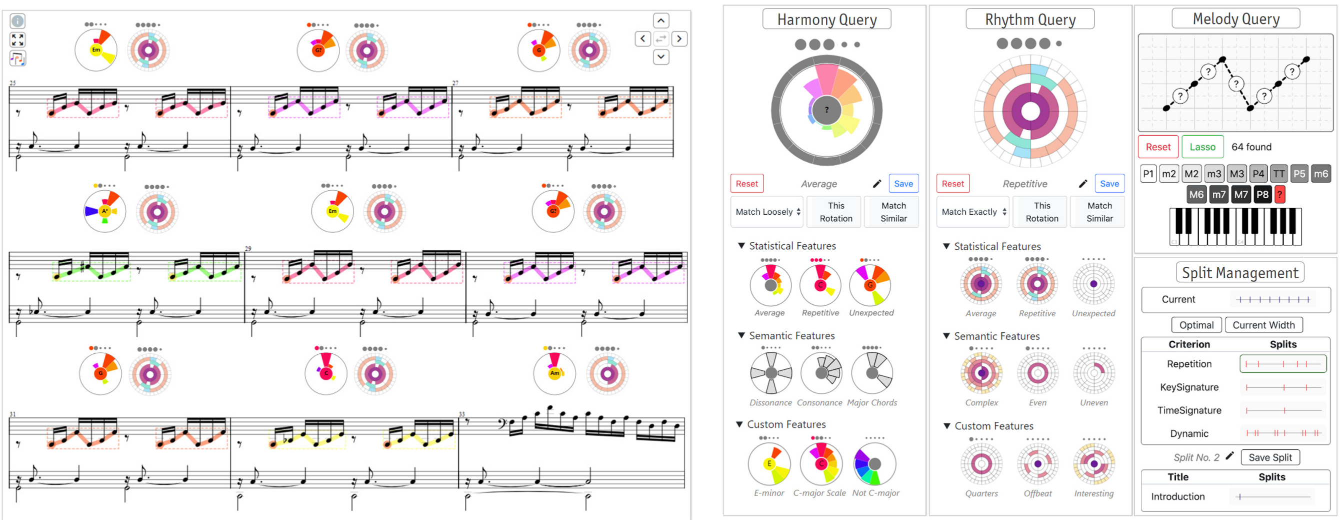
Augmenting Digital Sheet Music through Visual Analytics

Matthias Miller[1], Daniel Fürst[1], Hanna Hauptmann[1,2], Daniel A. Keim[1], and Mennatallah El-Assady[1,3]

[1]University of Konstanz, Germany [2]Utrecht University, The Netherlands [3]ETH AI Center, Zürich, Switzerland



(a) The **Composition View** enriches digital sheet music by juxtaposing harmonic and rhythmic fingerprints as inline glyphs. Colored melodic highlights emphasize repetitive occurrences of a musical *theme* as an overlay on common music notation.

(b) Flexible **Visual Feature Queries** enable analysts to explore harmonic, rhythmic, and melodic patterns. The split management facilitates the identification of a composition's structure based on semantic split criteria.

Figure 1: *MusicVis supports composition analysis by augmenting sheet music with feature glyphs (Figure 1(a)). These can be interactively explored through (Figure 1(b)) visual harmony, rhythm, and melody queries. Applying semantic split criteria facilitates structure analysis.*

Abstract

Music analysis tasks, such as structure identification and modulation detection, are tedious when performed manually due to the complexity of the common music notation (CMN). Fully automated analysis instead misses human intuition about relevance. Existing approaches use abstract data-driven visualizations to assist music analysis but lack a suitable connection to the CMN. Therefore, music analysts often prefer to remain in their familiar context. Our approach enhances the traditional analysis workflow by complementing CMN with interactive visualization entities as minimally intrusive augmentations. Gradual step-wise transitions empower analysts to retrace and comprehend the relationship between the CMN and abstract data representations. We leverage glyph-based visualizations for harmony, rhythm, and melody to demonstrate our technique's applicability. Design-driven visual query filters enable analysts to investigate statistical and semantic patterns on various abstraction levels. We conducted pair analytics sessions with 16 participants of different proficiency levels to gather qualitative feedback about the intuitiveness, traceability, and understandability of our approach. The results show that MusicVis supports music analysts in getting new insights about feature characteristics while increasing their engagement and willingness to explore.

CCS Concepts

• ***Human-centered computing** → **Visualization techniques; Visual analytics; Visualization systems and tools;** • **Information systems** → **Content analysis and feature selection; Document structure;***

1. Introduction

Music analysts and musicologists extract meaningful relationships from musical compositions regarding, for instance, style, epoch, or composers [KKM*20]. Besides meta-information, other features such as rhythm, dynamics, and harmony are crucial characteristics when comparing musical pieces [MSK*19]. Music analysts typically perform manual analysis of sheet music based on common music notation (CMN) that requires proficiency and years of training to understand underlying feature relationships. Such manual analysis tasks are segmentation, structure analysis [LJ83b], annotation and comparison of sheet music. Without the help of additional tools, these tasks are time-consuming and tedious due to the visual complexity of CMN [MHK*18]. Close reading is restricted to CMN which does not readily convey an overview of larger sections.

Multiple computational and visualization-based approaches have been proposed to facilitate music analysis tasks [Tym12, Wat02]. Integrating the analyst into the analysis process to support knowledge generation and verification [SSS*14] is necessary since exclusively computational approaches [Urb17] cannot model human intuition about semantic meaning. In the visualization domain, graph-based [HMM00], matrix-based [KS10], and projection-based [LS19] techniques have been proposed to facilitate the visual analysis at different levels of abstraction. Such data-driven, abstract visualizations are meant to assist the music analysis process but lack a suitable connection to the CMN. Consequently, analysts adhere to the familiar context of musical scores applying traditional analysis methods without benefiting from abstract visualizations.

We argue that seamless transitions between the familiar and abstract representations of sheet music increase the accessibility of analysis-centered music visualizations. This paper introduces a visual analytics technique for bridging CMN with glyph visualizations in a minimally intrusive manner. By addressing this gap, music analysts can benefit from both the CMN and abstract visualizations. We demonstrate the technique's capabilities by presenting an instantiation, called *MusicVis* (*visual-musicology.com/sheetmusicvis*). *MusicVis* uses existing visual designs such as the harmonic and rhythmic fingerprint from our previous work [MBEA19, FMK*20] to display separate musical aspects from simultaneously encoded information. We follow these visual designs to provide visual querying functionality, supporting tailored pattern detection based on the analysts' understanding and domain knowledge. Since the technique is independent of the specific visual encoding, these designs are interchangeable. Thus, *MusicVis* enables music analysts to perform close reading of the presented data, analogous to slow analytics [BSC19]. Our technique further enables distant reading, including corpus analysis, maintaining an overview of the abstract level following Shneiderman's "Visual Information Seeking Mantra" [Shn96].

Contributions – This paper contains the following contributions in the domain of sheet music analysis: (1) A concise description of the problem and data characteristics of sheet music. (2) A visual analytics technique enabling close and distant reading of sheet music at multiple abstraction levels that are tightly connected through staged animations [HR07]. (3) *An instantiation* of the technique (*MusicVis*) based on existing designs rooted in domain knowledge including interactive visual query for harmony, rhythm, and melody (see Figure 1). (4) An evaluation using a pair analytics study with 16 participants of different user groups to demonstrate its broad applicability.

2. Background and Related Work

Music analysis is a broad field dealing with various aspects such as audio recordings or sheet music comprising several heterogeneous data formats. Analyzing musical aspects requires suitable approaches to cope with existing challenges, such as the complexity of CMN. The musicological context of this paper is *Theory & Analysis* and *Education* [MSK*19]. Our work focuses on the analytical tasks *Information Retrieval*, *Segmentation*, *Structure Analysis*, and *Contextual Analysis* of the *Visual Musicology graph* [MSK*19]. Primarily, we are considering musical features such as structure, harmony, rhythm, and melody, which Orlovaitt describes as middle-term features [Orl13]. Donnini et al. describe substantial differences and emphasize the separate consideration of symbolic music and its interpreted version [Don86]. Similarly, we argue that sheet music deserves more attention, especially from the perspective of visual analytics [SSS*14]. Fully-automated approaches produce high-quality results but lack the opportunity for music analysts to step into the process, preventing them from influencing the output [DG08].

2.1. Visualization Techniques for Sheet Music Analysis

We discuss related work that addresses the issue of analyzing musical features based on sheet music. We focus on the compound musical features *rhythm*, *harmony*, and *melody*. Besides, we give an overview of existing approaches that enable structure analysis based on these features. We deliberately omit research concerning audio data, which we consider to be outside the scope of this work.

Harmony – *Harmony* is a central feature of music analysis which builds on top of the primitive feature *pitch* [Ler04]. Existing abstract visualization approaches, such as *Isochords* from Bergstrom et al. [BKH07a] and *Tonnetz* by Tymoczko et al. [Tym12] reveal the geometry of chords. Sapp proposed a visualization to support hierarchy-based tonality analysis [Sap05]. Mardirossian and Chew propose a music visualization tool to investigate tonal progressions of compositions at an abstract level while dropping CMN [MC07]. Malandrino et al. introduce a technique visualizing harmony to analyze the structure of a composition [MPZ18], which is limited to four-voice chorales and is only applicable at the sheet level. Exploiting the circle of fifths [Hei69], we use the abstract harmonic fingerprint design that we introduced in our previous work [MBEA19] to display and query harmonic patterns in sheet music.

Rhythm – Recent work by Pesek et al. comprises the analysis and visualization of rhythmic patterns based on unsupervised learning [PLM20]. Foote et al. present an analysis approach for rhythm structure grounding on self-similarity [FC01]. Regarding Ren et al., rhythm plays an essential role in automatic pattern extraction [RVSV18]. Their approach is limited to algorithmic concepts without showing visualizations that provide understanding for the analyst. Liu and Toussaint follow an approach to visualize rhythm using *clock* diagrams with a polygon notation to emphasize different rhythmic patterns [LT12]. Inspired by their approach, we create a radial design based on the *clock metaphor* to visualize the rhythmic content of symbolic music based on measure units (see Section 5.1). Similarly, Benadon exploits a circular plot to visualize and analyze rhythmic information [Ben07]. For a more detailed description of the complexity of the musical feature *rhythm*, we refer to Toussaint, who provides different concepts for its visualization [Tou19].

Melody – Various approaches exist for melody analysis in sheet music. Meek et al. provide an automatic thematic extractor to identify *themes* by automatically extracting melodic patterns while reducing the resulting space through suitable filters [MB03]. Urbano proposes the library MelodyShape that enables similarity search for melodies based on the shape [Urb15] but not on the sheet level. Unfortunately, these approaches reside either on the conceptual level or require programming skills that musicologists, domain experts, music teachers, and students might not have. For instance, Laurson et al. use a pattern-matching syntax that analysts must learn to create rules for analyzing musical scores [LKK08]. Thus, these approaches are not directly accessible for a broad audience that requires intuitive and interactive visualization. There are existing approaches that enable searching a musical corpus based on melody contours like Musipedia [Irw08]. Similar to these considerations, the design of our melody query enables analysts to quickly find matching melodic motives through the interactive definition of a melody [Lar05].

Segmentation Structure – Understanding music structure is a challenging task that depends on the ambiguous hierarchy of the underlying segmentation. Creating music visualizations that properly cope with this ambiguity should reveal the underlying semantic structure [DPMP*17, NF14]. For instance, Chan et al. proposed a visual solution to reveal the semantic structure in classical works using a layer braid and theme fabric prototype [CQM10]. The concept of structure is a salient characteristic in many different lower-level music features. For instance, Bergstrom et al. visualize the chordal structure through their *Isochords* concept [BKH07b]. ImproViz by Syndal et al. reveals structure by the melodic landscape but omits CMN as well [SH05]. To address the segmentation problem, Rafael et al. employ a genetic algorithm to provide a segmentation strategy for symbolic music whose quality highly depends on the parameters [ROAW09]. While these approaches yield promising results, they all lack a connection to the familiar CMN context. The horizontal structure enables us to create summary visualizations. With our approach, users can either interactively segment a piece manually or based on structural elements such as repetition or harmonic, rhythmic, and melodic features to explore structural ambiguities.

2.2. Methodology Transfer

Digital humanities concepts inspire this work. We conduct a methodology transfer from other fields, such as text analytics, to the music domain [MSK*19]. For example, Poemage is a technique that reveals the beauty of poems by extracting the sonic topology based on textual features enabling close reading [MLCM16]. El-Assady et al. combine close and distant reading views for text data to investigate forum discussions [ESKC18]. Keim and Oelke introduce a visualization to display textual features through literature fingerprints enabling distant reading on text data [KO07]. Similarly, we use musical fingerprints for the independent musical features pitch and timing. By juxtaposing sheet music with glyph visualizations, we combine close and distant reading, creating interactive visualizations to analyze music. Jänicke et al. highlight the importance of the combination of close and distant reading [JFCS15]. Similarly, this work aims at supporting sheet music analysis at different abstraction levels. We transfer the idea to music analysis by augmenting sheet music with visual elements in a minimally intrusive way to highlight relevant patterns. To tightly connect both concepts we use staged animations [HR07] enabling slow analytics [BSC19] for musicologists.

3. Problem and Data Characteristics

The notational complexity of the musical dimensions [MHK*18] requires reading proficiency to understand the underlying musical concepts. Below, we introduce the musical features from the smallest unit to compound concepts including *harmony*, *rhythm*, and *melody*. The requirements below have been extracted from existing literature by Weihs et al. [WJVR17] and Khulusi et al. [KKM*20] who provide a profound basis for music analysis tasks and data characteristics. While we have not consulted domain experts during this first step, we added a feedback loop at a later point to verify the design decisions that were made based on our previous literature research.

3.1. Musical Features

The process of combining primitive musical features leads to compound features and increases CMN's complexity. Since musical compositions combine several of these features, it is crucial to understand their relationship, dependencies, and differences. This analysis of musical features is a fundamental aspect of *Feature Analysis on Harmony, Rhythm, and Melody* [**T1**] and *Data Exploration* [**T2**].

Primitive Musical Features – CMN is built up by a composition of different *elementary musical features*. Every musical *note*, which is the smallest musical element, has a specific *pitch* that describes the frequency of sound, and a *duration* that indicates the temporal length of a note. The note's length depends on the tempo and meter usually defined by the composer. The note's *onset* defines its temporal placement in a composition. In contrast to notes, *rests* explicitly show the absence of sound providing temporal and structural information. Additionally, CMN encodes the perceived differences of *timbre* between instruments or voices. Musical scores define the instrumentation at the beginning. Some compositions have textual assignments to a sequence of notes indicating *lyrics* that can be sung using the underlying music as accompaniment. Moreover, instructions such as *dynamics* or *articulation* give additional information about the loudness and playing style. These instructions are *interpretable annotations* that are less exact than, for example, pitch and duration. Therefore, we differentiate between the symbol, the mere placement of a symbol inside CMN, and its meaning leading to performed actions.

Harmony is a compound feature of *pitch* in the vertical direction. Multiple simultaneous pitches result in a chord. There are tremendously many possible combinations, particularly when increasing the number of simultaneous notes. If onset and duration are added, we can describe chord progressions and more complex harmonic patterns that require theory knowledge to understand [LJ83a]. Chord progressions influence the perceived *tonality* or *modulation*.

Rhythm is another independent compound feature of *duration* and *onset* of equal importance. Rhythm defines the timing or horizontal placement of notes. Rhythmic patterns influence the style of musical composition, such as waltz, blues, or minuets. Consistency, coherence, and regularity of musical notes affect the listening experience.

Melody represents a specific notes sequence with different pitches and durations providing the required material for musical motifs. We distinguish between a single melody and a progression of chords. Any melodic instrument, such as the recorder, can create a melody, which differs from multiple simultaneously played notes. Based on this definition, we can define *motifs* and their *variations*. Melody is the most salient feature of these since it is the aspect that is usually remembered by the listener, e.g., influencing catchy tunes. Especially

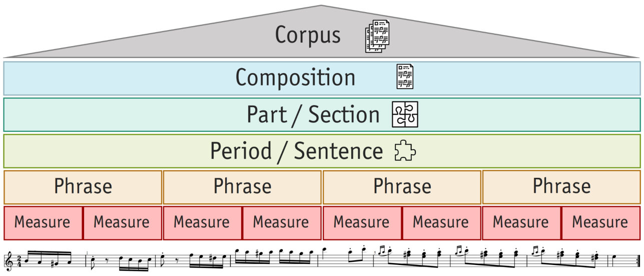


Figure 2: *The hierarchical segmentation of music compositions.*

the leitmotif [BS15] are of high interest for music analysts, which is the leading motive that affects the central theme of a musical piece.

Abstraction and Aggregation – Section 3.2 introduces different structural levels. We institute statistical abstractions to analyze features at any level. While losing detailed information, this abstraction facilitates analysis tasks such as comparison or interpretation.

3.2. Hierarchical Music Structure

Music contains hierarchical structures at different levels [LJ83b]. Weihs et al. introduce *Structure and Form Analysis* [**T3**] as a relevant high-level task [WJVR17]. We can identify such structures by applying segmentation on both the corpus level or the measure level. Many musical elements affect the structural ambiguity. The structure of music comprises a *horizontal* and a *vertical* dimension. For instance, the simultaneous performance of multiple instruments, e.g., orchestral scores, or instruments that can simultaneously play multiple pitches (e.g., the piano) represent the *vertical* dimension. The vertical structure separates melodic from harmonic progressions. Melodies comprise only single notes [Lou12]. The *horizontal* structure represents the temporal aspect. Figure 2 illustrates the different structural entities from the corpus to the measure level. Rhythm plays a significant role in the horizontal structure since it affects the temporal aspect of harmonic, rhythmic, and melodic patterns. The inherent musical structure enables *Comparative Analysis* [**T4**] at different levels [WJVR17].

3.2.1. Single Composition

Musical works have various properties, such as composer, style, and epoch. Analysts investigate single pieces at different levels.

Measure-level – The smallest unit that builds up a musical composition can be found at the measure-level, where a composer constructs notes, rests, articulation, beat, and dynamic details.

Phrase-level – Phrases, periods, and sections set the context and provide structure on top of the smallest musical entities. The concatenation of several phrases, sentences, or periods leads to even more complex structural relationships at higher levels. Time-span reduction builds up a hierarchy of stability, providing a tonal center [LJ83b].

Composition-level – Musical pieces may play a key role within their context. For instance, Beethoven's piano sonatas comprise up to four movements providing musical tension and relief [Dra00] to enhance the listener's experience. Musical compositions build up complete parts, movements, and full works that may comprise numerous smaller musical units at this higher level.

3.2.2. Multi Composition

Similar to text corpora [CWDH09], sheet music collections can be grouped by characteristics such as composer, epoch, style [KKM*20]. Comparing collections requires high abstraction levels as it is impossible to view every musical detail as low-level entities. The musical structure depends on the horizontal (temporal) dimension, which can be inferred from the compound features rhythm or harmony. Structuring music based on these musical entities requires analysts to extract relevant patterns that can be identified by their salience in a piece. Eventually, the reader must develop a complex mental model about the inner working of musical relationships.

3.3. Analysis Tasks

There are numerous analysis tasks that music readers and analysts can perform based on the vertical and horizontal structure in combination with the primitive or compound musical features [WJVR17].

In the following, we present an extensive task list of varying importance for different user types. Based on the horizontal structure levels introduced in the previous section, we provide a concise overview of different tasks that can be performed at each level. If we take a look at the musical features, then we can extract various tasks that can be the subject of music analysis.

Measure-Level – At the measure level, a common harmonic task is chord detection based on simultaneous pitches. Motif extraction is a melodic task enabling the identification of repetitive note sequences. Understanding the fundamental rhythm complexity is substantial. At the measure level, only low-level tasks can be executed.

Phrase-Level – Often, a phrase contains 4, 8, or 16 measures, combining multiple musical sentences or periods. At this level, analysts identify the primary melodic content and main themes of a composition. Understanding the differences and similarities of melodic movements between different voices is another essential task that can be performed at the phrase level. For instance, this task facilitates the understanding of melody sequences in a fugue that is shared and repeated at different transposed stages. Besides melody, understanding chord progressions and their recurrences are required to elicit modulations from the harmonic data characteristics.

Composition-Level – A fundamental task at the piece-level is the analysis of instrumentation and their usage throughout a composition. This is directly related to the interpretation of voice development. The extraction of tonalities, including the identification of modulation, tonal dominance, and unexpected occurrences of tones that are entirely foreign to the current tonal center, is crucial. Understanding the different rhythmic patterns used in combination with harmony is another fundamental aspect for a composition to convey the journey described by the composer.

Corpus-Level When considering a collection of different compositions that may have certain commonalities (e.g., composer, epoch, style, genre, instrumentation), it is interesting to compare the tonality and rhythmic complexity to reveal prevailing feature patterns. Understanding the similarity between pieces allows the analyst to draw conclusions about a composition regarding the date of creation or originality of the composer. Furthermore, analyzing sheet music at this stage enables us to understand the variety and complexity within a corpus. Repetitions or specific song segments such as verse or chorus can help to identify structural, semantic entities.

Level-Independent Numerous tasks can be performed at *each* analysis level. E.g., analysts can always add annotations to record semantic information. Comparative analysis can be performed at each stage as well. These only reveal different information that requires different interpretations. Of course, summarization of features is possible to show the underlying statistics scales to any analysis level as well as structure analysis. Analysts can be interested in the frequency or scarcity of a pattern which is independent of the analysis level.

4. Technique: Augmenting CMN for Multi-Level Analysis

Musical compositions displayed by CMN comprise numerous features. Music analysts aim to understand and interpret musical compositions on different abstraction levels. To support low-, mid-, and high-level analysis tasks, we propose a visual analytics technique to add glyph visualizations as minimally intrusive augmentations to CMN. Thus, analysts can explore sheet music at multiple abstraction levels [**T2**] from the sheet to the projection level by maintaining the intermediate connection through animated transitions [KCH19].

4.1. Embedding Glyph Visualizations into Sheet Music

Abstract visualizations complicate the understanding of the encoded information. To bridge this gap, our technique starts from CMN, which is the traditional and familiar context for music readers [JFCS15], including music analysts (see Figure 3(a)). CMN does not provide extra space, which requires suitable visual modifications before we can augment it with visualizations. The required additional space depends on the visual characteristics of a respective visual design. We consider glyph visualizations to be suitable due to their small size and the flexible application of visual attributes (color, size, etc.) [BKC*13]. Besides, glyphs can be placed independently, which allows revealing of relationships between single entities [**T1**]. Glyph-based designs have the power to communicate multivariate data while preserving the connection to the spatial context [BKC*13]. Composers employ the basic horizontal unit of measure to set single notes defining the musical structure [**T3**]. The technique applies the given measure boundaries as a minimum window of aggregation for the glyph representations. The technique positions the glyph by horizontally centering it closely to the top of the respective measure as displayed in Figure 3(b). The proximity to the CMN emphasizes the connection while ensuring that the glyphs do not overlap with the CMN. The layout of the feature glyphs requires a homogeneous design to support visual comparison [**T4**].

Horizontal and Vertical Segmentation – CMN comprises a horizontal segmentation defined by multiple split criteria such as key and time signature changes or phrase repetitions [**T3**]. These criteria can be leveraged to obtain different perspectives on sheet music. Such divisions build the foundation to keep the overview at more abstract levels, compare segments, and identify relevant patterns. While horizontal units provide the foundation for the musical section's temporal segmentation, vertical units separate single voices. For instance, conductors need to understand the simultaneous progression and interplay of several instruments. In this case, it would not be sufficient to encode parallel voices by a single glyph. Showing separate (parallel) glyphs would be required to support the differentiation of single instruments or instrument groups.

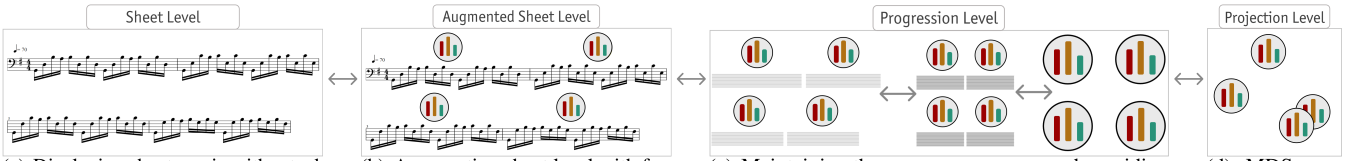
Extraction and Processing of Musical Features – The visual encoding of a glyph design depends on which musical feature(s) should be displayed and what analysis tasks should be supported. Therefore, the feature characteristics that are to be encoded should influence the glyph design [**T1**]. For instance, analysts could be interested in the duration of notes within single measures. In this scenario, a bar chart design could reveal the statistical distribution of note durations ordered by their length. Musical features have varying characteristics that need to be reflected by a corresponding glyph design. If possible, the use of discrete bins facilitate the comparability of visual representation of features (e.g., 12 pitch classes for the feature *harmony*). Statistical qualities such as mean, sum, min, and max can be extracted and serve as input to the glyph visualization. For example, there exist libraries such as jSymbolic [MCF18] that can be used to extract, e.g., melodic, rhythmic, and pitch features from symbolic sheet music representation such as MusicXML. These numerical features can be encoded in suitable glyph representations to visualize other relevant characteristics. Moreover, glyph designs can be replaced by simpler versions to improve understandability which may be subject to less information being encoded.

Aggregation of Adjacent Glyphs – Enabling analysts to aggregate musical information is an integral part of the feature analysis. Aggregating adjacent glyphs should create parent glyphs that contain the summarized content of their children. This aggregation allows for transitioning from a low (measure-level) to a high aggregation (composition-level) by keeping the CMN context. When reading the glyphs, it must be clear which underlying measures are included. Therefore visual elements must be used to indicate which structural entities are included in a glyph summary. At aggregated stages, analysts can compare larger sections to identify salient differences [**T4**].

Interactive Exploration and Querying – The query design should fit to the visualization design to support discovering feature patterns [**T1**]. Interactive filtering should likewise facilitate the search for similar and dissimilar patterns, letting analysts focus on subsets. For example, the analysts could be interested in identifying all measures with a specific chord root. The interactive exploration of musical features at the different extraction layers is required to support the generation and verification of hypotheses [**T2**]. Additionally, saving interesting queries with custom labels is crucial to capture valuable results and retrace the analysis process.

4.2. Analysis Abstraction Layers

Transforming separate features into glyphs provides an abstraction of the CMN. This provides multiple perspectives on data, which allows analysts to perform feature exploration [**T1**] and structural comparison [**T4**]. Integrating staged animations enables seamless transitions between abstraction layers. Thus, a visualization should guide analysts from close to distant reading between each abstraction layer. The only requirement is to change a few visual aspects of each layer to ensure that the analyst can trace the visual changes. The most abstract level of the technique is the projection layer. Every available layer covers a specific analytical purpose and does not only subserve for smooth transitions as described in Section 3.3. The requirement of displaying a single layer at a time tasks reduces the amount and complexity of simultaneously depicted information. Our technique comprises four core layers from the sheet to the projection layer, which are independent of the aggregation stage. We consider that the technique can be extended by additional layers, but we consider these four layers represent the minimum required layers for an understandable transition from the sheet to the projection layer.



(a) Displaying sheet music without additional information at the sheet level.

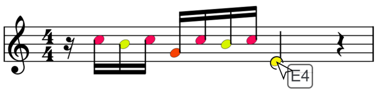
(b) Augmenting sheet level with feature fingerprints placed on top.

(c) Maintaining the measure sequence and providing staged animations [HR07] to omit CMN.

(d) MDS projection [CC08].

Figure 3: ***From Close to Distant Reading:*** *Four abstraction levels connected through staged animations [HR07] provide different perspectives. The sheet level (a) is the familiar CMN context. The augmented sheet level (b) embeds glyphs into CMN. The progression level (c) omits CMN while keeping the glyph sequence order. The projection level (d) rearranges the visualization applying cosine similarity-based MDS [CC08].*

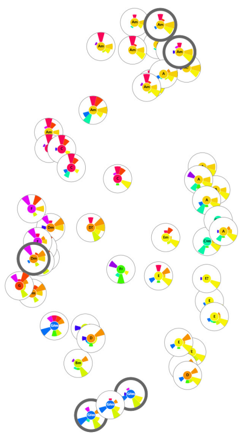
Sheet Layer – Analysts can read music as usual based on CMN at the *sheet-level* (see Figure 3(a)). Hitherto, no abstract information is displayed, allowing one to focus on the unaltered CMN. Optional note coloring can enhance different characteristics or tasks such as pitch class identification. Coloring notes also improves accessibility by mitigating reading difficulties [FTK16] and introduces cognitive pleasure [KOK*13].



Augmented Sheet Layer – The second layer (see Figure 3(b)) adds the glyph representations above the CMN connecting feature glyphs to single measures. Analysts can explore the relationship between the glyphs and the score through interactive highlighting. The vertical space must be adapted to avoid overlapping of the CMN.

Progression Layer – The augmented sheet layer occupies much screen space due to the uncondensed nature of CMN. Specifically, conductor scores with many instruments are space-consuming, requiring conductors to scroll through the score quickly or to learn it by heart. The *progression layer* (see Figure 3(c)) keeps the sequential order, but omits the CMN through collapsing all instruments through *staged animations* [HR07] to set the focus on the glyph visualizations. Figure 3(c) illustrates the transition from the augmented sheet layer. Eventually, the stave placeholders are removed to shift attention to the glyph progression, and the glyphs are rearranged to improve the layout of the progression layer.

Projection Layer – The highest abstraction layer is the *projection layer*. Applying data projections such as cosine similarity-based MDS [CC08] rearranges the glyphs' position based on the feature. The projection can reveal clusters and outliers. The example of the instantiation we introduce in Section 5 applies a MDS projection based on the harmonic glyph design that we published in our previous work [MBEA19]. Here, the outer stroke of the glyphs' circle indicates its frequency. Using data projection enables the user to explore the landscape of different or similar glyph groups clustered together, revealing a feature's complexity at a glance. Besides groups, outliers or infrequent glyphs that reside between clusters can be discovered.



5. Instantiation: *MusicVis*

The technique description builds the foundation for possible analysis applications. We implemented the technique as an exemplary instance (*MusicVis*) by using existing abstract visualization designs that we published in our previous work [FMK*20, MBEA19]. Figure 1 illustrates the different analysis components of *MusicVis*.

Design Rationale – *MusicVis* integrates interactive highlighting, linking & brushing, and details-on-demand to improve the connection of the glyph abstractions to the CMN as required by the technique. In previous work, we have introduced glyph designs for harmony [MBEA19] and rhythm [FMK*20]. Both visual fingerprint designs build on music domain knowledge. The fingerprints' design meets the requirement to summarize temporal windows of the CMN that are scalable to any aggregation level. The radial layout of the harmony and rhythm glyph designs allows them to be handled in *MusicVis* in a similar fashion, ensuring *design consistency*. *MusicVis* deliberately separates the analysis of harmony and rhythm to improve the analysis' focus on each feature. At any aggregation and abstraction stage, the analyst can interactively switch ( ) between analyzing harmony and rhythm. Besides, the instantiation provides the opportunity to add annotations that the analyst can apply to any used visual element. Thus, analysts can perform their analysis tasks as usual with additional abstract views that facilitate the generation and verification of interpretation hypotheses. *MusicVis* also enables flexible melody analysis. Relevant results are placed directly at the corresponding position as a CMN overlay since melodies comprise both harmonic and rhythmic information.

Composition View – The central widget that provides the different abstraction levels is the *Composition View*. The analyst can use the navigation elements at the top-right corner of the interface to aggregate the glyphs of adjacent measures or transition between the available abstraction levels. At any stage, analysts can interactively investigate single measures or get an overview of a complete composition. Several layout options allow for optimal exploitation of the available screen space via, e.g., zoom-to-fit. With that, users can apply desired line breaks if a piece is too large to view all measures at once.

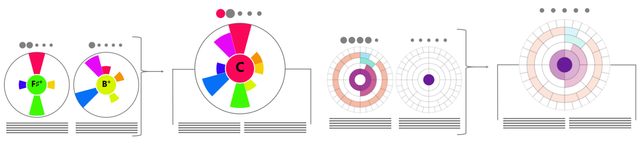
5.1. Feature Visualization

MusicVis uses different visualization designs to encode harmonic, rhythmic, and melodic information. The harmonic and rhythmic fingerprint glyphs cover respective problem/data characteristics and analysis tasks (see Section 3). As discussed in Section 4, the glyphs are modular; their visual design is arbitrarily interchangeable without violating the technique's requirements. The separate feature encoding enables analysts to focus on single characteristics.

Harmony – *MusicVis* uses the harmonic fingerprint design from previous work [MBEA19] to visualize pitch-class distributions. The *harmony glyph* enables analysts to compare the harmonic complexity. We refer to the original publication [MBEA19] regarding its visual encoding based on the *circle of fifths* [Hei69]. The proximity of arc segments indicate harmonic similarity and sound more consonantly together (e.g., C is closer to G than to B). Pitch classes that reside at the opposite glyph side are dissonant (e.g., tritone intervals).

Rhythm – Besides harmony, *rhythm* is the second fundamental dimension in music. Similarly, to provide separate visual fingerprints for rhythm, we use the visual encoding that we proposed in a previous publication employing the metaphor of a clock [FMK*20]. If the presence of notes and rests collide at a position in the glyph, the rhythmic fingerprint prioritizes the presence of rhythm over its absence (rests), which are more relevant for the perceived rhythm of the music [LGA*11]. The rhythmic fingerprint supports analysts in detecting the rhythmic similarities and differences of musical sections.

Melody – As a combination of pitch and rhythm, melody builds up musical motives of themes. For the visual presentation of melodic occurrences, Figure 7 illustrates how lines based on the respective melody notes' positions emphasize their contour. The dominant pitch or chord root of a melodic passage is emphasized using the harmonic color map. A dashed outline box indicates the boundaries of each melodic match. The melody overlay visualization is only available for analysis at the (augmented) sheet level. We exploit the harmonic fingerprint's color scheme to color notes enabling users to draw the connection between the sheet and the harmony glyphs.



(a) Harmony Glyph Aggregation (b) Rhythm Glyph Aggregation

Figure 4: *Summarizing neighboring harmonic and rhythmic glyphs to get a more abstract representation of the underlying features.*

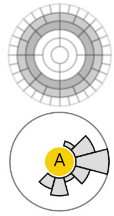
Aggregation of Harmony and Rhythm – As required by the technique, *MusicVis* allows analysts to aggregate and split the rhythm and harmony glyphs independent of the close and distant reading level. The aggregated glyphs provide statistical summaries about the encompassing measures. Figure 4 illustrates the aggregation of harmony and rhythm glyphs. The example for harmony shows how the pitch class distribution in the aggregated glyph reflects the underlying glyphs' content revealing the overall dominant pitch. While the harmonic summary glyph uses the *area* to encode the differences, the rhythmic summary shows the dominance of notes or rests using opacity, emphasizing the relative differences of the dominance of notes regarding their offset and duration. Aggregated glyphs are visually distinguished from single-measure glyphs by bracket lines (*whiskers*) showing which underlying measures are summarized by their parent.

5.2. Visual Querying

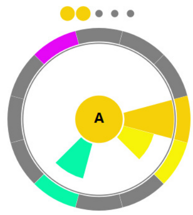
We integrated interactive visual query selectors to support exploration through user-defined, flexible search filters. Querying fundamental musical features is essential to reduce the analysts workload of identifying recurring patterns. Providing separate queries for each feature allows the user to focus on a single characteristic at a time.

5.2.1. Querying Harmony and Rhythm

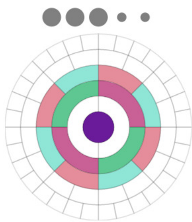
MusicVis provides visual query selectors tailored to the respective feature glyph design to enable flexible exploration. Consistent user interaction concepts support the intuitive handling of the filters. After querying, all non-matching glyphs (e.g., different root pitch) are grayed out. The shape difference is calculated by cosine-similarity which is then applied to the glyphs' opacity in the composition view. Rotation invariant queries (as indicated in the right figure) help to identify rotated examples of a search template. The queries can reverse the match opacity scales to emphasize the dissimilar glyphs, which is particularly helpful for identifying outliers.



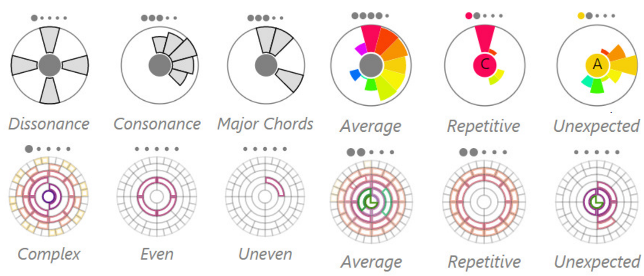
Harmony Query – The adjacent figure shows an exemplary A-Major query by filtering the pitch classes A, C#, and E while prohibiting the pitch class of F since the outer segment is selected, but the area is set to zero. The outer ring segments indicate required or optional pitches. Combining optional and mandatory pitches creates custom queries (Match Custom). The five dots above the query indicate the relevance score of the selected root pitch by its color and the query shape match by gray dots. If any rotation is allowed in the query, the technique matches the rotated glyph versions. The rotation invariance for matching similar glyph shapes corresponds to similar chords independent of their absolute pitch value. Thus, the glyph rotation corresponds to the *transposition* of, e.g., chords.



Rhythm Query – The visual rhythm query has the same layout as its glyph. The adjacent figure displays a search with a *whole* note, two *quarter* notes and rests, and four *eighth* notes and four rests equally distributed. The query does not contain $16^{\textbf{th}}$ or $32^{\textbf{th}}$ notes or rests. This relevance score at the top shows a relatively high shape match. Similar to the harmony query, analysts can interactively create flexible and custom queries. Focusing on the shape by changing the search from *Keep Offset* to *Any Offset* is possible, thus ignoring the offset of the notes highlighting all rotated matches in the composition view. Consequently, the visual rhythm glyph allows for custom queries of *complex* or *even* patterns.

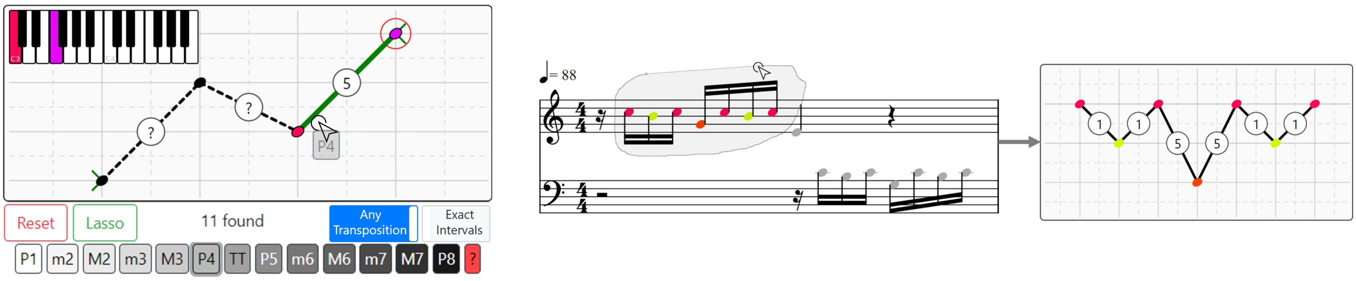


Semantic and Statistical Features – Music analysts can discover automatically extracted glyph templates, including the *most common*



(a) Semantic Query Templates (b) Statistical Query Templates

Figure 5: *Querying by semantic and statistical feature templates.*



(a) Interactive melody query user interface.

(b) The interactive lasso facilitates the direct search for melodies available through the CMN.

Figure 6: ***Melody Query.*** *The melody search interface (a) enables analysts to explore melodic patterns. Different parameters allow to create custom queries by assigning intervals or pitches. The lasso selector (b) allows for querying existing melody themes directly from the CMN.*

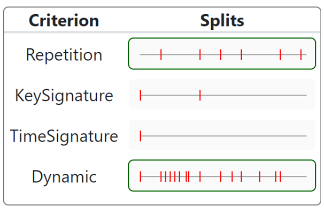
harmony and rhythm glyphs (see Figure 5). The *average* rhythm template reveals all notes that are present in a composition. Likewise, the *average* harmony template summarizes all used pitch classes and depicts their respective dominance. Users can exploit a predefined template as a starting point for further specifying and restricting the feature search. For instance, to detect for C-Major chord occurrences, the analyst could first select the *Major Chords* (see Figure 5(a)) template and then apply the root pitch C. On demand, analysts can also save glyph queries as custom templates to reuse them later in the analysis process [**T1**].

5.2.2. Querying Melody

An essential contribution of *MusicVis* is the separate visual query for melodic motifs (see Figure 6). The melody query allows analysts to explore occurrences of a melodic motif. Figure 6(a) shows an exemplary query consisting of *four notes* or *three intervals*: the query will detect melodies that match the interval contour ↗↘↗. The last interval must start with the pitch C and end with the pitch F which is a distance of *five semitones* or a *perfect fourth interval*. The analyst can set any interval ranging from 0 (*perfect unison*) to 12 semitones (*perfect octave*). To bypass the tedium of creating manual queries, analysts can apply an interactive *lasso* (see Figure 6(b)) to select a motif directly from the CMN. The lasso allows users to query any number of notes that compose a melody, excluding rests. Originating from this starting selection, the analysts can change or widen the search by removing the absolute pitch by Any Transposition to keep only the intervals or the directions. Therefore, the melody search is a powerful interactive instrument that complements the harmonic and rhythmic query.

5.3. Structure Management View

Manually restructuring a composition is time-consuming. *MusicVis* displays structural elements including time and key signatures, repetitions, and dynamics changes. Analysts can interactively apply these semantic split criteria to a composition in any combination (here *repetition* and *dynamics*).



Query-supported split criteria – The analyst can also use the harmony, rhythm, or melody query to apply a splitting based on the query results. For instance, if a new section in a musical piece often starts or ends with a specific chord, then the query template can be used to save a user-defined split criterion. Finally, the user can always manually adjust a splitting before and after any measure at the sheet level. *MusicVis* allows analysts to save and load interesting splittings. In this way, they can explore and compare different combinations of custom and predefined split criteria without the need to recreate expensive manual splittings repeatedly.

5.4. Initial Expert Feedback

Before conducting a detailed user study, we first gathered feedback from two musicology professors teaching university classes. The presented version of MusicVis puts a primary focus on the harmonic and rhythmic analysis while neglecting the melodic characteristics. They described the harmonic and rhythmic summaries as an intriguing way of "visual language" that supports the generation of different hypotheses, such as the dominance of a dominant rhythmic pattern. Nevertheless, the experts expressed a particular interest in an interactive method to explore melodic motives to identify variations and repetitions of themes in a composition. Additionally, the experts emphasized that harmonic, rhythmic, and melodic recurrences provide indications of potential segmentation besides obvious criteria such as repetition symbols. Based on this valuable feedback, we designed the melody analysis component to address the need to explore melodic content. Moreover, we adapted the prototype to leverage the interactive feature query components as segmentation criteria leading to the structure management view.

5.5. Technical Details

We provide *MusicVis*, the introduced instantiation of the proposed technique, as a web-based interaction workspace that is publicly accessible online under *visual-musicology.com/sheetmusicvis*. The implementation builds on a client-server architecture for the data loading mechanism and information processing. The frontend builds on web technologies (JS, HTML, CSS) and we use the JavaScript library D^3 [BOH11] to create visual elements and animations. The backend is implemented using Python with Flask [Gri14] and provides an API REST interface which sends analysis results such as feature query matches to the frontend. We use the python library music21 to compute analysis results such as the distribution of notes for measures which are used as input for the fingerprint visualizations, for instance. In addition, music21 supports the identification of separate voices that facilitates querying single melodic themes.

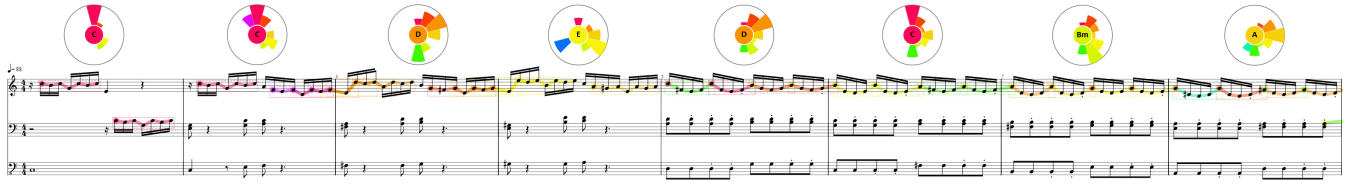
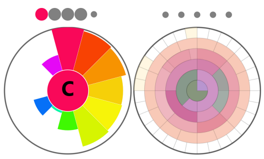


Figure 7: *The first eight measures of* BWV 553 *at the augmented sheet level reveal a harmonic modulation from* C*-Major to* E*-Major. The melodic content in this section consists of two motifs. The first motif is repeated by the second voice. The second motif uses circulatio. [Kir84]*

6. Exemplary Analysis Workflow

We take the *Prelude in C-Major (BWV 553)* [Mus20b] by J. S. Bach (only 29 measures) to showcase *MusicVis*' components. Figure 7 displays the first eight measures at the augmented sheet level, indicating the presence of four voices in three staves [**T2**]. The excerpt reveals how much space CMN requires to display musical information at the sheet level. Note the chord modulation from C to E via D in Figure 6(b) and back to C in measures 2–6. The melody highlighting indicates how the starting motif (red) transforms to a repetitive *Circulatio* pattern [Kir84] in measures 4-8 [**T1**].

To get an overview, the analyst *aggregates* all fingerprints to retrieve the *harmonic and rhythmic summary* of the whole piece at the *highest close and distant reading level*. The rhythmic summary fingerprint shows that the piece uses a diverse set of notes from whole to 32**nd** notes. The rests' shortest duration is displayed at the 8**th** ring. In contrast, the harmony summary glyph has a high shape match score (4 of 5 gray dots), indicating that about 80% of the piece has the root C. This glyph also reveals to the analyst that Bach's prelude primarily consists of the scale notes but also contains notes (F#, G#, C#) that do not belong to the C-Major scale (C, D, E, F, G, A, B) [**T1**].



The adjacent harmony query filters glyphs containing pitches that do not belong to the C-Major scale. The relevance score indicates that **65%** of all harmony glyphs match the query [**T4**]. The analyst is wondering about the distribution of pitches from other scales and switches to the *progression level* (see Figure 8) to realize that Bach only sticks to the C-Major scale notes if the *leitmotif* (see Figure 6(b)) is played or at the ending of the prelude [**T3**]. Consequently, the analyst quickly grasps that Bach's consistent employment of notes from other scales provides harmonic diversification, making this piece a really unique short composition.

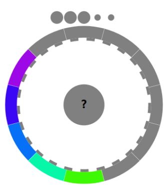


Figure 8: *65% of the harmony glyphs in Bach's BWV 553 contain notes that do not belong to the* C*-Major scale (*F#*,* C#*,* G#*).*

7. User Study

To evaluate the applicability, usability, and effectiveness of our visualization technique, we conducted a pair analytics study with different user types covering a wide range of application scenarios. We decided to use the pair analytics study method [KF14] due to the limited availability of qualified domain experts. Moreover, we consider that the numerous components may require additional explanations during the analysis process, which would be rather difficult to get acquainted with all at once. We use Sperrle et al.'s evaluation setup dimensions to describe the design, participants, and procedure of our study [SEAG*21].

7.1. Study Design

We used an online video conferencing software to present *MusicVis* to the participants via the study director's screen. We gained valuable insights into the challenges students, teachers, professionals, and hobby musicians face when dealing with sheet music. We gathered expectation feedback at the start and interactive session feedback. Final retrospective feedback was inquired revealing to what extent *MusicVis* met the initial expectations of the participants.

Methodology – We performed a combined *pair analytics* and *observation* study [AKGF11] to evaluate our approach. We conducted sixteen open-ended sessions (average duration of 2 hours and 45 minutes), each led by a **v**isual **a**nalytics **e**xpert (VAE) of our team. Every study session comprised six parts. After a short *introduction*, each participant completed a *demographic questionnaire* to provide background information (e.g., domain knowledge). We also gathered the participants' *expectations* of a system that supports score-based music analysis. Besides the analysis tasks, we introduced the harmony and rhythm glyph design, followed by a detailed user interface introduction (*walkthrough*). After the introduction, the participants received the control to start with *unguided exploration* [**T2**] before performing an *open-ended* analysis of two different musical pieces, one determined by us and the second of their own choice. During the analysis session, the VAE asked questions concerning the task guiding the analysis. We also encouraged them to think aloud and to verbalize their interactions. After the analysis, we conducted a final feedback session to reflect on each participant's experiences regarding traceability, functionality, and usability. Finally, we closed the study with a second questionnaire comprising 23 questions based on a five-point Likert scale assessing the available features.

Datasets and Tasks – For the study, we used the first 64 measures of the 3rd movement of Mozart's Piano Sonata No. 11 Allegretto Rondo Alla Turca [Mus20a] from KV. 331. The composition contains repetition signs, key signature changes, and recurring rhythms. Moreover, using a well-known score sparks interest and facilitates confirmatory analysis [**T3**][**T1**]. Consequently, we were able to qualitatively compare the analysis results between the participants who performed given analysis tasks (see Section 3.3). In addition, every participant selected a second piece that was not used by the other participants to show the general applicability of *MusicVis*. The used sheet music was based on the MusicXML format.

Participants – The participants had a wide range of domain expertise (from < 1 year to > 20 years), educational background (from secondary school students to professional musicians), and age (from 18 to over 55). Thus, the participants represent a quite heterogeneous group. All participants joined the study out of personal interest and were not financially compensated for their time. Two **d**ata **a**nalysts ($\mathbf{D_1}$, $\mathbf{D_2}$), three secondary school music students ($\mathbf{A_1}$–$\mathbf{A_3}$), two **m**usic **t**eachers ($\mathbf{MT_1}$, $\mathbf{MT_2}$), four **h**obby ($\mathbf{H_1}$–$\mathbf{H_4}$), and five **p**rofessional musicians ($\mathbf{P_1}$–$\mathbf{P_5}$) participated in the user study.

7.2. Study Results

Expectation Interview – At the start, we asked the participants to explain typical analysis tasks and describe what they expect of a system that supports performing music analysis. Multiple participants stated that understanding the underlying concepts in music and the structure of musical compositions regarding the musical features (e.g., harmony, rhythm, and motivic content) [**T1**] is a substantial aspect of music analysis ($\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{D_2}$, $\mathbf{A_1}$). Nevertheless, the participants had different approaches to investigate a composition. For instance, while $\mathbf{P_2}$ tries to first get an overall impression (top-down), $\mathbf{P_3}$ usually only performs close reading within the familiar context (i.e., sheet level) [**T2**]. $\mathbf{A_1}$ stated that he performs music analysis for entertainment, for instance, reading a musical composition while listening to a symphony to understand the instrument relationships within their context. $\mathbf{H_1}$ – $\mathbf{H_4}$ typically perform music analysis to learn a musical piece faster or remember the parts [**T3**], hence being extrinsically motivated. At the time of the study, $\mathbf{A_2}$ and $\mathbf{A_3}$ had two years of preparation for their final music examination in secondary school behind them. They readily stated many aspects that are relevant for music analysis such as *motives* [**T3**], repetitive *chord progressions* [**T1**], and *interpretation of semantic concepts*. $\mathbf{D_2}$ specifically stated that he is interested in understanding the concept of tension and relief of the harmony in jazz compositions [**T4**]. In summary, all participants except $\mathbf{P_3}$ stated that a music analysis application should automatically provide relevant suggestions by, for instance, using color for emphasis.

Introduction of Design Concepts – Understanding the harmony and rhythm glyph largely depends on previous music theory experience. For instance, $\mathbf{MT_2}$ explains rhythmic concepts with a less-detailed pizza metaphor similar to the radial display of the rhythm glyph to his students. All participants stated that the abstract glyphs are comprehensible but may be difficult to read. Before the study, they often had no idea for what they could be useful. Especially $\mathbf{H_1}$ – $\mathbf{H_4}$ were challenged to understand the abstract representations since they usually do not deal with data abstraction concepts.

Interactive Think-aloud Session – In the interactive analysis session, the participants were asked to word their thoughts. As expected, the participants first required some time and assistance to apply the available features. When asked to provide interpretations about the harmonic complexity in the piece, $\mathbf{P_1}$ could see [it] at first glance due to the dominant pitches encoded by color in the fingerprints, and so could $\mathbf{A_1}$. Likewise, $\mathbf{P_1}$ appreciated the insight that the rhythm glyph provides. Due to $\mathbf{P_4}$, the progression level helps to get a structured overview for single features by omitting other characteristics. $\mathbf{P_5}$ considered querying for melody [**T1**] as Google for Motifs but wished for an extension to support diatonic intervals. While the fingerprints already assist in understanding the structure [$\mathbf{MT_1}$], most participants used the automatic split criteria (e.g., repetitions) combined with manual splits to retrieve a semantic partition [**T3**]. $\mathbf{MT_1}$ used the distant reading projection level to get an overview over the harmonic complexity [**T2**] which helped her to confirm the belonging epoch of the present composition. Overall, she explained that at the projection level the occurrence of clusters helped her visually identifying dominant patterns for each musical feature.

Final Feedback – Besides $\mathbf{P_3}$, all participants found the provided features to be useful. Most participants stated that they would like to use *MusicVis* in the future for, e.g., performance preparation and teaching tasks. Many participants saw potential use for early music education to illustrate theoretical concepts. For instance, $\mathbf{MT_1}$ stated that for music analysis "using abstract levels help to illustrate complex matters", since it "surely can help to convey music theory concepts in teaching contexts". $\mathbf{MT_1}$ found the use of color "appealing and motivating" for the analysis tasks. Additionally, she stated the *MusicVis* may be useful for performance preparation of musicians who are learning an instrument. $\mathbf{P_2}$ remembered that conductors often use color to highlight score parts with many instruments to keep an overview. Also, with the harmonic fingerprints, [he is] faster in finding a specific chord based on the harmony query [**T1**]. In contrast to her expectations, $\mathbf{P_1}$ was convinced by *MusicVis* and found it helpful to [view] complex matters [...] in a more understandable way. $\mathbf{MT_1}$ and $\mathbf{P_5}$ could identify rhythmic changes using the glyphs [**T1**] and get an overview of the rhythmic and harmonic complexity by employing the distant reading views [**T2**]. $\mathbf{P_3}$ found the glyphs to be incomplete compared to CMN since they omit musical information at the abstract levels. He could not think of any use case of how he could utilize *MusicVis*, since he primarily faces performance tasks requiring to remain within the familiar CMN context.

Closing Questionnaire – After the study, 15 of the 16 participants answered 23 questions [Q1–Q23] (Likert scale: Strongly Disagree(1) – Strongly Agree(5)) to assess *MusicVis*' functionalities. [Q1–Q8] are general questions, [Q9–Q12] cover the harmony glyphs, query, and aggregations. Similarly, [Q13–Q16] address rhythm, [Q17, Q18] cover the melody query, and [Q19, Q20] the structure analysis. [Q21–Q23] deal with the close and distant reading levels, including the staged animations. We refer to the supplementary material for the exact questions that were asked in the questionnaire and the results that we obtained from the participants.

The questionnaire received a total average score of $\sim$3.65. Grouped by topic, the participants preferred the melody features (average: $\sim$3.93), followed by harmony (average: 3.85) and structure (average: $\sim$3.83) analysis. **Q17** (I can identify Leitmotifs by using the melody search) received the best average score: $\sim$**4.26**. **Q9** (The application supports me in understanding dominant key signatures) received the second best average score of $\sim$**4.06**. **Q5** (I have learned something new about music) had the lowest score of **3**, followed by **Q1** (The application is easy to use) with the score $\sim$**3.26**.

8. Discussion

The conducted study shows that *MusicVis* addresses the demand for analysis flexibility covering different use cases. The minimally intrusive augmentation of CMN combined with the gradual and interactive transitions facilitates users to trace the abstract visualizations to their origin in the familiar CMN context. The pair analytics study revealed that music analysts could learn more about compositions, even if they are already familiar with them, obtaining a more comprehensive understanding. Providing abstract and statistical information through visual representations of basic music concepts engages analysts to step into a more focused analysis of some otherwise disregarded features. With that, our approach facilitates both low and high-level interpretations. The participants could get in-depth knowledge and explore harmonic and rhythmic patterns through different visual queries without setting up preliminary hypotheses. The tight connection between the familiar context and the glyph visualizations helps users to interpret and understand them better. Based on the study feedback, we elicited that our approach can be profitable for different application contexts including music theory education, performance preparation, and music information retrieval. Thus, we see that music students and teachers, as well as musicians and musicologists will benefit from this work.

Limitations – We instantiated the visual analytics technique to support a wide range of analysis tasks. This approach favors flexibility and broad applicability resulting in a *steep learning curve* depending on the many provided interaction opportunities. Hence, some components and views may not be required by every analyst or to solve a certain task. Displaying musical pieces requires a large display due to the visual encoding of CMN. Consequently, an in-depth analysis is challenging on smaller displays, but music analysts would not expect to read a symphony score on a single printed page, either. In this work, we set a primary focus to demonstrate how users with diverse tasks and requirements could employ MusicVis to analyze various musical compositions. Consequently, the focus of our evaluation was the qualitative feedback for our technique's instantiation, which there is no state-of-the-art technique that we could perform a quantitative comparison with, to the best of our knowledge. While $\mathbf{MT_1}$ stated that *MusicVis* could be helpful for musicians in performance preparation tasks, our study does not explicitly address this use case. In the future, conducting separate studies to investigate the application of our approach over multiple weeks could reveal whether users get a better understanding of the system over time and can apply it for learning musical pieces.

Lessons Learned – Based on our study, we confirmed that it is essential to build from a visual data representation that users are familiar with. In our case, it is the CMN that is well known by domain experts. Similarly, the analysis should always start from a well-known representation before introducing new or abstract feature representation as our glyph designs. Most participants appreciated the level of automation provided by MusicVis and even asked for more automatic suggestions, e.g., for the melody search. Consequently, we learned that increasing automated suggestions is beneficial while still keeping the user in charge of the applied level of automation. The opportunity to store custom queries and structure splits engages the user to apply different exploration strategies and improves the verification of hypotheses and knowledge generation. Ensuring scalability and considering several data characteristics (e.g., multiple instruments) enlarges the amount of possible applications of a music analysis technique. As we expected, music analysts pursue different analysis goals. There is a significant trade-off between high flexibility supporting many different tasks and the users' learning curve. As our study showed, the current instantiation is differently suitable for different user types. While music experts easily understand the CMN, music beginners already struggle to understand the CMN. Yet, we included novices in the study, as our prototype provides additional aspects that help understand the CMN, such as the note coloring, which supports beginners in reading the notation. Our investigation showed that novices could benefit from music data abstraction even though they are challenged in reading the traditional representation, even though they face a steeper learning curve than domain experts. In the future it could be useful to employ more techniques that further facilitate the comprehension of the CMN. Simultaneously, omitting the abstract visualizations that introduce a higher visual complexity may be suitable for music beginners that are learning the CMN. Alternatively, additional visualizations that help explain the CMN would be suitable for novices to learn the notation, which is a typical use case for people who want to learn to play an instrument.

Research Opportunities – From the user feedback and current limitations, we identified further research opportunities. We recognized that *MusicVis* could potentially support collaborative analysis of compositions, similar to collaborative text editors [ABG*16]. This would not be feasible on printed sheet music, which is usually used for teaching. Moreover, gamification [SSB*19] poses a great opportunity: students could provide information that is automatically checked by the system, e.g., the chords of each measure. Providing a user experience score could increase the engagement level to learn musical concepts. Musicologists and linguists could benefit from glyph visualizations to analyze the relationship between lyrics and the accompanying music. Combining topic model visualizations [ESS*18] with sheet music would open a new perspective for textual music analysis. This study showed the potential of the feature glyphs for statistical and semantic insights of musical compositions. It would be valuable to investigate musical pieces' classification by genre, style, epoch, composer, and other musicological aspects through the fingerprints. Similar to Simonetta et al., this would enable the analysis of large music sheet collections [SCOR18]. As Borgo et al. investigated, there are several alternative glyph designs [BKC*13] that could be applied on musical data. With our technique, we demonstrate how glyphs can be used to supply aggregated information about musical characteristics. We argue that research at the interface of musicology and visualization holds great potential for interdisciplinary collaboration. In specific, it would be interesting to invent glyph designs that address different analysis tasks such as chord identification, instrument segregation, and motive detection. We plan to conduct further studies that are tailored to specific user groups and analysis tasks to fathom out how our work could be applied in different application scenarios.

9. Conclusion

We presented a new visual analytics technique to augment digital sheet music, applying a minimally intrusive strategy by gradually switching from CMN to abstract visualizations. Our approach utilizes existing visual designs rooted in domain knowledge for harmonic and rhythmic features that reveal musical characteristics on different aggregation and abstraction levels.

We present *MusicVis*, an instantiation of the technique, that integrates four tightly connected close and distant reading levels using staged animations. *MusicVis* offers different visual queries for enhanced exploration of harmonic, rhythmic, and melodic features. A structure management view supports the interactive analysis of compositions' structural organization through splitting criteria such as repetitions and key signature changes or manual segmentation.

We evaluated *MusicVis* through a pair analytics study with four hobby and five professional musicians, two teachers, three students, and two data analysts. Even though the results indicate a steep learning curve, it proves to be applicable in different scenarios. Notably, the visual features queries enable music analysts to explore dominant and uncommon harmonic, rhythmic, and melodic patterns. Based on the study feedback, we elicited further research opportunities, such as applying *MusicVis* in collaborative teaching contexts. *MusicVis* is available online under *visual-musicology.com/sheetmusicvis*.

References

[ABG*16] Attiya H., Burckhardt S., Gotsman A., Morrison A., Yang H., Zawirski M.: Specification and complexity of collaborative text editing. In *Proc. on Principles of Distributed Computing* (2016), ACM, pp. 259–268. doi:10.1145/2933057.2933090. 11

[AKGF11] Arias-Hernández R., Kaastra L. T., Green T. M., Fisher B. D.: Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics. In *Conf. on Syst. Sc.* (2011), pp. 1–10. doi:10.1109/HICSS.2011.339. 9

[Ben07] Benadon F.: A Circular Plot for Rhythm Visualization and Analysis. *Music Theory Online 13*, 3 (2007). doi:10.30535/mto.13.3.1. 2

[BKC*13] Borgo R., Kehrer J., Chung D. H. S., Maguire E., Laramee R. S., Hauser H., Ward M., Chen M.: Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. In *STAR* (2013), Eurographics Association, pp. 39–63. doi:10.2312/conf/EG2013/stars/039-063. 5, 11

[BKH07a] Bergstrom T., Karahalios K., Hart J. C.: Isochords: Visualizing Structure in Music. In *Proc. of Graph. Interface* (2007), ACM Press, pp. 297–304. doi:10.1145/1268517.1268565. 2

[BKH07b] Bergstrom T., Karahalios K., Hart J. C.: Isochords: Visualizing Structure in Music. In *Proc. of Graph. Interface* (2007), ACM, pp. 297–304. doi:10.1145/1268517.1268565. 3

[BOH11] Bostock M., Ogievetsky V., Heer J.: D^3 Data-Driven Documents. *IEEE Trans. Vis. Comput. Graph. 17*, 12 (2011), 2301–2309. doi:10.1109/TVCG.2011.185. 8

[BS15] Bribitzer-Stull M.: *Understanding the Leitmotif: From Wagner to Hollywood Film Music*. Cambridge University Press, 2015. doi:10.1017/CBO9781316161678. 4

[BSC19] Bradley A. J., Sawal V., Collins C.: Approaching Humanities Questions Using Slow Visual Search Interfaces. In *VIS4DH at IEEE VIS* (2019). 2, 3

[CC08] Cox M. A., Cox T. F.: Multidimensional Scaling. In *Handbook of Data Visualization*. Springer, 2008, pp. 315–347. doi:10.1007/978-3-540-33037-0_14. 6

[CQM10] Chan W., Qu H., Mak W.: Visualizing the Semantic Structure in Classical Music Works. *IEEE Trans. Vis. Comput. Graph. 16*, 1 (2010), 161–173. doi:10.1109/TVCG.2009.63. 3

[CWDH09] Chen Y., Wang L., Dong M., Hua J.: Exemplar-based Visualization of Large Document Corpus. *IEEE Trans. Vis. Comput. Graph. 15*, 6 (2009), 1161–1168. doi:10.1109/TVCG.2009.140. 4

[DG08] Dannenberg R. B., Goto M.: Music Structure Analysis from Acoustic Signals. In *Handbook of Signal Processing in Acoustics*. Springer, 2008, pp. 305–331. 2

[Don86] Donnini R.: The Visualization of Music: Symmetry and Asymmetry. *Comput. & Math. with Applic. 12*, 1-2 (1986), 435–463. 2

[DPMP*17] De Prisco R., Malandrino D., Pirozzi D., Zaccagnino G., Zaccagnino R.: Understanding the structure of musical compositions: Is visualization an effective approach? *Inf. Vis. 16*, 2 (2017), 139–152. doi:10.1177/1473871616655468. 3

[Dra00] Drake K.: *The Beethoven Sonatas and the Creative Experience*. Indiana University Press, 2000. 4

[ESKC18] El-Assady M., Sevastjanova R., Keim D. A., Collins C.: ThreadReconstructor: Modeling Reply-Chains to Untangle Conversational Text through Visual Analytics. *Comput. Graph. Forum 37*, 3 (2018), 351–365. doi:10.1111/cgf.13425. 3

[ESS*18] El-Assady M., Sevastjanova R., Sperrle F., Keim D. A., Collins C.: Progressive Learning of Topic Modeling Parameters: A Visual Analytics Framework. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2018), 382–391. doi:10.1109/TVCG.2017.2745080. 11

[FC01] Foote J., Cooper M. L.: Visualizing Musical Structure and Rhythm via Self-Similarity. In *Proc. of Int. Comput. Music Conf.* (2001), Int. Comp. Music Assoc., pp. 419–422. 2

[FMK*20] Fürst D., Miller M., Keim D. A., Bonnici A., Schäfer H., El-Assady M.: Augmenting Sheet Music with Rhythmic Fingerprints. *VIS4DH at IEEE VIS* (2020). 2, 6, 7

[FTK16] Flach N., Timmermans A., Korpershoek H.: Effects of the design of written music on the readability for children with dyslexia. *Int. J. of Music Education 34*, 2 (2016), 234–246. doi:10.1177/0255761414546245. 6

[Gri14] Grinberg M.: *Flask Web Development - Developing Web Applications with Python*. O'Reilly, 2014. URL: http://shop.oreilly.com/product/0636920031116.do. 8

[Hei69] Heinichen J. D.: *Der General-Bass in der Composition (Dresden, 1728)*. Georg Olms Verlag, 1969. 2, 7

[HMM00] Herman I., Melançon G., Marshall M. S.: Graph Visualization and Navigation in Information Visualization: A Survey. *IEEE Trans. Vis. Comput. Graph. 6*, 1 (2000), 24–43. doi:10.1109/2945.841119. 2

[HR07] Heer J., Robertson G.: Animated Transitions in Statistical Data Graphics. *IEEE Trans. Vis. Comput. Graph. 13*, 6 (2007), 1240–1247. doi:10.1109/TVCG.2007.70539. 2, 3, 6

[Irw08] Irwin K.: Musipedia: The Open Music Encyclopedia. *Reference Reviews 22*, 4 (2008), 45–46. doi:10.1108/09504120810872337. 3

[JFCS15] Jänicke S., Franzini G., Cheema M. F., Scheuermann G.: On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In *Eurogr.* (2015), pp. 83–103. doi:10.2312/eurovisstar.20151113. 3, 5

[KCH19] Kim Y., Correll M., Heer J.: Designing Animated Transitions to Convey Aggregate Operations. *Comput. Graph. Forum 38*, 3 (2019), 541–551. doi:10.1111/cgf.13709. 5

[KF14] Kaastra L. T., Fisher B. D.: Field experiment methodology for pair analytics. In *Proc. of the Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization* (2014), ACM, pp. 152–159. doi:10.1145/2669557.2669572. 9

[Kir84] Kirkendale W.: "Circulatio"-Tradition, "Maria Lactans", and Josquin as Musical Orator. *Acta Musicologica 56*, Fasc. 1 (1984), 69–92. doi:10.2307/932617. 9

[KKM*20] Khulusi R., Kusnick J., Meinecke C., Gillmann C., Focht J., Jänicke S.: A Survey on Visualizations for Musical Data. *Comput. Graph. Forum* (2020). doi:10.1111/cgf.13905. 2, 3, 4

[KO07] Keim D. A., Oelke D.: Literature Fingerprinting: A New Method for Visual Literary Analysis. In *IEEE Conf. on VAST* (2007), IEEE, pp. 115–122. doi:10.1109/VAST.2007.4389004. 3

[KOK*13] Kukka H., Oja H., Kostakos V., Gonçalves J., Ojala T.: What Makes You Click: Exploring Visual Signals to Entice Interaction on Public Displays. In *Conf. on Human Factors in Comp. Sys* (2013), ACM, pp. 1699–1708. doi:10.1145/2470654.2466225. 6

[KS10] KAISER F., SIKORA T.: Music Structure Discovery in Popular Music using Non-negative Matrix Factorization. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2010), pp. 429–434. 2

[Lar05] LARTILLOT O.: Efficient Extraction of Closed Motivic Patterns in Multi-Dimensional Symbolic Representations of Music. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2005), pp. 191–198. 3

[Ler04] LERDAHL F.: *Tonal Pitch Space*. Oxford University Press, 2004. 2

[LGA*11] LEVÉ F., GROULT R., ARNAUD G., SÉGUIN C., GAYMAY R., GIRAUD M.: Rhythm Extraction from Polyphonic Symbolic Music. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2011). 7

[LJ83a] LERDAHL F., JACKENDOFF R.: A Generative Theory of Tonal Music. *MIT Press, Cambridge, MA* (1983). 3

[LJ83b] LERDAHL F., JACKENDOFF R.: An Overview of Hierarchical Structure in Music. *Music Perception: An Interdisciplinary Journal 1*, 2 (1983), 229–252. doi:10.2307/40285257. 2, 4

[LKK08] LAURSON M., KUUSKANKARE M., KUITUNEN K.: The Visualization of Computer-Assisted Music Analysis Information in PWGL. *J. of New Music Research 37*, 1 (2008), 61–76. doi:10.1080/09298210802259595. 3

[Lou12] LOUI P.: Learning and Liking of Melody and Harmony: Further Studies in Artificial Grammar Learning. *Top. in Cogn. Sci. 4*, 4 (2012), 554–567. doi:10.1111/j.1756-8765.2012.01208.x. 4

[LS19] LOPEZ-RINCON O., STAROSTENKO O.: Music Visualization Based on Spherical Projection With Adjustable Metrics. *IEEE Access 7* (2019). doi:10.1109/ACCESS.2019.2944083. 2

[LT12] LIU Y., TOUSSAINT G. T.: Mathematical Notation, Representation, and Visualization of Musical Rhythm: A Comparative Perspective. *Int. J. of Machine Learning and Computing 2*, 3 (2012), 261. doi:10.7763/IJMLC.2012.V2.127. 2

[MB03] MEEK C., BIRMINGHAM W. P.: Automatic Thematic Extractor. *J. Intell. Inf. Syst. 21*, 1 (2003), 9–33. doi:10.1023/A:1023549700206. 3

[MBEA19] MILLER M., BONNICI A., EL-ASSADY M.: Augmenting Music Sheets with Harmonic Fingerprints. In *Proc. of ACM Symp. on Doc. Eng.* (2019), pp. 201–210. doi:10.1145/3342558.3345395. 2, 6, 7

[MC07] MARDIROSSIAN A., CHEW E.: Visualizing Music: Tonal Progressions and Distributions. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2007), pp. 189–194. 2

[MCF18] MCKAY C., CUMMING J., FUJINAGA I.: JSYMBOLIC 2.2: Extracting features from symbolic music for use in musicological and MIR research. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2018), Gómez E., Hu X., Humphrey E., Benetos E., (Eds.), pp. 348–354. 5

[MHK*18] MILLER M., HÄUSSLER J., KRAUS M., KEIM D., EL-ASSADY M.: Analyzing Visual Mappings of Traditional and Alternative Music Notation. *VIS4DH at IEEE VIS* (2018). 2, 3

[MLCM16] MCCURDY N., LEIN J., COLES K., MEYER M. D.: Poemage: Visualizing the Sonic Topology of a Poem. *IEEE Trans. Vis. Comput. Graph. 22*, 1 (2016), 439–448. doi:10.1109/TVCG.2015.2467811. 3

[MPZ18] MALANDRINO D., PIROZZI D., ZACCAGNINO R.: Visualization and Music Harmony: Design, Implementation, and Evaluation. In *Int. Conf. Inf. Vis.* (2018), pp. 498–503. doi:10.1109/iV.2018.00092. 2

[MSK*19] MILLER M., SCHÄFER H., KRAUS M., LEMAN M., KEIM D. A., EL-ASSADY M.: Framing Visual Musicology through Methodology Transfer. *VIS4DH at IEEE VIS* (2019). URL: https://visual-musicology.com/. 2, 3

[Mus20a] MUSESCORE: Piano Sonata No. 11 K. 331 3rd Movement, "Rondo alla Turca", 2020. Accessed Nov. 20, 2020 from https://musescore.com/classicman/scores/49143. 9

[Mus20b] MUSESCORE: Prelude in C-Major (BWV 553), 2020. Accessed Nov. 20, 2020 from https://musescore.com/user/28657557/scores/5390410. 9

[NF14] NIETO O., FARBOOD M.: Identifying Polyphonic Musical Patterns From Audio Recordings Using Music Segmentation Techniques. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2014), pp. 411–416. 3

[Orl13] ORLOVAITĖ E.: Visual Fingerprints: Identifying, Summarizing and Comparing Music. *Sum. and Compar. Mus.* (2013). doi:10.2139/ssrn.2373133. 2

[PLM20] PESEK M., LEONARDIS A., MAROLT M.: An Analysis of Rhythmic Patterns with Unsupervised Learning. *Applied Sciences 10*, 1 (2020), 178. doi:10.3390/app10010178. 2

[ROAW09] RAFAEL B., OERTL S., AFFENZELLER M., WAGNER S.: Using Heuristic Optimization for Segmentation of Symbolic Music. In *Computer Aided Systems Theory* (2009), vol. 5717 of *Lecture Notes in Computer Science*, Springer, pp. 641–648. doi:10.1007/978-3-642-04772-5_83. 3

[RVSV18] REN I. Y., VOLK A., SWIERSTRA W., VELTKAMP R. C.: Analysis by Classification: A Comparative Study of Annotated and Algorithmically Extracted Patterns in Symbolic Music Data. In *Proc. Int. Soc. of Music Inf. Retrieval Conf.* (2018), pp. 539–546. 2

[Sap05] SAPP C. S.: Visual Hierarchical Key Analysis. *Comp. in Entertainment 3*, 4 (2005), 1–19. doi:10.1145/1095534.1095544. 2

[SCOR18] SIMONETTA F., CARNOVALINI F., ORIO N., RODÀ A.: Symbolic Music Similarity through a Graph-Based Representation. In *Proc. of the Audio Mostly on Sound in Immersion and Emotion* (2018), ACM, pp. 26:1–26:7. doi:10.1145/3243274.3243301. 11

[SEAG*21] SPERRLE F., EL-ASSADY M., GUO G., BORGO R., CHAU D. H., ENDERT A., KEIM D.: A Survey of Human-Centered Evaluations in Human-Centered Machine Learning. *Comput. Graph. Forum* (2021). 9

[SH05] SNYDAL J., HEARST M. A.: ImproViz: Visual Explorations of Jazz Improvisations. In *Ext. Abstr. Proc. of the Conf. CHI* (2005), ACM, pp. 1805–1808. doi:10.1145/1056808.1057027. 3

[Shn96] SHNEIDERMAN B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proc. of the IEEE Symp. on Visual Languages* (1996), IEEE Comput. Soc. Press, pp. 336–343. doi:10.1109/VL.1996.545307. 2

[SSB*19] SEVASTJANOVA R., SCHÄFER H., BERNARD J., KEIM D. A., EL-ASSADY M.: Shall we play? – Extending the Visual Analytics Design Space through Gameful Design Concepts. In *MLUI for Visualization and Analytics, IEEE VIS Workshop* (2019). 11

[SSS*14] SACHA D., STOFFEL A., STOFFEL F., KWON B. C., ELLIS G., KEIM D. A.: Knowledge Generation Model for Visual Analytics. *IEEE Trans. Vis. Comput. Graph. 20*, 12 (Dec. 2014), 1604–1613. doi:10.1109/TVCG.2014.2346481. 2

[Tou19] TOUSSAINT G. T.: *The Geometry of Musical Rhythm: What Makes a "Good" Rhythm Good?* CRC Press, 2019. doi:10.1201/9781351247771. 2

[Tym12] TYMOCZKO D.: The Generalized Tonnetz. *J. of Music Theory* (2012), 1–52. doi:10.1215/00222909-1546958. 2

[Urb15] URBANO J.: MelodyShape: a Library and Tool for Symbolic Melodic Similarity based on Shape Similarity, 2015. Retrieved September 16, 2020 from https://github.com/julian-urbano/MelodyShape/. 3

[Urb17] URBERG M.: Pasts and Futures of Digital Humanities in Musicology: Moving Towards a "Bigger Tent". *Music Reference Services 20*, 3-4 (2017), 134–150. doi:10.1080/10588167.2017.1404301. 2

[Wat02] WATTENBERG M.: Arc Diagrams: Visualizing Structure in Strings. In *IEEE Symp. Information Visualization* (2002), IEEE Comput. Soc., pp. 110–116. doi:10.1109/INFVIS.2002.1173155. 2

[WJVR17] WEIHS C., JANNACH D., VATOLKIN I., RUDOLPH G.: *Music Data Analysis: Foundations and Applications*. CRC Press/Taylor & Francis Group, 2017. doi:10.1111/anzs.12244. 3, 4