Visualisation of Multi-scale Formal Diagrams for Music Analysis

Joséphine Calandra, Jean-Marc Chouvel, Myriam Desainte-Catherine, Erwan Michel

Visualisation of Multi-scale Formal Diagrams for Music Analysis

Joséphine Calandra
Institut de Recherche en Musicologie, IReMus
Sorbonne Université
Paris, France
josephine.calandra@labri.fr

Jean-Marc Chouvel
Institut de Recherche en Musicologie, IReMus
Sorbonne Université
City, Country
jeanmarc.chouvel@free.fr

Myriam Desainte-Catherine
Laboratoire Bordelais de Recherche en Informatique, LaBRI
Université de Bordeaux
Bordeaux, France
myriam@labri.fr

Erwan Michel
Université Toulouse Jean-Jaurès
Toulouse, France
erwan.michel.2023@pm.me

***Abstract*—This paper presents visualizations of a paradigmatic and syntagmatic dynamic multi-layered representation of musical pieces. We implemented a software called *Multiscale Oracle Representations For Organized Sounds (MORFOS)* that automatically builds formal diagrams from musical data, according to a cognitive simulation of the memory behavior while listening. A coherent visualization is a valuable aid for music analysis to represent simultaneously the different levels of the musical structure.**

Three visualizations are presented: the first one is the superposition of two-dimensional formal diagrams for all levels of structure. The second and the third ones superimpose in three dimensions the diagrams and emphasize in different ways the links between the structural levels. In such visualizations, the musical form, organization, and structure can be visualized as three projections on planes sharing the same time axis. Because of their coherence with cognitive procedures, all these visualizations are particularly useful to understand complex musical phenomena better.

***Index Terms*—Music analysis, musicology, computer sciences, multi-scale, visualization, cognitive algorithm, formal diagram**

I. Introduction

We study the multi-disciplinary application of automated music analysis using computer science, in order to better understand the cognitive processes involved in music listening. This article presents specifically a study of the visualization of the data structures produced by our algorithm which are multi-scale representations of the music organization. Indeed, music analysis has always looked for appropriate visualizations. We thus seek to improve the understanding of the musical phenomenon through its visualization, beyond its notation and its representation. We wonder then which representations highlight the musical processes and even more the cognitive implication of these musical processes through their temporal development.

Our algorithm takes as input audio music representation, but it can also take character strings and gives as output a music representation called multi-scale formal diagrams. We distinguish musical information through the notions of *materials* and *objects*. The materials are the paradigms, i.e. the models, the classes of objects, while an object is an occurrence in a musical work of a given material. For example, a material can be the concept of the note A, while the object will be the note A played at a precise moment.

Musical information is found at different temporal scales. The musical objects are situated between the microscopic, of the order of a millisecond, and the macroscopic, corresponding to the part of the work or even the complete musical work. Based on the Cognitive Algorithm described in [7], we propose a bottom-up modeling of data structuring: a given grouping of elements constitutes a higher-level element. For example, a grouping of notes constitutes a musical pattern, and a grouping of musical patterns constitutes a musical phrase. Moreover, we present a paradigmatic and syntagmatic analysis of the musical piece. A paradigmatic analysis is the study of the relation between the elements based on their similarity and a syntagmatic analysis is the study of the relation between elements based on their position in a sequence: therefore, we study the constituent materials of the work and the relationship between these materials. Thus, the notion of form such as we consider it corresponds to the visualization of this analysis, i.e. the relation between the materials of a given temporal scale in time. We distinguish the notion of form from the notion of structure, which highlights the hierarchical organization of the musical work in time.

Therefore, a representation of the different scales of musical information is presented: multi-scale formal diagrams. How can such a structure be visualized efficiently despite the amount of information? Which visualization allows us to represent all the levels and their overlaps efficiently?

Music by definition is temporal data: it makes sense by the order in which the data that constitute it appears. The representation of multi-scale formal diagrams takes into account the order of appearance of the elements. We suggest three axes that deal with three aspects of temporality. The first one is the usual timeline, the second one orders chronologically the musical materials and the third one corresponds to structural levels as durations of those materials. The creation of this visualization from these three axes allowed us to distinguish three planes: the time/material plane corresponding to the form, the time/levels plane corresponding to the structure, and the material/levels plane that underlines the relation between the materials of the different levels of structure: we call it the organization plan. This third plane in particular seems to us to be a rather innovative contribution to the musicological field and its potential has yet to be fully explored. One could also display the construction of the visualization at the same time as the diffusion of the musical work which might help better understand the structure.

In Section 2, we present a state of the art about music visualization: written music visualization, audio music visualization, and formal diagrams. In Section 3 we present the multi-scale diagram and the suggested visualizations: one two-dimensional and two three-dimensional visualizations, one using pavements and the other representing music elements by spheres. We will try to present an overview of the potential of the three different planes described above in the 3D visualization. These visualizations will be then illustrated with two examples extracted from Wolfgang Amadeus Mozart's *K545 Rondo* and Claude Debussy's *Hommage à Rameau* and discussed in Section 4.

II. State of the Art

Although music visualization is a widely expanded research domain, we precise that our work does not focus on the visualization of the audio content of the work to improve the music listening experience, but on the visualization of a musical representation (in this case, the formal diagram) that enables us to highlight the internal construction and temporal development of musical works, to offer a better musicological reading of them.

Furthermore, musical representations and their visualizations allow us not only to better understand the musical phenomenon but also help us to represent and conceive new musical forms. We show handwritten/audio notations and representations that lead to the structural visualization of music, then we present formal diagrams.

A. Handwritten Representations

One of the oldest representations of the musical work is the score. The score is both notation and visualization of music instructions for performers and represents pitch, duration, intensity, and timbre. Nevertheless, this notation indicates the characteristics of the elements which constitute the work but not the relation between these elements. Nicolas Ruwet proposes in [22] a first paradigmatic analysis of the musical work and introduces the notions of linguistics to the world of music. Lerdhal and Jackendhoff propose in their book *A Generative Theory of Tonal Music* [18] a tree-like representation of musical grammar. However, Chouvel says in [8] that music cannot be resolved to linguistics [16][1], and these different representations are constrained to works for which there are written notations such as classical music.

We then observe the first attempts to describe electro-acoustic music with the definition of *musical objects* by P. Schaeffer in [23]. This led to the creation of the *Temporal Semiotic Units* [10] that inspired Jean-Louis Di Santo to propose the *Acousmoscribe* [12], which is a music editor with a set of notations for acousmatic music.

However, it is difficult to extract with manual notations certain characteristics of the signal data and thus to represent and visualize these data.

B. Towards a Representation of Audio Music

The Information Theory presented by Claude Shannon [24] and the spectral analysis carried out with the Fourier transform [5] implicates access to the waveform and the sound spectrum. These analyses are used for new visualizations such as the *Sonogram* which represents the amplitude for each frequency as a function of time. From the *Sonogram*, the GRM (Research Musical Group) has proposed the *Acousmographe* [14] and P. Couprie developed *E-Analysis* [9] that both offer graphical representations of electro-acoustic music analyses.

However, these different representations do not emphasize the structural aspect of audio. Self-similarity matrices [13] overcome this by representing the similarity of the piece of music with itself over time. The Tonnetz of V. Gillot [15] based on the *System and Contrast Model* [3], highlights the structure, and the relations between the constitutive elements of the work are given *a priori*. Moreover, the temporal representation of the music is completely nonlinear, such as the visualization proposed by L. Bigo in [2]. This is not a dynamic visualization over time, but it allows us to easily compare two different excerpts. A. Marmoret proposes a representation of the classes of elements in the musical work through time in [19]. He provides a fixed number of element classes to be found in the work and automatically obtains a representation with neural networks.

These visualizations allow the representation of the unwritten music but also put forward the musical structure. These require fixing upstream the topological space of the work, where the formal diagrams are a visualization that creates the constitutive space of the work as the musical data is acquired.

C. The Formal Diagram

The formal diagram is a musical representation presented by one of the authors which is obtained from the analysis of a musical work represented as a succession of fragments. It is obtained based on the Algorithm 1, presented in [7]. This pseudo-code was firstly not implemented but applied "by hand" on the raw material which is the score or the audio sequence of the musical work. The word *"heared"* is then used metaphorically as the Cognitive Algorithm is a model of

[1]to cite Hanslick among others, he considers music as an autonomous phenomenon, which consists mainly of *"musical shapes in movement"*.

the music auditory process. It corresponds to the acquisition of the fragment, which might in some cases be actually made by hearing it.

Algorithm 1 Calculation of a formal diagram

for each fragment heard:
comparison with previous fragments;
if different:
    inscription in material memory;
in all cases: write in form memory;

This algorithm is one step of the Cognitive Algorithm [7] which is the *Paradigmatic Recognition Test* (similarity test). In the Cognitive Algorithm, a second test follows this one: the *Syntagmatic Recognition Test* (segmentation test), which determines if the group of fragments that have just been heard constitute or not a fragment of the upper level. If this is the case, the two tests are iterated again at the upper level, an upper level being a level in which fragments are longer, as a concatenation of fragments of the lower level. Thus, each level corresponds to the production of a different formal diagram. Let us specify again that this algorithm is a non-automatised algorithm, used as a process for musicologists. The automation of this algorithm is then proposed with the Multiscale Oracle Representations For Organized Sounds (MORFOS) by the authors, introduced in [6] and briefly reminded in section III.

The formal diagram allows to structure and carries out a paradigmatic analysis in the continuity of the linguistic theories of Ferdinand de Saussure [11]. This is inherited from the use of semiotics for music, introduced by J.J Nattiez [21]. We wish to automate this modeling from audio files and like the *Morphoscope* [20] to obtain the representation dependently on the temporal development of the work. It is also inspired by the cognitive musicology of O. Laske [17] which presents music as a reflection of the cognitive system, thus modeling cognition.

In formal diagrams, the x-axis corresponds to time and the y-axis to the *material memory*. *Material memory* is the dictionary of the *materials* constituting the musical work under analysis.

Each fragment is analyzed one by one. When a fragment is acquired, it is compared with all previously occured fragments, and classified accordingly. Classification involves creating a new line in the formal diagram if the class does not exist, or writing the object in the line corresponding to the appropriate material (in the material memory). In all cases, the new object is registered at the right time (in the form memory).

Let us take the character string $abacabac$. The successive formal diagrams obtained are shown in Fig. 1. Initially, the material memory is empty and so is the formal diagram (1). The algorithm acquires the first fragment a: this is the first fragment and therefore necessarily different from those which occurred previously, so a new line corresponding to material a is created in the formal diagram and the object labeled a is written at time 1 (2). the algorithm acquires then fragment b. This is different from a, so a new line corresponding to material b is also created, and the object labeled b is entered at time 2 (3). Next, we acquire object a again. This is similar to the first a, so the algorithm enter it in the same line at time 3 (4). This process continues until all the information has been acquired (5).

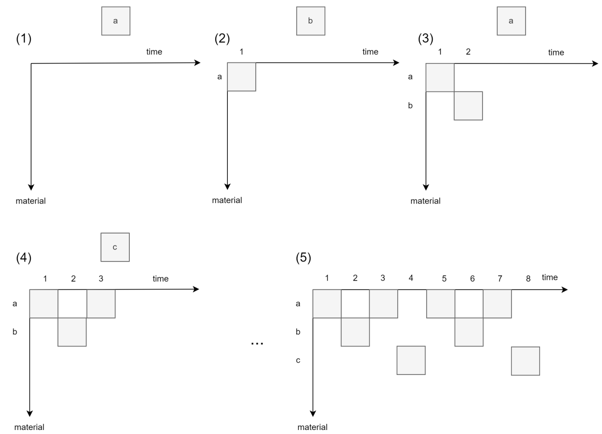

Fig. 1. Successive formal diagrams obtained while acquiring the character string *abacabac*.

Formal diagrams have been until then visualized in two dimensions. In the following section, we present MORFOS, that produces multi-scale formal diagrams.

III. The Multiscale Oracle Representations For Organized Sounds (MORFOS)

Before presenting our proposed visualizations of the multi-scale formal diagram, we present briefly MORFOS. The precise modelization is not the subject of this article and is not presented here. You may refer to [6] for more information.

We have shown how to obtain a single formal diagram, but the human brain is capable of structuring in time on different time scales: when we listen to a musical work, we can propose a segmentation both at the scale of the motif and at the scale of the musical part. Therefore, the materials can have different durations according to the scale at which one is situated (for example note, musical motif, musical phrase). We represent and construct simultaneously the formal diagrams corresponding to different temporal scales, which gives us a set of multi-scale formal diagrams.

MORFOS takes an audio, vector, or symbolic representation of the musical work and provides a multi-scale formal diagram.

It pre-processes the initial information to compute a representation of the musical work as a sequence of objects. This sequence of objects is then provided to the *automated Cognitive Algorithm*, which returns a multi-scale formal diagram. The automated Cognitive Algorithm is a recursive call of the *Cognitive Algorithm of level* n presented below.

The Cognitive Algorithm of level n takes as input a sequence of objects of a given level and provides a sequence of higher-level objects. Each object is processed successively in the Cognitive Algorithm of level n and undergoes two operations: a *segmentation operation* (syntagmatic recognition test), and a *classification operation* (paradigmatic recognition test). Classification and segmentation operations are defined by a set of rules implemented in MORFOS. The user selects a subset of the implemented rules, which are called up for each segmentation and classification test. The two tests of the cognitive algorithm of level n are shown in Fig. 2:

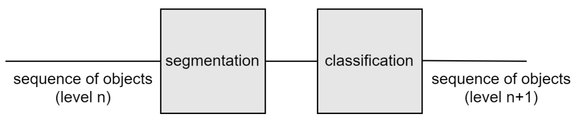

Fig. 2. The Cognitive Algorithm of level n.

Let's take the previous example $abacabac$. We consider the segmentation criterion in this example is the return to the beginning of a segmented word or a word under acquisition, and the classification criterion is strict equality between two words. So, as the example starts with the letter a, we segment each time a new a is acquired. When the third a is acquired, ab in aba is segmented as a higher-level object, and classified as a new material A. Then, when the fifth a is acquired, ac in $abaca$ is segmented and it is different from ab, so it is also classified as a new material B. When all the level 0 objects are acquired, the result is the level 1 representation $ABAB$, with $A = ab$ and $B = ac$. Then, we can represent the diagram induced from the character string $ABAB$ accordingly to the Algorithm 1.

MORFOS produces then a sequential representation of objects at all levels and each of the corresponding formal diagrams. This terminates when a single material is obtained in the level (the level containing a single material is then excluded as irrelevant) or when two successive levels are identical.

In the example provided, we obtain a third level $\mathcal{A}\mathcal{A}$ with $\mathcal{A} = AB$. The formal multiscale diagram made up of the 3 formal diagrams for the three different levels is shown in Fig. 3.

MORFOS is real-time as it builds the representation meanwhile the signal is acquired in the System. Although the dynamic aspect of the visualization of object segmentation and classification is not specifically discussed in this article, it is a mechanism that could help to understand cognitive mechanisms during music listening.

IV. The Multi-scale Formal Diagram and its Visualizations

Three distinct visualizations are presented to view and study the formal diagrams. The first one, used until now, is in two dimensions and allows a simple readability of the diagrams of each level. The two other representations in three dimensions emphasize the links between the objects of the different levels. MORFOS is fully automated as well as the 2D visualization, however, the automation of the two 3D visualizations still needs to be improved and we show handmade visualization based on the results of the Cognitive Algorithm.

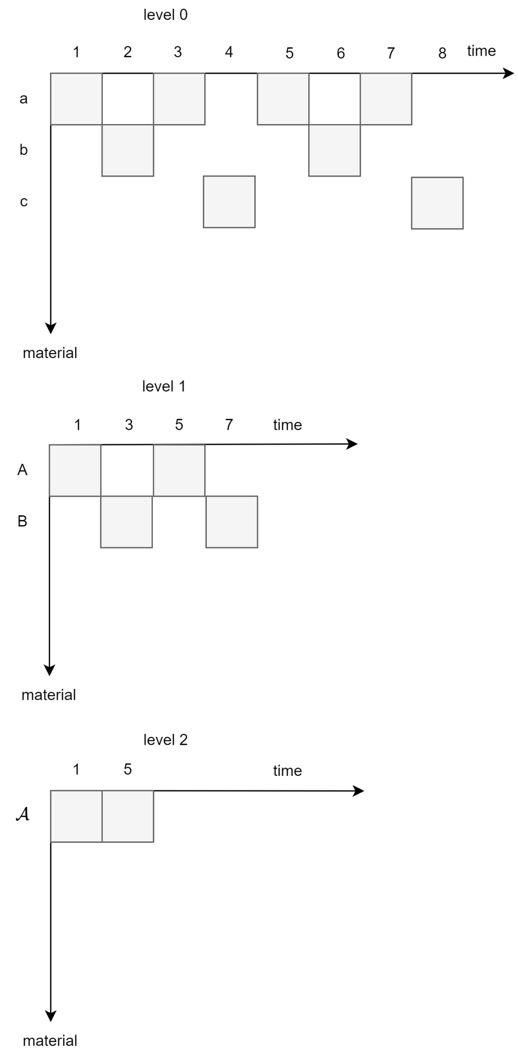

Fig. 3. Multi-scale diagram of the character string $abacabac$ given the segmentation criterion as the return to the beginning of a segmented word or a word under acquisition, and the classification criterion as strict equality between two words.

A. 2D Visualization

The first visualization presented by an example of the *Rondo K.545* of Wolfgang Amadeus Mozart in Fig. 4 is the two-dimensional representation, computed automatically. The whole process to obtain these diagrams is explained and compared with the score in [6]. The formal diagrams are superimposed and aligned along the temporal axis. Thus, the objects aligned on the same temporal segment belong to the same sound reality. The objects of smaller temporality constitute those of larger temporality on the same segment. For all levels, the first few milliseconds of each object are shown in black to distinguish two successive objects.

Each diagram has its own material axis, which corresponds to the materials of the same musical work but for a given time duration. The representation of the material compresses the information as we condense the recurrent information on a

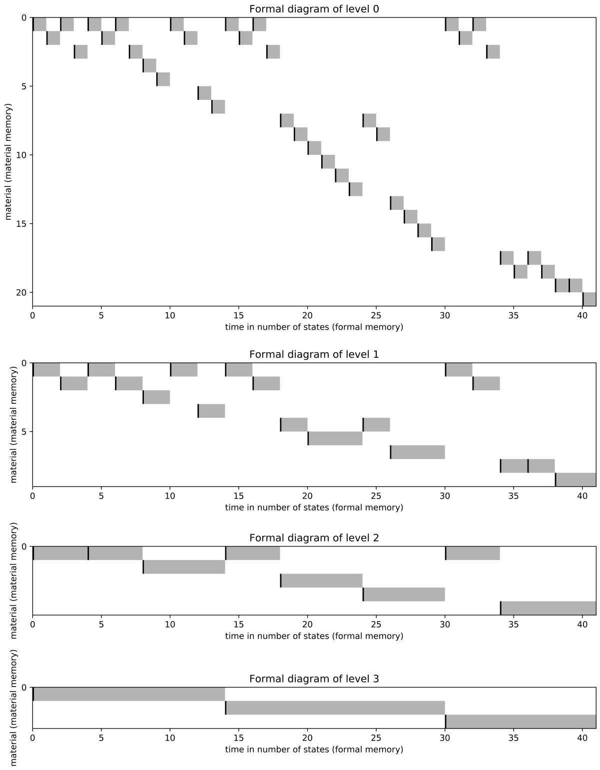

Fig. 4. 2D visualization of the multi-scale formal diagram automatically obtained from W.A. Mozart's *Rondo K545*: two objects are considered as the same material if the corresponding character strings are strictly similar, and segmentation rules are specified in [6].

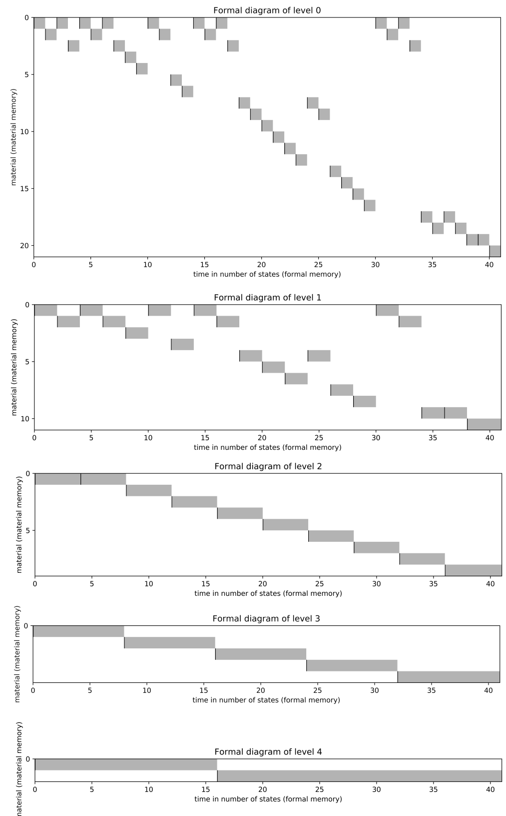

Fig. 5. 2D visualization of the multi-scale formal diagram automatically obtained from W.A. Mozart's *Rondo K545*: two objects are considered as the same material if the corresponding character strings are strictly similar, and segmentation rules are that every two objects are segmented by default.

single line and all the space necessary to represent the musical work is used and increased when new material appears.

Inspired by the *Gestalt Theory*, the analyst can find a correspondence between the musical patterns and the visual patterns and then find reproductions of forms independently of the material's contents. We also note the line that connects the last discovered materials, and that, according to its slopes, gives information on the quantity of information discovered during the piece.

Figure 5 presents a 2D visualization of another segmentation automatically obtained from W.A. Mozart's *Rondo K545* with another set of rules for similarity and segmentation computation: The similarity of two objects is validated is the corresponding character strings are strictly similar, and the segmentation test is validated by default every two objects, except if the last object is alone.

In spite of the alignment of the diagrams at the level of the temporal axis, this representation does not always make links easily between the materials of higher levels and those of lower levels. Three-dimensional representations overcome this.

B. 3D Block-Wise Visualization

The first proposal of visualization in three dimensions comes from the idea to superimpose the diagrams of the 2D representation. We represent the objects with blocks. The length of the block corresponds to the duration of the object, the depth is constant and the height is proportional to the level of the diagram. In this visualization, the position of the block in the z-axis is indexed by the level and not the duration of the objects.

Fig. 6 shows a representation of a multi-scale formal diagram for W.A. Mozart *Rondo K.545*, given the same rules as for figure 4. This representation is built manually with *Blender* [4].

In order to see simultaneously all the levels of the diagram, the lowest levels are in front of the diagram, so that the objects of lower levels that constitute the objects of higher levels are physically included in the larger blocks. In addition, a different color has been attributed to each level of structure, and the object of each level is represented with adequate color. However, we notice that the height of the diagrams induces a shift in height between the objects of a given level and the equivalent objects of lower levels, which makes them no longer aligned. We then propose another three-dimensional representation that aligns the object's height at the different levels on the material axis.

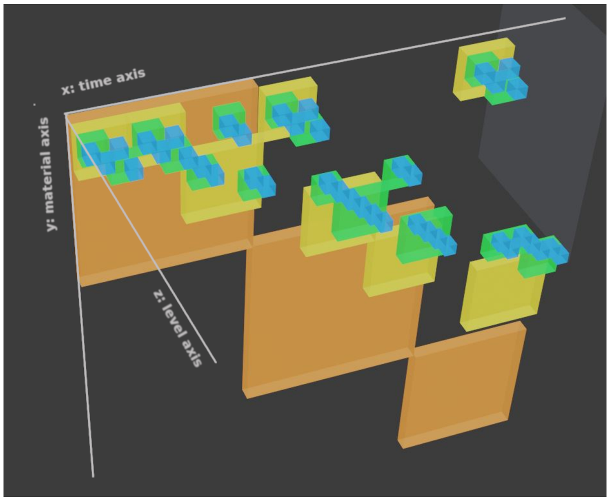

Fig. 6. Block-wise 3D visualization of an extract of W.A. Mozart's *Rondo K545*'s multi-scale formal diagram.

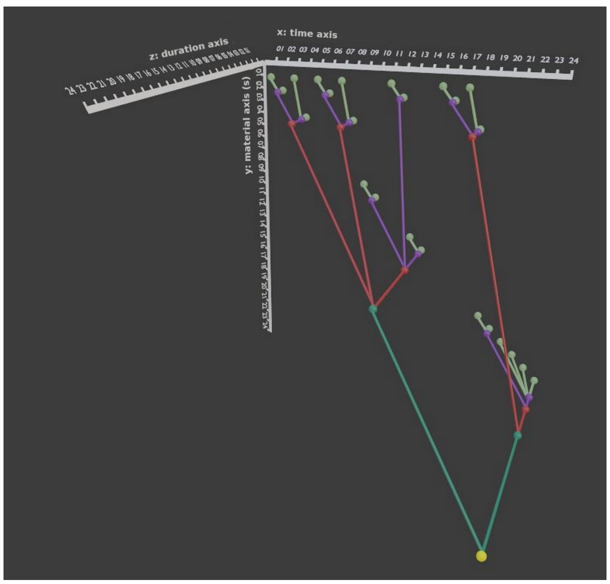

Fig. 7. Sphere-wise 3D visualization of an extract of W.A. Mozart's *Rondo K545*'s multi-scale formal diagram.

C. 3D Sphere-Wise Visualization

For this second 3D representation, we represent the objects by spheres of unique diameter whatever the level of the diagram and the duration of the object. Their center is at the instant t of the segmentation of the object in the time axis. This avoids the confusion that existed in the previous representations between two adjacent objects, and there is no longer a problem of height for materials that appear at the same time at different levels. Fig. 7 is a 3D representation with spheres built manually with *Blender* of the beginning of the *Rondo* shown previously.

Each axis is a representation of time: the first axis corresponds to the temporal course of the work (time axis), the second axis corresponds to the timestamp of the first appearance of the materials (material axis) and the third axis corresponds to the duration of the objects (duration axis). Displaying the timestamp of appearance of the material and not the label corresponding to the order of appearance of the material makes it a less compressed representation, it would be necessary to remove the useless temporalities to recompress this temporal axis.

We also propose to represent the filiation links between the objects of two different levels by physical links connecting the two objects.

As the objects are classified in the model by level, we still distinguish the levels by representing them with different colors. However, the objects for the same level are no longer aligned on the axis of the duration of the objects but have a certain flexibility. This ensures continuity between the different levels.

This superposition of objects allows a reading of the diagrams in three dimensions on three different planes as shown in Fig. 8: the time/material plane (form), the material/duration plane (organization), and the time/duration plane (structure).

The formal plane makes it possible to study the development of the musical language through time and to deduce from the use of the materials the effects produced while listening. In Mozart's example, we note the recurrent return of a refrain (materials first appearing at moments 01, 02, and 04), alternated with developments that we call discovery phases. This allows the listener to acquire an increasing amount of information and to avoid boredom, while regularly returning to known material might help to not be lost under a flood of new information.

The structural plane helps to determine the rhythmic organization between the different temporal scales. In Mozart's piece, we note an almost systematic subdivision of the objects of higher level (higher duration) into two objects of lower level (lower duration). We are thus on a very conventional binary structure, with sometimes sub-divisions in three or four elements allowing to bring a little variation and a surprise effect.

The organizational plane describes the dictionary of the language used within the musical piece and how it is articulated between the different levels of structure. It might be interpreted as a footprint of the musical piece. In Mozart's piece, three materials are used several times to constitute higher-level materials but the organization of the materials is globally serial, with materials appearing successively for the first time constituting a new higher-level material.

The study of these different planes might help to compare the visualizations between different works, different styles, or different composers. We could for example ask ourselves if the form can be recursive at the different temporal scales of the same musical work or if the structure is characteristic of a genre or an era. To go further, an analyst could observe how composers play with the codes of form and structure to make the musical work more refined. Concerning the organizational plane, one could wonder if this plane is similar according to

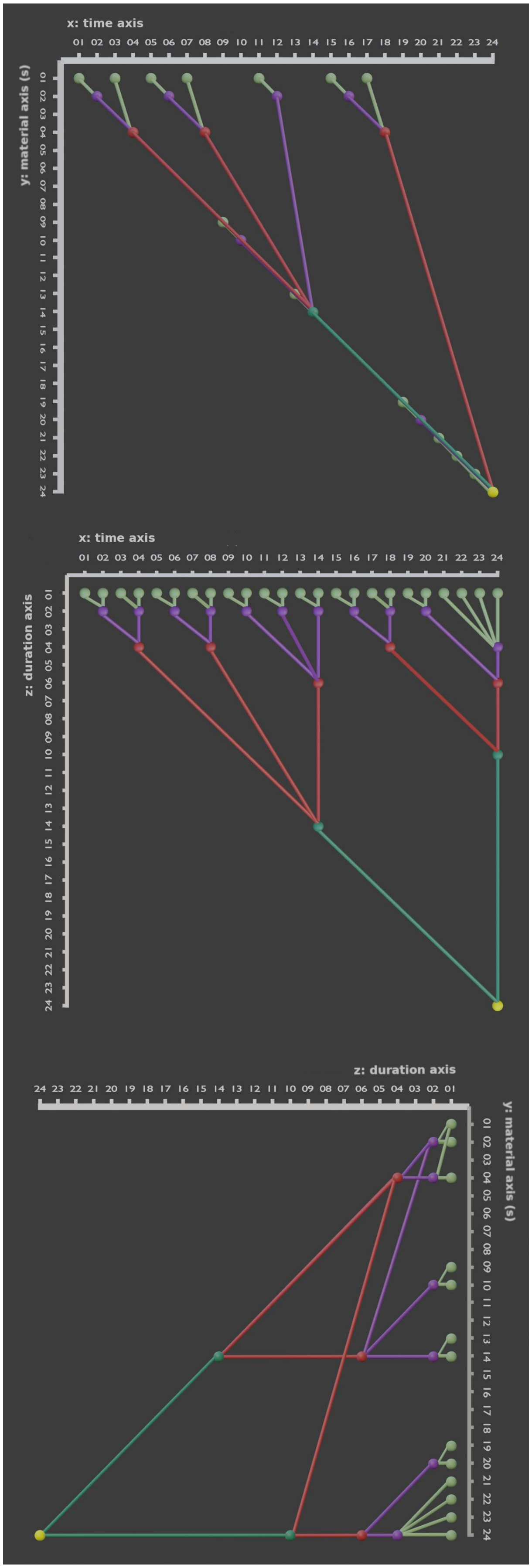

Fig. 8. Sphere-wise 3D visualization of an extract of W.A. Mozart's *Rondo K545*'s multi-scale formal diagram. From top to bottom: formal plane, structural plane, and organizational plane.

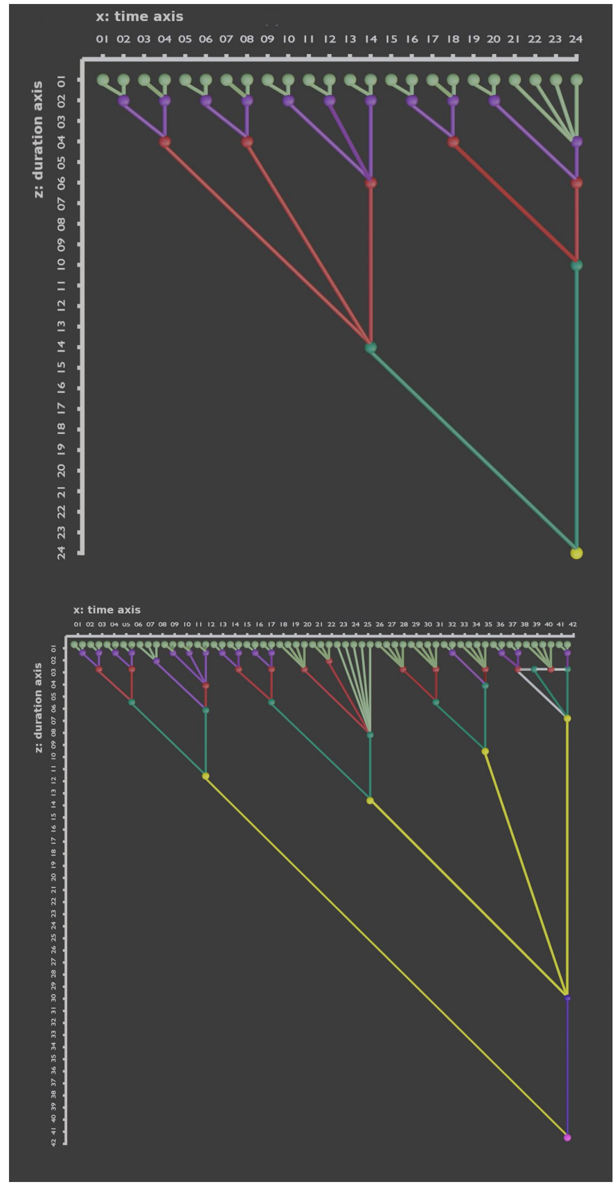

Fig. 9. Sphere-wise 3D visualization of an extract of W.A. Mozart's *Rondo K545* (top) and Debussy's *Hommage à Rameau* (bottom) multi-scale formal diagram: z/x view.

the composers or according to the characters of an Opera in different *Aria*.

To illustrate that, we present a comparative study of Mozart's piece with Claude Debussy's *Hommage à Rameau*, from a structural point of view.

Fig. 9 shows a time/duration (z/x) plane of Mozart's *Rondo K545* and Claude Debussy *Hommage à Rameau*'. As a reminder, the colors of the spheres correspond to a given level. The light green color corresponds to level 0, the smallest time scale, the purple corresponds to level 1, the red to level 2, the dark green to level 3, and so on. In Mozart's work, the spheres of the same color are globally aligned on the duration axis and the levels follow one another without ambiguity along this axis. As opposed to Mozart's z/x visualization, Debussy's z/x visualization shows an ambiguity between the levels of structures for some objects. Indeed, we notice that there is an ambiguity in the succession of levels for the Debussy piece. Indeed, for example, level 1 (in red) does not exist during moments 22 to 25: this might create a floating sensation during the listening of the musical work. Moreover, the objects of level 3 have a very variable duration, going from 10 to 20 objects of level 0. The last yellow object (instants 35 to 42) is considered as being part of level 3 whereas it is shorter than the fourth object (instants 17 to 25) of level 2. This maintains a rhythmic ambiguity in the work.

V. Discussion and Perspectives

Each visualization has its advantages from a user's point of view. We showed the 3D visualization to musicologists who are used to the cognitive algorithm theory and formal diagrams. Their feedbacks were that the 2D visualization is closer to what they are used to seeing and working with. As the theory of formal diagrams itself requires a significant effort of understanding, this visualization avoids an additional layer of understanding. On the other hand, the 3D representations perhaps offer a finer reading of the superposition of objects at different temporal scales. They also present new perspectives due to the rearrangement of the information between the level of structure, leading to the visualization of three distinct planes: form, structure, and organization. Except for the fact that the 3 separated plans are fully correlated (as a modification in one plane changes at least one of the two other planes) and that the 3D visualization makes it easier to move between them, the full potential of the 3D representation still has to be explored by the use of musicologists.

A. Visualization Improvements

The question of representing polyphony has yet to be studied. On 2D diagrams, it is possible to manually represent two materials present at the same temporal instant, but this is not managed in the actual automated modeling of formal diagrams.

This might be linked to the representation of the degree of similarity between materials. Nevertheless, even if the degree of similarity between materials is calculated in the model, it is not represented in the visualization. In the 2D representation, one could consider playing with the gray levels: at a given time t, each material would be darker if the current material is closer to the according material.

B. Interface and Interactions

An interface is being deployed as a cross-platform desktop application, although it does not yet take into account all possible parameters of the analysis proposed by the implemented algorithm. This interface is developed with the JavaScript framework *Electron*[2] and interacts with the algorithm implemented in Python.

A number of dynamic aspects and interactions are possible with this interface. First of all, as the hierarchical formal diagrams are produced as the material is acquired, the diagrams can be displayed in real-time and the associated audio work is simultaneously played. The objective is to allow a better understanding of the segmentation during the listening of the work, and of the attention mechanisms while listening to music.

To fully explore the 3D visualization and its three different planes it is also possible to move around in the representation and to display the 2D views which correspond to the form, the structure, and the relation between the objects of different levels.

In addition, the user can choose to hide certain elements of the visualization to make it less cumbersome such as links or ranges of diagrams. We can also highlight one or more links or objects by selecting them, which displays a number of statistics such as the element's coordinates. To help the musicologist have a better understanding of the diagram, an audio file corresponding to the selected object can also be played.

C. Validation

It remains to validate the visualization and the interface. A few propositions for future work can be suggested: we could present the visualizations and the interface to get feedback from the users with questionnaires and interviews. In addition, we would have to do user tests to evaluate the usability of the interface. For that, we could use for example the DEEP (as Design-Oriented Evaluation of Perceived Usability) questionnaire. The user would perform a task on the interface and this would evaluate the content, structure, information architecture, navigation, cognitive effort, consistency, and visual guidance. We could then adjust them accordingly. Finally, we could perform an inspection that follows the precise ergonomic criteria of Bastien and Scapin [1]. We could then use these results to make new versions of the 3D representation and the interface, improving the experience that users will have.

VI. Conclusion

Three visualizations of the hierarchical formal diagrams are presented. The first visualization in two dimensions represents the musical form of the work in as many diagrams as there are temporal levels. The diagrams are one on top of the other, aligned along the time axis. The second visualization, in 3D, consists in representing the objects by blocks and in superimposing all the diagrams on a third axis, the duration axis. The third visualization, in 3D, represents the objects as spheres of unique size whose center has as coordinates the timestamp of segmentation of the object, the timestamp of the first appearance of the object, and the duration of the object. Links between the objects then materialize the nesting between the objects of the different levels. This 3D visualization highlights the 3 planes of form, structure, and organization which allow us to bring a new look to the analysis of musical works. This last plane is what seems to us to be the most consequent contribution and would deserve further studies.

[2]https://www.electronjs.org/

The model is developed in such a way to manage attention phenomenon and in particular, hypotheses made about what can follow what has just been segmented. A proposal we make for future works is a dynamic display of the hypotheses made, and then to see when they are realized. Thus the spheres are currently arbitrarily associated at the moment of the segmentation, but they could be represented at the moment of recognition with a higher-level object, with the indication of the surprise effect if there is no confirmation of the hypotheses. Finally, we have proposed in this article some manually obtained notations, and we still have to exploit the automated visualizations that could be obtained with the interface presented above.

REFERENCES

[1] Bach, C., Scapin, D. L. Critères Ergonomiques pour les Interactions Homme-Environnements Virtuels : définitions, justifications et exemples. [Ergonomic Criteria for Human-Virtual Environment Interactions: definitions, justifications and examples: Research report] RR-5531, INRIA. 2005, pp.47. ffinria-00070476v2f

[2] Bigo, L. & Andreatta, M. Filtration of Pitch-Class Sets Complexes. *7th International Conference, MCM 2019*. pp. 213-226

[3] Bimbot, F., Deruty, E., Sargent, G. & Vincent, E. System & contrast: a polymorphous model of the inner organization of structural segments within music pieces. *Music Perception, University Of California Press*. **33** pp. 631-661 (2016)

[4] Blender Community - a 3D modelling and rendering package. (Blender Foundation,2018), http://www.blender.org

[5] Bracewell, R. The Fourier transform and its applications. (McGraw-Hill New York,1965)

[6] Calandra J., Chouvel J.M., Desainte-Catherine M., Multi-Scale Oracle and Automated Representation of Formal Diagrams Based on the Cognitive Algorithm. *Int. Conf. On Technologies For Music Notation And Representation - TENOR 2021*. (2021)

[7] Chouvel J.M., Musical form, from a model of hearing to an analytic procedure. *Interface*. **22** pp. 99-117 (1993)

[8] Chouvel, J.M. *Analyse Musicale, sémiologie et cognition des formes temporelles*, L'Harmattan, Paris, 2006.

[9] Couprie, P. EAnalysis: Developing a Sound Based Music Analytical Tool. *S. Emmerson, L. Landy (éd.), Expanding The Horizon Of Electroacoustic Music Analysis, Cambridge, Cambridge University Press,*. **27** pp. 170-194 (2016)

[10] Delalande, F., Formosa, M., Frémiot, M., Gobin, P., Malbosc, P., Mandelbrojt, J., & Pedler, E. Les Unités sémiotiques temporelles : éléments nouveaux d'analyse musicale [Temporal semiotic units: New elements for musical analysis; book-CD]. *Marseille: MIM/Documents Musurgia,* (1996)

[11] De Saussure, F. Cours de linguistique générale. (Payot,1906)

[12] Di Santo, J. A SIGN TO WRITE ACOUSMATIC SCORES. *Int. Conf. On Technologies For Music Notation And Representation - TENOR 2015*. (2015)

[13] Foote, J. Visualizing music and audio using self-similarity. *Proc. Of ACM Multimedia*. pp. 77–80 (1999)

[14] Geslin, Y. & Lefevre, A. Sound and musical representation: the Acousmographe software. *" In Proceedings Of The 2004 International Computer Music Conference (ICMC 2004)*. (2004)

[15] Gillot, V. & Bimbot, F. Polytopic reconfiguration: a graph-based scheme for the multiscale transformation of music segments and its perceptual assessment. *SMC 2019-16th Sound & Music Computing Conference*. pp. 1-8 (2019)

[16] Hanslick, Edward, Vom Musikalisck-Schönen (1854) ; English trad., The beautiful in music, The Library of Liberal Art, London, 1891

[17] Laske, O. Music, Memory and Thought, explorations in Cognitive Musicology. *University Microfilms International, Ann Arbor (MI)*. (1977)

[18] Lerdhal, F. & Jackendoff, R. A generative theory of tonal music. (Cambridge, Mass.: MIT Press,1983)

[19] Marmoret, A., Cohen, J., Bertin, N. & Bimbot, F. Uncovering Audio Patterns in Music with Nonnegative Tucker Decomposition for Structural Segmentation. *ISMIR 2020 - 21st International Society For Music Information Retrieval*. **33** pp. 1-7

[20] Mesnage, M. Morphoscope, a computer system for music analysis. *Interface*. **22, Issue 2** pp. 119-131 (1993)

[21] Nattiez, J. Musicologie générale et sémiologie. (Christian Bourgois,1987)

[22] Ruwet, N. Méthodes d'analyse en musicologie. *Revue Belge De Musicologie / Belgisch Tijdschrift Voor Muziekwetenschap*. **Vol. 20, No. 1/4** pp. 65-90 (1966)

[23] Schaeffer, P. Traité des objets musicaux : essai interdisciplines. (Éditions du seuil,1966)

[24] Shannon, C. A Mathematical Theory of Communication. *The Bell System Technical Journal*. **27** pp. 379-423 (1948)