ENHANCING MUSIC VISUALIZATION WITH HAPTIC FEEDBACK TO EASE PERCEPTION FOR DHH LISTENERS

Original
ENHANCING MUSIC VISUALIZATION WITH HAPTIC FEEDBACK TO EASE PERCEPTION FOR DHH LISTENERS / Mauri, Noemi; Sacchetto, Matteo; Bagnus, Piera; Nicora, Chiara; Zanoni, Massimiliano; Bianco, Andrea; Rottondi, Cristina. - ELETTRONICO. - (2025), pp. 2913-2920. (11th Convention of the European Acoustics Association Forum Acusticum / EuroNoise 2025 Malaga (Esp) June 23 - 26, 2025) [10.61782/fa.2025.0344].

Availability:
This version is available at: 11583/3006707 since: 2026-01-19T17:04:39Z

Publisher:
Sociedad Española de Acústica

Published
DOI:10.61782/fa.2025.0344

DOI: 10.61782/fa.2025.0344

ENHANCING MUSIC VISUALIZATION WITH HAPTIC FEEDBACK TO EASE PERCEPTION FOR DHH LISTENERS

Noemi Mauri[1] **Matteo Sacchetto**[1*] **Piera Bagnus**[2] **Chiara Nicora**[3]
Massimiliano Zanoni[4] **Andrea Bianco**[1] **Cristina Rottondi**[1]
[1] Dept. of Electronics and Telecommunications, Politecnico di Torino, Italy
[2] Conservatorio Giuseppe Verdi, Como, Italy
[3] Conservatorio Guido Cantelli, Novara, Italy
[4] Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, Italy

ABSTRACT

This study introduces a novel integration of visual and haptic feedback to improve music perception for Deaf and Hard-of-Hearing (DHH) individuals. The proposed prototype, named MusicTanvas, leverages a touchscreen augmented with electroadhesion-based haptic technology, enabling dynamic overlay of tactile textures on visual content. The system processes pre-recorded audio tracks to extract spectral, harmonic, and rhythmic features, which are mapped to both visual and haptic representations. More in detail, key features such as chroma, Mel-Frequency Cepstral Coefficients (MFCCs), and rhythmic figures, are translated into corresponding visual attributes (e.g., color, shape, and size) and tactile textures. Users can explore music in real-time through these multisensory outputs, which evolve synchronously with the audio playback. A preliminary user assessment demonstrates the system's potential to enhance musical interaction by combining visual-haptic feedback.

Keywords: *Music Visualization; Haptic Feedback; DHH Listeners.*

1. INTRODUCTION

Music perception for Deaf and Hard-of-Hearing (DHH) subjects is challenging due to the unavoidable involvement of auditory perception. Hence, the reliance on extra-auditory cues is essential to ease music conveyance to DHH listeners. Among those, visual representations [1] and tactile feedback [2] are the most widely adopted ones.

Various visual techniques exist (see [3] for a thorough overview), including color coding (i.e., the association of musical elements such as notes, scales, and chords with specific colors) and graphical notation (i.e., the use of symbols, shapes, and patterns to represent musical concepts such as rhythm, dynamics, and phrasing). Meanwhile, tactile feedback has been provided through wearables (e.g., vibrating wristbands) synchronized with musical events [4], to convey to DHH users the sense of rhythm, beat, and dynamics. Moreover, tactile graphics can be exploited to create representations of musical scores using raised textures or embossed materials, to be explored while listening to a recording of the pieces.

The joint adoption of visual representations and tactile feedback can be supported by programmable touchscreens enhanced with electroadhesion-based haptic technology. Such hardware has already been explored for the conveyance of the notion of "sonic grain", i.e., a form of "tactilization" that allows for a cross-modal connection between the texture of sound and material surfaces, thus bridging tactile and sonic experiences [5].

In this paper, we adopt the TanvasTouch[1] hardware

---

*[*] ***Corresponding author**: matteo.sacchetto@polito.it.*

[1] Tanvas. Surface Haptics. `https://tanvas.co/`

support to offer a time-varying visual and tactile representation of one or multiple audio tracks [2] constituting a musical piece. Our developed prototype, named MusicTanvas, processes pre-recorded audio files to extract the evolution over time of their spectral characterization in terms of Mel-Frequency Cepstral Coefficients (MFCCs), chromas, spectral bands, entropy, their rhythmical content, and timbral features. Once the pre-processing phase is completed, a mapping between the above-mentioned features and the choice of colors, shapes, luminosity, and haptic textures to be associated to each track is identified and proposed to the listener during the audio reproduction [3] .

Results acquired through a preliminary usage assessment of MusicTanvas by a pool of 20 subjects testify to the usefulness of joint visual-tactile stimuli in enhancing music perception.

The remainder of the paper is organized as follows: Sec. 2 provides a review of the related literature. Sec. 3 introduces the TanvasTouch technology, describing its principles. Sec. 4 describes the characteristics and functionalities of our proposed implementation. The chosen visual and haptic representations are discussed in Sec. 5. The performance assessment of MusicTanvas is discussed in Sec. 6. Finally, Sec. 7 ends the paper.

2. RELATED WORK

2.1 Visual Representations

Visual representations of music play a crucial role in translating auditory information into visual stimuli for DHH individuals. Historically, Aristotle [6], Newton [7], Graves [8], Garner [9], and Kandinsky [10] explored correlations between colors and sounds. These insights have shaped interdisciplinary research, suggesting the use of color and shapes to convey musical tones and rhythmic elements [11]. From these studies, it emerges that the most common approach is to associate pitch or timbre with colors and shapes, whereas the intensity of the sound is often represented through the variation of shapes' sizes.

Smith and Williams [12] introduced a visualization technique in which the velocity of a single tone designates its loudness or softness. Velocity is represented by the relative size of a displayed sphere, with its radius ranging from 0 to 127, reflecting the minimum and maximum MIDI velocity values. Moreover, instruments with similar timbres are grouped and differentiated by color, according to the MIDI specification. Other, more recent options, such as the one proposed by Lima et al. [13], include a visualization prototype called *SongVis* that uses emojis, colors, lines, and shapes to represent different features such as danceability, mood, tempo, music genre, and instrument.

Malandrino et al. [14] proposed *VisualHarmony*, a visualization tool able to represent harmonic structures. It provides a music editor for the composition of 4-voices music with the possibility to highlight the harmonic structure with different colors, based on the circle of fifths. Furthermore, special attention is given to inclusiveness and accessibility, since VisualHarmony gives users the possibility to customize the circle of fifths to be discernible also by people with color perception deficiencies.

Deja et al. [2] introduced *ViTune*, a visualizer prototype built to allow DHH individuals to experience music by displaying a real-time visual representation of the played sound. ViTune performs a harmonic and percussive source separation once users upload their music files. The harmonic part is divided into four ranges, each one visually represented through bars and stars, with size changing based on the signal strength. Conversely, saturation, brightness, and opacity of the visualized components depend on the extracted chromagram. In our prototype, we will leverage a similar approach to visualize either chromagrams or MFCCs.

Another approach to visually representing musical features involves using shapes and sizes: Cruz et al. [15] use polygons, where the number of sides corresponds to different octaves. Higher octaves, representing treble sounds, are associated with polygons with more sides, whereas lower octaves, associated with bass sounds, are associated with polygons having fewer sides. Moreover, the volume (or velocity) of the note affects the size of the polygon: larger polygons indicate higher velocities. Our prototype adopts the same approach to visualize sound loudness.

2.2 Haptic Representations

In recent years, vibrotactile stimuli have gained attention as a method to convey musical elements, such as rhythm and melody, to DHH individuals. Research has increasingly explored how tactile feedback can enhance musical experience by converting audio signals into vibrations per-

---

`technology`

[2] A track is a single audio recording that may contain one or multiple musical stems (e.g., vocals, melody, bass, drumset).

[3] A video demonstrating the prototype usage is available on YouTube: `https://youtu.be/HEz74yxcOZ8`

ceived through the skin. Most research has focused on vibrating stimuli rather than more advanced haptic technologies, such as those based on the variation of electroadhesivity properties of surfaces [16], to create tactile sensations. Our prototype is based on the latter technique.

Furthermore, researchers have investigated how the tactile sense can convey information about a stimulus's roughness. Brown et al. [17] characterized roughness using a sinusoidal stimulus modulated by another signal, demonstrating that increasing the modulation frequency results in a rougher perceived signal. It is worth mentioning that auditory and tactile perceptions have significantly different dynamic ranges [4]. Hearing can perceive a dynamic range of approximately 130 dB, whereas tactile perception can discriminate only about 50 dB. Furthermore, audible sounds for humans range from about 20 to $20,000$ Hz, with peak sensitivity between 500 and $4,000$ Hz. In contrast, tactile sensitivity ranges from 1 to $1,000$ Hz, with peak sensitivity around 200–300 Hz, depending on the body part. Despite these disparities, both auditory and vibrotactile senses exhibit high temporal resolution, enabling accurate perception of rhythm.

Alves Araujo et al. [18] conceived the *Auris System* to enable subjects with hearing loss to experience music through other senses. Musical features such as melody, rhythm, and harmony are reproduced by the combination of two different devices, the *Auris Chair* and the *Auris Bracelet*. Furthermore, Fontana et al. [19] examined how vibrations can be leveraged to experience musical harmony through vibrotactile stimuli.

3. BACKGROUND ON TANVASTOUCH

The TanvasTouch device leverages *electroadhesion* [20], a technique that uses an electric field to create friction between surfaces. It can produce friction at 256 distinct intensity levels, enabling users to feel a wide range of textures on a 2D surface [16] [21].

The TanvasTouch device is composed of two parts: hardware and software [22]. The hardware includes a specialized touchscreen that dynamically adjusts its surface friction in response to user interaction. Custom electronics detects touch input and modifies friction forces in real time, creating tactile sensations as users slide their fingers across the screen.

The software part is composed of the TanvasTouch Engine, which controls the electrostatic forces by interpreting finger movements on the screen and adjusting the electrostatic field to provide tactile feedback. Furthermore, the TanvasTouch API allows developers to develop touch-based applications that simulate various textures and tactile sensations.

When developing an application that utilizes TanvasTouch, developers can align the graphical and haptic representations with the corresponding displayed figures to ensure that tactile feedback is delivered in the correct area of the screen. However, there could be scenarios where haptic feedback is used to offer multiple textures within a single visual representation.

The images used for haptic feedback generation need to be in gray scale with values 0-255 per pixel and in .png file format. Essentially, the Tanvas device lets users perceive different textures through the contrast between black and white images: white areas represent high friction, creating a rough sensation, while black areas represent low friction, resulting in a smooth feel. As users slide their finger across the screen, the differences in friction between the black and white areas allow them to feel varying textures.

4. PROTOTYPE IMPLEMENTATION

Our prototype[4] was created using the C# language, within the Windows Presentation Foundation (WPF) User Interface (UI) framework, part of the .NET framework. To incorporate haptic feedback, we employed the TanvasTouch API from the TanvasTouch Software Development Kit[5]. The WPF framework uses XAML (Extensible Application Markup Language) for the declarative specification of UI components, while C# is used to implement their corresponding functionalities. In our implementation, the XAML script specifies the collection of UI components such as buttons, images, and layout configurations. Conversely, the C# script manages the display state of these UI elements, toggling image visibility, facilitating page navigation, initiating audio playback, extracting audio features, and handling haptic feedback. Within the XAML document, images are integrated using the `Image` tag, and haptic feedback functionality is achieved through the `TSprite` tag, a component of the TanvasTouch API.

Besides managing UI components and haptic feedback, the application also incorporates advanced audio

---

[4] The source code, along with audio clips, is publicly available in the following GitHub repository: `https://github.com/moeni27/MusicTanvas`

[5] Related documentation can be found at `https://api-docs.tanvas.co/tanvastouch/dotnet/5.0.1/api/index.html`

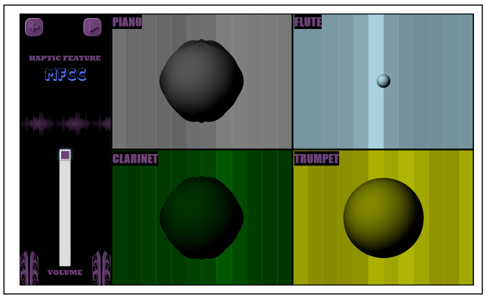

Figure 1: MusicTanvas UI, with 4 sections and MFCC values associated to background stripes.

feature extraction. This functionality is essential for providing meaningful audio analysis and improving user interactions based on audio content. The NAudio library[6], integrated into the .NET framework, facilitates audio playback, stop, and amplitude retrieval across multiple audio files concurrently. Meanwhile, feature extraction is achieved through Python and various libraries such as librosa[7] and numpy, specifically via the pythonnet library[8]. Key features extracted include spectral centroid, MFCCs, chromagram, and onset detection, which are crucial for analysis and interaction with the audio content.

5. VISUAL AND HAPTIC REPRESENTATIONS

5.1 Graphical User Interface Structure

The MusicTanvas UI is structured into up to four distinct sections, each representing a prominent instrument group (see Fig. 1). Users can explore the app in two modes: *Multi-Track Mode* and *Single-Track Mode*. In Multi-Track Mode, the user can upload up to four audio tracks (i.e., files) simultaneously. The UI dynamically adjusts based on the number of tracks uploaded, displaying a distinct section per track. All tracks must have the same duration to be reproduced or stopped simultaneously. Conversely, the Single-Track Mode utilizes Artificial Intelligence (AI) for audio processing. By uploading a single audio track, the app uses *Spleeter* [23][9] to split it into four stems: vocals, bass, percussion, and other instruments. Different studies [24] [23] evaluated Spleeter's performance by comparing it with other state-of-the-art systems, showcasing its high-quality source separation capabilities.

5.2 Audio Processing Algorithms

To offer meaningful visualization of audio content, our application relies on a number of audio processing algorithms. More in detail, we implemented a window-based extraction procedure for chromagram, MFCC, spectral centroid (a descriptor of the perceived brightness of a sound [25]) and loudness. Such window-based extraction process is applied separately to each track. This ensures that the features are computed independently for each audio source, maintaining the integrity of the individual tracks. Chromagrams, MFCCs, spectral centroids, and onsets are extracted using the *librosa* Python library and stored in NumPy arrays. The extraction uses a window length of 1024 samples with an overlap of 512 samples. Given a sample rate of 44.1 kHz, this setup results in an update time for feature extraction of $512/44100 = 0.0116 \simeq 0.012$ seconds per frame, where each frame represents a brief time interval used for feature extraction and analysis. Furthermore, the application uses the NAudio stream properties to extract the amplitude values from audio segments. Furtermore, the application employs two complementary algorithms: as previously mentioned, in Single-track mode it uses *Spleeter* to segment the audio track into four predefined stems, whereas in Multi-track mode it implements an instrument recognition algorithm. This algorithm leverages a Convolutional Neural Network (CNN) and a K-Nearest Neighbors (KNN) classifier based on MFCCs, chroma, and spectral centroid acoustic features, to identify and categorize the primary musical instrument present in each uploaded audio track [26]. Based on the classification output, each track is then assigned to the corresponding section and color according to the associations proposed by Kandinsky [10] (see Table 1).

Finally, to quantify rhythmic complexity, we use an onset extraction algorithm [27], that starts by transforming the audio signal into a time-frequency representation using the Short-time Fourier Transform (STFT). This process breaks down the audio signal into overlapping frames (with a 50% overlap) and derives the spectral characterization for each frame, capturing the energy distribution of the signal. The overlapping frames are crucial to enhance the temporal resolution of the analysis, as they ensure that each frame shares information with adjacent frames, capturing smooth transitions, preserving temporal relationships in the audio signal, and reducing the influence of

---

[6] https://github.com/naudio/NAudio

[7] https://github.com/librosa/librosa

[8] https://github.com/pythonnet/pythonnet

[9] An AI source separation tool that uses pre-trained models to extract individual stems from audio tracks, where each stem corresponds to a specific instrument

Table 1: Musical Instrument and Color Mapping

Musical Instrument	Color Name	Visual Representation
Cello	Deep Blue	🟦
Clarinet	Dark Green	🟩
Flute	Light Blue	🟦
Acoustic Guitar	Orange	🟧
Electric Guitar	Violet	🟪
Organ	Deep Purple	🟪
Piano	Grey	⬜
Saxophone	Brown	🟫
Trumpet	Light Yellow	🟨
Violin	Light Violet	🟪
Vocals	Gold	🟨
Bass	Blue	🟦
Drum	Red	🟥
Other	Green	🟩

windowing artifacts.

From the STFT output, the spectrogram of the audio track is computed, which represents the magnitude of the audio signal over time. We apply a logarithmic compression to the spectrogram's magnitude, which enhances weak spectral components and simulates human perception, highlighting smaller variations that might otherwise be overlooked. Subsequently, a novelty spectrum is derived from the spectrogram, highlighting abrupt energy changes between consecutive frames. This spectrum is obtained by computing the difference between adjacent columns of the spectrogram: a positive difference indicates significant changes in the audio signal, hinting to potential onsets. The novelty spectrum represents such energy changes and helps detect new events. To enhance detection accuracy and minimize noise, a local average is computed across the spectrum. This smoothing process helps differentiate genuine onsets from minor signal fluctuations. In the final onset detection stage, each spectrum value is adjusted by subtracting its respective local average. Values surpassing a predefined threshold are identified as significant onset indications.

5.3 Mappings of Audio Features to Visual and Haptic Representations

Tab. 2 shows how specific audio features are mapped to their corresponding visual and haptic representation in our interface. More in detail, a sphere (i.e., the shape chosen to represent musical features) is visually represented in the foreground of each section (as in [12]). The sphere's radius dynamically adjusts based on the loudness of the corresponding audio track. This adjustment reflects amplitude variations within the track and is influenced by a volume slider controlling the overall audio output of the application (located on the left side of Fig. 1). If the tracks have different initial audio levels, the application performs normalization to ensure consistent volume output. The sphere's color brightness is linked to the spectral centroid feature. As the spectral centroid shifts across the audio track, the sphere's color brightness dynamically adapts, visually reflecting changes in tonal brightness. Moreover, the output of the onset detection algorithm influences the surface of the sphere, visually accentuating event changes in the audio content (Fig. 2). This representation gives users the possibility of visually appreciating the temporal structure and intensity variation within the musical composition. Spherical representations were chosen for their smooth visual appearance, complemented by dynamic changes to their surfaces to reflect audio characteristics effectively.

Table 2: Mapping of Audio Features to Visual Representations

Audio Feature	Visual Representation	Haptic Representation
MFCC or Chromagram	Color brightness of background stripes	Texture roughness of background stripes
Rhythmic complexity	Sphere's surface	-
Spectral centroid	Sphere's color brightness	-
Loudness	Sphere's radius	-

Regarding the background, inspired by [10], each section is adorned with a specific color palette, chosen according to the instrument automatically detected in the corresponding audio track. Moreover, in each section, the background is divided into several vertical stripes. Each stripe symbolizes either one of the 7 MFCCs derived from the feature extraction process or one of the 12 pitch classes of the chromagram. The choice between 7 and 12 stripes is determined by the feature being represented. The stripes' colors dynamically transition from lighter to darker hues, reflecting the normalized values of the corresponding MFCCs/chromas within a specific temporal context. Besides the visual representation, each

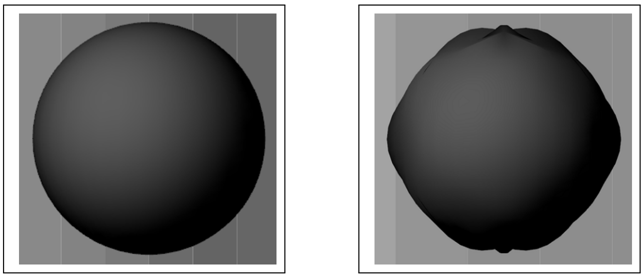

Figure 2: MusicTanvas GUI showing onset detection: smooth surface indicates no onset (left), non-smooth surface indicates onset detection (right).

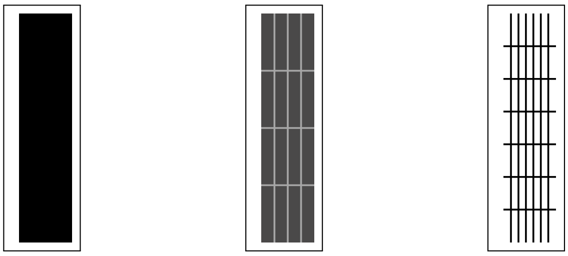

Figure 3: Haptic images representing feature value ranges: $0 \leq x < 0.1$ (left), $0.4 \leq x < 0.5$ (center) and $0.9 \leq x \leq 1$ (right).

stripe also features a corresponding haptic feedback representation. This haptic representation consists of a gradient stripe, transitioning from black to white through various shades of gray. Specifically, ten distinct haptic images are generated, ranging from black to white via 8 intermediate shades of gray.

Furthermore, vertical and horizontal lines are added, forming a grid-like pattern. For darker shades of gray, the grid consists of a few horizontal and vertical lines. As the shade lightens, more lines are progressively added, increasing the density of the grid. This granularity variation enhances the haptic experience, allowing users to better distinguish between different values through tactile feedback. These shades correspond to MFCC/chroma values ranging from 0 to 1. For instance, an MFCC value between 0 and 0.1 triggers a black haptic image, a value between 0.1 and 0.2 triggers a dark gray image, whereas a white image is associated with values between 0.9 and 1 (see Fig. 3). This approach allows users to perceive tactile feedback corresponding to the MFCC/chromagram variations by swiping their fingers over the screen. MFCC and chroma values are generated approximately every 12 milliseconds, but this interval is too brief for users to perceive the tactile feedback effectively. To overcome this issue, the system aggregates the values over a 5-second window and computes their average, thereby smoothing the tactile feedback into a more coherent and perceptible signal.

Table 3: Validation Accuracy vs. Number of Categories.

Number of Categories	2	3	11
Validation Accuracy	90%	75%	60%

Table 4: Processing Times for Feature Extraction and Source Separation

Audio Length (s)	10	40	120
Extraction Time (s)	0.17	0.68	2.17
Separation Time (s)	25.93	31.28	51.60

6. PERFORMANCE ASSESSMENT

The performance assessment of MusicTanvas evaluates instrument recognition accuracy, computational efficiency and user experience.

The IRMAS dataset [28] was used to train and test the instrument recognition model. The dataset includes 6705 audio recordings from 11 instrument classes. The dataset was split into 80% for training and 20% for testing. Moreover, a 5-fold cross-validation approach was performed to ensure robustness evaluation of the model. As shown in Tab. 3, the accuracy drops from 90% when classifying two instruments to 60.70% for 11 categories. This highlights the challenge of distinguishing between instruments with simlar audio features.

The system's computational efficiency was evaluated by measuring processing times for audio feature extraction and source separation on tracks of varying lengths, using *Spleeter*. As summarized in Tab. 4, feature extraction time and source separation time grow approximately linearly with the track length.

Due to the difficulty of recruiting DHH participants, a preliminary user assessment study was conducted with 20 non-impaired participants aged between 20 and 60 (see Tab. 5 for an overview of the participants' musical experience level). To simulate deafness, participants interacted with MusicTanvas without any auditory feedback, and all tasks were evaluated without sound. Participants were given a brief training session before the evaluation.

The user experience assessment evaluated the color-

Table 5: Musical Experience Level (scale from 1 = beginner to 5 = expert)

Experience Level	1	2	3	4	5
Frequency	2	4	10	4	0
Percentage	10%	20%	50%	20%	0%

Table 6: Enjoyment of MusicTanvas (scale from 1 = not enjoyable to 5 = very enjoyable)

Enjoyment	1	2	3	4	5
Frequency	0	0	0	12	8
Percentage	0%	0%	0%	60%	40%

instrument mapping and the impact of tactile and visual feedback on musical perception. Results show that color-instrument associations proposed by participants largely aligned with the MusicTanvas mapping. Furthermore, 85% of users preferred the multi-quadrant display for multi-instrument recognition, citing improved instrument differentiation.

Tactile and visual feedback were appreciated for enhancing the perception of individual musical parts, although the tactile feedback alone received mixed ratings regarding its effectiveness. Moreover, participants reported no negative experiences with tactile feedback, and all of them found the visualization helpful for instrument recognition. MusicTanvas was also appreciated for its ease of use (see Tab. 6). All participants agreed that it has great potential to help DHH subjects.

7. CONCLUSIONS

This paper presented MusicTanvas, a prototype application that utilizes a touchscreen enhanced with electroadhesion haptic technology, which enables the overlay of tactile textures onto displayed content that dynamically changes over time. The prototype incorporates a set of algorithms that pre-process single or multiple pre-recorded audio tracks to extract key spectral, harmonic, and rhythmic elements within set time intervals. These extracted musical features are stored as arrays or lists. During playback, the system retrieves the pre-extracted feature values, ensuring that they evolve synchronously with the audio. These values are then used to control both visual and haptic feedback, to ease music perception for DHH individuals with a multisensory interaction that matches the auditory playback.

As future work, we aim to develop a customizable feature mapping that will allow each user to define the sensory feedback that suits them best to perceive a given musical characteristic. Moreover, enhancing the system to process live audio in real-time would allow users to experience music as it is being performed.

Finally, this technology may prove useful also for blind and visually impaired individuals, if all visual elements are properly replaced with a haptic overlay, thus conveying a wider variety of musical features exclusively via tactile feedback.

8. ACKNOWLEDGMENTS

We acknowledge the support of the “Musical Metaverse: an inclusive Extended Reality platform for networked musical interactions” project (grant n. 2022CZWWKP) – funded by European Union – Next Generation EU within the PRIN 2022 program (D.D. 104 - 02/02/2022 Ministero dell’Università e della Ricerca) and of the MUSMET project funded by the EIC Pathfinder Open scheme of the European Commission (grant agreement n. 101184379). This manuscript reflects only the authors’ views and opinions and the Ministry, the European Union or the European Innovation Council cannot be considered responsible for them.

9. REFERENCES

[1] M. D. Fletcher, “Can haptic stimulation enhance music perception in hearing-impaired listeners?,” *Frontiers in Neuroscience*, vol. 15, p. 723877, 2021.

[2] J. A. Deja, A. Dela Torre, H. J. Lee, J. F. Ciriaco IV, and C. M. Eroles, “Vitune: A visualizer tool to allow the deaf and hard of hearing to see music with their eyes,” in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2020.

[3] H. B. Lima, C. G. D. Santos, and B. S. Meiguins, “A survey of music visualization techniques,” *ACM Computing Surveys (CSUR)*, vol. 54, no. 7, pp. 1–29, 2021.

[4] S. Papetti and C. Saitis, *Musical haptics*. Springer Nature, 2018.

[5] M. Sacchetto, M. Sangüesa, P. Bagnus, C. Nicora, and C. Rottondi, “Collection of design directions for the realization of a visual interface with haptic feedback to convey the notion of sonic grain to dhh students,” in *2023 4th International Symposium on the Internet of Sounds*, pp. 1–7, IEEE, 2023.

[6] A. Platt, J. A. Smith, W. D. Ross, *et al.*, *The Works of Aristotle*. Clarendon Press, 1912.

[7] R. Shamey, "Newton,(sir) isaac," in *Encyclopedia of Color Science and Technology*, pp. 1225–1230, Springer, 2023.

[8] M. Graves, *The Art of Color and Design*. McGraw-Hill Book Company, Incorporated, 1941.

[9] B. Hero and W. Garner, "On the relationship between color and music," *Leonardo*, vol. 12, no. 1, pp. 87–88, 1979.

[10] W. Kandinsky and M. Sadleir, *The Art of Spiritual Harmony*. Constable limited, 1914.

[11] J. L. Caivano, "Color and sound: Physical and psychophysical relations," *Color Research & Application*, vol. 19, no. 2, pp. 126–133, 1994.

[12] S. M. Smith and G. N. Williams, "A visualization of music," in *Proceedings. Visualization'97 (Cat. No. 97CB36155)*, pp. 499–503, IEEE, 1997.

[13] H. Lima, C. Santos, and B. Meiguins, "Visualizing the semantics of music," in *2019 23rd International Conference Information Visualisation (IV)*, pp. 352–357, IEEE, 2019.

[14] D. Malandrino, D. Pirozzi, and R. Zaccagnino, "Visualization and music harmony: Design, implementation, and evaluation," in *2018 22nd International Conference Information Visualisation (IV)*, pp. 498–503, IEEE, 2018.

[15] L. Cruz, V. Rolla, J. Kestenberg, and L. Velho, "Visual representations for music understanding improvement," in *Proceeedings of Music Technology with Swing: 13th International Symposium, CMMR 2017*, pp. 468–476, Springer, 2018.

[16] S. Feitl, J. Kreimeier, and T. Götzelmann, "Accessible electrostatic surface haptics: towards an interactive audiotactile map interface for people with visual impairments," in *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments*, pp. 522–531, 2022.

[17] L. M. Brown, S. A. Brewster, and H. C. Purchase, "A first investigation into the effectiveness of tactons," in *First joint eurohaptics conference and symposium on haptic interfaces for virtual environment and teleoperator systems. world haptics conference*, pp. 167–176, IEEE, 2005.

[18] F. Alves Araujo, F. Lima Brasil, A. Candido Lima Santos, L. de Sousa Batista Junior, S. Pereira Fonseca Dutra, and C. Eduardo Coelho Freire Batista, "Auris system: Providing vibrotactile feedback for hearing impaired population," *BioMed research international*, vol. 2017, no. 1, p. 2181380, 2017.

[19] F. Fontana, I. Camponogara, M. Vallicella, M. Ruzzenente, P. Cesari, *et al.*, "An exploration on whole-body and foot-based vibrotactile sensitivity to melodic consonance," in *Proceedings of the 13th International Conference in Sound and Music Computing, SMC 2016*, pp. 143–150, 2016.

[20] Tanvas, "Tanvastouch technology," 2024. Accessed: 2025-01-11. Available at `https://tanvas.co/technology`.

[21] A. Krishna, A. W. Luangrath, and J. Peck, "A review of touch research in consumer psychology," *Journal of Consumer Psychology*, vol. 34, no. 2, pp. 359–381, 2024.

[22] Tanvas, "Tanvastouch engine," 2024. Accessed: 2025-01-10. Available at `https://tanvas.co/resources/tanvastouch-basics`.

[23] R. Hennequin, A. Khlif, F. Voituret, and M. Moussallam, "Spleeter: a fast and efficient music source separation tool with pre-trained models," *Journal of Open Source Software*, vol. 5, no. 50, p. 2154, 2020.

[24] L. A. CA and C. Ancy, "Research on dnn methods in music source separation tools with emphasis to spleeter," *International Research Journal on Advanced Science Hub*, vol. 3, no. 6S, pp. 24–28, 2021.

[25] E. Schubert and J. Wolfe, "Does timbral brightness scale with frequency and spectral centroid?," *Acta acustica united with acustica*, vol. 92, no. 5, pp. 820–825, 2006.

[26] S. Prabavathy, V. Rathikarani, and P. Dhanalakshmi, "Classification of musical instruments using SVM and KNN," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 7, pp. 1186–1190, 2020.

[27] M. Müller, *Fundamentals of music processing: Audio, analysis, algorithms, applications*, vol. 5. Springer, 2015.

[28] Irmas, "Irmasdataset," 2024. Accessed: 2025-01-07. Available at: `https://zenodo.org/records/1290750#.WzCwSRyxXMU`.