiment 1 (see Fig. 2 bottom row). However, the octave above the damped frequency displays a clustering pattern that spans across both principal components of the SOM extensively, while the clustering from the octave below focuses only on one part of the map (bottom right), suggesting that its variance is mostly explained by only one of the principal components. Concerning the octave around the damped frequency, although it also coincides with the heat map, the SOM allocates some data points outside the overall pattern (see Fig. 4, second row, first column, top right), resulting in a decrease in how accurately the SOM matches the heat map. At t = 50 ms the octave below shows a general cluster distribution that matches the changes in centroid seen on the heat maps, while the octaves around and above cluster data mostly based on the damping coefficients with no pattern regarding the centroid, although the intra-cluster distance increases with lower rate of decay. This fuzzy clustering is also observed at t = 100 ms in the octave below the damped frequency.

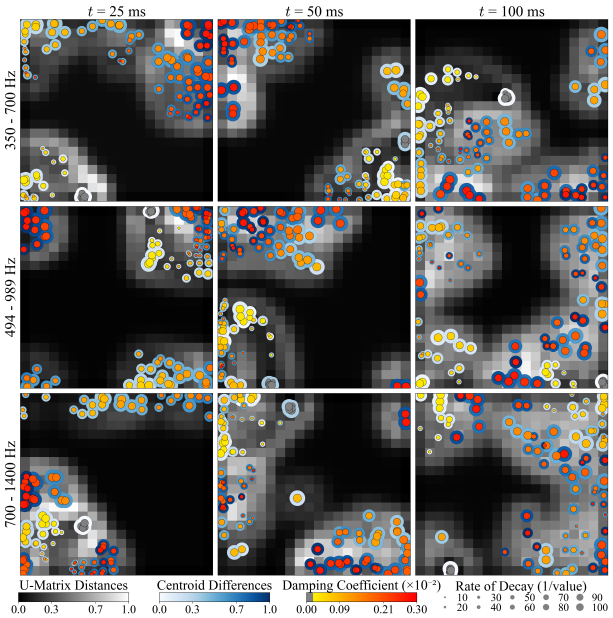

Figure 3: Physics-Informed SOM for VD$_{freq}$ = 700 Hz. Each row corresponds to a frequency range—one octave below (top), around (middle), and above (bottom) VD$_{freq}$—and each column to a time point (t = 25, 50, 100 ms). The U-Matrix displays local dissimilarities between neurons (lighter = more dissimilar). Outer color indicates spectral centroid difference (darker = higher difference), using the same color scale as in Fig. 2. Inner color shows the damping coefficient $\Re\{E(s)\}$ (redder = higher; gray = no damping). Point size reflects decay rate γ (larger = slower decay). In the octave below, a clear relationship emerges between damping coefficient, decay rate, and centroid difference—most evident at t = 25 and 50 ms—mirroring the patterns in Fig. 2. At t = 100 ms, patterns become more diffuse but the trend persists. In the octave around, data is grouped according to centroid differences, especially at t = 50 ms, though clustering is less consistent than in the lower octave. At t = 100 ms, no clear pattern emerges. In the octave above, no clustering by damping coefficient is observed, though at t = 50 ms, regions of low and high centroid difference are weakly distinguished.

4. DISCUSSION

Results suggest that how VD influences sound perception largely depends on the proximity of the viscoelastically damped frequency to the fundamental frequency. This proximity has consequences in the frequency domain that translate to the time domain and ultimately to the perception of the sound. In the frequency domain, when the damped frequency is close to the fundamental——as in the 700 Hz case—the influenced frequency region includes

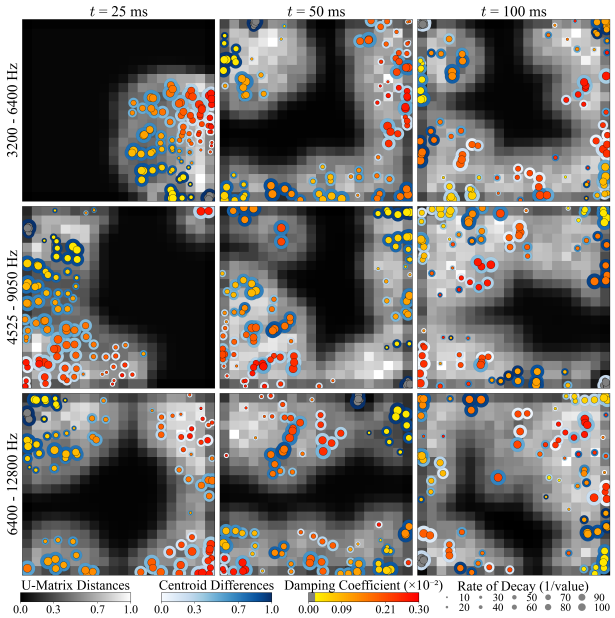

Figure 4: Physics-Informed SOM for VD$_{freq}$ = 6400 Hz. Visualization parameters (color and size coding) are consistent with those used in Fig. 3 for comparability. At t = 25 ms, clustering reflects a gradual increase in damping coefficient across all frequency ranges, aligning with centroid difference heat maps (Fig. 2). The lower octave shows compact clustering driven by the second principal component, while the upper octave spans both components. In the middle octave, some outliers reduce alignment with the heat map. At t = 50 ms, clustering in the lower octave matches trends observed in centroid differences, whereas middle and upper octaves group mainly by damping coefficient, with fuzzier patterns and greater intra-cluster distance for lower decay rates. At t = 100 ms, only the lower octave shows weak clustering; patterns in the higher ranges become increasingly ambiguous.

eigenmodes close to the fundamental frequency. Now, since lower eigenmodes carry more energy than higher ones, the memory effect stores and injects more energy in this frequency region than in a higher one. However, this is not a linear process due to the way the damping coefficient, which acts as both energy supplier and damper, and the rate of decay, which administrates the speed with which the stored energy is inserted back into the system, work together, specially when both have a middle to high value: the higher the damping coefficient, the more energy is available, but also the stronger this new injected energy is damped, on the other hand, if the stored energy decays slowly during the integration time of the memory effect, energy decays nonlinearly in the long run. According to the analysis framework, how long this decay lasts decides wether its the timbre, the pitch or the harmonicity/roughness of the sound what is influenced by the VD. In the 700 Hz case, results show that the spectral dissimilarities caused by VD may persist for 100 ms or more, affecting both the perception of pitch and the sensation of harmonicity. This phenomenon is observed in musical instruments such as the pat wain, where VD enhances pitch perception by influencing the lower modes [2]. Conversely, when the damped frequency is farther from the fundamental, as in the 6400 Hz case, the effect remains perceivable until approximately 50 ms, as shown by the clustering patterns found until this time point, sufficient to influence the timbre but unlikely to impact pitch perception or harmonicity.

Results also suggest that the size of the frequency region influenced by VD decreases over time, gradually shifting towards the region below the damped frequency. Although this size is still frequency dependent because of the different energy levels present in different parts of the spectrum as discussed above, this hypothesis is based on the absence of overall clustering patterns found in higher frequency ranges at longer time points as well as on the continuous decrease in spectral centroid found in the centroid analysis of 6400 Hz case. With regard to the clustering patterns, the clustering of data based on the damping coefficient aligns with the fact that, despite the nonlinear behavior, the damping coefficient is the driving parameter of VD [9].

In relation to the method, the proposed features (centroid differences and spectral flux per Mel band) offer a flexible tool for transient analysis in percussion instrumens. Together, they provide a framework for examining the impact of VD on perceived sound characteristics using psychoacoustically relevant metrics. Specifically, decibels measure the loudness of dissimilarities, determining whether they should be perceptible; quarter-tone Mel bands capture frequency changes that remain perceptible depending on the frequency range; and the spectral centroid serves as a well-established psychoacoustic metric for comparison. Furthermore, focusing on time intervals that are relevant to both the perception of musical instrument sounds and the physical evolution of VD provides insight into the research questions posed in the introduction—namely, how VD influences the perception of sounds produced by musical membranes, how the frequency spectrum evolves over time, and the psychoacoustic implications of the observed patterns. Additionally, each preprocessing step contributes to reduce the dimensionality of the feature vector based on psychoacoustics characteristics of the data, yielding clear results while preserving essential information.

5. CONCLUSSION

A method for analyzing the influence of viscoelastic damping on the perceived sound of musical membranes has been presented. This approach is grounded in the physical characteristics of viscoelastic damping and as well as in temporal cues relevant in the perception of musical sounds. The influence of viscoelastic damping was found to depend largely on the location along the frequency spectrum where damping occurs, specifically in relation to the distance from the fundamental frequency. A decrease in the size of the frequency region affected by viscoelastic damping was also found, with the region gradually shifting towards frequencies below the initially damped area over time. Further studies are needed to investigate additional parts of the spectrum beyond those addressed in the present work, as well as to examine scenarios involving simultaneous damping across multiple frequency regions. Also, future work should incorporate alternative time–frequency representations with higher frequency resolution—such as wavelet transforms—allowing for a more detailed analysis of both the extent and development of the affected regions. Finally, the results demonstrate that physics-informed self-organizing maps offer a powerful tool for analyzing viscoelastic damping given the complexity and high dimensionality of its parameter space.

6. REFERENCES

[1] Alexandraki, C., Starakis, M., Zervas, P., & Bader, R.: Inferring drumhead damping and tuning from sound using finite difference time domain (FDTD) models. *Acoustics* 5 (3) (2023), 798–816. https://doi.org/10.3390/acoustics5030047

[2] Bader, R.: Finite-Difference model of mode shape changes of the Myanmar pat wain drum circle using tuning paste. *Proc. Mtgs. Acoust.* 29 (1) (2016). https://doi.org/10.1121/2.0000450

[3] Bader, R.: Spectrally-Shaping Viscoelastic Finite-Difference Time Domain (FDTD) model of a membrane. *Int. J. Informatics Society* 12 (2) (2020), 81–93.

[4] Böck, S., & Widmer, G.: Maximum Filter Vibrato Suppression for Onset Detection. *Proc. 16th Int. Conf. Digital Audio Effects (DAFx)*, 55–61 (2013).

[5] Choi, K., & Cho, K.: Deep unsupervised drum transcription. *Proc. Int. Soc. Music Information Retrieval Conf. (ISMIR)* (2019), 183–191.

[6] Christensen, R. M.: Theory of viscoelasticity. An introduction. 2nd edition. Academic Press, New York, 1982.

[7] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L.: Physics-informed machine learning. *Nat. Rev. Phys.* 3 (6) (2021), 422–440. https://doi.org/10.1038/s42254-021-00314-5

[8] Kohonen, T.: *Self-Organizing Maps*. Vol. 30. Springer, Berlin Heidelberg, 2001. https://doi.org/10.1007/978-3-642-56927-2

[9] Martínez Orellanos, C. F., & Bader, R.: Analysis of nonlinear behavior of viscoelastic damping in musical membranes using physics-informed self-organizing maps. *APL Mach. Learn.* 3 (1) (2025). https://doi.org/10.1063/5.0242985

[10] Martínez, C. F., & Bader, R.: Physics- and Psychoacoustics-Based Analysis of Viscoelastically Damped Membranes. *Proc. Fortschritte Der Akustik – DAS DAGA*, to appear (2025).

[11] Masri, P.: Computer Modeling of Sound for Transformation and Synthesis of Musical Signals. Ph.D. thesis, University of Bristol, 1996.

[12] Reuter, C.: Der Einschwingvorgang nichtperkussiver Musikinstrumente. Vol. 148. Peter Lang, 1995.

[13] Schubert, E., Wolfe, J., Tarnopolsky, A., & others: Spectral centroid and timbre in complex, multiple instrumental textures. *Proc. Int. Conf. Music Perception Cognition*, Northwestern Univ., Illinois (2004), 112–116.

[14] Shete, S., & Deshmukh, S.: Automatic Tabla Stroke Source Separation Using Machine Learning. In *Adv. Signal Data Process.* (2021), 234–243. https://doi.org/10.1007/978-3-030-81462-5_22

[15] Siedenburg, K.: Specifying the perceptual relevance of onset transients for musical instrument identification. *J. Acoust. Soc. Am.* 145 (2) (2019), 1078–1087. https://doi.org/10.1121/1.5091778

[16] Vogl, R., Dorfer, M., Widmer, G., & Knees, P.: Drum transcription via joint beat and drum modeling using convolutional recurrent neural networks. *Proc. 18th Int. Soc. Music Information Retrieval Conf. (ISMIR)* (2017), 150–157.

[17] Worland, R., & Miyahira, W.: Physics of musical drum head damping using externally applied products. *Proc. Acoust. Soc. Am.* 35 (2019), 035004. https://doi.org/10.1121/2.0001011

[18] Zwicker, E., & Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. 2nd edition. S. Hirzel, 1967.