Visual Representations for Improvement of Music Understanding

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

[1] VISGRAF Laboratory - IMPA
{vitorgr, lvelho}@impa.br
[2] ISR - Universidade de Coimbra
lmvcruz@isr.uc.pt
[3] Universidade Federal do Rio de Janeiro
julianok@impa.br

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 and, ultimately, of the whole composition. Although our approach is pedagogical/educational, the visual representations (of musical elements) we use to enhance the comprehension of a composition could be extended to performing arts context. It could be adopted as motion graphics during live musical performances. We have developed several videos to illustrate our method. They are available at *https://eden.dei.uc.pt/∼vitorgr/VLMU.html*.

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

1 Introduction

Visual representations for music has 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 have the clear premise about the relationship between visual elements and musical elements. This is the main objective of this research. We define visual representations which are clearly associated with musical elements. Such a visual representations improves the perception of the musical elements. It facilitates the understanding of musical phenomena. This pedagogical goal is also 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 their ear to be able to understand and identify the contrivance utilized by the composer.

For this purpose, visual representation can enhance the perception of musical elements occurrence, and thus, they can improve the understanding of some musical phenomena. In this paper, we will present an approach for enhance the perception of a note and its properties like tonal pitch, octave and volume. The proposed visual representation are polygons, and they are related to the notes through an appropriate association of visual properties like color and shape with the note properties.

Our intention is to create a visual language able to represent complex musical phenomena, like rhythm, harmony, melody, etc. However, in this first step, we are proposing the basis of this language by the analysis of the most important element for musical studies: the note. From our visualization approach, we can highlight phenomena related to, besides the note perception (very useful for ear training), the occurrence of chord e the timbre perception. Furthermore, we can observe the phenomenon of the relation of dominance and accompaniment.

The music visualization can be an image (a global representation) or a video (a temporal representation). The local approach presents a big picture of the music. In this image, we will represents global phenomena (i.e., phenomena that happen throughout the music). The local approach is a video with the length equals to the music, and where the visual representations are exhibited synchronized to the respective musical elements). In our case, for each note in the music, we will show a correspondent polygon with the respective color and shape throughout the time the note is played.

This sync of the local music visualization creates a synesthesia between the sonorous stimulus (music) and the visual stimulus (video). From this relation, the video improve the understanding of the music, creating the pedagogical effect which we argue that our representation brings to the musical education.

In Section 2, we show other music visualizations types, and compare them to our approach. In Section 3, we show the association of colors and note tonal pitch. Moreover, we show how this association can be used to observe the dominance and accompaniment phenomena in a music of chamber. Following, in Section 4, we show how to relate the shape of the representation with other musical elements. Finally, in Section 5, we present our conclusions and future works.

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 to use it to represent 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 relate them to musical elements such as rhythm and harmony. Finally, it also highlights the use of movements and forms 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 simpler and easier to learn. It can be used as a pedagogical tool in the study of musical phenomena.

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 sound tones to colour tones mapping is proposed. Miyazaki et al. (2003) propose a method to create a 3D interactive picture of MIDI files. Pitch, volume, and tempo of a note are encoded as height, diameter, and colour 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) present 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 the pitch and time domains. Their main goal were to visualize the music *motif* by analysing 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, they also claim that it is possible to visually detect familiar *motif* patterns with Isochords. Ciuha et al. (2010) present a different colour mapping for visualizing a group of concurrent tones. The authors propose a continuous colour wheel for mapping, different from most sound tones to colour tones mapping, which are commonly discrete. Their emphasis are on modelling harmony, ergo the affinity of tones and consonance with colours. Dissonance is represented with low colour saturation, while consonance is represented with high colour 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. So, the visual language proposed by Chan et al. (2010) is highly complex, and thus, it is not interesting in pedagogical terms. The visual representation proposed in this paper is designed to evince specific characteristics of music, such as: timbre, chord, note tone, and volume. Though, the proposed visual language can also help the user to grasp some information related to: 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. It is also designed to be pedagogically appealing.

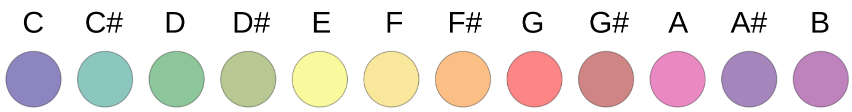


Fig. 1. Sound tones to colour tones according to Louis Bertrand Castel.

3 Colors

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

In this work, we are using an association between the tonal pitch of a note with a specific color in a color palette. The construction of this palette is arbitrary, ie it is possible to use several types of palettes for specific purposes (for example, creating a view with a particular color aesthetics). But note that the colors of this palette should be easily distinguishable.

In our experiments, we used the palette of colors proposed by Louis Bertrand Castel. Such a palette presents a continuous chromatic transition, maintaining sufficient contrast between them. In this way, the colors present a coherent flow (and therefore facilitate learning), as well as, their colors are easily recognized. There are some studies that show how the relation of this palette of colors with 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, that is, an image that represents music from beginning to end. In particular, we will present two very simple visual representations that can be used to perceive the relationship of dominance and accompaniment in a song. We will compare this approach with the temporal process of listening to a song, or look at a table 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 table 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 within a principal theme, which are repeated along the piece by different instruments. Variations on volume, rhythm, and melody are utilized to introduce form and interest to the song. Thus, the instruments may have different roles in a music composition. For example, an instrument may play the role of accompaniment or the role of

	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)

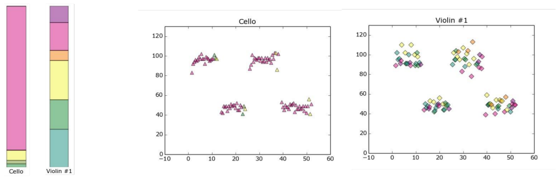


Fig. 2. Perception of Dominance and Accompaniment by Color Analysis.

dominance. When the instrument has the role of accompaniment, its composition part is played in a supporting manner to 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 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 play only a small set of notes (which indicates that they play a accompaniment role) and which play a larger variety of notes (which indicates that they play a dominant role). However, careful reading of several numbers is necessary to realize such a pattern.

Although it is possible to observe the instrument accompaniment and dominance relationship within the table. It is not as intuitive as color-based representations. For example, Figure 2(b) shows two bars with proportion of occurrence of each note played by different instruments (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 is the accompaniment (T4).

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

In this second visual representation it is not easy to notice the proportions of the notes in the bar chart, because in the representation of any song with a reasonable amount of notes we will have an overlap of shapes along the graph. However, we can still have a general perception of whether or not there is a variety of colors (more perception of variety than perception of

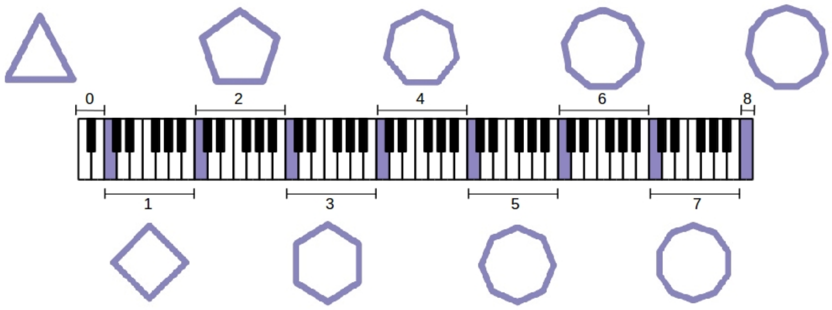


Fig. 3. Visual representation of octaves.

which colors are present in the graph), as well as perceiving the varieties and velocities of notes. In this way, we can also perceive the dominance relationship (many colors of notes played and many velocity) and accompaniments (little variation of colors and velocity).

4 Shapes

In addition to colors, visual representations can also be more easily understood than musical elements in a song. In this section, we will present approaches to define the shape of these objects, and we will present the relation between the shape and musical elements. Thus, the simultaneous occurrence between the visual representation and the 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 on a such element. In addition to the association of tone notes with a specific color, 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 is the velocity, the larger is 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 translated, as illustrated in Figure 4.

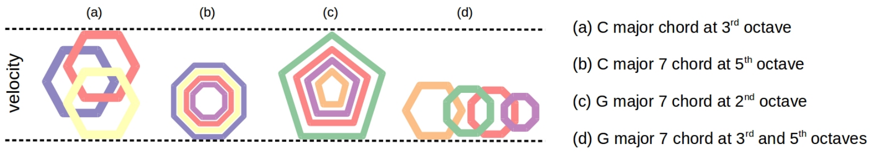


Fig. 4. Chords are represented by nested or shifted polygons.

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 the velocity of a note. The velocity is related to the 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 is characterized by a musical instrument which is playing a certain note. 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 song 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 the shape of the representation. These features are easily observed, and thus, they produce a more efficient synesthetic effect (i.e., they highlight the occurrence of some musical phenomena in a satisfactory way). Moreover, we can also introduce a behavior in the representation and relate it with some musical phenomenon. In this section, we will talk about two simple behaviors and their meaning for music visualization. These behaviors are motion and blurring.

The motion can be a changing of the representation size (a scaling), a rotation of the representation, a changing of the representation position (a translation), etc. 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 representation occurrence.

The other effect that we apply is the blurring. After a note (or chord) is played, it disappears slowly blurred. The following notes are then shown in the

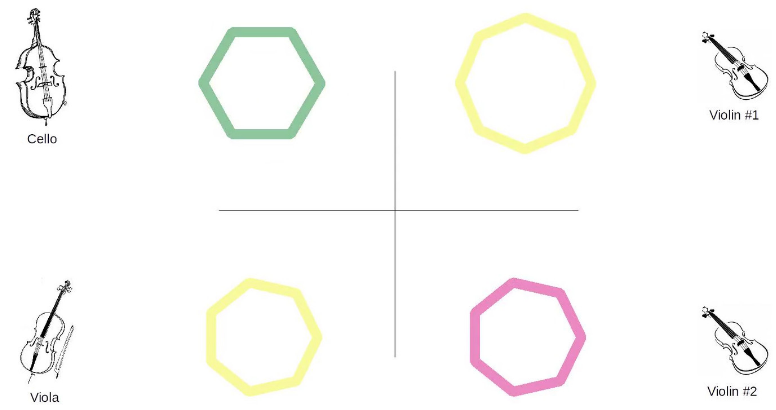


Fig. 5. For timbres of a String quartet.

front. This effect lets us perceive notes that have been played recently, exercising our short-term memory.

Despite of we do not approached rotations and translations, we believe that they can be used for representing some musical phenomena (for example, rhythm or harmony). More complex musical phenomena are beyond of the scope of this paper (we kept the focus in basic phenomena), but it is a future work of this research.

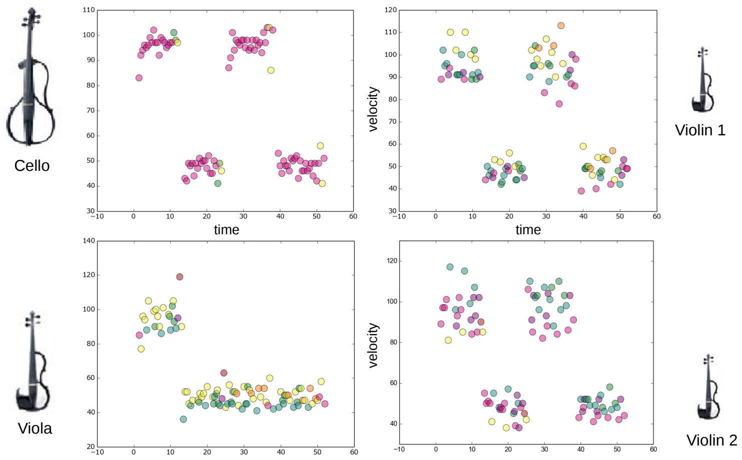
Furthermore, in the film An Optical Poem, it is presented a changing of perspective of the camera (what also can be considered a motion). It introduces a nice visual dynamics. However, as far we have evaluated, this motion does not enhance the music comprehension (it is only an aesthetic feature). Thus, we do not included this behavior in our visual representation proposal.

5 Conclusion and Future Work

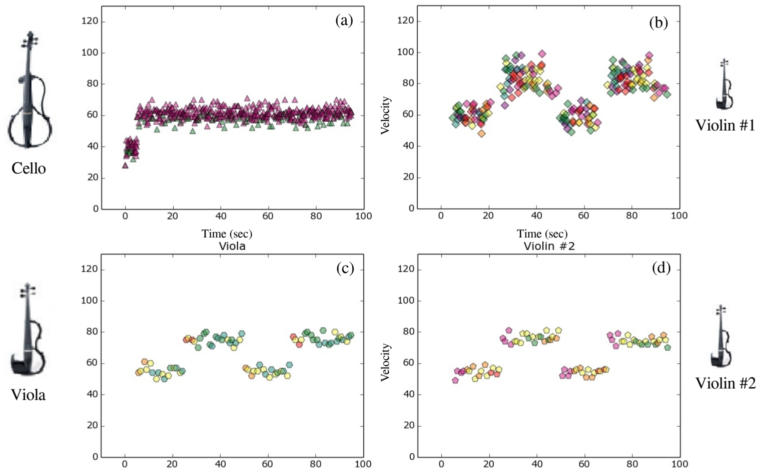
This work demonstrates the potential of using visual representations to improve the understanding of different musical elements. We present a method on how to use colors, shapes, movements, and blurring associated with musical notes and their respective properties. With our visual representations we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

We intend to expand the set of visually represented musical elements. In this way, we can create more complex visual representations of musical phenomena, based on the analysis of rhythm and harmony.

A research direction with great potential is to explore polygon position and movements, and also appearance and disappearance, and blurring effects. Another extension is to work on instruments grouping in a way that would let us represent a whole orchestra.



(a) 1st Movement of *La primavera* (Spring) - *Allegro* - The Four Seasons, Antonio Vivaldi



(b) Ravel - Boléro

Fig. 6. Examples of Color Analysis

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 on 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: VISGRAF Lab at IMPA, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Instituto de Sistemas e Robótica, and finally Instituto Nacional Casa da Moeda (Portugal).

References

Fred Collopy (2000) Color, Form, and Motion: Dimensions of a Musical Art of Light. Journal Leonardo

Ohmi K (2007) Music Visualization in Style and Structure. Journal of Visualization 10-3:257-258. doi:10.1007/BF03181691

Gareth L, (1985) Musicians Make a Standard: The MIDI Phenomenon. Comput. Music J. 9:8-26. doi:10.2307/3679619

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

Brougher K, Strick J, Wiseman A, Zilczer J (2005) Visual Music: Synaesthesia in Art and Music Since 1900. Thames & Hudson, London.

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

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

Ciuha P, Klemenc B, Solina F (2010) Visualization of concurrent tones in music with colours. In Proceedings of the 18th ACM international conference on Multimedia (MM '10). ACM, New York, NY, USA, 1677-1680. doi:10.1145/1873951.1874320

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

Sapp C (2003) Harmonic visualizations of tonal music. In Proceedings of International Computer Music Conference, Havana, Cuba, pages 423–430.

Cohn R (1997) Neo-Riemannian Operations, Parsimonious Trichords, and their Tonnetz Representations. Journal of Music Theory, 41/1 , 1–66.

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

Dannenberg R (1993) A brief survey of music representation issues. Computer Music Journal, Vol. 17, Pages 20-30.

Gardner R (1964) Music notation: a manual of modern practice. Allyn & Bacon, Boston.