# Visual Representations for Music Understanding Improvement

Leandro Cruz[1], Vitor Rolla[2], Juliano Kestenberg[3], and Luiz Velho[2]

[1] Universidade de Coimbra
lmvcruz@isr.uc.pt
[2] Instituto Nacional de Matemtica Pura e Aplicada
{vitorgr, lvelho}@impa.br
[3] Universidade Federal do Rio de Janeiro
info@julianokestenberg.com

**Abstract.** Classical music appreciation is non trivial. Visual representation can aid music teaching and learning processes. In this sense, we propose a set of visual representations based on musical notes features, such as: note type, octave, velocity and timbre. In our system, the visual elements appear along with their corresponding musical elements, in order to improve the perception of musical structures. The visual representations we use to enhance the comprehension of a composition could be extended to performing arts scenarios. It could be adopted as motion graphics during live musical performances. We have developed several videos to illustrate our method. We have also developed an ear training quiz and a research questionnaire. This material is available at http://www.impa.br/~vitorgr/VLMU/home.html.

**Keywords:** music visualization, geometric shapes, pedagogy and music education

## 1 Introduction

Visual representations of music have been studied since 1938, when Oskar Fischinger presented the film "An Optical Poem". It is based on the 2nd Hungarian Rhapsody (by Franz Liszt). The movie associates movements of geometric shapes and musical elements. The definition of shapes is not clear, but the video is completely synchronized with musical elements. Two years later, in 1940, Disney's movie "Fantasy" featured eight cartoon segments in which the characters movement is also synchronized with excerpts from classical pieces of music.

Both films give evidence of the relationship between visual and musical elements. This is the main fundamental path of this research. We define visual representations which are clearly associated with musical elements in order to improve their perception. It facilitates the understanding of musical phenomena. This pedagogical goal is the main objective of this research. Ordinarily, the full appreciation of classical music demands a substantial amount of time to learn the basics of music theory, as well as training the ear to be able to understand and identify the contrivances utilized by the composer.

We will present an approach for enhancing the perception of a note and its properties like tonal pitch, octave and volume. The proposed visual representation is based on polygons, and they are related to the notes through an appropriate association of visual properties like color and shape.

Our intention is to create a visual language that is able to represent complex musical phenomena, like rhythm, harmony, and melody. However, in this first step, we are proposing the basis of this language with the analysis of the most essential element for musical studies: the note. From our visualization approach, we can then highlight the occurrence of chords and also timbre. Furthermore, we can observe the relationship between dominance and accompaniment.

The music visualization can be an image (a global representation) or a video (a temporal representation). In the image, we will represent global phenomena providing a big picture of the composition. The local approach is a video with the same duration of the music, where the visual representations are synchronized with the respective musical elements. In our case, for each note in the music, we will show a corresponding polygon with the respective color and shape during the time the note is being played.

In Section 2, we describe other music visualizations types, and compare them to our approach. In Section 3, we present the association of colors and tonal pitch. We demonstrate how this association can be used to observe the dominance and accompaniment phenomena. In Section 4, we present our approach. Finally, in Section 5, we conclude with final considerations and future work.

## 2 Related Work

Coloppy (2000) proposed a conceptualization for music visualizations. In this work, the author presents a complete study on how to represent colors and how they can be related to elements of a song. In addition to colors, he presented a conceptualization about the use of shapes, such as points, lines, curves, polygons and free shapes, and how to associate them with musical elements such as rhythm and harmony. Finally, it also highlights the use of movements and shapes to create a relationship between the representations which are displayed. The conceptualization presented by Collopy is quite broad, and very associated with the use of procedural methods to control elements of a song. Our work goes in another direction, because we use visual elements that are more simple and can be used as a pedagogical tool in music learning environments.

Chan et al. (2010) propose an innovative representation to reveal the semantic structure of classical compositions. They propose an analogy of weaving in textile art to construct pictures. Such pictures demonstrate correlations and occurrences of the music *motif* within different sets of instruments. Sapp (2007) proposes two types of diagrams to view the harmonic structure and relationships between key regions in a musical composition. A discrete mapping of sound tones into color tones is proposed. Miyazaki et al. (2003) propose a method to create a 3D interactive picture from Musical Instrument Digital Interface (MIDI) files. Pitch, volume, and tempo of a note are encoded as height, diameter, and color

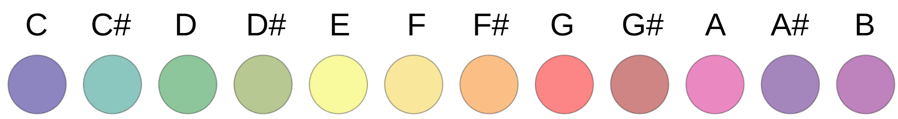



**Fig. 1.** Sound tones to color tones according to Louis Bertrand Castel.

saturation in the interactive illustration. The users can take advantage of 3D perspective view and illumination to navigate within the MIDI file data. These works present a global music visualization approach.

Fonteles et al. (2013) introduce a particle system to generate real-time animated particle emitter fountains choreographed by classical music. The authors describe music as highly organized sounds which exhibit time-varying structures in pitch and time domains. Their main goal was to visualize the music *motif* by analyzing such time-varying structures. Bergstrom et al. (2007) propose a method called Isochords, for visualizing the chord progression structure of musical compositions. A triangular isometric grid called Tonnetz (Cohn, 1997), which was invented by Euler, is utilized to make the quality of chords and intervals more salient to the audience. The authors claim that their proposal shows harmony as it changes overtime and they believe it is possible to visually detect familiar *motif* patterns with Isochords. Ciuha et al. (2010) present a different color mapping for visualizing a group of concurrent tones. The authors propose a continuous color wheel for mapping. Their method is different from most sound tones to color tones mapping which are commonly discrete. Dissonance is represented with low color saturation, while consonance is represented with high color saturation. These works present a local music visualization approach.

Most of the related work is mainly concerned with: harmonic analysis, chord progression structure, *motif* identification, or 3D interactive illustration. The only exception is the paper published by Chan et al. (2010). In such paper, the authors are concerned with the role of instruments inside an orchestra. An orchestra is a large instrumental ensemble. Thus, the visual language proposed by Chan et al. (2010) is highly complex, turning out to be inadequate to pedagogical purposes. The visual representation proposed in this paper is designed to evince specific characteristics of music, such as: timbre, chord, note tone, and volume. Although the proposed visual language can also help the user to grasp some information about chord progression and *motif* identification, the focus is to apply the visual language on small ensembles, i.e., on compositions arranged for chamber music. The proposed visual language can be utilized to create global and local views of music.

## 3 Colors

Commonly, it is easier to recognize colors in a high contrast palette compared to distinguishing musical notes. In this way, we use an appropriate coloration of our visual representations to highlight the occurrence of each type of note played. In this section, we present a very simple association between colors and notes. We also show an example of how it is easy to extend a perception of the relationship between instrument dominance and accompaniment.

In this work, we are using an association between the tonal pitch of a note with a specific color from a palette. The construction of this palette is arbitrary, i.e., it is possible to use several types of palettes for specific purposes. But note that the colors of this palette should be easily distinguishable.

In our experiments, we used the colors palette proposed by Louis Bertrand Castel which presents a continuous chromatic transition, maintaining sufficient contrast between the tones adopted. In this way, the colors have a coherent flow, being easily recognized and facilitating learning processes. There are studies that show how the relationship between this colors palette and musical notes is intuitively perceived by the human brain (Brougher et al., 2005).

From the combination of colors and notes, we can create global visual representations of music. In particular, we will present two very simple visual representations that can be used to perceive dominance and accompaniment in a song. We will compare this approach with the temporal process of listening to a musical piece, or looking at a diagram containing the precise distribution of the notes for each timbre, showing how the colors make a great contribution to the perception of the dominance and accompaniment relationship. The Figure 2 illustrates the diagram with the distributions of the notes and the two proposed visual representations.

This study was performed with chamber music, which is a type of classical music commonly played by a small group of musicians, which could be positioned in the chamber of a palace. In particular, we will analyze two interpretations of classical music played by string quartets.

Classical music composers usually build a piece of music with a principal theme, which is repeated along the piece by different instruments. Variations on volume, rhythm, and melody are utilized to introduce form and interest to the musical piece. Thus, the instruments may have different roles in a music composition. When the instrument has the accompaniment role, its composition part supports the dominant instrument.

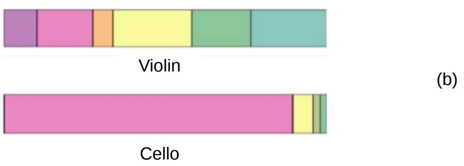
Accompaniment instruments usually provide harmonic background and/or rhythmic structure to dominant instruments. Often, the harmonic background is a chord progression related to the song theme, while the rhythmic structure is a regular recurrence in time of a small set of notes. Instrument dominance can be characterized by higher volumes, higher variance of notes, and complex melodies.

For instance, in a string quartet, the cello commonly provides the rhythmic structure in a lower volume, while the first violin provides the melody in a higher volume. In this case, the cello is the accompaniment instrument, while the first violin is the dominant instrument.

The Figure 2(a) shows a table with the number of notes played by each instrument in a song. This is a fairly accurate representation, objectively presenting such quantities. That way, with careful analysis, we can see which instruments

| | C | C# | D | D# | E | F | F# | G | G# | A | A# | B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| T1 | 0 | 23 | 18 | 0 | 24 | 0 | 6 | 0 | 0 | 17 | 0 | 10 |
| T2 | 0 | 26 | 8 | 0 | 6 | 0 | 0 | 0 | 2 | 32 | 0 | 20 |
| T3 | 0 | 20 | 16 | 0 | 42 | 0 | 8 | 0 | 2 | 4 | 0 | 2 |
| T4 | 0 | 0 | 2 | 2 | 6 | 0 | 0 | 0 | 0 | 84 | 0 | 0 |

(a)



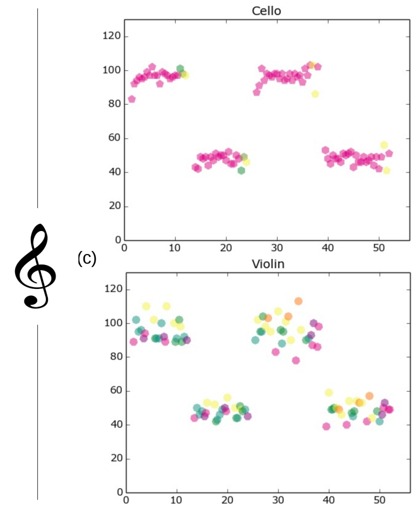




**Fig. 2.** Color Analysis: Dominance vs. Accompaniment.

play only a small set of notes (which indicates that they perform an accompaniment role) and which play a larger variety of notes (which indicates that they perform a dominant role). However, careful reading of several numbers is necessary to realize the existence of this pattern.

Although it is possible to observe the instrument accompaniment and dominance relationship based on the table, it is not as intuitive as color-based representations. For example, Figure 2(b) shows two bars with the proportion of occurrence of each note played by different instruments (timbre: T4 and T1). This visual property quickly shows us the variety of notes played by each instrument, allowing us to easily define which is the dominant (T1) and which performs the accompaniment (T4).

The volume of a note is associated with the intensity in which a note is played by the instrumentalist. Figure 2(c) shows a complete view of a song. It simultaneously represents a small geometric shape for each note played. The shape position is determined according to the moment in which the note was touched (x-axis) and also to its volume (y-axis).

In Figure 2(c), it is not easy to notice the proportions of the notes like in the bar chart, because in the representation of a song with a reasonable amount of notes we will have an superposition of colors throughout the graph. However, we can still have a general perception of whether or not there is a variety of colors, as well as perceiving the velocities (volume) of notes. In this way, we can also perceive the dominance (more colors with varying velocities) and accompaniment (few colors with continuous velocities).

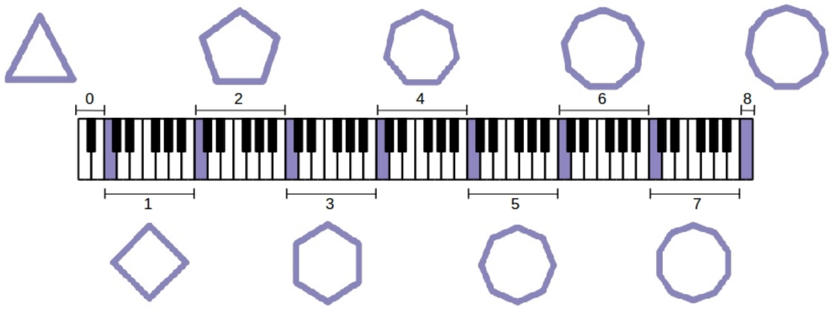



**Fig. 3.** Octaves visual representation.

## 4 Shapes

In this section, we will present the relationship between shapes and musical elements. The simultaneous occurrence of the visual representation and its respective musical element creates a synesthetic effect.

### 4.1 Note Perception

A note is the most fundamental musical element in a song. Therefore, we focus our approach on it. In addition to the association of note tones and specific colors, we will also show how other note properties can be highlighted.

Whenever a note is played, it must be visually represented by a geometric shape, in our case a polygon. As mentioned above, the pitch of the note is represented in accordance with a color palette already presented in Figure 1. We will represent the octaves according to the number of sides of a polygon, as shown in Figure 3. Higher octaves mean treble sounds and more polygon sides. Lower octaves mean bass sounds and less polygon sides. In addition, the note velocity is associated with the size of a polygon. The larger the velocity, the larger the size of the polygon in the visual representation.

### 4.2 Chord Perception

Chords are sets of notes played at the same time. They are represented by a group of polygons associated with the respective color notes. For example, representations may be nested or shifted, as illustrated in Figure 4.

For instance, an instrument playing chord (a) would be playing notes "C" , "E", and "G" at the third octave (hexagon). An instrument playing chord (b) would be playing notes "C" , "E", "G", and "B" at the fifth octave (octagon). An instrument playing chord (c) would be playing notes "D" , "G", "B", and "F#" at the second octave (pentagon). An instrument playing chord (d) would be playing notes "D" , "G", and "B" at the fifth octave (octagon), and also "F#" at the third octave (hexagon). The height of a polygon is proportional to

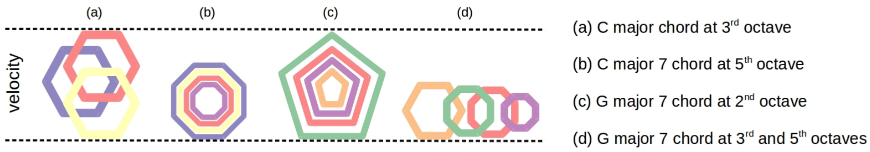



**Fig. 4.** Chords are represented by nested or shifted polygons.

the velocity (volume) of a note. For instance, the green note in chord (d) has lower volume when compared to the green note in chord (c).

### 4.3 Timbre Perception

The timbre of a musical instrument enables a listener to judge that two non-identical sounds, similarly presented and having the same loudness and pitch, are dissimilar. An approach that can be used to represent different timbres in a composition is to draw the polygons in different areas of the image accordingly. Figure 5 is a picture of a musical piece played by four instruments: a cello, two violins and a viola. The visual representation of each timbre is placed in a quadrant of the frame.

### 4.4 Motion And Blurring

In this paper, we kept the focus on the visual properties like color and shape. These features are easily observed, and thus, they produce a more efficient synesthetic effect. In this section, we will talk about two simple behaviors and their meaning for music visualization. These behaviors are motion and blurring.

When a note (or chord) is played, its corresponding polygon appears on the screen and grows reaching its maximum size at half of the note duration. For this purpose, we have observed that an appropriate use of the scaling (an arising and disappearing of the representation) has highlighted the musical note occurrence.

The other effect that we apply is blurring. After a note (or chord) is played, it disappears slowly blurred. The following notes are then shown in the front. This effect lets us perceive notes that have been played recently. Although we do not approach rotations and translations, we believe that they can be used further to represent rhythm or harmony.

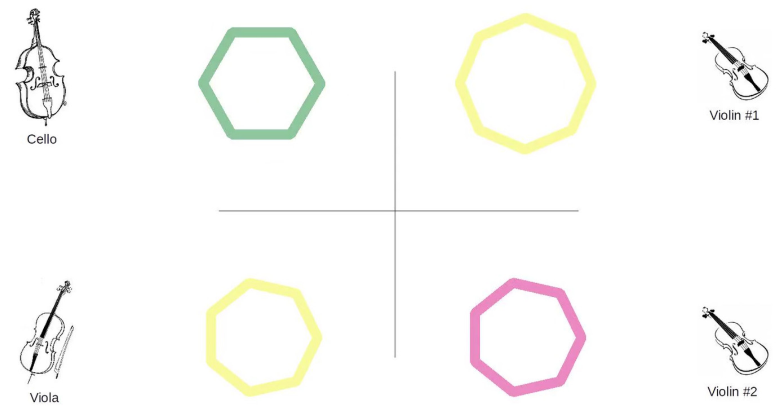



**Fig. 5.** Four timbres of a String quartet.

## 5 Conclusion and Future Work

This work demonstrates the potential of using visual representations to improve the understanding of musical elements. We present a method on how to use colors, shapes, movements, and blurring associated with musical notes and their respective properties. With our visual representations we show examples of dominance and accompaniment between different musical instruments.

Currently, we have made available an ear training quiz and a research questionnaire at the project website: `http://www.impa.br/~vitorgr/VLMU/home.html`. Our goals with this evaluation are (i) to understand the level of expertise necessary to use the system; and (ii) to receive feedback from users with different background experiences that might help us improve our method.

We intend to expand the set of visually represented musical elements. In this way, we can create more complex visual representations based on the analysis of rhythm and harmony. A research direction with great potential is to explore polygon positioning and movements, and also appearance and disappearance effects. Another extension is to work on instruments grouping in a way that would let us represent a whole orchestra.

A visual language for music understanding is proposed to facilitate common listeners and music students to understand and visualize important structures of classical music. The presence of *motifs* and instrument roles were visually identified in the music illustrated videos. MIDI music visualization has the potential to become a powerful pedagogical tool for teaching and learning classical music.

**Acknowledgement:**
The authors would like to thank the following institutions for their support: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and finally Instituto Nacional Casa da Moeda (Portugal).

## References


1. Fred Collopy (2000) Color, Form, and Motion: Dimensions of a Musical Art of Light. Journal Leonardo
2. Ohmi K (2007) Music Visualization in Style and Structure. Journal of Visualization 10-3:257-258. doi:10.1007/BF03181691
3. Gareth L, (1985) Musicians Make a Standard: The MIDI Phenomenon. Comput. Music J. 9:8-26. doi:10.2307/3679619
4. Fonteles JH, Rodrigues MAF, Basso VED (2013) Creating and evaluating a particle system for music visualization. J. Vis. Lang. Comput. 24:472-482. doi:10.1016/j.jvlc.2013.10.002
5. Brougher K, Strick J, Wiseman A, Zilczer J (2005) Visual Music: Synaesthesia in Art and Music Since 1900. Thames & Hudson, London.
6. Chan WY, Qu H, Mak WH (2010) Visualizing the semantic structure in classical music works. IEEE Trans. Vis. Comput. Graph. 16:161-173. doi:10.1109/TVCG.2009.63
7. Bergstrom T, Karahalios K, Hart J (2007) Isochords: Visualizing Structure in Music. In Proceedings of Graphics Interface 2007 (GI '07). ACM, New York, NY, USA, 297-304. doi:10.1145/1268517.1268565
8. Ciuha P, Klemenc B, Solina F (2010) Visualization of concurrent tones in music with colors. In Proceedings of the 18th ACM international conference on Multimedia (MM '10). ACM, New York, NY, USA, 1677-1680. doi:10.1145/1873951.1874320
9. Miyazaki R, Fujishiro I, Hiraga R (2003) Exploring MIDI datasets. In ACM SIGGRAPH 2003 Sketches & Applications (SIGGRAPH '03). ACM, New York, NY, USA, 1-1. doi:10.1145/965400.965453
10. Sapp C (2003) Harmonic visualizations of tonal music. In Proceedings of International Computer Music Conference, Havana, Cuba, pages 423430.
11. Cohn R (1997) Neo-Riemannian Operations, Parsimonious Trichords, and their Tonnetz Representations. Journal of Music Theory, 41/1 , 166.
12. Zhu J, Wang Y (2008) Complexity-Scalable Beat Detection with MP3 Audio Bitstreams. Computer Music Journal, Spring 2008, Vol. 32, No. 1, Pages 71-87. doi:10.1162/comj.2008.32.1.71
13. Dannenberg R (1993) A brief survey of music representation issues. Computer Music Journal, Vol. 17, Pages 20-30.
14. Gardner R (1964) Music notation: a manual of modern practice. Allyn & Bacon, Boston.