Visual Sheet Music Analytics

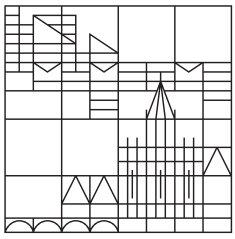
Dissertation zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

vorgelegt von

Matthias Miller

an der

Universität Konstanz



Mathematisch-Naturwissenschaftliche Sektion

Fachbereich Informatik und Informationswissenschaft

Konstanz, 2024

Dissertation der Universität Konstanz
Tag der mündlichen Prüfung: 24. Mai 2024

1. Referent: Prof. Dr. Daniel A. Keim, University of Konstanz

2. Referent: Prof. Dr. Mennatallah El-Assady, ETH Zürich

Acknowledgements

There are so many people I want to thank genuinely for their help and encouragement in completing this thesis. Without your continuous support, this work would have truly been impossible, and I could have never reached this important milestone!

First, I want to thank my mentor, *Daniel A. Keim*, for the exceptional opportunity to be part of your outstanding research group for the last six years. I am really grateful for the ability to work on diverse projects, including some related to my passion for music. Thank you, *Daniel*, for always having an open door and providing constructive advice and guidance for my research endeavors. Thank you for creating and enabling a pleasant and productive working atmosphere. I also cherish our regular BBQ and Christmas gatherings. A big thank you also goes to *Ilse* for making them possible.

I would like to thank *Mennatallah El-Assady*, my second supervisor, for her valuable insights and knowledgeable support all this time, helping me focus on the most relevant aspects during the conceptualization, preparation, and submission of several publications that are also presented in this dissertation.

A big thank you also goes to *Michael Grossniklaus* for his regular feedback. Your detailed questions helped me to focus appropriately on individual publications. Additionally, thank you to *Oliver Deussen*: thank you for stepping in as the chair of my oral examination on short notice, allowing me to complete it in a timely manner.

I also want to thank my PhD supervisors, *Dominik Sacha*, *Dominik Jäckle*, *Halldór Janetzko*, *Daniel Seebacher*, and *Manuel Stein*, who laid the foundations when I was still a student and motivated me to pursue a PhD. Thanks to all Data Analysis & Visualization Group members and colleagues at D3 and ZT in Konstanz for making my time as a doctoral student interesting and fun through valuable conversations, prolific collaborations, multiple retreats with exciting social events, and even occasional Skat and Job Jumble tournaments. Thank you, *Matthias Kraus*, *Frederik L. Dennig*, *Thilo Spinner*, *Fabian Sperrle*, *Rebecca Kehlbeck*, *Hanna Hauptmann*, *Rita Sevastjanova*, *Wolfgang Jentner*, *Michael Blumenschein*, *Johannes Fuchs*, *Yannick Metz*, *Juri Buchmüller*, *Sabine Kuhr*, *Maximilian T. Fischer*, *Lucas Joos*, *Dirk Streeb*, *Raphael Buchmüller*, *Bastian Jäckl*, *Johannes Häußler*, *Martin Schall*, *Sara Di Bartolomeo*, *Evanthia Dimara*, *Michael Behrisch*, *Daniela Blumberg*, *Eren Cakmak*, *Giovanna Ratini*, *Matt Sharinghousen*, *Benedikt Bäumle*, and *Philipp Traber*.

Furthermore, *Daniel Fürst* and *Julius Rauscher*, thank you for your reliable support in developing sheet music analysis applications as student assistants, but also as part of your BA and MA theses. I learned a lot from your programming skills and enjoyed working with you a lot. I want to also thank all my co-authors from outside our group for the exciting discussions. Thank you, *Alexandra Bonnici*, *Victor Schetinger*, *Andrew McNutt*, *Jane Adams*, and *Marc Leman*, for our prolific collaborations. I also want to thank all my friends who I met over the years. Thank you all for your companionship at the University in Konstanz throughout the years.

Menna, I want to specifically thank you for continuously motivating me to pursue projects bridging musicology and InfoVis. *Hanna Hauptmann*, thanks to you as well for our discussions and your practical help in paper writing and designing music analysis approaches. Thanks to both of you, I am convinced this thesis would not have been possible without your regular supervision and pushing me to keep a deadline when I was eager to postpone a submission.

Of course, I want to thank my amazing family for their persevering prayers, help, and mental support during my time as a Ph.D. student. *Mum* and *Dad*, you are the best parents I could wish for. I know I can always count on you; you only want the best for me. Thank you for sparking my interest in music and computers when I was young and raising and educating me to be who I am today. Because of you, I know the real value and long-term effects of education in our lives. Thank you for believing in me and for building my Christian faith through integrity, kindness, and support. You taught me how to trust in God, who always wants the best for us, and who gave me the ability to achieve this milestone.

Nelli, *Anja*, and *Veronika*: thank you for being my wonderful and supportive sisters over all the years. You are amazing teachers who also helped me to focus on what really counts in life. And *Nelli*, I want to give you a special thanks for accompanying me during my years of studies in Konstanz. I would have never come to Konstanz if it weren't for you. Also, *Daniel*, it is great to have such an intriguing brother who is also an outstanding researcher and excellent musician! I had tons of fun working with you on research projects combining quantum physics and visual analytics. I can't wait to see what God has in store for you.

Finally, I want to genuinely thank my gorgeous and ravishing wife ***Céline***. I love you with all my heart and all that I am. Thank you for everything. You are the love of my life! Going through thick and thin with you is the best that could have ever happened to me. I can't imagine it any better, nor would I want it to be any different. I do not take everything you do for our little family all the time for granted, supporting me in so many ways. Without your ongoing and persistent support and help, I would not be where I am today. My little, beautiful daughter *Skye*, I want to also thank you for enriching every single day of my life, from the day I first held you as a baby in my arms. I really hope that you take over my passion for music and someday become even better and more creative than I am or will ever be. Thank you for being so funny and cute. You always remind me why I am doing what I do. Through the both of you, I am truly the happiest and luckiest man alive.

Abstract

Through integrating visual interactive data analysis with sheet music, this thesis addresses a new interdisciplinary field: Visual Musicology. This work bridges the gap between information visualization and musicology, paving the way for new methods to analyze and interpret music data, specifically focusing on sheet music. The Western common music notation (CMN), used in sheet music, is uncontested among musicologists, educators, students, and musicians. Nevertheless, CMN's intricate and complex design challenges interpreting and analyzing large amounts of sheet music collections, particularly for those who are no visualization experts. We address these challenges by introducing novel visual analytics solutions that extend beyond the scope of existing research in the field. The initial foundational pillar establishes Visual Musicology, where we introduce the Visual Musicology Graph and the Methodology Transfer Model as organizational frameworks stimulating interdisciplinary research collaboration. They facilitate solution transfer across domains, as demonstrated through various musicological scenarios such as embodied music interaction. We also examine existing music notation techniques, discovering opportunities for unapplied designs that replace or connect CMN with advanced visualization techniques. Then, visualization designs for augmenting digital sheet music with harmonic and rhythmic fingerprints based on domain-specific concepts are presented, including the circle of fifths and rhythm tree. We show how to employ melodic operators to trace the progression of melodies in compositions, enabling analysts to understand the melodic relationship of the motivic and thematic material. By augmenting instead of replacing CMN, our research shows that these designs improve the accessibility and understandability of sheet music for people with varying musical and visual analytic expertise. Through active collaboration with musicologists and comprehensive qualitative user studies, our research assesses the applicability and effectiveness of the augmentation strategy. Building on these previous two foundational parts, the thesis proposes two approaches for sheet music analysis: Bottom-Up Composition Analysis (MusicVis) and Top-Down Corpus Analysis (CorpusVis). MusicVis enriches digital sheet music by extending the harmonic, rhythmic, and melodic analysis through flexible visual interactive queries and seamless transitions to abstract analysis levels. Conversely, CorpusVis facilitates macro-level exploration of sheet music collections. Qualitative user studies show that the proposed techniques support the interpretation and understanding of complex musical structures at different levels of abstraction while, for instance, scalability remains a limitation. The thesis concludes by reflecting on lessons learned, such as building on familiar concepts and separating independent data characteristics for analysis. We also address challenges encountered during our conducted Visual Sheet Music Analytics research, like transforming abstract musical concepts into intuitive visual representations while factoring in user requirements. Furthermore, we emphasize specific future research projects in Visual Musicology, such as integrating our proposed fingerprint designs into AR platforms, thus offering a hands-on approach for musicians to interact with augmented sheet music while learning and playing their instruments.

Zusammenfassung

Durch die Integration visuell-interaktiver Datenanalyse auf Notenebene widmet sich diese Arbeit einem neuen interdisziplinären Gebiet: *Visual Musicology*. Diese Arbeit verbindet Informationsvisualisierung und Musikwissenschaft und ebnet den Weg für neue Methoden zur Analyse von Musikdaten, wobei der Schwerpunkt dieser Doktorarbeit auf dem digitalem Notenmaterial liegt. Die Western Common Music Notation (CMN) ist unter Musikwissenschaftlern, Pädagogen, Studenten und Musikern unumstritten. Dennoch stellt das komplexe Design der CMN eine Herausforderung für die Interpretation großer Notensammlungen dar, insbesondere für diejenigen, die keine Visualisierungsexperten oder Datenanalysten sind. Wir adressieren diese Herausforderungen mithilfe neuer visueller Analysemethoden, die über den Rahmen der bisherigen Forschung hinausgehen. Der erste Grundpfeiler ist die Visual Musicology, in der wir den Visual Musicology Graph und das Methodology Transfer Model als organisatorische Methoden einführen, die die interdisziplinäre Forschungszusammenarbeit fördern. Sie ermöglichen interdisziplinäre Lösungen, wie anhand verschiedener musikwissenschaftlicher Szenarien wie der Embodied Music Interaction gezeigt wird. Wir untersuchen auch bestehende Musiknotationstechniken und entwickeln neue Methoden, die CMN mit fortgeschrittenen Visualisierungen verbinden. Anschließend werden Visualisierungsdesigns zur Erweiterung digitaler Noten mit harmonischen und rhythmischen Fingerprints vorgestellt, die auf domänenspezifischen Konzepten basieren, darunter der Quintenzirkel und der Rhythmusbaum. Wir zeigen, wie man melodische Operatoren einsetzt, um den Verlauf von Melodien nachzuvollziehen und die melodische Beziehung des motivischen Materials zu verstehen. Unsere Forschung zeigt, dass diese Designs die Zugänglichkeit und Verständlichkeit von Noten für Menschen mit unterschiedlichen musikalischen und visuellen Analysefähigkeiten verbessern, indem sie die CMN ergänzen, anstatt sie zu ersetzen. Durch die aktive Zusammenarbeit mit Musikwissenschaftlern und qualitative Nutzerstudien untersucht diese Arbeit die Effektivität der Erweiterungsstrategie. Aufbauend auf den beiden vorangegangenen Grundlagenteilen werden zwei Ansätze für die Analyse von Musiknoten vorgeschlagen: Bottom-Up Kompositionsanalyse (MusicVis) und Top-Down Korpusanalyse (CorpusVis). MusicVis bereichert digitale Noten, indem es die harmonische, rhythmische und melodische Analyse durch flexible visuelle interaktive Abfragen und nahtlose Übergänge zu abstrakten Analyseebenen ermöglicht. Umgekehrt erleichtert CorpusVis die Erkundung von Notensammlungen auf Makroebene. Qualitative Nutzerstudien zeigen, dass die vorgeschlagenen Techniken das Verständnis komplexer musikalischer Strukturen auf verschiedenen Abstraktionsebenen unterstützen, während die Skalierbarkeit der Visualisierungen und Algorithmen im Hintergrund in Zukunft weiter verbessert werden muss. Die Arbeit reflektiert die gewonnenen Erkenntnisse, wie die Verwendung bekannter Konzepte und geht auch auf die Herausforderungen ein, auf die wir während der Visual Sheet Music Analytics-Forschung gestoßen sind, wie die Umwandlung musikalischer Konzepte in intuitive Visualisierungen. Wir heben auch spezifische zukünftige Forschungsprojekte in der visuellen Musikwissenschaft hervor, wie z.B. die Integration unserer vorgeschlagenen Fingerabdruck-Designs in AR-Plattformen, um Musikern einen praktischen Ansatz für die Interaktion mit erweiterten Noten zu bieten, um den Instrumentenlernprozess zu optimieren.

Contents

Acknowledgements **v**

Abstract **ix**

1 Introduction **1**
1.1 A Motivation for Sheet Music Analytics 1
1.2 Research Objectives 3
1.3 Structure and Contributions of the Thesis 4
1.4 Publications and Contributions 10

I Visual Musicology Foundations **17**

2 Framing Visual Musicology through Methodology Transfer **19**
2.1 Introduction 20
2.2 Background 21
2.3 Methodology Transfer 22
2.4 Visual Musicology 23
2.5 Use Cases 24
2.6 Research Implications and Opportunities 25
2.7 Summary 27

3 Analyzing Visual Mappings of Music Notations **29**
3.1 Introduction 30
3.2 Background 31
3.3 Music Notation Design Space 32
3.3.1 Musical Meta-Features and Notation Dimensions 32
3.3.2 Users and Tasks 34
3.3.3 Visual Mapping of Musical Dimensions 35
3.4 Traditional and Alternative Music Notations 35
3.5 Research Implications and Opportunities 36
3.6 Summary 37

II Sheet Music Analysis for Harmony, Rhythm, and Melody **39**

4 Augmenting Music Sheets With Harmonic Fingerprints **41**
4.1 Introduction 42
4.2 Related Work 43

4.2.1 Music Harmony Visualization . . . 44
4.2.2 Annotating the CMN Score . . . 45
4.3 Harmony Fingerprint Design . . . 48
4.3.1 Circle of Fifths . . . 48
4.3.2 Design Rationale . . . 48
4.3.3 Creating the Glyph . . . 49
4.4 Adding Harmonic Fingerprints to the Score . . . 50
4.5 Evaluation . . . 51
4.5.1 User Study Methodology . . . 51
4.5.2 Use Cases . . . 53
4.5.3 User Feedback . . . 53
4.6 Discussion and Lessons Learned . . . 56
4.6.1 Educational Aspects . . . 56
4.6.2 Staff Separation . . . 56
4.6.3 Harmonic Invariance . . . 57
4.6.4 Limitations and Future Work . . . 57
4.7 Chord Identity Stability in Rotational Transformation . . . 58
4.8 Summary . . . 61

5 Augmenting Sheet Music With Rhythmic Fingerprints 63
5.1 Introduction . . . 64
5.2 Related Work . . . 65
5.2.1 Music Structure Analysis . . . 65
5.2.2 Radial Hierarchy Visualization . . . 65
5.2.3 Rhythm Visualization . . . 66
5.2.4 Augmenting Sheet Music . . . 67
5.3 Rhythmic Fingerprint Design . . . 68
5.3.1 Characteristics of Rhythm . . . 68
5.3.2 Design Rationale . . . 69
5.3.3 Extending the Score With Rhythmic Fingerprints . . . 71
5.4 Use Cases . . . 71
5.4.1 Interpretation of Rhythm . . . 71
5.4.2 Comparative Analysis of Rhythm . . . 72
5.5 Evaluation . . . 72
5.5.1 Study Methodology . . . 74
5.5.2 User Feedback Results . . . 75
5.6 Discussion . . . 76
5.6.1 Lessons Learned . . . 76
5.6.2 Limitations and Future Work . . . 77
5.7 Summary . . . 79

6 MelodyVis: Visual Analytics for Melodic Patterns in Sheet Music 81
6.1 Introduction . . . 82
6.2 Background and Related Work . . . 84
6.2.1 Sheet Music Retrieval Based on Melodic Sequences . . . 84
6.2.2 Melodic Pattern Mining and Extraction . . . 85
6.2.3 Applying Methodology Transfer . . . 86

6.3 Structure of Melodic Patterns . . . 87
6.3.1 Atomic Melodic Operators . . . 88
6.3.2 Complex Chaining of Melodic Operators . . . 91
6.3.3 Analysis Tasks and Requirements . . . 92
6.4 MelodyVis – Visual Interface for Melody Analysis . . . 93
6.4.1 Data Preparation . . . 93
6.4.2 Interaction Components . . . 93
6.4.3 Visual Interaction Pipeline . . . 95
6.5 Use Cases . . . 96
6.5.1 The Four Seasons: Spring – by Antonio Vivaldi . . . 96
6.5.2 BWV 846: Prelude in C-Major – by J. S. Bach . . . 97
6.5.3 Tetris Theme – by Nikolai Alexejewitsch Nekrassow . . . 98
6.6 Online User Study of MelodyVis . . . 98
6.6.1 Study Results . . . 99
6.7 Discussion . . . 101
6.7.1 Take-Home Messages . . . 101
6.7.2 Limitations and Potential Extensions . . . 102
6.7.3 Open Research Opportunities . . . 102
6.8 Summary . . . 103

III Multi-level Sheet Music Analysis Techniques 105

7 Bottom-Up | Augmenting Sheet Music through Visual Analytics 107
7.1 Introduction . . . 109
7.2 Background and Related Work . . . 110
7.2.1 Visualization Techniques for Sheet Music Analysis . . . 110
7.2.2 Methodology Transfer . . . 111
7.3 Problem and Data Characteristics . . . 112
7.3.1 Musical Features . . . 112
7.3.2 Hierarchical Music Structure . . . 113
7.3.3 Single and Multi Composition . . . 113
7.3.4 Analysis Tasks . . . 114
7.4 Technique: Augment CMN for Multi-Level Analysis . . . 115
7.4.1 Embedding Glyph Visualizations into Sheet Music . . . 115
7.4.2 Analysis Abstraction Layers . . . 117
7.5 Instantiation: MusicVis . . . 118
7.5.1 Feature Visualization . . . 119
7.5.2 Visual Querying . . . 120
7.5.3 Querying Harmony and Rhythm . . . 120
7.5.4 Querying Melody . . . 122
7.5.5 Structure Management View . . . 123
7.5.6 Initial Expert Feedback . . . 123
7.5.7 Technical Details . . . 123
7.6 Exemplary Analysis Workflow . . . 124
7.7 User Study . . . 125
7.7.1 Study Design . . . 125

7.7.2 Study Results . . . 127
7.8 Discussion & Summary . . . 128

8 Top-Down | CorpusVis: Visual Analysis of Music Collections 131
8.1 Introduction . . . 132
8.2 Related Work . . . 133
8.3 Problem Description . . . 135
8.4 Visual Interactive Analysis Workspace . . . 138
8.4.1 CorpusVis – User Interface . . . 138
8.4.2 Analysis Use Cases . . . 143
8.5 Pair Analytics Study . . . 144
8.5.1 Study Design . . . 145
8.5.2 Evaluation Results . . . 147
8.6 Discussion . . . 149
8.6.1 Limitations . . . 150
8.6.2 Take-Home Messages . . . 151
8.6.3 Future Work and Research Opportunities . . . 152
8.7 Summary . . . 152

9 Conclusion and Future Perspectives 155
9.1 Summary of Research Contributions . . . 155
9.2 Lessons Learned and Experiences . . . 158
9.3 Future Directions and Outlook . . . 162
9.4 The Value of Visual Musicology . . . 165
9.5 Concluding Reflections and Final Remarks . . . 166

A Appendix 169
A.1 Systematic Investigation of the Harmonic Geometry . . . 170
A.1.1 Compact Harmonic Fingerprints Summary Overview . . . 172
A.1.2 Harmonic Fingerprints for the Major Scale Degree Chords . . . 174
A.2 Systematic Exploration of the Rhythmic Fingerprints . . . 178
A.3 Comparing Aggregated Feature Glyphs . . . 180

List of Figures 185

List of Tables 187

Nomenclature and Acronyms 189

Bibliography 191

1 Introduction

> *Music is a higher revelation than all wisdom.*
>
> **— Ludwig van Beethoven**
> German Composer

The first chapter of this thesis introduces the research objectives I am addressing as a primary goal. It provides an overview of the problems by motivating why I postulate that more research is needed at the intersection of musicology and information visualization. The introduction serves as a backbone for the remainder of this work by presenting the overall contributions of the separate chapters. In this chapter, I also list all relevant and used publications in this thesis while clarifying which parts of each work I contributed as an author to each paper.

1.1 A Motivation for Sheet Music Analytics

Music is a universal language that people of all cultures have enjoyed for centuries [128]. It is an effective means to evoke emotions, from joy to sorrow, excitement to relaxation [52]. Because of these characteristics, music is used in almost any area of our lives that we can think of. To name a few examples, these areas range from entertainment [12], background music of advertisements [355], using it for enhancing concentration in educational contexts [18], reflecting the temporal evolution of cultural self-awareness, to critical but often also satirical confrontation with the content of political developments and agendas [250]. In summary, music is also a form of art that people can use to express themselves, connect with others, tell stories, and positively impact health in society [6].
Typically, when we mention the term "music", we refer to the auditory part we can hear and experience when listening to music [334]. But in addition to the auditorial aspects, music can also be described by its visual notation counterpart, often referred to as **"Sheet Music"** [85]. Traditionally, *sheet music* means written or printed music on paper. Within this thesis, we use the term with an extended meaning by adding "digital" to make a direct connection to its digital counterpart. Sheet music enables the visual representation of sound-based characteristics, which includes the most important musical features, such as notes comprising harmony [204] and rhythm [326] and the timbre of the instruments [261] that make up a piece of music [213]. Thus, sheet music could also be considered as the detailed recipe of how musical compositions are composed, providing musicians and composers a means of communication to produce and share music [240]. In this thesis, we discuss the encoding between the visual and the auditorial characteristics [281] and how this relates to information visualization in general. Finding a faithful and suitable mapping between these two realms is a hard challenge, which is already a century-long process that has influenced the type of visual encodings. One example

of this visual encoding is the Common Music Notation (CMN), which has evolved over a long period based on the emergence of new instruments and the decrease of older instruments that are no longer used or played nowadays as newer devices replaced them. Besides musical notation or musical scores, other visualizations that reflect musical content could also be considered abstract representations of music, such as diagrams or illustrations. For example, Wattenberg uses an Arc Diagram visualization to visually reveal the structure within sequence data, such as biological genome data and musical scores [346].

Music's auditory and visual parts are fundamental for its understanding and appreciation [164]. The auditory part provides music's emotional and perceptive impact, while the visual part provides a structural understanding of music, facilitating an overview of musical compositions [346]. Therefore, the auditory and visual aspects must be considered together to achieve a more complete musical experience. However, audio has achieved much more attention than its visual counterpart, sheet music based on CMN. This is because large companies such as Spotify and Deezer offer music services to a much larger target audience. On the contrary, the group of people interested in reading music and wanting to understand it better or even play it for themselves, such as musicians, music experts, and musicologists, is much smaller. Despite that, services such as OKTAV [169] emerge and target, e.g., musicians to provide sheet music as CMN.

Consequently, in this thesis, I focus on the research on sheet music, the visual notation of music, and how to use information visualization to improve our understanding by enabling the visual analysis of music in this format. As a major contribution of this thesis, I explicitly frame the research area of **"Visual Musicology"** at the intersection of musicology and information visualization [229], a growing field that encompasses the study of the potential of visual analytics for musicology. To achieve this, I systematically address multiple characteristics of musical data, such as harmony [222], rhythm [103], and melody [223], in addition to analysis tasks, considering the needs of potential users and analysts. The overarching research objectives in this area consider using visualization techniques to explore and analyze musical data and create new ways of experiencing and understanding music regarding the needs of music analysts, readers, and other people interested in music. As a result, I postulate that *visual musicology* has the potential to bridge the gap between musicology and visualization and to promote the use of visualization in music education and research. I am also convinced that continuing research in this overlapping research area offers the potential for both musicologists who benefit from technological advances and information visualization researchers who can work on musicological challenges by developing new visual methods and techniques supporting the field of musicology.

Hence, in this thesis, we will focus on the visual part of music and explore how visualization can improve our understanding and appreciation of music. I will present visualization techniques that can be used to investigate, explore, and analyze sheet music based on fundamental musical low- and high-level features. Additionally, we will discuss the results of qualitative user evaluations to demonstrate the benefits and drawbacks of these techniques. We will also regard the potential of visual musicology to bridge the gap between musicology and information visualization and promote visualization in music education and research.

1.2 Research Objectives

In this thesis, my primary aim is to investigate how information visualization research can be beneficial for the field of musicology with a specific focus on sheet music. I use the format of (digital) sheet music, which builds on Common Music Notation (CMN) as a starting point since it is the most basic visual format and the natural connecting point between these research areas.
To pursue this, I address the overarching research question:

How can we employ visual analytics to support the understanding and investigation of sheet music?

I subdivided this thesis into several building blocks that build on top of each other or even influence each other to some extent.

Visual Musicology – The first challenge comprises the explicit consideration of musicology with all its subdomains, data characteristics, as well as analytic and visualization tasks. This part helps identify and understand the connections between the field of musicology and information visualization, which also enables us to identify and transfer possible existing solutions for problems of other domains to musicology or the opposite direction. Understanding the data characteristics and musical specifics of sheet music and existing alternative music notations is necessary to compare the advantages and drawbacks of existing approaches, revealing suitable visual mappings and the potential use of Gestalt laws and visual variables for visual designs that help the analysis of sheet music data.

Visual Analysis of Harmony, Rhythm, and Melody – The complexity of the simultaneous encoding of multiple musical features, such as harmony, rhythm, and melody, complicates the analysis and understanding of sheet music. Therefore, as a second step, we examine employing alternative visual designs that facilitate the separate analysis of harmony, rhythm, and melody, which hold the potential for upscaling the analysis processes due to the abstract nature of, e.g., the harmonic and rhythmic fingerprints. We consider the melodic feature to be intertwined with harmony and rhythm, which is why abstraction is more challenging.

Bottom-up and Top-down Composition Analysis – The final challenge addresses bottom-up and top-down approaches, which build on the foundation, that typical domain users must be able to connect abstract visualization and statistical features with the familiar music notation format to understand more abstract visualizations of the underlying sheet music data. The primary objective is to use the result from the previous challenge to augment existing and well-known sheet music encoding with the newly introduced visual designs based on domain-specific foundations such as the harmonic circle of fifths or the rhythm tree. Finally, interconnecting multiple levels of abstraction and aggregation of diverse musical features enables a thorough investigation of single musical compositions or even sheet music collections.
As a result, I present several approaches, methods, visual designs, and user evaluations to address the challenges mentioned before.

1.3 Structure and Contributions of the Thesis

This section outlines the structure and main contributions of my thesis based on the research question introduced in Section 1.2. As depicted in Figure 1.1, I organized it into three major parts comprising two chapters in Part I and Part III (Chapter 2 & 3, Chapter 7 & 8) and three chapters in Part II (Chapter 4, 5, & 6).

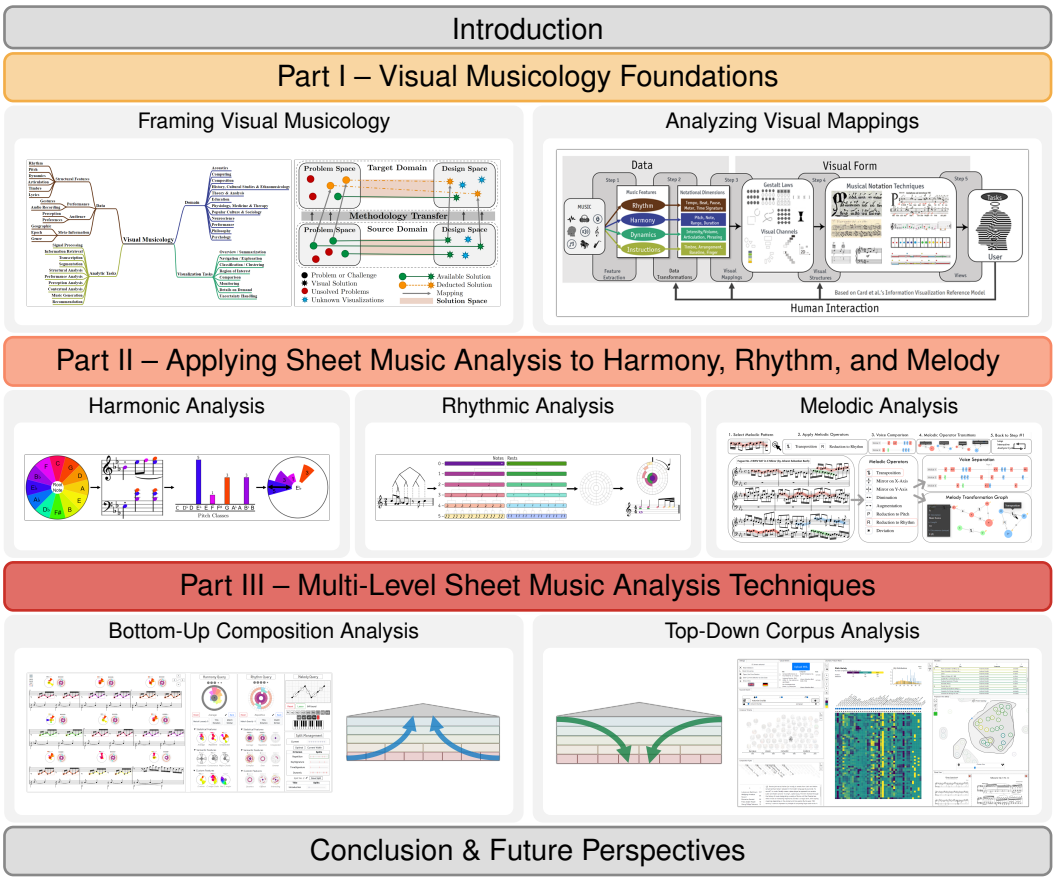


Fig. 1.1.: This thesis is structured into three major parts, with nine chapters in total. Part I introduces *Visual Musicology* and its foundations. We cover the characteristic musical features of harmony, rhythm, and melody in Part II. Part III contributes the investigation for two opposite analysis directions, bottom-up and top-down, by presenting two approaches demonstrating the use of Visual Analytics for CMN.

Overall, this thesis contributes a specification of the interface between information visualization and musicology by describing typical tasks of both domains and where they meet. For this, the Visual Musicology Graph (VMG) (illustrated in Figure 2.1) supplies these details and provides a framework. I use it in this thesis by providing an overview figure at the start of each chapter to categorize and embed my research content into this framework by explicitly stating which VMG branches and leaves I address with each work. For instance, Figure 1.2 shows how the four main branches are used to extract the four major categories *Data Characteristics*, *Domains*, *Analytic Tasks*, and *Visualization Tasks* for Chapter 3. Analogously, I use the same boxed representation to maintain the connection to the VMG and the underlying embedding for each chapter throughout my thesis.

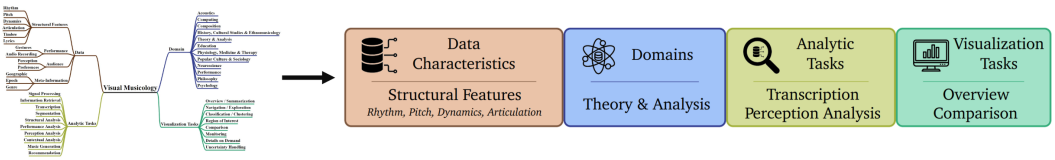


Fig. 1.2.: I use the *Visual Musicology Graph* (see more details in Figure 2.1) as a framework to classify and embed each chapter (except Chapter 2, which introduces the framework itself) into the field of visual musicology. This figure shows the extracted keywords exemplarily for Chapter 3. Similarly, the VMG classification for the remaining chapter is provided at the start of each chapter.

Chapter 1

The first chapter introduces the main research topic *Visual Sheet Music Analytics* of this thesis and contains a motivation for interdisciplinary research at the intersection of information visualization and musicology with a specific focus on visual analytics. Additionally, the introduction presents research problems, an overview of the major contribution of this work, such as the visualization designs for musical features, and lists all relevant publications used in this thesis accompanied by a detailed description of which parts I contributed to each work.

Part I – Visual Musicology Foundations

The two chapters in the first part describe the required foundations that span the remainder of this work. The first part establishes the bedrock for an interdisciplinary exploration and consideration of musicology and information visualization by merging them into the combined intersecting field "Visual Musicology". The encompassing research question (RQ) of the first part is

Research Question – Part I
*What are the **synergistic potentials** at the **intersection of musicology and information visualization**, and how can they be realized?*

Chapter 2

The second chapter institutes the framework of *Visual Musicology* at the intersection of information visualization and musicology, highlighting the potential for interdisciplinary research and motivating mutual synergies between researchers of both fields. Based on the research question

Research Question: Part I – Chapter 2
*What are musicology's **data, task, and domain characteristics**, and **how do they relate** to information visualization?*

the chapter gives a comprehensive background on the diverse application domains within musicology. A major contribution is the Visual Musicology Graph (VMG), which is accessible as an interactive website and enables interested readers to explore more details about each branch at `https://visual-musicology.com/graph/`. The chapter also introduces the general Methodology Transfer Model (MTM) that

can be applied to transfer existing solutions from other domains, such as text or geospatial analysis, to musicological problems or the opposite direction, identifying similarities between the data and task characteristics. By clearly describing the problem space, considering data and domain challenges, and the design space comprising the visualization and analytic tasks, we can elicit a solution space encompassing potential solution transfer between two fields. Eventually, the chapter contains multiple use cases to demonstrate how this methodology transfer can be executed by considering tasks such as the visual analysis of harmony, embodied music interaction, and the investigation of musical features. The chapter also contributes seven research implications and opportunities within the field of Visual Musicology, thus highlighting the potential of technological progress to address complex data challenges, further encouraging interdisciplinary research work. Chapter 2 lays the groundwork by establishing the framework of Visual Musicology, exploring the interdisciplinary synergy between information visualization and musicology.

Chapter 3

Building on the foundation from Chapter 2, Chapter 3 delves into the *visual design space of music notation,* examining existing techniques and extending our understanding of how these visual methodologies enhance our interpretation and creation of music. This natural progression from theoretical framework to practical application exemplifies the thesis's exploratory journey through Visual Musicology.

Research Question: Part I – Chapter 3

*What is the **visual design space** of music notation?*

Chapter 3 contributes an investigation of existing music notation techniques considering foundational information visualization concepts, including Gestalt laws and visual channels of human perception. This chapter presents a *Music Notation Visualization Pipeline* that can be used to understand and compare existing music notation techniques but also provides a structured scope to design new ones, catering them to different target groups such as composers or musicians. Such processing and transformation pipelines, originating from data and progressing to visual representations, are a common strategy to illustrate data transformation processes within the field of *Information Visualization* (InfoVis). Besides music feature extraction and data transformation on the data processing side, the pipeline also provides visual mappings onto visual structures to lay out music notation techniques. Eventually, an interactive loop illustrates that human users could be given flexible options to manipulate a visual design to their needs. As a result of the pipeline, an overview table of the assignment of existing music notations shows tendencies of prominent and well-known visual encodings. The table also reveals potential gaps for unexplored visual design mappings that may open up research opportunities. The comparison of the 15 different music notation designs provides benefits and drawbacks of the varying approaches, revealing a bias toward the classical CMN, the most dominant visual representation in musicology.

Chapter 3 completes the first part and provides the required foundation for the second part, which dives into investigating the salient features of sheet music: harmony, rhythm, and melody. The remainder of the doctoral thesis reuses and adopts the discussed task and data characteristics presented in *Visual Musicology Foundations*.

Part II – Sheet Music Analysis for Harmony, Rhythm, & Melody

The second part of my thesis focuses on the three musical aspects: harmony, rhythm, and melody. Altogether, within this part, I pursue the RQ

Research Question – Part II

*How can visual analytics techniques **enhance and augment** the analysis of **harmony, rhythm, and melody** in sheet music?*

For each of the three topics, I present a separate chapter introducing new visual designs inspired by domain knowledge that enable visual interactive sheet music analysis through augmentation, juxtaposition, or visual overlays by extending the sheet music accordingly. I evaluated the presented approaches through qualitative user studies to demonstrate their advantages and drawbacks.

Chapter 4

The fourth chapter deals with the musical aspect of harmony. I introduce a visual design grounded on the circle of fifths, a fundamental musical concept.

Research Question: Part II – Chapter 4

*How can we use the **circle of fifths** as a visual metaphor for **analyzing harmony** by **augmenting** sheet music?*

In this chapter, we discuss how we can extract and separate harmonic information into a visual representation to reduce the complexity of other visual information typically simultaneously encoded in CMN. We employ the foundational design of the circle of fifths to extract weighted statistics from notes that occur in sheet music. This leads to a radial visual design similar to a pie or Nightingale chart [36] that encodes harmonic information, which we call *harmonic fingerprint*. Then, juxtaposing the harmonic fingerprints on top of the sheet music enables analysts, learners, and music readers to understand the connection between both representations and focus only on the harmonic information when needed. While we start the consideration by extracting single fingerprints from single measures, we leverage the same visual abstraction in Chapter 7 to aggregate neighboring measures. This strategy allows analysts to get a harmonic overview of compositions at different levels of abstraction and aggregation. In this chapter, we demonstrate how harmonic fingerprints can be used by both novices and experts who can identify harmonic patterns, even if they are not literate enough to read and understand the familiar CMN accurately.

Chapter 5

Similar to Chapter 4, we consider the rhythmic or temporal dimension encoded in CMN. Analog to the circle of fifths for harmonic characteristics, we employ the rhythm tree concept, which builds on music's repetitive rhythmic nature.

Research Question: Part II – Chapter 5

*What visual designs **support the rhythmic analysis** considering the **temporal** and **repetitive dimension**?*

For the evaluation of the proposed *rhythmic fingerprints*, we follow the same evaluation approach as we already introduced for the harmonic fingerprints. The temporal

dimension of musical information encoded in sheet music provides special challenges regarding simple and compound meters in combination with varying beats. For instance, exploring multiple simultaneous voices poses a severe challenge when aggregating the common rhythmic aspects, as it often leads to irreversible visual representations. Depending on the musical tasks, this could be problematic if the analyst needs to understand the underlying details. Therefore, we discuss the limitations and demonstrate the benefits and drawbacks of the different visual designs.

Chapter 6

While the previous two chapters addressed the two orthogonal musical dimensions of harmony and rhythm, we must also pay attention to the combination of pitch and rhythm. Therefore, in Chapter 6, we have a detailed consideration of how the melodic material evolves based on sheet music by following the RQ

Research Question: Part II – Chapter 6

*How can we **employ visual analytics** to enable the **semi-automatic investigation of melodic patterns** based on sheet music?*

To do so, I propose eight melodic operators to understand the similarity and interrelationship between melodic themes. I present the prototype MelodyVis, which provides an interactive analysis workflow that enables melodic pattern exploration in combination with a *Voice Separation View* and a *Melodic Transformation Graph*. Three use cases showcase how my approach supports melodic analysis using CMN. In contrast to the harmonic and rhythmic fingerprints, we evaluated MelodyVis through an asynchronous web-based user study. The evaluation details are presented in this chapter, demonstrating the effectiveness of the operator-based strategy.

Part II concludes by combining the two orthogonal dimensions of pitch and rhythm, which have been separately addressed in Chapter 4 and 5. The harmonic and rhythmic fingerprints and the melodic analysis provide the foundation for Part III.

Part III – Multi-level Sheet Music Analysis Techniques

While Part II is limited to a separate investigation of harmony, rhythm, and melody, Part III progresses by combining the previous chapters into a single technique. Thus, enabling the integrated exploration of diverse characteristics to support the structural investigation of CMN. Part III's pursues a bottom-up and top-down approach in dealing with sheet music analysis from opposite directions.

Research Question – Part III

*How can **interactive visualization** techniques facilitate a bridge between **individual musical piece** analysis and **corpus-level exploration** in digital sheet music, enabling **multi-level analysis** in musicology?*

While understanding single compositions to see how a composer designed a musical piece (Chapter 7), musicologists also want to view and investigate a collection of musical pieces, which requires a higher level of abstraction (Chapter 8). Therefore, I discuss both analysis direction approaches separately and describe their common interface to show how they mutually benefit from each other.

Chapter 7

To start with the bottom-up composition analysis, I first present our MusicVis technique, which employs the harmonic and rhythmic fingerprints as an intermediate step for seamless transitions between low-level investigation at the sheet level to high-level exploration of patterns on the project level using the MDS dimension reduction technique. The main RQ I am addressing in this chapter is

Research Question: Part III – Chapter 7

*How to facilitate **interactive analysis** of digital sheet music at different levels of **abstraction** and **aggregation**?*

While reusing the visual designs already introduced in Part II, I diligently prepared visual query methods on top of the harmonic and rhythmic fingerprint design, facilitating visual exploration based on templates and custom user interactions. Besides the visual queries at every stage of abstraction and aggregation, I developed a melodic query that is a highly effective tool for identifying transpositions and similar melodic patterns through a flexible input grid. The concepts of abstraction and aggregation are independently implemented and thus orthogonally controllable.

Chapter 8

The opposite top-down corpus analysis consideration is the focus in Chapter 8. In the previous chapter, the analysis entity is always a single composition, starting the analysis process at the sheet level, the familiar representation for music readers. In contrast, CorpusVis is designed to support the comparative and exploratory analysis tasks of larger sheet music datasets, addressing the following RQ

Research Question: Part III – Chapter 8

*How can music analysts **explore**, **investigate**, and **compare** sheet music **collections** using **interactive visualization**?*

The introduced visual analytics technique extracts many high-level musical features, providing statistical information about melodic, rhythmic, and pitch-related characteristics but also considers metadata such as composer, epoch, and composition type. Through an in-depth pair-analytics study, we gathered qualitative feedback from music analysts, which showed that the approach enables exploratory and confirmatory analyses. Thus, CorpusVis provides valuable insights into the differences between composers or types based on the material available in digital sheet music.

Part IV – Conclusion and Future Perspectives

Chapter 9 concludes this doctoral thesis by reflecting on the current status of *Visual Musicology* at the intersection of information visualization and musicology research. I intend this last chapter to review the contributions of my approaches by considering the respective research objectives. Moreover, I summarize my main contributions, discuss lessons learned, and emphasize promising future research opportunities. I also indicate how to continue research within the realm of *Visual Musicology*, as there are vast possibilities to promote interdisciplinary research.

1.4 Publications and Contributions

This section provides an outline of multiple journal and conference publications that I published during my time as a doctoral researcher. I was responsible for the writing process as well as the editing and revision of the content. Subsequently, I list all publications I included in this dissertation, delineating my contributions and those of my co-authors for each paper project separately. The order of the papers reflects their appearance in the chapters of this thesis.

Part I – Visual Musicology Foundations

[229] **Matthias Miller**, Hanna Schäfer, Matthias Kraus, Marc Leman, Daniel A. Keim, and Mennatallah El-Assady. **"Framing Visual Musicology through Methodology Transfer"**. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2019.* *(Chapter 2)*

Contribution clarification. Together with Mennatallah El-Assady and Hanna Hauptmann (née Schäfer), I identified the research question: "*What are the data, task, and domain characteristics of musicology, and how does it relate to Information Visualization?*" to focus on the intersection of information visualization and musicology. This publication resulted from a collaboration of multiple authors. It focuses on the problem and design space of musicology, addressing problem-design mappings and solutions. With this paper, we contribute *Visual Musicology*, addressing any aspect of musicology concerning visualization. To achieve this, we designed the methodology transfer model and show how it can be applied to transfer existing solutions from other domains to musicological issues. For instance, Marc Leman, as an expert on embodied music interaction, provided useful material for an exemplary use case. I set up the Visual Musicology Graph (VMG) and assigned all the respective characteristics to each branch, which resulted from multiple discussions with Matthias Kraus, Hanna Hauptmann, and Mennatallah El-Assady. Matthias Kraus helped assign the solution space to match the branches of the VMG to other domains and select suitable use cases. El-Assady gave useful feedback based on her experience in the digital humanities field, which helped to set up the publication structure. She also emphasized the importance of the paper's research implications. During the writing phase, she and Hanna Hauptmann supported me with the introduction and background section to achieve a seamless order of arguments. Daniel Keim provided valuable comments on the paper draft and the abstract. Overall, I wrote and edited all sections, which is why I use them without further citation marks in Chapter 2.

[225] **Matthias Miller**, Johannes Häußler, Matthias Kraus, Daniel A. Keim, and Mennatallah El-Assady. **"Analyzing Visual Mappings of Traditional and Alternative Music Notation"**. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2018.* *(Chapter 3)*

Contribution clarification. This publication contributes a design and visualization pipeline for music notation as a multi-step process to transform musical data into visual structures. It summarizes existing music notations by mapping musical notation features to visual variables. To understand traditional

and alternative music notations, I devised the idea of applying a structured analysis to compare these notations and identify potential research gaps. Johannes Häußler and Matthias Kraus used the coding scheme from iterative joint discussions to classify fifteen music notation techniques. The initial paper idea was improved by Daniel Keim's comments, who also reviewed the paper throughout the drafting process. Additionally, Mennatallah El-Assady provided significant guidance on the paper's organization, structure, and content. As the project leader, I was responsible for the publication and authored all sections. I repeatedly edited and updated every paper section. Consequently, I reuse the paper's text in Chapter 3 without further citation indicators.

Part II – Sheet Music Analysis for Harmony, Rhythm, & Melody

[222] **Matthias Miller**, Alexandra Bonnici, and Mennatallah El-Assady. ***"Augmenting Music Sheets with Harmonic Fingerprints"***. *Proc. of the ACM Symposium on Document Engineering, 2019, pp. 1–10.* ***Best Paper Award***. *(Chapter 4)*

Contribution clarification. In cooperation with Mennatallah El-Assady, I identified the research question: "*How can we use the circle of fifths as a visual metaphor for analyzing harmony by augmenting sheet music?*". For this publication, I contribute the harmonic fingerprint design based on the circle of fifths and the implementation of the visualization augmentation for digital sheet music. I also developed the interactive web prototype that provides an example of the composition, of which I used a printed version for the user study. Together with Mennatallah El-Assady and Alexandra Bonnici, we designed the qualitative study. Alexandra Bonnici provided valuable connections to domain experts who participated in the study and with whom she completed study trials. Mennatallah El-Assady provided her experience in the digital humanities field regarding close and distant reading. My co-authors provided several comments and feedback on the paper draft, such as pointing out relevant literature that we considered in the paper. While I conducted the user study with half of all participants, I also wrote and revised all sections of the paper. Hence, I use all material without citation marks in Chapter 4.

[103] Daniel Fürst, **Matthias Miller**, Daniel A. Keim, Alexandra Bonnici, Hanna Schäfer, and Mennatallah El-Assady. ***"Augmenting Sheet Music with Rhythmic Fingerprints"***. *IEEE Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2020.* *(Chapter 5)*

Contribution clarification. In collaboration with the co-authors, I identified the research question: "*How can we augment sheet music with rhythmic fingerprints for an improved analysis of rhythmic patterns?*" Inspired by our previous work, "Augmenting Music Sheets with Harmonic Fingerprints," Daniel Fürst implemented a rhythmic fingerprint design as part of a web-based prototype. Daniel Fürst worked on this project as part of his bachelor thesis, which I supervised within the scope of this paper idea. Thus, he wrote the major content as the paper's first author. He also worked on the second alternative rhythmic fingerprint design, of which I added illustrative figures for comparison to the chapter but was not part of the referenced publication. The hierarchical radial design of the rhythmic fingerprint emerged from multiple design discussions

with co-authors. I provided supervision for the project and crucial feedback on the paper draft, as well as applying multiple edits and rewrites to the paper. I also supported the final design of all figures illustrated in the paper. Alexandra Bonnici offered high-level feedback on the draft and using the rhythmic fingerprint for the sight-reading of sheet music and for obtaining an overview of rhythmic patterns in digital sheet music. Hanna Hauptmann (née Schäfer) contributed with her comments on the paper draft during the writing process. Mennatallah El-Assady also provided guidance and feedback on the paper iterations and the content of the paper sections. During multiple joint discussions, she helped maintain the focus on the design rationale of the rhythmic fingerprint and discussed design alternatives. Overall, I contributed to successfully developing and evaluating the rhythmic fingerprint. Consequently, the material in Chapter 5 is built upon this paper, and I rewrote most of the section reflecting the collaborative efforts of the authors.

[223] **Matthias Miller**, Daniel Fürst, Maximilian T. Fischer, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. **"MelodyVis: Visual Analytics for Melodic Patterns in Sheet Music"**. *To be published. (Chapter 6)*

Contribution clarification. This paper resulted from a close collaboration between Daniel Fürst and me. I supervised his master's project and thesis, during which he implemented the research prototype under my supervision. Daniel Fürst was primarily responsible for implementing MelodyVis based on my initial research question "*How can we employ Visual Analytics to enable the semi-automatic investigation of melodic patterns based on sheet music?*". I managed the overall project and provided the main ideas, such as using melodic operators to identify relevant reoccurring patterns based on exemplary melodic patterns. Maximilian T. Fischer provided helpful feedback about the understandability of the concepts explained in the paper from an information visualization perspective. Daniel A. Keim helped revise the paper abstract and clarify the contribution of the paper. Hanna Hauptmann (née Schäfer), Maximilian T. Fischer, and Mennatallah El-Assady commented on the paper draft and helped revise and restructure the paper to facilitate the reading flow. I wrote all paper sections and revised them multiple times based on the co-authors' feedback. Hence, I will reuse the paper's content in Chapter 6. This paper has undergone a review cycle of the IEEE VIS 2023 conference but has received a major revision. Hence, we plan to resubmit this paper based on the reviewer feedback we already integrated in this chapter.

Part III – Multi-level Sheet Music Analysis Techniques

[224] **Matthias Miller**, Daniel Fürst, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. **"Augmenting Digital Sheet Music through Visual Analytics"**. *Computer Graphics Forum, Vol. 41, No. 1, 2022, pp. 301–316. (Chapter 7)*

Contribution clarification. To continue on the original ideas of the previous publications based on the harmonic and rhythmic fingerprint, I had the main idea to generalize these concepts as a visualization technique by pursuing the research question "*How to facilitate interactive analysis of digital sheet music at different levels of abstraction and aggregation?*". In an initial step, I

presented an early version of the interactive melodic and harmonic analysis as a late-breaking demo (poster) at the ISMIR conference in Delft in 2019 [228]. Afterward, I continued implementing the major parts of the application MUSICVIS that we contribute in this paper. MUSICVIS serves as an instantiation of the general concepts of providing abstract visualizations of musical features to a target audience who is only familiar with the CMN but not with more advanced visualization concepts. With the support of Mennatallah El-Assady, I framed the main contributions and the overall structure, and I prepared the required figures of the paper, such as considering this approach as a bottom-up approach starting from the well-known CMN and then seamlessly progressing to abstract representations. Hanna Hauptmann (née Schäfer) provided several ideas on the design rationale and the interplay of the visualization components. Daniel Fürst was primarily responsible for supporting the implementation of the application, such as preparing the seamless transitions between the abstraction layers and aggregating the glyph visualizations for the two layers of harmony and rhythm. Daniel A. Keim provided valuable feedback on the prototype's functionality, helped edit the paper abstract, and emphasized the paper's major contributions. Mennatallah El-Assady contributed her invaluable experience from the digital humanities field, which helped to integrate concepts such as close and distant reading that we applied on the multiple abstraction levels supporting different music analysis tasks. Eventually, I wrote the first draft and edited it multiple times based on the co-authors' comments. Therefore, I will use the publication with minor changes in Chapter 7.

[227] **Matthias Miller**, Julius Rauscher, Daniel A. Keim, and Mennatallah El-Assady. **"CorpusVis: Visual Analysis of Digital Sheet Music Collections"**. *Computer Graphics Forum, Vol. 41, No. 3, 2022, pp. 283–294.* *(Chapter 8)*

Contribution clarification. For this final publication I reuse in this thesis, I contributed the key design ideas for the CORPUSVIS visual analysis workspace leveraging sheet music collections. Besides gathering the used data set and defining the seamless connections between the prototype's components, I also selected the musical features. Based on this selection, Julius Rauscher prepared the similarity analysis and computed the data projections during the implementation phase of the prototype as part of his master project and thesis, which I supervised. He also conducted the study with nine participants based on a study design defined together. While I came up with the three discussed use cases based on existing literature, J. Rauscher made them easily accessible for reference in CORPUSVIS. Through multiple iterations, Mennatallah El-Assady helped complete the paper draft, providing guidance in structuring the sections and figures, increasing the overall understandability and reading flow. She also elaborated on the take-home messages relevant to the visualization community to make our contributions more explicit. Daniel A. Keim commented on the paper draft, provided valuable feedback on the paper abstract, and helped focus on the visualization contribution of our approach. Since I wrote all major sections and went through several editing iterations, I use the content of this publication with only minor modifications in Chapter 8.

In addition to these seven papers, I contributed as a co-author to 15 further publications. They only indirectly influenced my research, which is why I did not further include them as content in this thesis. Subsequently, I list these projects in chronological, descending order.

[91] Maximilian T. Fischer, Yannick Metz, Lucas Joos, **Matthias Miller**, and Daniel A. Keim. **"MULTI-CASE: A Transformer-based Ethics-aware Multimodal Investigative Intelligence Framework"**. *arXiv:2401.01955,* 2024.

[270] Julius Rauscher, Raphael Buchmüller, Daniel A. Keim, and **Matthias Miller**. **"SkiVis: Visual Exploration and Route Planning in Ski Resorts"**. *IEEE Transactions on Visualization and Computer Graphics,* 2023.

[76] Frederik L. Dennig, **Matthias Miller**, Daniel A. Keim, Mennatallah El-Assady. **"FS/DS: A Theoretical Framework for the Dual Analysis of Feature Space and Data Space"**. *IEEE Trans. on Visualization and Computer Graphics,* 2023.

[294] Victor Schetinger, Sara Di Bartolomeo, Mennatallah El-Assady, Andrew McNutt, **Matthias Miller**, João Paulo Apolinário Passos, and Jane Adams. **"Doom or Deliciousness: Challenges and Opportunities for Visualization in the Age of Generative Models"**. *Computer Graphics Forum, Vol. 42, No. 3,* 2022.

[90] Wolfgang Jentner, Fabian Sperrle, Daniel Seebacher, Matthias Kraus, Rita Sevastjanova, Maximilian T. Fischer, Udo Schlegel, Dirk Streeb, **Matthias Miller**, Thilo Spinner, Eren Cakmak, Matthew Sharinghousen, Philipp Meschenmoser, Jochen Görtler, Oliver Deussen, Florian Stoffel, Hans-Joachim Kabitz, Daniel A. Keim, Mennatallah El-Assady, Juri F. Buchmüller. **"Visualisierung der COVID-19-Inzidenzen und Behandlungskapazitäten mit CoronaVis"**. *In: Resilienz und Pandemie – Handlungsempfehlungen anhand von Erfahrungen mit COVID-19, Stuttgart, Kohlhammer, pp. 176–189. ISBN 978-3-17-039930-3,* 2022.

[148] Maximilian T. Fischer, Simon David Hirsbrunner, Wolfgang Jentner, **Matthias Miller**, Daniel A. Keim, and Paula Helm. **"Promoting Ethical Awareness in Communication Analysis: Investigating Potentials and Limits of Visual Analytics for Intelligence Applications"**. *Proceedings of FAcct '22: ACM Conference on Fairness, Accountability, and Transparency,* 2022.

[226] **Matthias Miller** and Daniel Miller. **"GraphStateVis: Interactive Visual Analysis of Qubit Graph States and their Stabilizer Groups"**. *IEEE International Conference on Quantum Computing and Engineering (QCE),* 2021.

[102] Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, **Matthias Miller**, and Daniel A. Keim. **"Teaching Clustering Algorithms With EduClust: Experience Report and Future Directions"**. *IEEE Computer Graphics and Applications (CG&A), Vol. 40, No. 2, pp. 98–102,* 2020.

[174] Matthias Kraus, Thomas Pollok, **Matthias Miller**, Timon Kilian, Tobias Moritz, Daniel Schweitzer, Jürgen Beyerer, Daniel Keim, Chengchao Qu, and Wolfgang Jentner. **"Toward Mass Video Data Analysis: Interactive and Immersive 4D Scene Reconstruction"**. *Sensors, Vol. 20, No. 18:5426,* 2020.

[173] Matthias Kraus, **Matthias Miller**, Juri Buchmüller, Manuel Stein, Niklas Weiler, Daniel Keim, and Mennatallah El-Assady. **"Breaking the Curse of Visual Analytics: Accommodating Virtual Reality in the Visualization Pipeline"**. *VISIGRAPP 2019: Computer Vision, Imaging and Computer Graphics Theory and Applications, pp. 253—284,* 2020.

[266] Thomas Pollok, Matthias Kraus, Chengchao Qu, **Matthias Miller**, Tobias Moritz, Timon Kilian, Daniel Keim, and Wolfgang Jentner. **"Computer Vision Meets Visual Analytics: Enabling 4D Crime Scene Investigation from Image and Video Data"**. *Int. Conf. on Imaging for Crime Detection and Prevention,* 2019.

[230] **Matthias Miller**, Xuan Zhang, Johannes Fuchs, and Michael Blumenschein. **"Evaluating Ordering Strategies of Star Glyph Axes"**. *IEEE Visualization Conference (VIS), pp. 91–95,* 2019.

[301] Daniel Seebacher, **Matthias Miller**, Tom Polk, Johannes Fuchs, and Daniel A. Keim. **"Visual Analytics of Volunteered Geographic Information: Detection and Investigation of Urban Heat Islands"**. *IEEE Computer Graphics and Applications (CG&A), Vol. 39, No. 5, pp. 83–95,* 2019.

[101] Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, **Matthias Miller**, and Daniel Keim. **"EduClust - A Visualization Application for Teaching Clustering Algorithms"**. *Eurographics (EG) - Education Papers,* 2019.

[41] Eren Cakmak, Udo Schlegel, **Matthias Miller**, Juri Buchmüller, Wolfgang Jentner, Daniel Keim. **"Interactive Classification Using Spectrograms and Audio Glyphs"**. *IEEE Conference on Visual Analytics Science and Technology (VAST Challenge MC1),* 2018.

Part I

Visual Musicology Foundations

2 Framing Visual Musicology through Methodology Transfer

> *Music expresses that which cannot be put into words and that which cannot remain silent.*
>
> **— Victor Hugo**
> French Romantic Writer and Politician



In this chapter, we frame the field of *Visual Musicology* by providing an overview of well-established musicological sub-domains and their corresponding analytic and visualization tasks. We discuss relevant data and domain characteristics to foster collaborative, interdisciplinary research. We give a concise description of the *problem space*, as well as the *design space* of musicology, and discuss how existing problem-design mappings or solutions from other fields can be transferred to musicology. Through *methodology transfer*, we argue that established methods can be exploited to solve current musicological problems and show exemplary mappings from analytics fields related to text, geospatial, time-series, and other high-dimensional data to musicology. Finally, we address open challenges, discuss research gaps, and highlight opportunities for future research, some of which are further considered in the remainder of this thesis. More information about future research direction is provided in the conclusion (see Chapter 9).

Chapter 2 is based on the following publication. Please refer to Section 1.4 for the detailed contribution clarification.

[229] **Matthias Miller**, Hanna Hauptmann (Schäfer), Matthias Kraus, Marc Leman, Daniel A. Keim, and Mennatallah El-Assady. **"Framing Visual Musicology through Methodology Transfer"**. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2019.*

2.1 Introduction

With its long history, music is an omnipresent element of cultures and a primary aspect of societal identity, deeply rooted in psychology, art, and entertainment [128, 355]. The current technological progress has a strong influence on music research, such as applying new methods for optical music recognition [271], music recommendation systems (e.g., Spotify [264]), and music information retrieval [153]. As a subfield of the humanities, *musicology* comprises any research-based analysis or study which is related to music [18]. In musicology, not only basic musical elements (e.g., rhythm or harmony) are of interest, but also *ethnomusicological* considerations including gestures and movement [18]. Music itself is only one aspect of musicology. Musicologists study a variety of topics such as genres [14], epochs [61], composers [209], and music theory [18], but also psychological effects and emotions [52]. Often, vast amounts of data and complex relationships within them require computational solutions to enable domain experts to extract interesting patterns.

To achieve this, Visual Analytics (VA) helps in the automatic extraction and processing of valuable information while considering the needs of analysts [156]. Incorporating the computational power enables us to deal with the complexity of large musicological data sources [345]. E.g., using large data sources increases the quality of music transcription [20] or improves music retrieval and recommendation [153]. In the information visualization community, many published approaches support domain-specific challenges, especially for analyzing text [178], time [4], geography [172], and high-dimensional data [195]. McNabb [217] provides an overview of all surveys conducted by the visualization community until 2017. A search by the keywords *music*, *sound*, *audio*, or *harmony* using the SoS Literature Browser yields no results [216]. Still, a few individual applications exist, such as the web service *Songle* proposed by Goto et al. [114] for automatic chord extraction to support active music listeners and musicians by utilizing voluntarily shared information to improve the automatic results.

Thus, we argue that music visualization, combined with visual analytics, is an emerging and under-researched field with many open research opportunities compared to other domains. In this work, we propose the term *Visual Musicology* to describe the interdisciplinary research field at the intersection of musicology and visual analytics. We argue that visual musicology will strongly benefit from transferring existing visualization techniques. Visualization researchers can contribute to unsolved challenges of the music domain, and the uniqueness of musicology offers them the opportunity to develop new visualization techniques and applications [33].

This chapter explains how to identify direct or indirect similarities between the music domain and other established visualization research domains, including text, geo, time, and high-dimensional data. Based on this knowledge, we introduce how to apply *methodology transfer* to map existing visual solutions to benefit from previous research work. Moreover, we recommend developing new music visualization techniques to cover current research gaps. We propose "*Visual Musicology*" to encapsulate information visualization research that addresses any aspect of music or musicology. It encompasses all design models and techniques in the visualization domain that provide a solution for problems related to musicology.

In this chapter, we first characterize the *problem* and the *design space* of *visual musicology* similar to TextVis [5]. Second, we describe parallels between the musicology problem space comprising data characteristics, users, domains, and goals and problem spaces of other application domains. Third, we divide the *visual musicology design space* into existing solutions and solutions that can be transferred or adapted from other domains, thus highlighting research gaps. Fourth, we conduct the *methodology transfer* process by presenting three exemplary use cases to emphasize the potential of transferring existing visual solutions.

We postulate that collaboration between musicologists and visualization researchers may lead to mutually beneficial synergies: Visualization researchers can design and evaluate new methods by solving complex and diverse tasks of musicologists. Musicologists can address previously unsolved questions in a new way. The primary purpose of this chapter is to motivate and initiate this kind of collaboration and to create a framework to embed the remaining chapters of this thesis.

2.2 Background

Musicology comprises any research effort covering aspects of music [160]. Thus, not only analysis of performance data such as audio, sheet music, or digital symbolic data such as MusicXML [112] and MIDI are of interest. Moreover, psychological aspects, such as the emotions of music audiences in addition to gestures and body

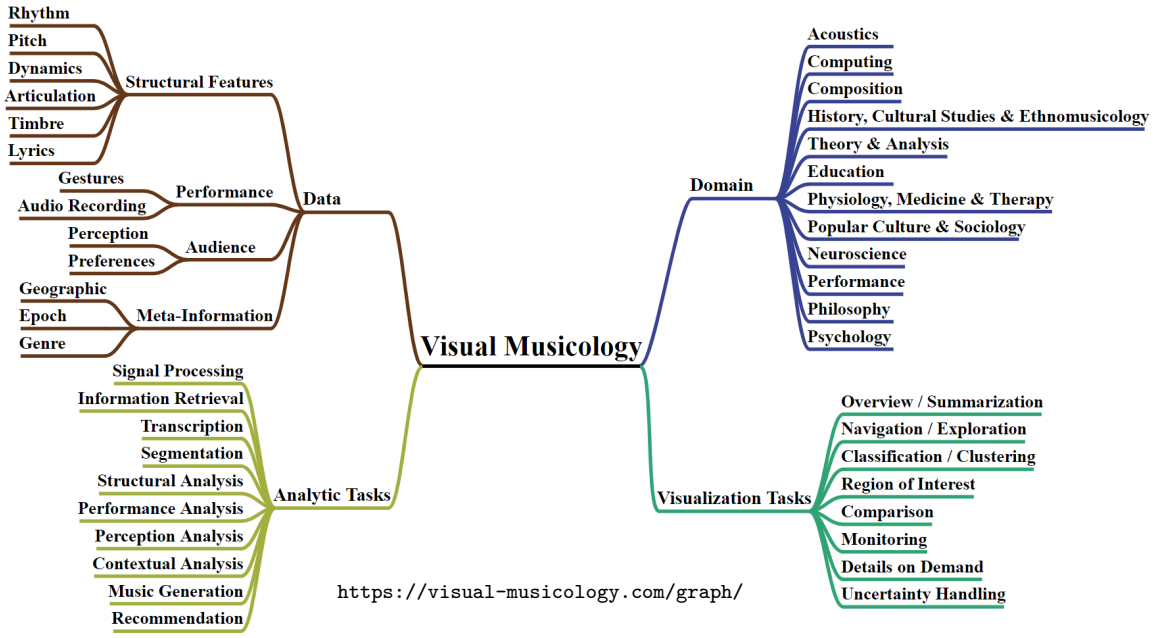


Fig. 2.1.: Visual Musicology encompasses different *data* types and various *domains* covering psychological, physical, physiological, and societal aspects. *Visualization tasks* can support musicologists in answering domain problems, which can be described by *computational tasks*. Strengthening cross-disciplinary and domain-driven research has the potential for original and synergistic research, paving the way for novel research collaborations between musicology and information visualization.

movement analysis of performing musicians, are relevant research areas within the realm of musicology. Consequently, musicology covers various application domains that we state in Figure 2.1.

We emphasize that the "Visual Musicology" graph in Figure 2.1 provides a diverse list of domains we extracted from the Journal of Interdisciplinary Music Studies [258]. The domain list is subject to extension since there are many specialized subdomains that we cannot consider within the scope of this work. For instance, besides music data, meta-information such as the places where composers lived and worked and how the styles of different composers influenced each other are typical tasks in the sub-domain of *history* or *cultural studies*. Nevertheless, the diversity of application domains leads to many complex challenges that require tailored solutions to cope with them. Not every sub-domain of musicology requires analytical techniques or is concerned about music experience [3]. While visualization is not a solution for every problem in the music domain, we argue that many open research gaps wait to be addressed by interdisciplinary researchers.

Music Visualization often is only concerned with analyzing structural features [350] in audio or performance data [356] or on the notation level [225]. Using information visualization methods and models for music data is not a new endeavor. More than two decades ago, Smith and Williams proposed a different way for music notation to visualize music [311]. Many alternative music notations have been proposed to deal with the legibility of the CMN, which primarily consists of the features *rhythm*, *harmony*, *dynamics*, *timbre*, and *articulation* [225]. In Chapter 3, I analyze multiple mappings in music notation from visual variables to musical features [225]. Besides alternative visual mappings of musical features, harmony [287] and structure analysis [346] are often the subject of visualization models [332, 190]. While the more obvious visualization models in the music domain are mainly rooted in direct features of music data, many more music-related topics are encompassed by *musicology*. Additionally, models were often blindly applied without clear musicological motivation or goals during the early stages of computational musicology [333]. At the Dagstuhl seminar, "Multimodal music processing" (11041), many topics regarding music processing of sheet and audio data were discussed by leading musicologists [237]. One of the results was a recommendation for interdisciplinary research efforts to improve the understanding of complex relationships between music and its effect on the human mind or body. Therefore, bringing music domain experts and information visualization experts into a conversation may yield further fruitful collaborations [307], resulting in valuable solutions for both sides.

2.3 Methodology Transfer

Methodology Transfer refers to the action of utilizing available models that provide solutions to existing and unsolved problems [40]. We propose a *Methodology Transfer Model (MTM)* which states that every domain contains a *problem space* and a *design space*, as depicted in Figure 2.2. Simon et al. [307] explain that problem spaces are defined by the *data* and *task* characteristics as well as *domain-specific user needs*. The design space comprises all visual techniques that serve as a solution for issues from the problem space. We consider the mapping between these two spaces as

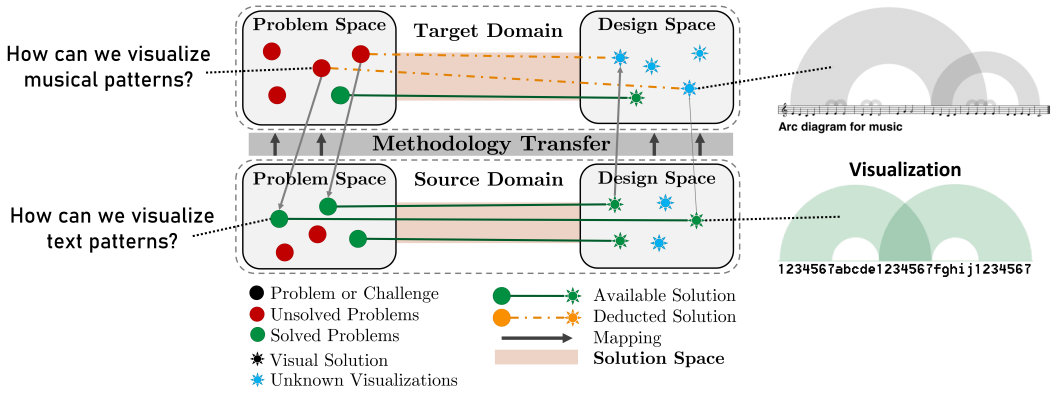


Fig. 2.2.: The methodology transfer model (MTM) maps problems and designs between two different domains to deduce new solutions.

the *solution space*. Visual models from the design space may be applied to different problems. At the same time, a single visualization may cover various tasks from the problem space. To cope with new challenges, we need to find still unknown problem-design mappings located in the *solution space*. Then, they can be transferred, e.g., to the *musicology domain*, by adapting solutions to comply with the requirements of the existing models. Figure 2.2 outlines our methodology transfer model inspired by Simon et al.'s *Liaison* [307]. Depending on how well the original solution aligned its problem with its design and how similar the problems in both domains are, the design must integrate more or fewer adaptations. The overall target of applying the MTM is to enlarge the solution space through available knowledge of every domain suitable for realizing methodology transfer.

2.4 Visual Musicology

We define the term *Visual Musicology* to cover all characteristics of the musicology and visualization domains relevant to building *Visual Musicology* systems. To conduct a methodology transfer as described in section 2.3, we first need to span the problem and design spaces of *Visual Musicology*.
In general, any visualization problem depends on the combination of domain users, data types, and tasks that users want to accomplish [221]. Figure 2.1 displays the core components of the *Visual Musicology* problem and design spaces, giving examples for each branch. These branches provide an in-depth description of the musicological interface for transferring to and from other domains. The problem space of *Visual Musicology* is described by the different application *Domains*, and their corresponding *Data*. For the domain branch, we consulted the research tracks suggested by the Journal of Interdisciplinary Music Studies (JIMS) [258]. For the data branch, based on initial feedback from a musicology expert, we extracted exemplary types of data to be analyzed in each of the different application domains of the graph. The design space of *Visual Musicology* is split into the design for *Analytic Tasks* and for *Visualization Tasks*. For the visualization branch, we adopted the visualization tasks from Kucher and Kerren [178], which are partly based on Shneiderman's information visualization mantra [306]. The analytic branch comprised the topics

covered by Weihs's book on music data analysis [347]. The solution space will be explored according to the methodology transfer (section 2.3) from previously established domain definitions such as time-oriented visualizations by Aigner et al. [4] as presented in the "TimeViz Browser" [324] and text visualization by Alharbi and Laramee [5] or by Kucher and Kerren's [178].

Next, we discuss the problem and design space within *Visual Musicology* and apply the MTM (Section 2.3) for the domains *Text*, *Geo*, *Time*, and *HD*.

2.5 Use Cases

We provide three different use cases of applying methodology transfer to derive a visual musicology solution.

(a) Visual Analysis of Harmony – Schroer [298] exploited the radial layout of the circos graph [246] to visualize the chord harmony in their respective tonal system of Western Art Music. Originally, Naquin et al. proposed the circos graph to enable visual analysis of structural genome variations. Therefore, Schroer was inspired by a visualization design for biological data to display dominant chord relationships by mapping tonal degrees to color. Figure 2.3(a) shows that multiple adaptations were applied to the original model, but the general idea is preserved. The results for single pieces of different epochs show a dominant chord in classical music (e.g., Mozart), highlighted in red. In contrast, Prokofiev's exemplary modern music piece contains an almost equal distribution of every degree. Consequently, this use case exemplifies successfully mapping an available problem - design space mapping.

(b) Embodied Music Interaction – Current research on *Embodied Music Interaction* [185, 191] focuses on understanding the complex relationship between music, performers, audience, and contexts. These aspects comprise a complex dynamic system with stable, in-between interaction states that change over time. Figure 2.3(b) shows how automatic analysis of video footage allows body recognition to analyze

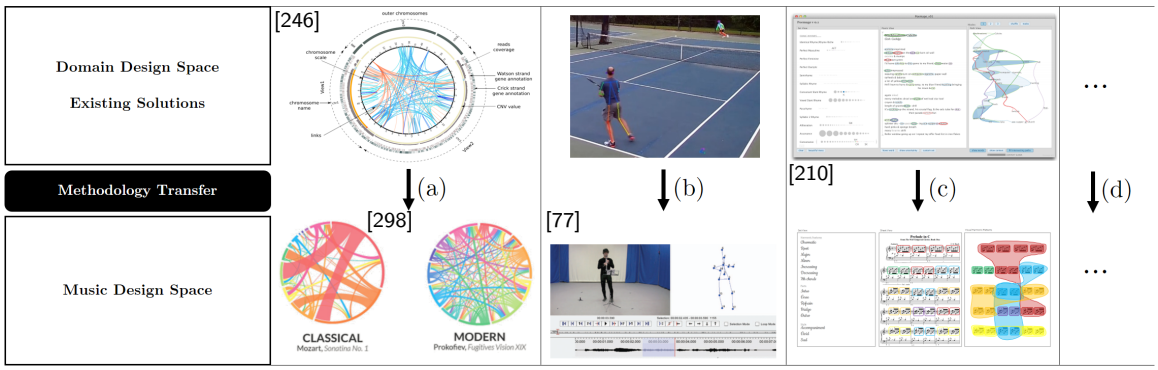


Fig. 2.3.: Exemplary use cases for Methodology Transfer from different domains into Visual Musicology: (a) Exploiting the Circos visualization to identify harmonic relationships. (b) Analysis of embodied music interaction: Transferring computational methods for tennis movements to elicit new information from musicians during a performance. (c) Mapping the idea of analyzing the typology of poems from Poemage [210] to analyzing the structure of musical features in music pieces.

the performance in tennis. Applying such methods to the body movement of a clarinet player [77] could help to identify recurring patterns of musicians' gestures. Music scores can be conceived as instruction sets for human actors whose gesturing underlies music playing. Of course, this use case requires looking at audio data, which might not be as relevant in tennis. Therefore, an additional component that supports such analysis tasks is necessary to relate body movement data with performance recordings.

(c) Investigating Musical Features – *Poemage* is a user interface to visualize the sonic topology of poems [210]. Linguists and other digital humanities researchers who analyze textual data can use Poemage for *close reading* to retrieve in-depth information about poems. With this application, text analysts can execute an exhaustive text investigation with the support of automatic highlighting and semantic relationships, such as rhymes by the use of sound in poetry. Visualizations help reveal hidden relationships through the employment of well-polished designs to cope with the challenges of textual close reading. We argue that Poemage is a suitable source of inspiration to apply similar or equal approaches on single music documents to enable *close reading* [145] based on different musical features such as the role of certain chords in their context or modulation. Figure 2.3 (c) indicates how the visual interface of Poemage can be adjusted to visualize the *harmonic topology of a piano composition*, for instance. Detail views and colored highlighting can be employed to emphasize interesting harmonic progressions, and a feature panel on the left could provide analysts with powerful tools to support in-depth analysis of musical pieces. Of course, the application must be adapted for musical data, and different methods to extract relevant information from the piece have to be integrated. Nevertheless, this application transfer could motivate music researchers to use such an application, since a similar approach has been successfully introduced to another digital humanities community.

We emphasize that there is still much potential for numerous new methodology transfer applications (see Figure 2.3(d)). To identify such research opportunities, close collaboration with domain experts and musicologists is required [307]. For instance, use case (b) describes a scenario that has not been applied yet and may serve as a starting point to apply body movement based on skeleton analysis models to the movement and gestures of musicians during performances. Of course, factors like cultural context and biological aspects are relevant for understanding the dynamics of music interaction such as in higher-level music training [48, 7], interactive multimedia art [341, 203], sports [200], and medical rehabilitation[232].

2.6 Research Implications and Opportunities

Our research introduces how musicologists and visualization researchers can initiate collaboration by applying *methodology transfer* for *Visual Musicology*. Furthermore, *Visual Musicology* provides countless opportunities for original research. In summary, we make seven claims designed to motivate musicologists and visualization researchers alike to increase mutual collaboration.

(1) Visual Musicology is more than visualizing music data – There is more to Visual Musicology than merely visualizing music. Figure 2.1 lists multiple musicology domains covering a diverse range of data characteristics and analytic tasks. For instance, *psychological, physiological, and philosophical* considerations are required to address relevant musicology topics and to understand the complexity of research questions in this domain. Often, data analysts only apply structural analysis on music data which we consider to be only a subfield of visual musicology reflected by *Theory & Analysis* and *Computing* as part of the domain branch in Figure 2.1.

(2) Visual Musicology is an under-researched field – Multiple challenges and problems that musicologists are facing still require special attention to find suitable solutions [236]. But even for existing solutions, information visualization research could lead to more effective and efficient results. For example, during Dagstuhl Seminar 16092, numerous possible research ideas regarding *Computational Music Structure Analysis* were discussed [236]. Musicology provides an extensive collection of diverse data sources. Different data analysis techniques, such as machine learning, e.g., with probabilistic models, and human-based analysis methods, e.g., visual analytics, can be applied and evaluated using these data sources.

(3) Musicology is expanding through technological progress – Technological advances, such as the development of virtual reality environments, the creation of new instruments, and the influence of music from different cultures due to the internet, lead to new considerations of musicological questions. The possibility of gathering data from brain activity in neuroscience enables research in different domains, which was impossible decades ago. *Psychological, physical, and physiological* factors are interesting aspects for musicologists. Automatic music generation using artificial intelligence is rising and gaining popularity. Employing visualization to facilitate understanding of such complex relationships can lead to a better comprehension of musicology topics. Technological progress expands musicology, raises new research questions, opens new data sources, and requires new solutions.

(4) Musicology offers rich collections of data – The emergence of improved digital format standards such as MusicXML [112], the invention of new instruments, and the increased usage of streaming music services such as Spotify or YouTube provide large data sources. Also, big data collections like IMSLP [120], "The Million Song Dataset" [24], BMLO [93], or MUSICI and MusMig [253] are available today. Such data sources not only provide music data, but often additional information about the life and work of (early) composers, their relationships, or genres. The new amount and diversity of available data pose challenges to musicology but also open up chances for interdisciplinary solutions such as the application of visual analytics.

(5) Complex musicological data requires analytic and visual support – Musicologists could benefit from visual and analytic solutions to cope with challenges in their domain. Conducting manual considerations on extensive datasets is often not feasible. Hence, they cannot realize their datasets' full potential without computational approaches and interdisciplinary collaboration. E.g., there are experimental musicological scenarios such as embodied music cognition [184] or embodied music interaction [203] in which domain experts must deal with complex body movement data. Adapting available visualization techniques could reveal previously unknown patterns and enable knowledge generation.

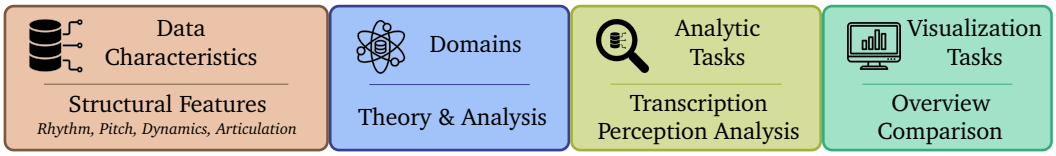
(6) Applying methodology transfer can open up new fields of interdisciplinary research – We propose a methodology transfer model (MTM) (see section 2.3) to present open challenges and motivate interdisciplinary research. We postulate that visualization research is a valuable source to support digital humanities researchers. Similarly, musicologists have the required domain knowledge to guide problem-oriented collaboration by providing data and real-world applications. Implementing the MTM instead of creating solutions from scratch saves precious research work. The MTM also allows focusing on issues that require new solutions due to identifying research gaps during the methodology transfer process.
(7) The digital humanities (DH) offer valuable insights for information visualization research – Visualization researchers typically work on methodologies to be applied to the problems of other domains. However, due to the experience and long history of humanities, there is much value in applying methodologies from DH research for information visualization [33]. While other domains benefit from close cooperation with visualization experts, they should not hesitate to learn from the humanities' research experience of many centuries. Instead of considering the DH only as data providers, we should start considering them as equal partners with different views on the world that we should learn from [33].

2.7 Summary

In this chapter, we proposed how to utilize available visualization techniques and applications with musicology. Thus, we frame *visual musicology* by providing a diligent overview of substantial musicological domains such as ethnomusicology or education. We describe typical computational *domain tasks* and *data characteristics* to define visual musicology's design and problem space. We emphasize that many musicology problems can be addressed by tailoring existing *visualization techniques* to support musicologists. We see high potential for mutual benefits of such cooperation since musicology provides many complex issues that require innovative solutions. We demonstrate in the subsequent chapters different use cases on how visualization can be employed to support music analysis tasks. This chapter also explains how *methodology transfer* can be strategically executed to profit from existing solutions. Eventually, motivated visualization and musicology researchers are both needed to develop and generate new approaches and knowledge through close interdisciplinary collaboration. Based on the proposed VMG, the remaining chapters of this doctoral thesis are categorized and put into context to each other in Figure 9.1. This is done to reveal which aspects of the VMG are covered or at least touched within the context of my thesis. In the conclusion, we then display the entirety of all three parts and the underlying chapters in a single figure (see Figure 9.1) to emphasize the topics that were addressed throughout this thesis.

3 Analyzing Visual Mappings of Music Notations

> *Learning to read music in Braille helped me develop a good memory.*
>
> **— Ray Charles**
> American Soul Composer and Pianist



In this chapter, we postulate that combining the domains of information visualization and music studies paves the ground for a more structured analysis of the design space of music notation, enabling the creation of alternative music notations tailored to different users and their tasks. Hence, we discuss the instantiation of a *design and visualization pipeline for music notation* that follows a structured approach based on the fundamental concepts of information and data visualization. This enables practitioners and researchers of digital humanities and information visualization alike to conceptualize, create, and analyze novel music notation methods. Based on analyzing relevant stakeholders and their usage of music notation as a means of communication, we identify relevant features typically encoded in different annotations and encodings used by interpreters, performers, and music readers. We analyze the visual mappings of musical dimensions for varying notation methods to highlight gaps and frequent usages of encodings, visual channels, and Gestalt laws. This detailed analysis leads us to conclude that such an under-researched area in information visualization holds the potential for fundamental research. This chapter also discusses possible research opportunities, open challenges, and arguments that can be pursued in analyzing, improving, or rethinking existing music notation systems and techniques. More general future research directions that originate from the entirety of this thesis are discussed in Section 9.3.

Chapter 3 is based on the following publication. Please refer to Section 1.4 for the specific contributions of each co-author.

[225] **Matthias Miller**, Johannes Häußler, Matthias Kraus, Daniel A. Keim, and Mennatallah El-Assady. **"Analyzing Visual Mappings of Traditional and Alternative Music Notation"**. *Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2018.*

3.1 Introduction

Ever since the earliest recordings of ancient cultures and societies, music has been an evident and integral part of the cultural experience and societal identity [128]. Musical instruments and tools have been discovered in archaeological sites, while paintings, scriptures, and folktale records form an indisputable trail highlighting music's development through human history [283]. However, as evident as the human desire to express feelings and emotions through music is part of our history [110], preserving the melodies of our past has been a continually changing, unguided process, influenced by many factors [186].

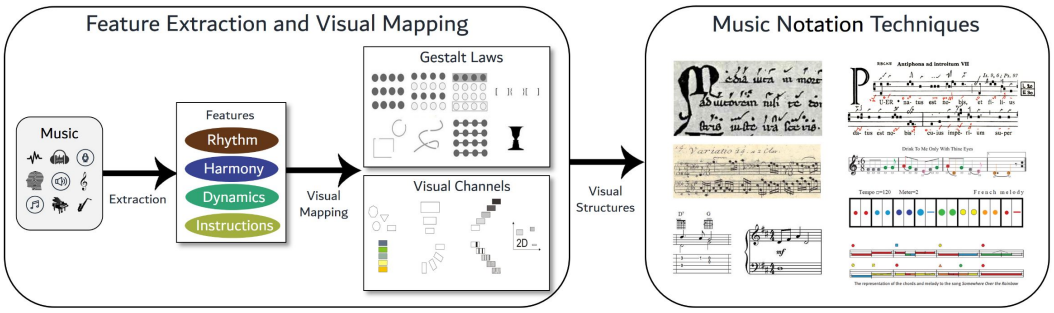


Fig. 3.1.: Segmenting music into its basic features to design and apply task-dependent visual encodings, based on visual channels [238] and Gestalt laws [254]. Exploiting the whole range of the *music notation design space* bears the potential for various novel music notation techniques.

To preserve music as part of their culture and to form a distinct identity, human societies strove to find methods that would allow them to pass such knowledge on to later generations. Music was initially only transmitted orally [329]. However, due to the fading collective memories over centuries, there always existed a desire for a more accurate method of capturing and preserving musical creations [260]. Consequently, the development of the first musical notations emerged and started with a simple approach; curved lines, also called neumes [343], were displayed over text to indicate the relative pitch for singers. Over time, the methods for music notation were altered and refined, depending on the purpose and objective of the music notation [318]. This serendipitous development in different cultures and times led to diverse notation systems for various application areas and instruments. Musical notation standardizations emerged due to the predominance of some representations, which finally led to the western, *traditional* music notation that has been established as the norm we are used to all over the world. This well-known **Common Music Notation** (CMN) [297] allows experienced musicians to understand how a composition should sound solely by reading musical expressions [138]. Such expressions are commonly represented as notes on paper structured in a predefined notation system. Thus, one central purpose of music notation is to assist musicians with accurate communication, interpretation, and reproducibility of compositions [51]. CMN enables capturing musical aspects for creators and composers alike to conserve their ideas unambiguously. Hence, composers use music notation to communicate with performers, providing instructions on how a composition should be musically interpreted and reproduced [159].

However, creating and standardizing a seemingly universal notation is challenging [69]. For example, depending on the instrument, musicians use different notations [104]. Due to the varying pitch ranges between instruments, different clef types exist to discriminate music notations between instruments. Likewise, singers require other musical representations than instrumentalists to understand their parts.

Due to the gradual development of notation over multiple centuries, CMN resulted from a long-term, historical development process. For instance, one bottleneck for a wide distribution was the lack of color at the beginning of mass printing to reproduce music scores on paper. Similarly, technological progress and digitization have not yet reached the point of using interactivity or motion. From an InfoVis research perspective, these historically grounded limits on the design space of music notation may have led to an established representation that may not exploit the full potential of all available techniques and encodings. For example, according to Flach et al., dyslexic people benefit from using an alternative notation exploiting other visual cues, enabling them to better distinguish the note pitches, which is more difficult using the CMN [92]. Current information visualization research provides a means for a more structured approach to exploring music notation designs.

Hence, we postulate that bridging the domains of information visualization and musicology opens up space for a structured analysis of music notation and a thorough exploration of its visual design space. Such an endeavor opens paths to novel and original research, tailoring music notation to different users and tasks.

Considering different application fields and existing music notation techniques, this chapter contributes a structured analysis of the visual design space of music notation. We define the span of the design space to integrate and analyze the state-of-the-art visualization techniques of various music notations. In doing so, we obtain means to identify the benefits and drawbacks of different visual encodings. In addition, our approach enables a systematic and comparative analysis of different concepts concerning their effectiveness by applying Gestalt laws [254]. Finally, these considerations reveal potential research gaps and opportunities for novel music notation visualization design concepts.

3.2 Background

Depicting music in a visual form is a complex task containing several representation issues [69]. The existence of a wide variety of different music notations shows that it is overly complicated to encode all musical features into a single and consistent system because of the limited number of visual encoding channels. Designing a notation system requires making fundamental design decisions that depend on the target audience and the application area. Making such design decisions limits the range of possible applications, which is acceptable if some use cases are well-supported even though other tasks are not. One reason why specialized notation is restricted is the loss of information. For instance, tailoring a notation system to a specific instrument makes reading more difficult for musicians who are not acquainted with the instrument. A further reason is that music is subject to continuing progress due to the development of new instruments, genres, and even rearrangements of existing works, making it impossible to take features into account that will be introduced

later. The range of applications of music notation can be divided into categories (examples are given by reference), which contain but are not limited to (live) performance [96, 322], analysis [208], art [265, 154], education [359], instrument support [279], composition [25], and entertainment [94].

Dannenberg states that the diversity between the single categories requires one to view music and its structure based on different levels of representation since each notation design possibly contains information not available in other notation systems [69]. For instance, to analyze music, abstract transformations often provide means of comparison that enable analysts to detect patterns and differences in musical pieces instead of focusing on single details [287]. Music analysts are interested in understanding a musical piece's harmonic progressions and relationships. Malandrino et al. propose a visualization approach to emphasize the harmonic structure of a composition by employing color to indicate tonal progressions [205]. Many approaches that propose an instrument-oriented design can also be assigned to the category 'Education' since visual metaphors based on the respective instrument's visual appearance provide a high potential for learning a new instrument. For example, Dascălue et al. utilize the keyboard design to create an instrument teaching platform for adult learners [72].

To summarize the broad range of music notation systems into a single framework, we first span a structured *Music Notation Design Space*.

3.3 Music Notation Design Space

Visually representing music requires a suitable mapping from heterogeneous input formats to visual channels. If music exists in audio format, we must extract structural features before applying visual encoding. Since the notational representation of music is ambiguous, even existing visual representations can be examined and altered to be restructured to support varying situations. Music is an intricate form of art containing both structural, such as harmony or rhythm, and non-structural properties, like emotion and imagination. The former is mathematically formalizable. Nevertheless, multiple attributes of music can hardly be communicated using formal notation. For instance, when interpreted, performing music often involves a specific level of expressiveness that is difficult to formalize or encode visually. Consequently, dealing with music as notation entails representational issues. This representation problem is responsible for the existing variety of musical representations. To analyze existing music notation specifications, we must focus on musical features commonly used to represent music visually. Due to the hierarchical structure of music, we subdivide music into four meta-features that comprise concrete musical dimensions, each of which can be subject to visual representation in musical notation systems.

3.3.1 Musical Meta-Features and Notation Dimensions

We suggest a set of meta-features that can be extracted from music, each characterizing a unique musical attribute. Every meta-feature, such as rhythm and harmony, contains multiple dimensions that describe its specific musical aspects (e.g., pitch,

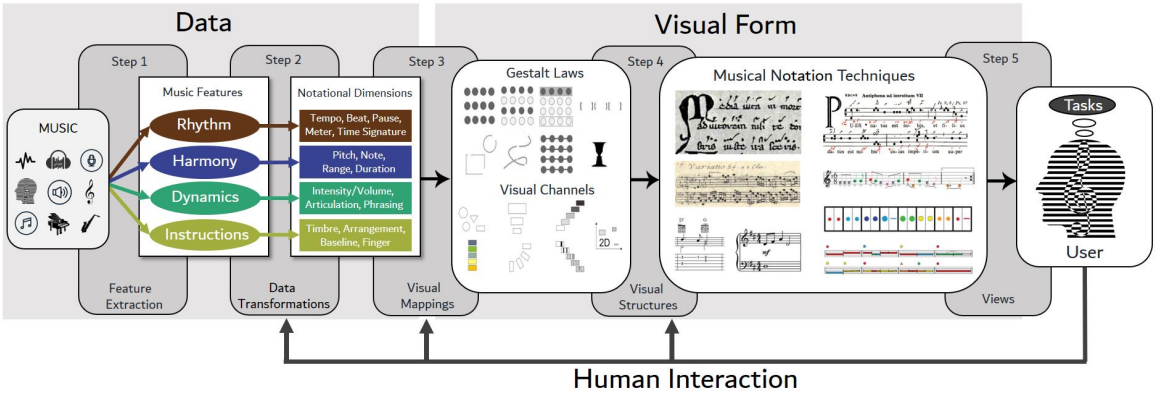


Fig. 3.2.: Our proposed *Music Notation Visualization Pipeline* builds on Card et al.'s Information Visualization Reference Model [44]. Transforming musical data into visual music notation is a multi-step process, including extraction of musical features (Step 1), data transformation (Step 2), visual mapping (Step 3), and encoding into visual structures (Step 4). Moreover, allowing users to interact with the system and changing the notation enables customization on views (Step 5) and enables notation improvement to fit the users' tasks.

tempo, or timbre) in detail. Subsequently, we consider music to be generally a composition of ***rhythm***, ***harmony***, ***dynamics***, and ***instructions*** (see Step 1 in Figure 3.2).

Rhythm describes the tempo (beats per minute) or speed of music including meter and time signatures. Pauses also belong to *rhythm* since they influence the rhythmic movement or flow in music. *Rhythm* is the cause of why music can be categorized and analyzed as time-series data and is fundamental in the design of music notation. A more detailed consideration of the rhythmic features is given in Chapter 5.

Harmony depends on how multiple tones simultaneously compose specific pitch levels. This co-occurrence of notes determines the harmonic progression of a musical piece. In music notation, one must differentiate between accidental and normal notes. Moreover, the range of a given part is defined by the octaves or height of a note. We assign duration to be part of every note since varying tone lengths directly affect the harmonic behavior. Often, dynamics is used to describe the loudness or the development of intensity, including the transitions between different volume levels, accents, and abrupt changes in the musical progression. Please read Chapter 4 for more details about harmony in sheet music.

Dynamics comprises volume (or intensity), articulation, and phrasing since phrasing and volume are frequently combined to partition music into connected segments. Besides, articulation also shapes musical dynamics with intensity and phrasing.

Instructions encapsulates contextual notation information such as timbre, arrangement, baseline, and which finger should be used to play a specific note. "Arrangement" includes the structure of a musical piece comprising repetitions and instrumentation. The instrumentation influences the timbre and is regarded as a separate dimension often implicitly represented. Depending on the instrument, the CMN uses different baselines encoded by respective clefs at the beginning of the notation, indicating the pitch of the displayed notes [293].

3.3.2 Users and Tasks

Different types of users need suitable music representations: music composers must be able to expressively communicate their thoughts through symbols, signs, and other instructions. They achieve this by mapping musical thoughts to be structured and represented so that the composed music can be reconstructed as imagined by the author. In contrast, performers should be provided with a musical notation that is easy to understand and suitable for the instrument they are playing. The notation of a musical piece can vary between different instruments to meet task-oriented requirements. Imagine a musical director who must know all the instrument parts of a whole orchestra. Providing an overview of all involved instruments, in this case, is beneficial. On the contrary, single instrumentalists do not profit from such an orchestra notation but could even be distracted while playing their part. This example emphasizes the role and importance of different music notations for diverse tasks. Learning an instrument is another field of application: Understanding the CMN can be challenging. Tailoring music notation to the user using intuitive representations such as visual instrument metaphors supports novices during the learning process. Rather abstract representations are helpful for music analysts who are more interested in patterns, progressions, and general differences between musical pieces than in specific details.

		Rhythm			Harmony					Dynamics			Instructions			
	Visual Variable	Tempo / Beat	Meter / Time Signature	Pauses / Breaks	Tones (Pitch / Frequency)	Note	Range (Octaves)	Accidental / Normal	Duration	Intensity / Volume	Articulation	Phrasing	Timbre (Instrument)	Arrangement	Baseline / Clef	Finger
	Text (Semantic Channel) [29]	♫, [161]	♫, [161, 278, 267]				♫, [161, 267]			♫, [267]	♫		♫, [267, 65]			♫, [267]
Magnitude Channels	Position on common scale	[245, 59, 311]	♫, [180, 287, 245]	♫, [161, 278, 267]	[311, 267, 59]	[21]	♫, [359, 161, 279, 267]			[65]		♫, [267]	[65, 59, 311, 267]			
	Position on unaligned scale	[21]				♫, [161, 267]									♫, [161, 267]	
	Length (1D)		[359]				[180]		[359]							
	Tilt / Angle															
	Area (2D size)						[21]		[65]	[180, 65]						
	Depth (3D position)															
	Color Luminance		[311]						[65]	[59]						
	Color Saturation					[59]										
	Texture					[359]		[359]						[180]		
	Curvature															
	Volume (3D size)									[311]						
Identity Channels	Spatial Region						♫									
	Color Hue				[287, 59, 323]	[180, 359, 161, 65, 59]		[161]	[323]				[267, 311, 245]			[279]
	Motion	[279, 245, 323]	[65]	[279]					[279, 65]							
	Shape	[108]		♫, [161, 278, 180, 267]	[108]	[359]	[65]	♫, [180, 161, 267]	♫, [180, 161, 278, 267, 311]		♫, [279, 267]		[65]	♫, [180, 267]	♫, [161, 267]	
Gestalt Laws	Proximity	♫			♫, [279]	[65]	[180]	♫			♫					♫
	Similarity	[180, 245]	[359]	♫	♫, [65, 59]	[161, 59, 65]	[180, 65, 359]	[180, 359]	♫		[279]		[267, 245]			
	Enclosure	♫	♫, [245, 180]		[267]		♫									
	Closure								[180]							
	Continuity	[245]			♫				♫		[279]		[267]			
	Connection		[180]		[21]	♫	[267]		♫, [359]		[279]	♫		♫		

Tab. 3.1.: An overview of mapping musical notation features (Rhythm, Harmony, Dynamics, Instructions) to visual variables: magnitude channel, identity channel, and Gestalt laws. The two-note symbol (♫) is used to indicate how the Common Music Notation (CMN) encodes music variables.

3.3.3 Visual Mapping of Musical Dimensions

Simultaneously representing multiple musical features by visual cues is substantial when providing musical notation information. During the design process of music notation, considering the *Gestalt Laws* helps estimate the effectiveness of a visual cue. Since visual variables are the foundation for information visualization, which musical notation is a subset of, we take Bertin's seven visual variables into account: *position, size, shape, value, color, orientation*, and *texture* [23]. Munzner extends this list by *motion, curvature, volume,* and *spatial region*. Munzener divides the visual channels into *magnitude channels*, which comprise ordered attributes, and *identity channels* that should be used to encode rather categorical data attributes which do not have an implicit and natural ordering [238].

It is useful to order the magnitude channels by effectiveness to reasonably apply visual channels in designing visualizations [46]. Based on Munzner's discussion, we ordered the different visual magnitude channels from best to least starting with '*position on common scale*' (top) to '*Volume (3D size)*' (bottom) in Table 3.1. This sorting facilitates the comparison of different visual encodings by effectiveness. Moreover, readers can get a better distribution overview of how musical features are visually encoded. Applying different visual mappings between concepts avoids confusion when reading a designed music notation.

3.4 Traditional and Alternative Music Notations

Besides CMN, various approaches have been proposed in application domains, such as research or teaching to reduce the CMN's drawbacks. These somewhat experimental notation concepts often use different visual channels to exploit the potential of hitherto unused visual variables. In Table 3.1, we provide an overview of existing musical notations based on visual channels and Gestalt laws introduced in section 3.3. We aim to indicate possible visual mappings not used in a music notation system, highlighted by empty cells with a gray background. We point out that the given overview does not claim to be exhaustive since many ideas and approaches are outside academia (e.g., Pitch Bracket Notation [108]). The overview table lists the CMN using a two-note symbol (♫) to emphasize differences between the listed music notation approaches. Moreover, we selected the included references to be reproducible by a performer or computer instead of adding all methods that visualize any musical feature for analysis or entertainment. Frequently, music notation analysis visualizations apply abstract methods that do not take fundamental notation features into account[193, 346].

In Table 3.1, we classify fifteen existing music notation techniques. We emphasize that this does not represent all notation techniques in the literature. It is structured to facilitate the understanding of which reference of a music representation method uses a visual variable (rows) grouped by Munzners visual channel categorization [238] and Gestalt law to encode a musical feature (columns). The characteristic musical meta-features group the columns introduced in subsection 3.3.1. Some music notation concepts use different encodings for notes (a tone within an octave: A-G), ranges, and pitch. We consider pitch to be composed of note and octave

information. Mostly, a performer requires the exact position to play a musical piece precisely. Therefore, we decided to list them separately to highlight explicit representations of this musical dimension. Music notation often includes textual descriptions to provide information about instrumentation, volume, tempo, and meter signature. Borgo et al. subdivide all visual channels into four categories: the *Semantic Channels* containing text, numbers, symbols, signs, icons, and others [29]. We consider this category appropriate for describing the contextual music notation information (first row in the overview table). Seldomly complete chords are directly mapped on visual variables instead of the respective single notes. For instance, Malandrino et al. enrich the CMN with a colored background to indicate the current pitch class at a particular position in the score [205]. We omitted a separate column for the chord dimension for simplicity and clarity due to the scarcity of such techniques.

3.5 Research Implications and Opportunities

Our research shows how music notations can be compared regarding their visual features. To the best of our knowledge, Table 3.1 comprises varying music notations but is incomplete and may serve as a starting point for a survey. The introduced visualization pipeline (Fig. 3.2) helps to both understand existing visualization techniques and develop new ones. Since the pipeline unifies the design process on a rather conceptual level, it may be required to consider specific musical dimensions in more detail. Nevertheless, we argue that any music notation visualization can be categorized using our proposed pipeline that can be extended to cover further musical dimensions if necessary. In doing so, weaknesses of a specific representation and its capability to visually map musical features for a given task can be revealed, considering relevant notational dimensions. Vice versa, in the design process of new music notation visualizations, the pipeline, in combination with the overview table, supports the consideration of different notational dimensions. The line-up of mappings between musical features and visual variables can also inspire new higher-level visualizations supporting content-based visual analysis of music.

Our research can be seen as a starting point for in-detail investigations of certain combinations of mappings and users or tasks. Specific pairs of visual variables and musical features could particularly fit certain user groups (e.g., learners, composers) or tasks (e.g., high-level analysis of a musical piece or performance). We assume that readers of the classical musical notation system are inherently biased due to early familiarization with the standard notation. Due to the familiarity, the popularity of CMN, and the historical contingency, developing a music notation that will finally replace CMN is extremely challenging.

By now, we can exploit visual channels requiring technology that was not available when CMN was developed, such as motion, color, or volume (3D). Since CMN has disadvantages, such as differentiating tone pitches, new techniques could address these drawbacks by rethinking how to apply visual variables. During the design process of new music visualization methods, it is necessary to consider the respective application scenario. Some mappings are more intuitive in a specific situation and convey precisely, while others provide abstract feature information.

We argue that our design and visualization pipeline for music notation (see Figure 3.2) is not only applicable to music notation but also to develop methods for music fingerprinting, to compare music regarding different epochs and styles or to compare the visual encoding of music visualizations. For instance, a well-established mapping for some notational feature could be deployed in a higher-level visualization, showing, for example, the trend of a musical feature throughout the entire musical piece. Visualization techniques such as pixel visualizations, glyphs, or others could be applied to visually analyze larger amounts of musical data and enable musical performance comparisons.

3.6 Summary

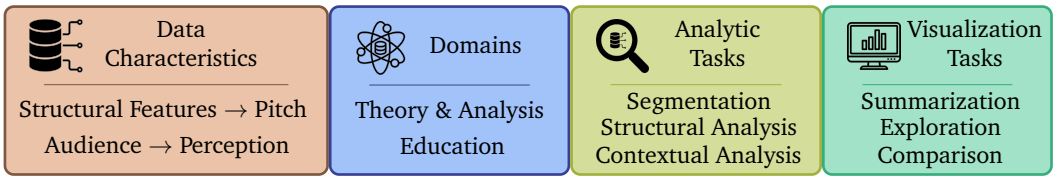
This chapter introduced a design and visualization pipeline for music notation based on fundamental information visualization research. This approach can be used as a foundation to compare existing music notation designs and develop new techniques tailored to different users and their tasks. Thus, we bridge the research field of data visualization and the field of music representation in a structured way. An exemplary classification of fifteen notation systems indicates how musical features can be encoded. The table can be a starting point for a full survey of existing music notation techniques. Simultaneously, the existing gaps in Table 3.1 could indicate possible opportunities for future research. In the remaining chapters of this thesis, we are reconsidering the different musical features and dimensions, with a special focus on harmony, rhythm, and melody in Part II. Eventually, animations and interactive visualization are implemented in Part III, bringing the initial concepts of this chapter together for a more integrated user experience in music analysis tasks.

Part II

Sheet Music Analysis for Harmony, Rhythm, and Melody

4 Augmenting Music Sheets With Harmonic Fingerprints

> *Music is a higher revelation than all wisdom.*
>
> **— Ludwig van Beethoven**
> German Composer



Common Music Notation (CMN) is the well-established foundation for the written communication of musical information, such as rhythm or harmony. CMN suffers from the complexity of its visual encoding and the need for extensive training to acquire proficiency and legibility. Although alternative notations that use additional visual variables (e.g., color to improve pitch identification) have been proposed, the community does not readily accept notation systems that vary widely from the CMN. Therefore, to support student musicians in understanding harmonic relationships, instead of replacing the CMN, we present a visualization technique that augments digital sheet music with a *harmonic fingerprint glyph*. Our design exploits the circle of fifths, a fundamental concept in music theory, as a visual metaphor. By attaching such glyphs to each bar of a composition, we provide additional information about the salient harmonic features available in a musical piece. We conducted a user study to analyze the performance of experts and non-experts in an identification and comparison task of recurring patterns. The evaluation shows that the harmonic fingerprint supports these tasks without the need for close reading compared to a not-annotated music sheet.

Chapter 4 is based on the following publication. Please refer to Section 1.4 for contribution clarifications.

[222] **Matthias Miller**, Alexandra Bonnici, and Mennatallah El-Assady. **"Augmenting Music Sheets with Harmonic Fingerprints"**. *Proc. of the ACM Symposium on Document Engineering, 2019, pp. 1–10.* ***Best Paper Award.***

4.1 Introduction

Isidore of Seville recounts how until the 7th century, music was only conserved as an auditory memory as there was yet no means of notating music [17]. By the 9th century, neumes came into use as visual aids to indicate the relative pitch and melody direction in Gregorian chant [318]. Over several centuries this neumatic notation was adapted to reflect the evolution of musical instruments and different music composition styles to create the Conventional Music Notation (CMN) used today. Despite the widespread use of the CMN, the notation faces criticism on two counts, namely that the notation is inadequate to represent musical notation and does not adequately represent all information of a musical score. The CMN is considered inadequate as a musical notation because it is mainly based on using a five-line staff such that only nine vertical spaces are available to represent the twelve notes of the chromatic scale. Thus the CMN introduces sharp and flat signs to represent all twelve notes, unnecessarily complicating the visualization of the note pitches. Likewise, the notation of the note durations does not intuitively represent the score's rhythmic properties.

These criticisms led to alternative music notation systems to facilitate music reading, especially for novice learners. Among these alternative systems, we may find notation systems such as Klavarskribo [150], Hummingbird [353], Dodeka [277] and Pizzicato [244] among others. These alternative notation systems may include the introduction of different staff systems to capture equally spaced chromatic intervals [150], the use of different symbols to represent note pitches and durations [353, 180], piano-roll style music representations [277], and numeric-based notations [244].

However, these alternative notations have not been adopted by the music community due to the lack of accessibility of music scores written in the different notations and because the number of musicians who can read the CMN by far outnumbers those who can read alternative notations. Thus, the problem of reading and quickly understanding the CMN remains a struggle for novices [206].

Technology, notably screen applications, has changed the way musicians are sharing and using musical scores [300]. Music scores are no longer restricted to print media. Advances in optical music recognition and the automated conversion of musical scores into different file formats led to the emergence of alternative notations. Instead of providing a complete overhaul of the CMN, additional score annotations are often employed as additional teaching aids [75]. Furthermore, the digital representation of music as web or tablet-based apps makes it possible to have both conventional representations and new notations available for the same document so that the student can easily switch between notation systems.

In general, sheet music encodes musical features such as rhythm, dynamics, harmony, and other instructions as visual symbols which specify how a performer can reproduce a composition [225]. The notation, however, omits the explicit representation of the relationship between the individual notes and their location within the overarching musical structure [139]. While understanding these complex relationships is of immediate interest for the musician [73], such an understanding typically requires analyzing the score and, thus, a certain level of musical expertise [206, 75]. Understanding the harmonic structure of the score allows for faster learning of new
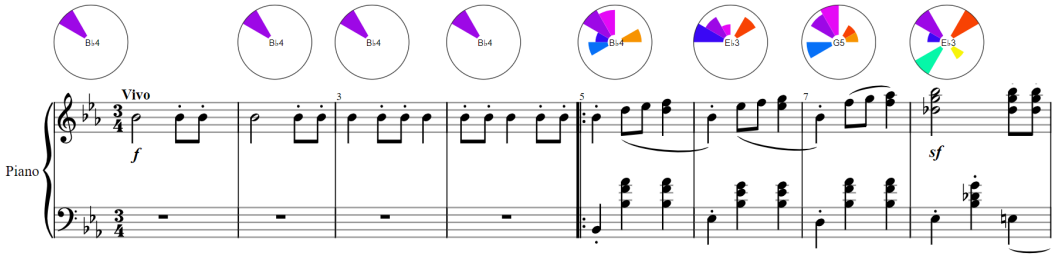


Fig. 4.1.: Augmenting piano sheet music with *harmonic fingerprint glyphs* facilitates identifying harmonic patterns and comparing musical parts to understand the note distribution. This excerpt from Chopin's '*Grande Valse Brillante*' is augmented with the fingerprints showing a recurring pattern in the first four glyphs.

musical pieces and enhances musical memory, and should ideally be more accessible to novices. Annotating the score with a visual representation of the hidden harmonic patterns would, therefore, help to convey this musical knowledge leading to a deeper understanding of the underlying musical structure [206].

In this chapter, we contribute a *harmonic fingerprint glyph* visualization to augment the distribution of notes in pre-defined windows within musical pieces as illustrated in Figure 4.1. This avoids the introduction of new music notations that differ from conventional music notation. We use the circle of fifths [125] as a visual metaphor to design the harmonic fingerprint. By attaching this glyph, we add valuable information about the harmonic content to digital sheet music, allowing readers to *efficiently* identify and compare harmonic relationships across a document. To test the performance of the fingerprint, we conducted a qualitative user study, collecting comparative feedback from domain experts and non-experts. We concentrate on the following three research questions to emphasize the benefits and drawbacks of our approach:

Q1 Does the proposed fingerprint annotation method support the *identification* of recurring and long musical patterns?

Q2 Does the fingerprint annotation influence the harmony *analysis* of the score?

Q3 Does the reader's musical *expertise* level influence the effectiveness of the fingerprint for performing harmony analysis?

4.2 Related Work

Understanding harmonic relationships and tonal progressions of a musical composition is fundamental to music analysis [290]. To reveal these connections of harmony, musicians must analyze the music to identify the chords and patterns based on music theory knowledge. Performing such a task requires knowledge and expertise, which are not necessarily within reach of beginner musicians [206]. Harmony visualization tools make harmonic analysis more accessible by providing visual aids highlighting non-explicit information about the score. Visualization algorithms typically follow Shneiderman's *Visual Information Seeking Mantra*, that is, follow the three rules: overview first; zoom and filter; then details-on-demand [306].

4.2.1 Music Harmony Visualization

There are different ways to visualize harmonic relationships in musical works. Abstract music visualizations enable an overview of whole musical pieces, but readers can no longer look at the details. In contrast, providing original details of the music score facilitates understanding the same harmonic and rhythmic relationships on a lower level. Still, the reader can not readily recognize the overall structure. Jänicke et al. explain that combining close and distant reading allows users to find regions of interest while still providing all details instead of only presenting an abstract representation. In this way, readers are directed by distant reading while investigating the underlying details afterward [145]. We transfer Jänicke et al.'s close and distant reading concept from the text domain to music visualization by showing its importance for the visualization harmony in music [145].

Distant Reading – Analyzing particular attributes of structural data and visually providing them is a typical approach to revealing unknown patterns. For example, Keim and Oelke presents a method to apply visual and literary analysis for different types of features that are characteristic of text [157]. The proposed visual literature fingerprinting method has proven to be a practical approach to showing meaningful differences between several text sections while including various text features such as average sentence length for creating the fingerprint visualization. Keim and Oelke do not include close reading, so users who want to investigate the details in interesting areas of the visualization are not supported.

One approach to visualize the harmonic content of a musical piece is to create a representation that displays the overall form of the music. For example, Wattenberg introduces an *arc diagram* visualization to reveal sequential melodic patterns. Sequential repetitions of melody notes are highlighted by arcs connected to the notes. The arc's radius and width indicate the distance and similarity between recurring melodic sequences [346]. Similar approaches include the representation of the score using *similarity matrix* visualizations [357], or the use of the Tonnetz grid to create an *Isochord* visualization of the score [21]. Using a similar approach, Schroer uses *circos graphs* to reveal patterns in genomic data. Using these graphs, Schroer represents the harmonic relationship between all the twelve tones of the chromatic scale by linking the chord root with other notes played simultaneously [298]. Figure 4.2 illustrates these based on compositions from varying artists, revealing high-level insights regarding the differences between the respective musical epochs.

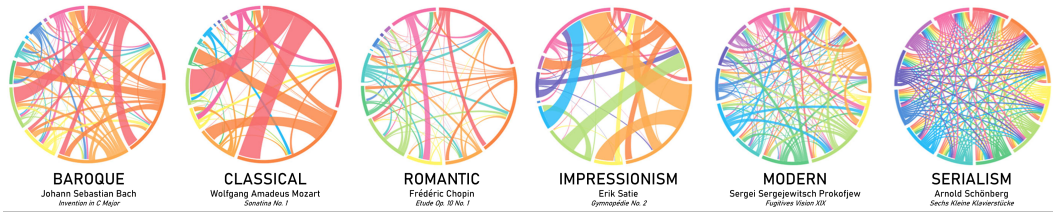


Fig. 4.2.: These illustrations adopted from Schroer [298] show harmonic differences and similarities between specific compositions ranging from baroque (J.S. Bach) to serialism epoch (A. Schönberg).

Such diagram representations of the score emphasize the commonalities between whole sections of the musical piece. Therefore, these visualizations can highlight similarities in the structure between different musical scores that are typical for a genre. However, these approaches do not retain the relation to the original score notation. While it is easy to visualize the overall structure of the score, the spatial position of these highlighted structures gets lost. Enabling music readers to retrieve the detailed CMN on demand is essential to support understanding of the underlying harmonies. This requires combining distant reading with close reading to exploit the advantages of both methods.

Close Reading – Local visualizations of the score retain the spatial information of the highlighted structure concerning the complete score. Smith and Williams exploit three-dimensional space and color to visualize typical musical information based on MIDI files [311]. They propose to map tone data described by pitch, volume, and timbre to colored spheres in their visualization. The aim of Smith and Williams is to visually present music to listeners unfamiliar with reading music notation. Rather than providing a static visualization, their model uses a dynamic visual model, which decreases the color intensity of the notes with time progress, thus allowing the listener to distinguish between different tones played at different points. While such an approach is visually pleasing, it does not allow the listener to understand the underlying harmonic sections or recurring patterns easily.

Algorithms such as that described by Snydal and Hearst create melodic landscapes and harmonic palettes from transcriptions of jazz improvisations [312], while Miyazaki et al. use cylinders whose height, diameter and color represent pitch, volume and duration information of the notes in the score. In the latter algorithm, the visualization can be represented at the overview and detailed levels according to user preference [231]. Ciuha et al. describe an approach to visualize concurrent tones, using color to highlight the harmony [59]. In this visualization, the horizontal axis represents the temporal progress of a single musical piece, while the vertical axis encodes the note pitches. Depending on the temporal segment size, the visualization is blurred based on tonal unambiguousness. In this way, this method conveys the harmonic journey of a piece while using an appealing representation to encode the affinity of notes by applying a color wheel on the circle of fifths to reflect the dissonance of intervals.

Sapp proposes a visual method that emphasizes the key strength at each section [290]. This model uses color to encode the dominant notes or keys at several hierarchical levels. At the top of the hierarchy, the model represents the most significant key of the entire composition. The lower levels of the hierarchy represent detailed tonal progressions, resulting in a triangular representation of the keys used in the music. Thus, this model is advantageous in revealing the tonal variations in the piece.

4.2.2 Annotating the CMN Score

The discussed visualization strategies detached from CMN, making it challenging to keep the connection between CMN and the visualization. One strategy to resolve this problem is to maintain the original notation format and annotate the score. An example of such an approach commonly used in music notation is the placement of chord symbols to indicate the accompanying harmonies the musician needs to

perform. However, these chord symbols are a rather general description of the harmony progressions and do not reflect the complexity of specific musical parts in detail. Hence, while scores can be annotated using a similar analogy, an annotated visualization would let readers focus on the harmonic aspects.

Addressing these concerns, Cuthbert and Ariza describe `music21`, a tool that provides several features for visual score descriptions. Among these, the harmonic analysis of the score is shown as chords written in closed form. The metric analysis is represented as asterisk signs beneath the notes, while a plot of the pitch class against the bar number illustrates the pitch usage over the progression of the score.

The *VisualHarmony* tool described by Malandrino et al. overlays visual information about the tonal information directly onto the musical score. The tool aims to support music composition by highlighting score parts not complying with classical music theory rules [206]. The tool facilitates identifying chord tonality and the respective scale degrees to display melodic errors supporting music composition tasks. In VisualHarmony, the visual information is shown as colored rectangular boxes enclosing the individual chords. Thus, the system preserves the spatial location of the harmonic patterns. However, the overlapping visualization was found to be too distracting, particularly for directly playing the music, and thus, De Prisco et al. suggest that the visual information should be displayed above the staff [74].

From this overview of the related work, we note that visualization algorithms that represent the melodic patterns in the music do so on a global level without preserving the spatial location of these patterns. While visualizations described by [206] preserve the spatial location, these focus on analyzing the music's harmony. While this is important, the harmonic analysis does not capture the melodic patterns that can also be present in the score. Thus, an annotation system that captures both the harmonic patterns and the melodic patterns would be of benefit to the community. For this reason, I address the exploration and identification of melodic patterns in Chapter 6 and also in Chapter 7.

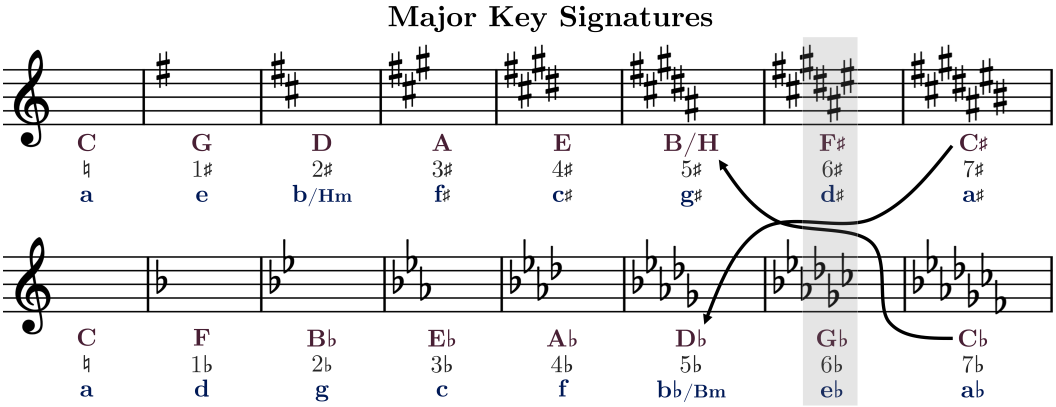
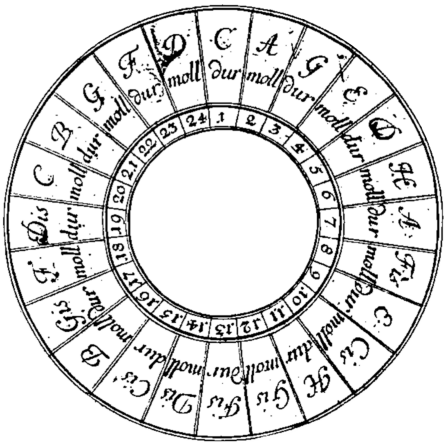
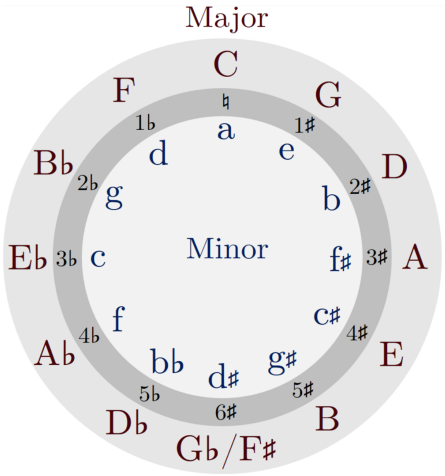


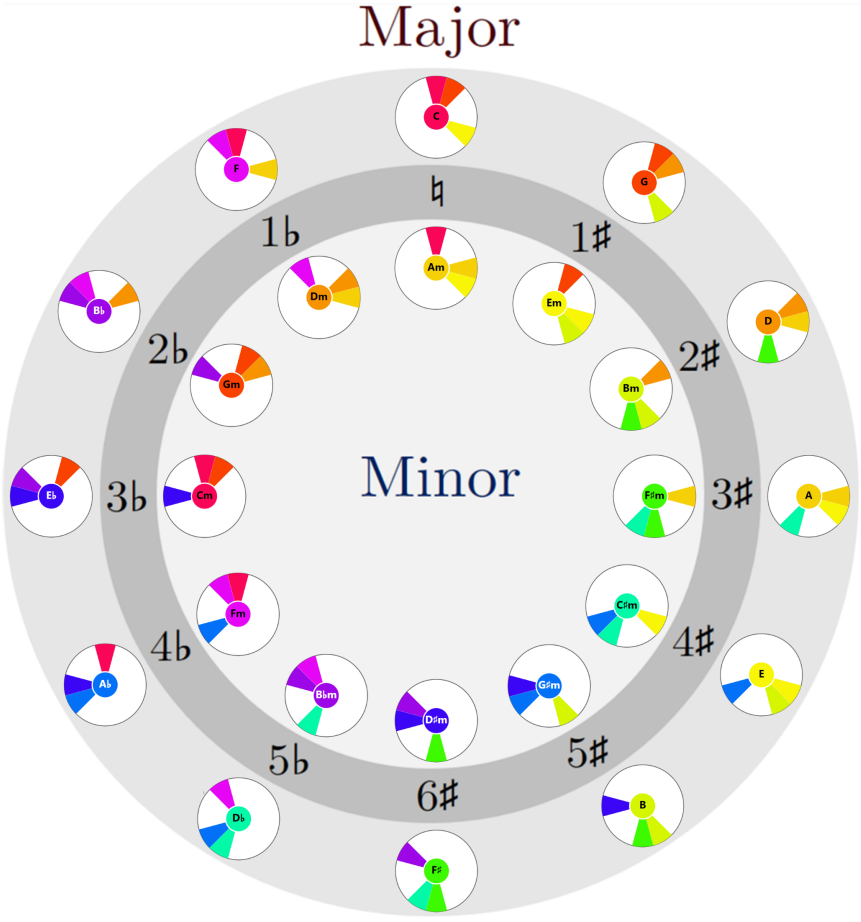
Fig. 4.3.: While there are only 12 pitch classes, in practice, even more key signatures are used, leading to overlapping enharmonic equivalents such as F♯/G♭, C♯/D♭, and B/C♭. In contrast to the English system, the German musical notation represents B♮ as 'H' and B♭ as 'B'. Thus, the third measure at the top system shows Hm as an alternative for b and Bm for b♭ (5^{th} measure in the second staff system).



(a) The musical circle originally published by Heinichen in 1728 [125].



(b) Separation of major and minor chords and their corresponding accidentals.



(c) Illustrating the COF (b) with our proposed harmonic fingerprints.

Fig. 4.4.: The circle of fifths shows the relationships between the 24 available minor and major keys in Western music. (a) The minor and major keys are arranged in a single circle, while in (b) the inner part represents the minor keys and the outer ring the major keys. The accidentals (♭ and ♯) are shown on the ring in between. (c) illustrates how to arrange the harmonic fingerprint design in the circle of fifths based on the initial circular design from (b). Color facilitates the identification of shared pitch classes between the different keys and chords.

4.3 Harmony Fingerprint Design

This section describes the proposed harmonic fingerprint visualization. This visualization approach will capture the harmonic and melodic content of each bar of the musical score. It exploits the circle of fifths as a visual metaphor for its design. While highlighting the harmonic information within a bar, the glyph also represents tonal information in a way that musicians may already be familiar with.

4.3.1 Circle of Fifths

Typically, we refer to key signatures based on the number of pitch classes, which are twelve. In practice, three additional key signatures (G♭, C♯, C♭) are used, leading to 15 key signatures displayed in Figure 4.3. In this figure, the first key signature in each row represents *C Major* which has no accidentals (♮).
The *circle of fifths* (COF) is a practical concept to explain the geometric structure of the chord relationships between all twelve chromatic pitch classes. It is often used by musicians and in music pedagogy [320] to visualize the relationships between pitch classes. In 1728, Heinichen augmented the circle of fifths to introduce the relationship between major and minor keys, representing the COF with 24 segments as shown in Figure 4.4(a). Figure 4.4(b) facilitates understanding the parallel minor and major chords while providing the number of accidentals for each key. The distance of the pitch classes in the circle measures the tonal similarity of the keys. For example, if *C Major* is the tonic, then *F Major* or *G Major* are the most similar chords in terms of auditorial perception and direct neighbors in the COF. We used the same coloring between the two figures to facilitate the direct connection between the Major (dark red) and Minor keys (dark blue). The two arrows in Figure 4.3 show where the COF is overlapping while F♯ and G♭ are enharmonically equivalent, which is emphasized in the illustration with the slight rectangular background behind the seventh measure. Based on this design, Figure 4.4(c) illustrates how the COF layout can be used to arrange the respective chord qualities in the inner and outer circle representing the minor and major chords. In contrast to Figure 4.4(a) and Figure 4.4(b), this extension makes it easy to spot pitch class overlapping through the categorical color encoding of the harmonic fingerprint design.

4.3.2 Design Rationale

In the glyph's design, we take into consideration multiple musical characteristics that determine the annotated visualization and the used annotation type:

The Region of Interest – In the design of the harmonic fingerprint, we capture the harmonic variations at a detailed level while retaining a general overview of the piece. For this reason, we choose single bars as the region of interest for which we automatically create the glyph. Since the tonal center remains the same within a bar, particularly for music with dance-like structure [320]. Representing harmonic content per bar reasonably compromises detail and a general score overview.

Representation of Harmonic Content – To represent the harmonic fingerprint of the bar, we use the chord's root note within the bar if it is identifiable. The root note is a common way of describing a chord in music theory; thus, the glyph uses a familiar description to musicians. Other alternatives from music theory would include describing the chord using Roman numerals or *figured bass* [320]. Still, this would be less meaningful to beginners or non-musical experts. The fingerprint allows for distant reading of harmonic relationships in single bars, summarizing all notes into a radial chromatic histogram.

Representation of Melodic Content – The glyph captures the note pitches, which also form the melodic content. Since the glyph's role is to provide an overview of the harmonic material, we represent this through a radial normalized histogram of the pitch class of notes within the bar. Notes with the same pitch name but in different octaves are grouped in the same histogram bin, allowing us to identify harmonic similarities across a piece. Hence, bars with similar melodic patterns written in different octaves or for which there is rhythmic variance will obtain the same histogram. At an overview level, this is desirable as it captures the broad similarities between bars while keeping close reading possible by maintaining the CMN.

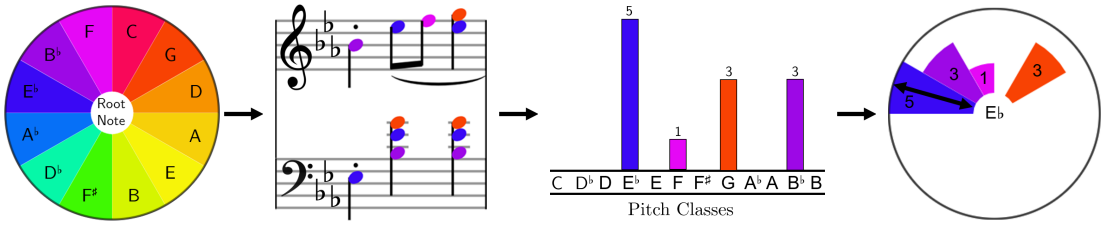
Capturing Harmonic Relationships Between Notes – The glyph visualization provides an overview of the harmonic relationship between notes in the bar. This overview should allow musicians to understand the presence of consonances or dissonances between notes in the respective bar at a glance. Neighboring segments in the fingerprint visualization are the most similar notes in music harmony theory. In contrast, segments placed on the other side of the circle are the most dissimilar relationships. In combination, the reader can efficiently identify the dominant notes for each bar. The circle of fifths metaphor enables us to illustrate the harmonic relationship using a musical concept familiar to the musician.

4.3.3 Creating the Glyph

We illustrate in Figure 4.5(d) the specific glyph shape using the 6^{th} bar from Chopin's waltz "Grande Valse Brillante" as an example. A histogram of the pitch classes, displayed in Figure 4.5(c), captures the information about the notes forming the melodic and harmonic content. This representation is relatively sparse and, as a result, would clutter the annotation. We alter the histogram by rearranging the pitch classes radially such that each of the twelve pitch classes from the chromatic scale is displayed by subtending angles of 30° (360°/ 12 = 30°) as shown in Figure 4.5(a). This allows us to represent the histogram's content more compactly. The number of note occurrences within the bar is encoded by the segment's radius, resulting in a visualization similar to a non-stacked Nightingale Rose Chart [36]. Moreover, rearranging the pitch classes into the circle of fifths format highlights the harmonic relations within the predefined music sheet window. Hence, the harmonic fingerprint can be considered a statistical overview of the number of notes from a single bar. The reading complexity of the glyph remains unchanged due to its independence from the original music document complexity, enhancing scalability.

To further improve the readability of the visual fingerprint, we use the color scale from Ciuha et al., which we show in Figure 4.5(a). This color scheme reflects pitch class distances in the color space [59]. The visual double encoding of the pitch classes in color and position simplifies the comparison of fingerprints.

Finally, to further emphasize the harmonic content of the bar, we display the chord's root formed by the notes in the bar at the glyph center as illustrated in Figure 4.5(a). The resulting shape structure of the glyph is similar to that shown in Figure 4.5(d).



(a) A radial design color mapping for the pitch classes according to the circle of fifths.

(b) Bar 6 taken from Chopin's "Grande Valse Brillante" using the note color encoding from (a).

(c) We calculate a colored histogram based on the chromatic pitch classes of all the notes that are present within a bar.

(d) The final fingerprint encodes the notes distribution of a bar by segment area size and color.

Fig. 4.5.: We use the circle of fifths as a visual metaphor to encode the distribution of notes in a bar (b) by mapping all notes to their corresponding pitch class (c). Each class's notes are then encoded by the size of the area segments, as shown in (d). The pitch classes are encoded both using color and the position of the segments inside the fingerprint (a).

4.4 Adding Harmonic Fingerprints to the Score

For the scope of this work, we assume that sheet music is readily available in MusicXML file format [112]. MusicXML is an XML-based digital sheet music interchange format designed to provide a universal format for CMN. The use of MusicXML is increasing, and music writing software such as *MuseScore*[1], *Sibelius*[2] among many others, support MusicXML representations of new works. Likewise, digital libraries such as the Mutopia Project[3] provide MusicXML sources for classical works.

To obtain the statistical data required for the glyph, the MusicXML file is preprocessed using the *music21* Python libary [67]. Through this library, we compute the number of occurrences of each pitch class and identify the chord and the chord's root note for each bar in the score.

To display the annotated CMN with the harmonic fingerprint, we use *OpenSheetMusicDisplay (OSMD)*, an open-source project to render MusicXML in a browser [109].

Similar to De Prisco et al.'s observations, we place the glyphs directly above the bars so the reader can intuitively detect the corresponding fingerprint [74]. To ensure that the annotations do not overlap with the CMN symbols, we expand the space between each system for situating the fingerprint annotations. We introduce this additional

[1]https://musescore.org/
[2]https://www.avid.com/sibelius-ultimate
[3]https://www.mutopiaproject.org/

space by increasing the standard distance between systems in the OSMD application by an amount proportional to the diameter of the glyphs. The allocation of this extra space ensures that both the harmonic fingerprint and the sheet music are legible.

4.5 Evaluation

To evaluate the performance of the fingerprint annotations, we conducted a user study to analyze music theme pattern identification tasks on different sheet music to investigate the users' performance with and without the fingerprint annotations. We specifically focused on revealing differences between domain experts and non-experts to assess our approach's intuitiveness and identify tasks that can be supported by the fingerprint visualization introduced in the previous section.

4.5.1 User Study Methodology

The study assessment is based on the overall satisfaction reported by the users and their performance in finding the patterns of the provided music sheet sections.

To perform the evaluation, we conducted a within-subject evaluation, giving each participant two scores, one with and the other without annotations. We instructed the participants to mark any patterns they could identify on the printed music score without explicitly telling them to perform harmony analysis. To reduce the potential for bias in the evaluation, we selected the scores as equal in length and complexity. Moreover, we prepared an annotated and a non-annotated version of both scores, randomizing the order in which we presented the scores.

Dataset and Controls – We applied the fingerprint annotations to the Chopin's "*Grande Valse Brillante*" [142], for which we processed a ground truth musical harmony analysis beforehand. This waltz has 311 bars of music arranged into seven distinct themes, which can be distinguished from the melodic and harmonic variations between the themes. Figure 4.6 shows the first 68 bars, highlighting the

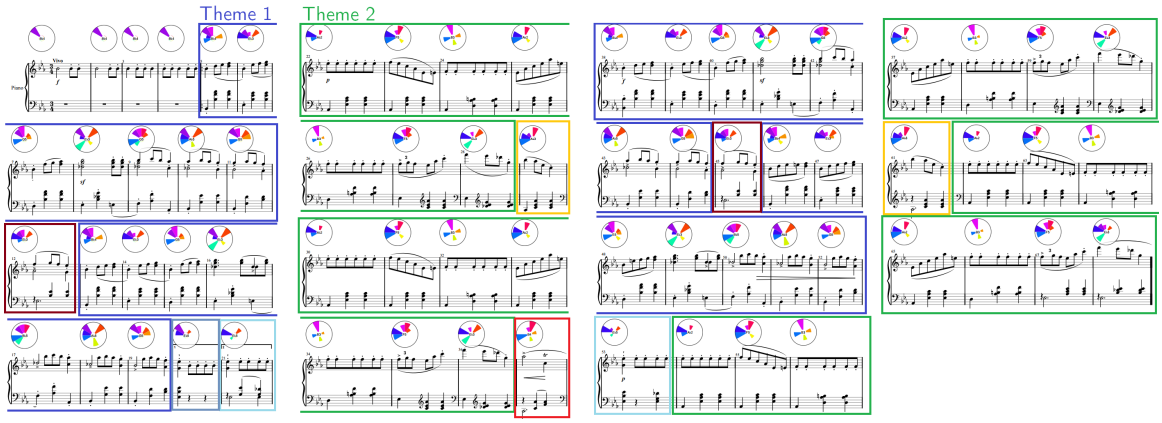


Fig. 4.6.: The first 68 bars of the 'Grande Valse Brillante' (forming MS1). This consists of two major themes enclosed by the blue and green rectangles. Bars 12 and 45, enclosed in yellow and red, divide the two themes into two parts, with the second part of each theme consisting of rhythmic variations to the first occurrence of the theme.

themes. By comparing the annotations obtained from the annotated fingerprints, it is possible to determine whether it represents these thematic divisions.

Since the music score is sufficiently long, we use selected parts in the evaluation as discussed hereunder. We extracted two suitable music-sheet sections, which we denote as MS1 and MS2, respectively, with MS1 consisting of the first 68 bars of the waltz while MS2 consists of 64 bars, from bar 69 up to bar 132. Bar 68 has a natural break in the music dividing MS1 and MS2 into two different musical parts. MS1 and MS2 consist of recurring patterns: MS1 consists of two major themes, as shown in Figure 4.6, while MS2 consists of a very similar structure but has three different themes instead of two. The full length of the themes varies between 7 or 15 consecutive bars, with the latter containing four-bar sub-patterns. These themes were used as ground truth to estimate the participants' performance.

Participants – We evaluated the fingerprint annotation approach with eight participants from diverse backgrounds who completed all phases of the user study. We selected the participants by dividing them into two equal-sized groups.

The first group consists of four domain non-experts (N1–N4) who have little to no music background (mean score of 1.5) and have never performed a harmony analysis (mean score of 1.0). However, these non-experts are experts with a post-graduate degree in data analysis and visualization with a mean age of 34 ± 9.
The second group contains people playing instruments daily or having fundamental to solid knowledge about the harmonic relationships in sheet music (mean score of 3.75). Here, we refer to the second group as our domain experts (E1–E4). These four participants have intermediate musical knowledge (mean score of 3.25). Three expert participants practice music as their hobby on a daily basis, while the last is singing in a choir but is aware of the harmonic relationship of multiple voices.

Study Design – We conducted the study in three phases. In the preliminary phase, we collected general demographic information from the participants, namely their age, gender, level of musical background, and familiarity with harmony analysis. To collect the information related to their musical background, we asked participants to rate their knowledge on a five-point Likert scale, from *novice* (1) to *expert* (5). Likewise, we asked the participants to rate their familiarity with harmony analysis from *no knowledge* (1) to *very familiar* (5).

In the study's second phase, we asked the participants to perform the pattern identification task by adding handwritten annotations to highlight patterns on the musical score. The participants completed the task twice, once with a score containing fingerprints and once with the CMN only. We recorded the time the participants took to perform the tasks, advising them that the analysis should not exceed 30 minutes. We deliberately decided to exclude a specific training task to investigate the accessibility to the music document on different knowledge levels without further explanation of the musical scores or the harmonic fingerprint.

The final phase of the study consisted of a short interview to elicit user feedback, helping us to answer the research questions stated at the end of the introduction section. We asked participants to verbalize their opinion on the harmonic fingerprint' usefulness and their strategy to fulfill the given pattern identification task.

4.5.2 Use Cases

Before presenting the qualitative feedback from the user study interviews, we describe two use cases supported by our harmonic fingerprint annotations.

Identify Octave-Invariant Harmony Relationships – Identifying harmony requires understanding how multiple notes represent a tonal center. This chord depends only on the notes' pitch classes of the chromatic scale and is independent of the octavation. Consequently, the absolute tone pitch does not change the harmony if the pitch class remains unchanged. To identify such similarities, a reader has to identify the line for each note and apply the accidentals of the current key to extract the underlying chord. This is a tedious task if the number of notes increases.

For example, the chord identification task becomes more complex if the note C is available three times simultaneously (e.g., C2, C4, C5) in addition to other notes. The fingerprint annotation merges all pitch class occurrences, making it easier for analysts to focus on the composition of pitch classes to decide which harmonic aspects are most dominant in a bar and define the chord.

Musical Theme Extraction – Depending on the composition, the CMN complexity complicates finding recurring musical themes. Using a unique color scale to represent each pitch class, the annotated fingerprint enables readers to easily skim a document for similarity. Since color is easier to differentiate than black and white, readers can visually filter for similar harmony patterns. Moreover, the layout of the CMN symbols, such as stem direction, note duration, or note divisions, has no visible influence on the fingerprint and provides a trustworthy foundation to identify harmonic differences and commonalities. Figure 4.6 displays how music themes can be identified by comparing the fingerprint visualizations across a score. Due to distant reading, even slight differences in pitch classes can be detected. By keeping the CMN, close reading remains possible to compare rhythmic details within the bars. Besides, the fingerprint enables identifying the most frequent note in each bar at ease.

4.5.3 User Feedback

Figure 4.6 shows the annotation results for bars 5–68 of Chopin's "Grande Valse Brillante" are shown superimposed on the ground truth thematic analysis of this waltz. The second occurrence of theme 1 has similar fingerprints as the first occurrence. There are some rhythmic differences in the melody line of the second occurrence, which are not reflected in the fingerprint visualization. The fingerprints for the second theme are considerably different from that of the first theme, highlighting the shift in the key. Recurring themes have very similar fingerprints, as expected.

To understand the effect of the harmonic fingerprint glyph on music readers at different music expertise levels, we reported the qualitative results and elicited feedback on the usefulness of our visualization from the participants. Based on this feedback and the users' performance in finding the themes, we will address the research questions Q1, Q2, and Q3 from the introduction. We compared the differences between the participants' annotations and the ground truth of the harmonic patterns included in the datasets. We highlight the participants' statements using italic and double quotation marks in the following paragraphs.

Usefulness of the Fingerprint – Except for N3, all participants stated that they would like to use our proposed fingerprint annotations in the future in pattern identification tasks for different reasons. For example, N1 "*used the fingerprint as an anchor to keep the overview, [but] got lost in the CMN because of the notes' optical similarity.*" N2 explained that by using the glyph, she could "*compare images*" to identify the pattern because it is easier since the complexity of the CMN makes it hard to identify differences. First, N3 performed the pattern identification task without the fingerprint and was confused about the fingerprint because the "*notes were easier to compare*" and the fingerprint seemed to not match with the CMN. For N4, it was "*much easier with the fingerprint [...] to see patterns based on the color and location [...] than trying to scan and compare the notes.*"

E1 argued that "*to practice a piano piece, it might be helpful to spot parts which are equal or have small differences.*" E3 found it "*an interesting way of putting a visual aspect to combining notes indicated within the score.*" E4 stated that it is "*easier to find the patterns with the visualization,*" but was unsure whether she might have missed some relevant details because she did not further consider the CMN. E4 could "*recognize how the piece develops*" and "*obtain an overview over the piece easier.*"

Pattern Identification Strategy – Depending on whether the fingerprint was annotated to the sheet, the user groups applied different strategies to identify repeating patterns. All non-experts, except N3, preferred the annotated condition since they found it easier to find matchings. N1 "identified differences and similarities by comparing first the presence of the glyph's segments and afterward its size without looking at the CMN anymore if the glyphs matched." N4 stated: "*I started with the first icon [...] to find another that matched until the next.*" N2 and N4 only highlighted the patterns by adding handwritten annotations by grouping the fingerprints instead of using the CMN. N1 "*identified differences and similarities by comparing first the presence of the glyph's segments and afterward its size without looking at the CMN anymore if the glyphs matched.*" N1 and N3 incrementally summarized subpatterns without fully identifying a theme.

In the first round, E1 was provided with the non-annotated document where he textually added the chords symbol to the first 12 bars. E1 continued by highlighting the melodies throughout the document without further looking at the harmony. *To give melodic characteristics due consideration, I refer to Chapter 6 where I provide a detailed discussion about melodic patterns and their relationships based on functional operators.* E1 found the fingerprint helpful "*to spot bars which are the same*" and was "*gathering information by switching focus between glyphs and [the] music sheet.*" At first, E2 "*skimmed through the whole piece to gain an overall picture of the piece [and then] moved on to analyzing bar by bar [to] compare phrases to each other to find where repetitions occurred.*" Figure 4.7 displays E3's handwritten annotations: E3 divided *Theme 1* in MS1 as two subpatterns, excluding bars 9, 12, and 17, and completely identified *Theme 2*.

To find similarities, E3 "*tried to merge the harmonic fingerprint glyph with the melodic patterns to identify the sequences present*" and "*followed the labeling of the chords as well as the note patterns [...] at a different pitch level.*" To identify recurring "*melodic patterns*" in MS2, E3 applied the same strategy as for MS1 with the glyphs by "*marking the tonal and real sequences [...] through identifying the patterns where the note intervals are different [...], but the pitch is moving in the same direction.*"

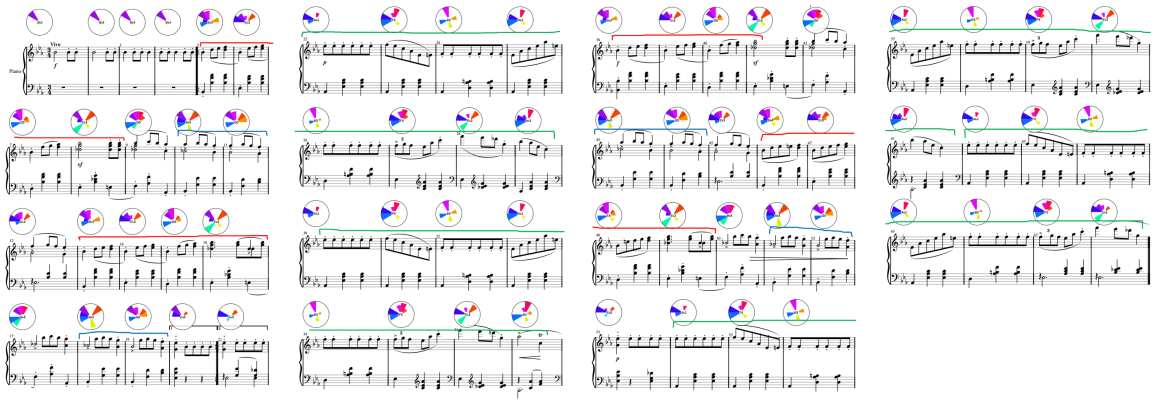


Fig. 4.7.: We told the participants to add handwritten annotations to the printed sheet music. The figure shows how E3 highlighted different patterns. E3 was provided with MS1, including fingerprint visualization. She highlighted the first half of *Theme 1* in red and its last two bars using blue color. She identified *Theme 2* completely by marking it in green.

Challenges – The participants faced various challenges in executing the assigned task. For example, N1 stated: "*I had to go back to the bars that I wanted to compare more often because I couldn't remember all the details as well as with the fingerprint.*" N1 argued that he "*found differences in the lower system, while the upper system did not change within a bar sequence.*" N1 mainly annotated shorter patterns and could not identify a complete theme. Similarly, N3 did not find complete harmonic theme patterns without the fingerprint but often highlighted subpatterns of two subsequent bars. N1 also gave the feedback that "*the fingerprint is easier to remember how they look compared to the original music sheet, especially if patterns are farther away.*"

N2 found it difficult to "*identify patterns that have optics jumps [because of] new rows continuing on the next page.*" N3 was the only participant who considered the fingerprint confusing since it indicated similarity, whereas the music sheet did not reflect this. Hence, N3 found the glyph visualization misleading in finding exact patterns. N3, who was provided with MS2 without the fingerprint, indicated that it is "*difficult to see exactly which line the notes are placed and if they are similar to the others.*" In the second round, N4 had to find the patterns without the glyphs. Due to their absence, N4 perceived it to be "*more tedious to look more closely at the single notes,*" especially if the patterns broke across multiple systems.

For E1, the most challenging aspect of the analysis task was in "*reading the sheet and deciding which overall chord it is [because] it is time-consuming.*" One of the most challenging aspects for E2 and E4 was "*to analyze the piece and find similarities [...] without having heard the [...] music being played, as this would make such repetitions extremely evident.*" E3 "*was not quite sure what the colors within the harmonic fingerprint glyph represented.*" Due to the bar-based distribution of the fingerprints, E4 declared to be "*too focused on single bars at the beginning*" requiring more time to find larger patterns in the sheet.

4.6 Discussion and Lessons Learned

The qualitative user study that we conducted to analyze our fingerprint annotation's performance revealed benefits and drawbacks compared to the CMN only. Except for N3, all non-experts could identify many parts or full patterns of the themes without the need to look at the CMN details. Thus, the fingerprints can facilitate access to sheet music for people with little to no knowledge about music theory or harmony rules because they encode salient harmonical characteristics (Q1). In this way, novices can become more motivated to look at the sheet music to understand the relationship between the notes in a bar and the annotated visualization. N2 and N4 found the visualization visually compelling, increasing their interest in looking through the music sheet (Q2). Novices and music learners often become overwhelmed by the complexity of the CMN and give up early due to the steep learning curve in understanding the CMN.

4.6.1 Educational Aspects

Since the circle of fifths is the foundation of harmonic relationships in Western music, the fingerprint is suitable for teaching learners about the harmonic structure in sheet music. It provides a quick overview of bars and facilitates the differentiation of bars that appear in different sections of a music sheet. The fingerprint allows for efficiently identifying bars with the same visual melodic pattern but written at a different pitch. Consequently, these bars can be distinguished from bars only transposed by octaves without the need for close reading of the single notes in CMN. To check rhythmic differences, users can still look at the original notation since we did not change the CMN in the processing step of the fingerprint annotation algorithm. Participants who performed the pattern search task without the fingerprint primarily focused on the pitch class and the rhythmic characteristics. The fingerprint visualization aggregates note pitches while inherently filtering out rhythmic details, thereby emphasizing the harmonic attributes in the sheet music.

4.6.2 Staff Separation

An interesting finding was that non-expert users who first worked through the task of having the fingerprints attached to the sheet tended to combine the upper and lower staves in the second condition, even when the fingerprint was not given anymore (Q2). Due to typical structures in the different staves, Two participants of each group identified the patterns without combining the staves. Therefore, they did not differentiate between these parts. Even when combined, the resulting harmony is different. They analyzed the upper and the lower stave separately to find harmonic similarities when the fingerprint was not present by starting with patterns with a length of single bars and extending them until they identified differences. The remaining study participants recognized the fingerprint as a summary of the bars and seldom looked further into the details of the CMN (Q2). We call these contrasting strategies top-down (with fingerprints), and bottom-up (no fingerprints) analysis approaches. The top-down approach was mainly applied by N1–N4 by including the fingerprint when adding handwritten annotations to the printed music sheet.

4.6.3 Harmonic Invariance

In the third theme of MS2, all notes from bars 119–124 are transposed by an octave higher in bars 127–132. To identify this similarity, the reader has to read the single notes to see the harmonic similarity. N2 and N3 were provided with MS2 without the fingerprint and could not identify this harmonic similarity. In comparison, N1 and N4 used the harmonic fingerprint to see the similarities without knowing the underlying relationships. With their domain knowledge, E2 and E3 could detect this harmonic similarity even without being provided with the visual annotation. E4 marked this harmonic similarity by highlighting the fingerprints, whereas E1 highlighted the pattern in the CMN. Hence, we argue that the fingerprint supports extracting such similarities more efficiently when readers want to find harmonic patterns of all simultaneous notes.

Another anomaly that we extracted from the feedback of E1–E3 is that they tend to ignore the fingerprint in favor of the CMN because of its familiarity (Q3). The fingerprint representations are less intuitive for this user group than the CMN. Since we did not explain the visualization's design decisions to the participants, we elicited that users who do not try to understand the detailed harmonic relationship trust the fingerprint without questioning. One music student replied that she was not sure about the color coding in the glyph. Nonetheless, she still found it a promising way to put a visual aspect of the combination of notes indicated within the music score.

4.6.4 Limitations and Future Work

The fingerprint visualization we designed to reveal harmonic relationships in sheet music has proved helpful to readers in some scenarios, including harmonic pattern identification tasks. Nevertheless, we identified some issues in the current design.

First, only one color map can be selected to encode the pitch class that is difficult or impossible to read for visually impaired people. Secondly, we combine all notes in a bar to extract the root of the dominant chord. Therefore, the fingerprints certainly show the exact notes histogram for each bar, which is equally considered in calculating the tonal center for the bars. Consequently, the root note displayed at the fingerprint's center is sometimes incorrect. For example, due to the tonal relationship of keys, the root note of the other chord mode of the relative key, or even a single note of a melody sequence that is not even part of the actual chord, is selected as the root by `music21` [67]. In the future, we plan to improve the accuracy by weighting the notes by their duration and excluding underrepresented notes from the calculation of the root note. Similarly, we want to additionally reflect the dominance of notes by their duration in the fingerprint. For example, if a note is played only once but has a long duration, it is currently underrepresented by the harmonic fingerprint and should be weighted according to its tone length.

Both musicians and music analysts are interested in the distribution of notes in a bar and want to know if the tonal center is a minor or major chord. We will include this information in the fingerprint to save readers time in extracting this information from the glyph visualization, which is currently required. We think it can be an interesting experiment to see how differences in music styles are reflected in the harmonic fingerprint that can visualize the salient characteristics of music.

Based on the qualitative user feedback, we elicited valuable feedback regarding the information encoded in the fingerprint visualization. Without understanding the underlying music theory, music readers require an introduction to the fingerprint because one may assume it encodes all musical information. Afterward, they can better estimate whether the visualization suits a given task. This has led to confusion for some of our non-expert participants, who did not trust the visualization after identifying presumed differences due to the harmonic invariance between the fingerprint and the CMN. When conducting a quantitative evaluation, we recommend explaining the visualization to the participants in advance to enhance the performance in harmony analysis tasks by improving their understanding. It will be useful to investigate the suitability for different tasks to elicit in which situation the fingerprint supports the reader and when it is better to show only the CMN.
Eventually, understanding the meaning of our fingerprint visualization is not readily intuitive, even for domain experts. This can result from the familiarity with CMN and their definition of a pattern, which was not prescribed in the user study.

4.7 Chord Identity Stability in Rotational Transformation

The harmonic fingerprint design provides a good basis for a deeper visual consideration of the relationships between different chords based on the COF. Similar chord qualities (e.g., minor and major) have resembling structures, and the overall template to identify similar harmonic fingerprints is rotationally invariant. Figure 4.8 illustrates this relationship by starting from the root notes in the first row (C G D A E B F♯ D♭ A♭ E♭ B♭ F) progressing through power chords (which is the 1st and 5th note of the respective scale). Note that the chords are arranged such that each column belongs to the stated root note (pitch class) at the top of the column. The power chord based on the root note C comprises C and G. In the COF, G is the next root note in the clockwise direction (or, in this table, the next column to the right). The third (*sus4*) to last (*parallel minor*) row in Figure 4.8 represent *chord triads*, a composition of chords built by exactly three different notes of different pitch classes. The question mark (?) after the pitch class at the center of the fingerprint indicates that the chord does not provide a clear chord quality, such as for *Root Notes*, *Power Chords*, and *Suspended* $4^{th}/2^{nd}$ *Triads*.
In the appendices, we continue the list for more complex and intriguing chords composed by four up to six different pitch classes (see Figure A.4). Diminished and augmented triads and chords with more than three pitch classes are often used, for instance, in Jazz music [316]. Characteristically, pitch classes on the opposite sides (such as A and E♭) in the COF have harmonically the farthest distance leading to dissonant sounds, whereas neighboring segments (such as *Power Chords*) are considered consonant, which is often perceived as sounding more pleasant compared to dissonant ones.

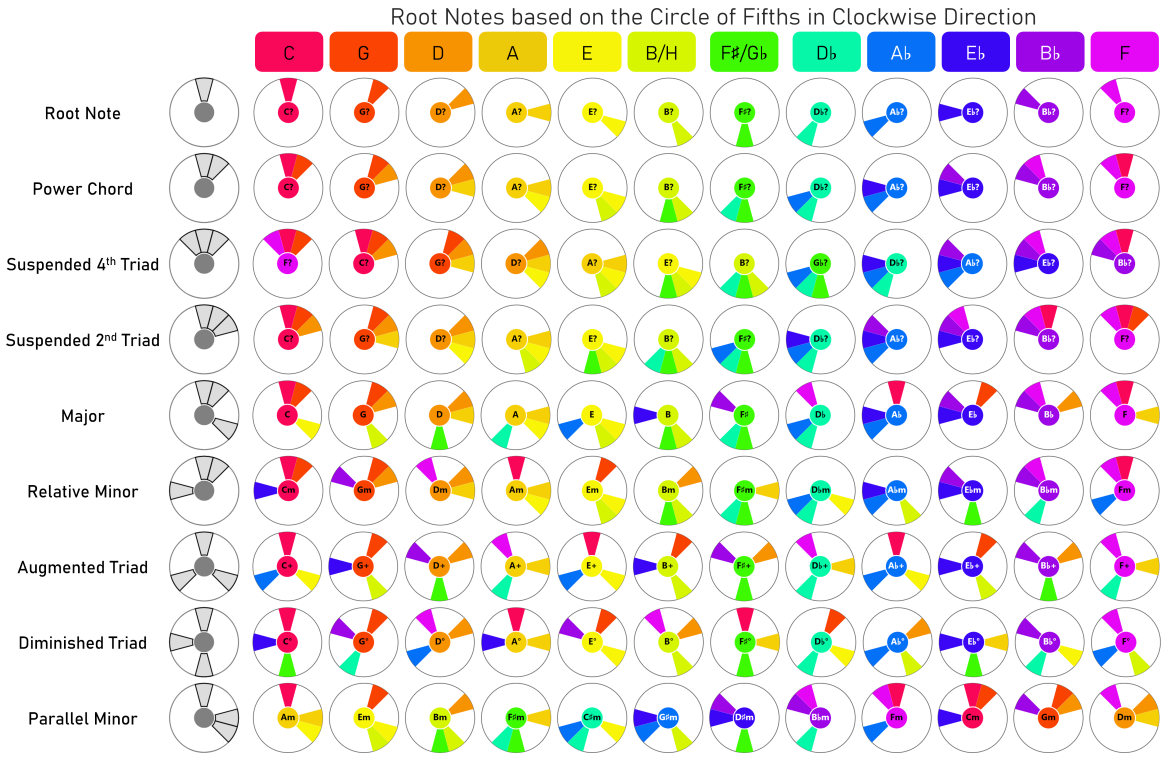


Fig. 4.8.: This harmonic fingerprint mapping to chord shapes reveals their respective connections. The first row with the colored pitch classes is arranged based on the circle of fifths in clockwise order, starting at the top center C. More details and extended chord shape overview, such as seventh chords, are given in Figure A.3.

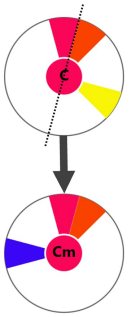
In addition, Figure A.1 and Figure A.2 provide the CMN details based on what chords the fingerprints are constructed. There, for simplicity and to enhance readability, every measure only contains whole notes. Based on the detailed overview, we can identify how harmonic fingerprints are constructed through combinations in the upper rows. For example, a *D♭-Major Triad* can be reconstructed by its subsets such as the *D♭-Power Chord* and the *F-Root Note*. Other combinations are also possible, leading to the same or similar chords. In general, you can always extend a triad that is a subset of the seventh chord (four-notes-chord) by adding the missing *Root Note* from the first row.

Angular and Rotational Consistency

Figure 4.8 illustrates the progression of chord shapes based on clockwise order in the COF. Since the COF comprises twelve segments, one for each pitch class, we can rotate the shapes by 30° to obtain the chords in the next right column. This visual representation reveals the rotational consistency within a row such that the shape templates on the left are abstract representatives for all instances in each row. We leverage this directional equivalence to prepare interactive visual queries that we further discuss in Subsection 7.5.3 to explore harmonic relationships.

Mirroring at Axes and Shifting Pitch Class Segments

The center of each of the twelve pitch class segments or the line between two segments can be used as a mirroring axis. For example, when we consider the C *-Major Triad* in the 5th row in Figure 4.8 compared to the C *-Relative Minor Triad* below (row 6), we can identify that the relative minor triad is the result of a mirroring operation at the axis between the C and G segments (see wrapfigure). This mirror transformation replaces the pitch class E with E♭. But there are more symmetric relationships such as the shift of the inner neighbor segment between a C *-Major Triad* and its A *-Parallel Minor* shifting the pitch class from G to A (major second interval). The parallel minor chords in the last row of Figure 4.8 actually have the root note A and not C, but due to its parallel minor connection to C, we arrange it as illustrated.



Due to their symmetric layout, the *Augmented Triads* as well as *Fully-Diminished-7th Chords* are exceptional cases compared to the other chords. Figure 4.9 illustrates their radial symmetry while leading to "wheel"-shaped harmonic fingerprints. Their special structure leads to a repetition of the respective fingerprints after four (augmented triads) or three (fully diminished 7th) 30°-rotations. For example, the augmented triads based on the root notes C, E, and A♭ share the same pitch classes (as emphasized in Figure 4.9), only the lowest pitch is changing at the CMN-level which does not affect the fingerprints design. Similarly, the fully diminished 7th chords are equivalent for the root notes C, A, F♯, and E♭.

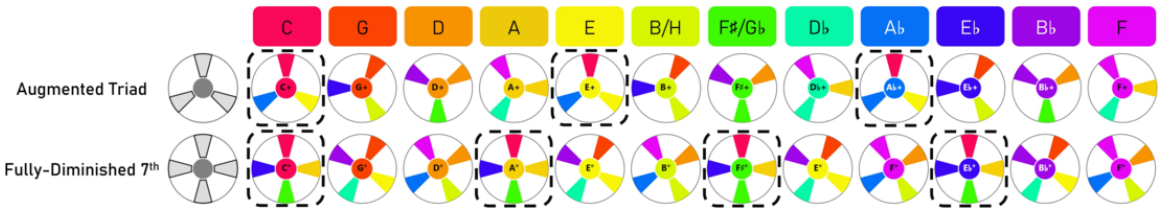


Fig. 4.9.: The repetitive character of *Augmented Triads* and *Fully-Diminished 7th* chords.

We consider what the shape of the harmonic fingerprints tells about consonance and dissonance in Subsection 7.5.1. Starting from harmonic fingerprints that represent *Triads* (exactly three notes), there are only nine remaining empty pitch class segments which can lead to *Four-Note-Chords*, such as adding the major seventh B on the root note C, adding the outer (in clockwise direction) neighbor of E resulting in the two pairs C-G and E-B in a 120°distance.

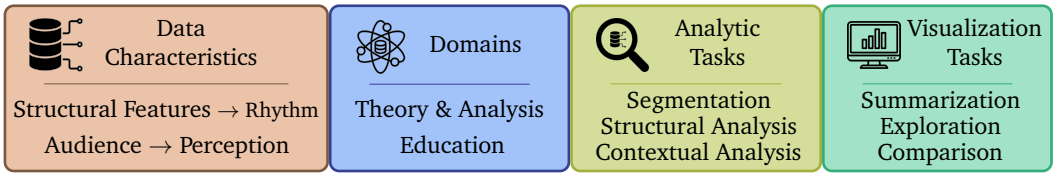
4.8 Summary

We introduced a visualization method to encode salient harmonic characteristics as a fingerprint. For the design of the *harmonic fingerprint glyphs*, [4] we exploit the circle of fifths, which is the foundation for harmony relationships in Western music, as a visual metaphor to augment digital music sheet with additional harmonic information. To analyze the usefulness of our approach, we conducted a qualitative user study with four domain experts and four participants without a musical background. The evaluation revealed the fingerprint's potential for identifying recurring harmony progressions and understanding the underlying harmonic structures in sheet music, even if the optical appearance suggests dissimilarity. Moreover, our method is suitable to support distant and close reading in sheet music exposing harmonic relationships on a more abstract level while keeping the original music notation unchanged. Consequently, readers can view sheet music on different levels of detail. In the future, we aim to enhance our approach by integrating user feedback and setting up a visual analysis system for music scores. Thus, we aim to combine close and distant reading into an interactive visual analysis system. Consequently, this web application will enable fast access to suitable music analysis tools for any sheet music available in MusicXML. Based on the findings of this approach, we designed an evaluated application called MusicVis. I present this follow-up approach in Chapter 7, where we investigated the use of the visual design for different music analysis tasks, such as structural analysis and more in-depth harmonic exploration at a more abstract level. In contrast to the study presented in this chapter, we provide detailed explanations about the visualization's design to make it more comprehensive.

[4] Our system is available under `https://musicvis.dbvis.de/app/fingerprint`

5 Augmenting Sheet Music With Rhythmic Fingerprints

> *Rhythm and harmony find their way into the inward places of the soul.*
>
> **— Plato – The Republic**
> Greek Philosopher



This chapter focuses on using a specific visualization design for rhythm analysis tasks. Similar to the harmonic analysis in the previous chapter, these are tedious due to the complex visual encoding of the well-established Common Music Notation (CMN). Thus, analog to the harmonic fingerprints, we augment sheet music with rhythmic fingerprints to mitigate the complexity originating from the simultaneous encoding of musical features. The proposed visual design exploits music theory concepts such as the rhythm tree to facilitate understanding rhythmic information provided through digital sheet music. Juxtaposing these rhythmic fingerprints and sheet music maintains the connection to the familiar representation. To investigate the usefulness of the rhythmic fingerprint design to identify and compare rhythmic patterns, we conducted a controlled user study with four experts and four novices. The design of this user study is analogous to the study already discussed in the previous chapter regarding the harmonic fingerprints. Similarly, the results indicate that the rhythmic fingerprints enable novice users to recognize rhythmic patterns that only experts could identify using nonaugmented sheet music.

Chapter 5 is based on the following publication. Please refer to Section 1.4 for contribution clarifications.

[103] Daniel Fürst, **Matthias Miller**, Daniel A. Keim, Alexandra Bonnici, Hanna Hauptmann (Schäfer), and Mennatallah El-Assady. **"Augmenting Sheet Music with Rhythmic Fingerprints"**. *IEEE Workshop on Visualization for the Digital Humanities (VIS4DH) at IEEE VIS, 2020.*

5.1 Introduction

As discussed before, CMN results from a century-long development [318] to encode musical information visually. While alternative notations have been proposed (see Chapter 3), the community has not adopted any to replace CMN. Hence, musicians, composers, and music analysts must be proficient in reading CMN [138]. Inexperienced music readers face a steep learning curve to read sheet music. Moreover, music analysis tasks require knowledge beyond the reading of individual notes. Such tasks involve finding harmonic progressions [206], melodic motifs, and rhythmic patterns [327], which helps understand musical structure and interpret a composition. While melody, harmony, and rhythm are simultaneously present in music, analysts often must focus on the different features separately, as discussed in Chapter 3. In this chapter, we concentrate on *rhythm*, a compound feature building on the primitive attributes meter, onset, and duration [317]. Rhythmic characteristics build the foundation of compositions and significantly influence their organizational structure [327]. Besides, rhythm is crucial in several music analysis tasks, including comparative and structure analysis.

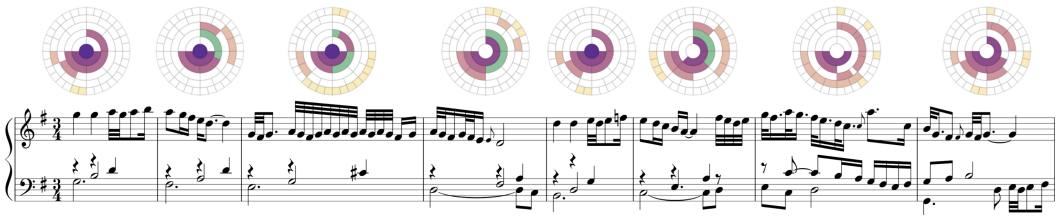


Fig. 5.1.: Augmenting sheet music with rhythmic fingerprints allows viewing only a composition's rhythmic aspects. This excerpt shows the beginning of the *Aria* from Bach's *Goldberg Variations* conveying a similar rhythm in the first and fifth measure.

Information Visualization has proven effective for music analysis tasks, such as identifying similar patterns or highlighting relevant aspects [162]. Precedingly, I demonstrated how augmenting sheet music with *harmonic* fingerprints helps to uncover such patterns. The study results confirmed that even novice users could uncover harmonic patterns that only music experts could see before. Harmony and *rhythm* are essential for music analysis, so we argue that rhythm deserves separate consideration. Hence, as a direct follow-up, we introduce a rhythmic fingerprint to extend our work on augmenting sheet music.

In this chapter, we similarly introduce a *rhythmic fingerprint* design to augment sheet music (see Figure 5.1) with entities that visually represent rhythmic characteristics. Using a radial tree layout to reflect the hierarchical structure of note durations, we exploit the visual metaphor of a clock. Through the augmentation, we provide additional information that supports the identification of rhythmic relations in a composition. To evaluate the approach introduced and the usability of the rhythmic fingerprint, we conducted a user study with both novices and music experts. We orient the study's design on the evaluation already presented in Chapter 4.

5.2 Related Work

Rhythm is one of the essential features of music [175, 116, 321]. Understanding its characteristics and relations is crucial for music analysis tasks such as structure analysis, pattern identification, and interpretation [319]. During such analytical processes, the structure of music can reveal rhythmic relations. However, analyzing rhythm and its reflection in the music's structure requires proficiency in music theory, a prerequisite that challenges novices and even intermediately-experienced musicians [50]. Rhythm visualizations offer one way to overcome this obstacle, exploiting the human's visual cognition ability. The augmentation of CMN with abstract visualizations enables close and distant reading [145]. In Chapter 3, we applied such a combination to visually aid harmony analysis with *harmonic fingerprints*.

5.2.1 Music Structure Analysis

A common *music structure analysis* task is to extract the temporal sections, often referred to by lettered labels (i.e., $A, B, C, ...$) [234]. These sections correspond to composition parts such as the introduction, exposition, or coda. The so-derived music segmentation depends on features such as harmony, rhythm, and melody [234]. While key changes and modulations provide an obvious partitioning [71], music tends to decelerate toward phrase endings, particularly of major sections, i.e., note durations become longer. Moreover, musical phrases often start with similar rhythmic patterns that indicate suitable segmentation split points. Consequently, the rhythmic qualities of a piece of music merit particular attention. Hence, it is often possible to utilize rhythm to infer the structure of a musical piece [192]. Boundaries that constitute the temporal segments in the music structure frequently correlate with rhythmic changes as in the occurrence of novel rhythmic patterns [147, 234]. On the other hand, the repetition of rhythmic patterns helps to identify the musical contents that belong to one section [234]. Radial hierarchy visualizations are particularly suitable to aid analysts in locating novelty and repetitions due to the hierarchy exhibited by rhythm and its cyclical appearance [198, 302, 82].

5.2.2 Radial Hierarchy Visualization

To visualize the hierarchical relationship between different durations occurring in rhythmic patterns, Longuet-Higgins introduced the *rhythm tree* [198]. The tree's hierarchical data structure is often visualized through *tree maps* [149] or *Icicle plots* [177]. While these approaches effectively utilize space through their squared appearance [149, 354], they are not suitable to accurately represent the rhythm's temporal aspect. To classify joint visualizations of hierarchical and time-series data, Draper et al. devise a taxonomy for radial hierarchy visualizations [82]. Before distinguishing them according to their exhibited design pattern, visualizations are divided into three categories: *Polar Plot*, *Ring*, and *Space Filling*. The *Polar Plot* radiates lines that convey semantics from an origin to display the relationships between these branches. In contrast, *Ring* visualizations display nodes on the circumference of a circle and highlight their relationships through connecting lines.

For the design proposed in this chapter, we employ a space-filling visualization to convey the semantics denoted to the hierarchy. Schulz et al. concisely describe these requirements to generate a tree layout identical to a *fan chart* [82] [299]. Draper et al. assign the visualization to their third category, *Space-Filling*, within the *concentric pattern* [82] similar to *Sunburst* visualizations [315]. We transfer the fan charts concept to the duration hierarchy since their structure is almost identical.

5.2.3 Rhythm Visualization

It is possible to visualize and analyze rhythmic aspects from musical compositions on several levels of abstraction [145]. While aggregated visualization techniques facilitate retrieving an overview of the underlying data, they struggle with conveying the details. In digital humanities, scholars often carefully examine data to understand its meaning and underlying relations on the close and distant reading level [54].

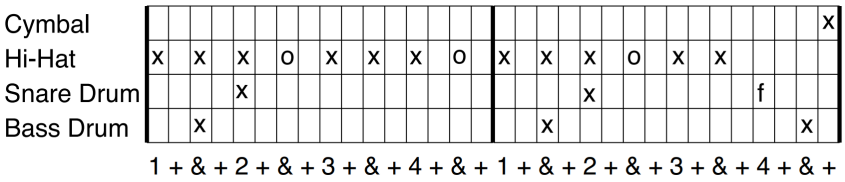


Fig. 5.2.: . This figure provided by Sethares shows a *Drum Tablature*, with percussion instruments listed on the left, indicating when and how to hit each drum [302]. The time grid and numbering below guide the rhythm. Symbols like 'x', 'o', and 'f' represent different strokes. A legend typically explains these symbols.

Typically, music readers perform close reading using the CMN that encodes all rhythmic details [225]. E.g., the *binary string* as a rudimentary visual representation can aid rhythm analysis by considering the presence or absence of sound [196]. This idea has been improved by the *Time Unit Box System* [326] and its generalization, *Drum Tablature* [302], which is illustrated in Figure 5.2 on how it can be applied.
Alternative representations allow music analysts to focus on rhythmic characteristics at a close reading level. Detaching such encodings hinders users from connecting to CMN. Another approach by Robledo exploits texture in the background of CMN, highlighting rhythmic units [276] to improve separation. These concepts use a linear sequence to visualize the rhythmic aspects but do not support the cyclic nature of rhythm [302] or their symmetry [196]. Circular representations are more suitable for these two characteristics [302, 196].
For instance, the *necklace* notation (see Figure 5.3) displays rhythmic events on the circumference of a circle [302]. By exploiting the metaphor of a clock [119], it supports the intuition of the connection between rhythmic repetitions and time. This presentation is suitable for unveiling the regularity inherent to many rhythmic patterns [302]. Similarly, the *Polygon* notation represents beats in music by connected lines. These connections build up polygons, enabling a comparison of rhythmic patterns. Such abstract visualizations facilitate understanding underlying patterns but aggravate the connection to the original representation. In the end, both close and distant reading is essential when reading music. Consequently, it is desirable to find solutions that support both concepts without sacrificing each other's advantages.

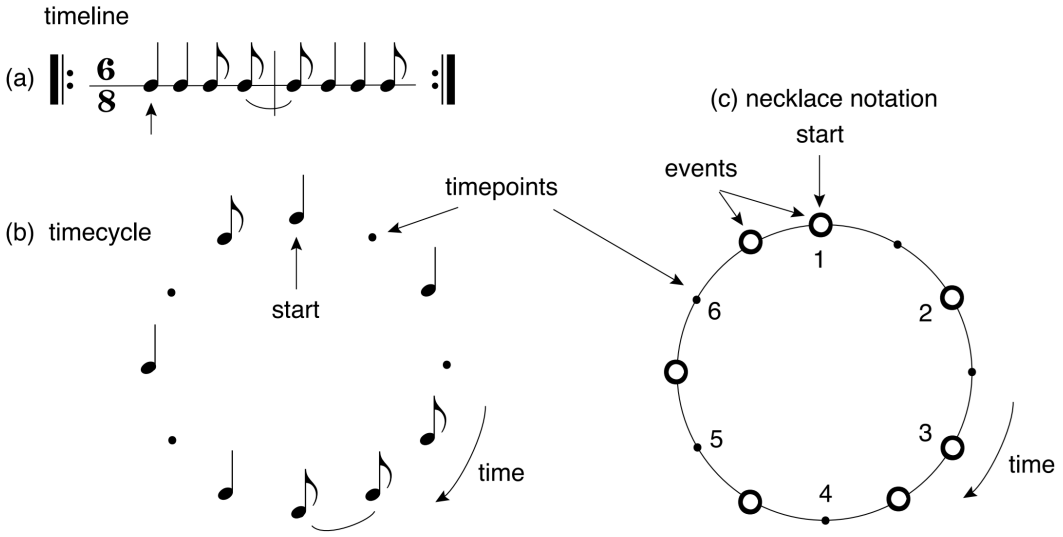


Fig. 5.3.: This illustration by Sethares shows rhythmic pattern representations: Standard Notation (a), Timecycle Notation (b) with Implicit Repetition, and Necklace Notation (c) with Donut-Shaped Events [302]. Both (b) and (c) use a twelve-point tatum grid, inferred from the $\frac{6}{8}$ time signature in (a).

5.2.4 Augmenting Sheet Music

In a review of visualizations for textual data, Jänicke et al. emphasize the importance of combined close and distant reading [145]. While the latter cannot replace the former, distant reading can point the reader toward interesting spots, e.g., patterns in the data. Regarding the discussed abstract visualizations, there is a lack of such a connection between them and the CMN. Musicians usually achieve such a combination by augmenting visualizations above or below the sheet music systems, for example, through syllables [68, 89]. In the *Generative Theory of Tonal Music* (GTTM), Lerdahl et al. provide numerous annotations through geometrical shapes, brackets, and trees to visualize rhythm features, such as articulation and meter [189]. While the elaborated approaches provide an abstraction over the CMN, they are hard to compare due to their lack of conventional structure and color, using only black and white [189, 68, 89]. Besides, their abstraction is either too narrow (e.g., syllables) or too general (e.g., multi-measure tree-like structures in GTTM). To cope with this challenge, we propose a compromise by introducing an abstract fingerprint visualization for the musical feature of rhythm, which can be attached to sheet music to ensure the visual connection to the CMN.

5.3 Rhythmic Fingerprint Design

Our proposed *rhythmic fingerprint* exploits the clock metaphor [119], a psychological anchor [340] in human cognition, and the rhythm tree [198], a music theory concept, to reflect rhythmic aspects.

5.3.1 Characteristics of Rhythm

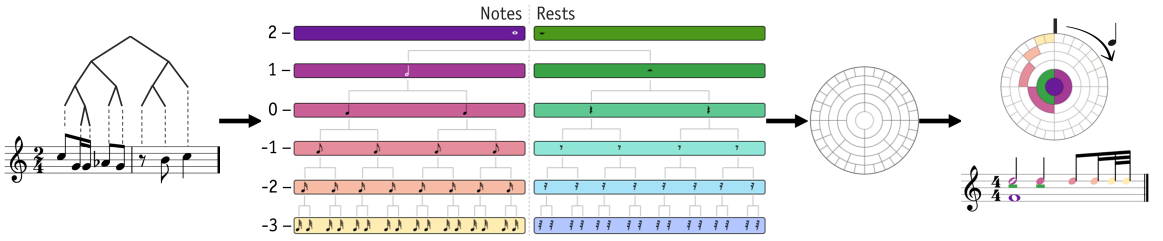
Occasionally, the term *rhythm* synonymously represents the phenomena of meter, accent, and timing [197]. In this work, we define rhythm as referring to two primary dimensions: *onset* and *duration* [175]. We consider notes as the presence of sound and its absence as rests. To quantify onset, we operate within the context of a single measure and calculate according to a predefined tatum [235] of 32$^{\text{nd}}$ notes. Typically, the duration of a note scales with a predefined tactus [235] of quarter notes depending on the time signature.

There are diverse types of time signatures, such as simple, compound, complex, and additive meter [269]. To address all specific attributes of each time signature type exceeds the scope of this chapter. The following detailed description addresses the *simple time signature*, but we later explain in section 5.6 how we could extend our approach to deal with other time signatures like *compound meter*.

A fundamental quality of duration is hierarchy [198]. This characteristic emerges in the rhythm tree as depicted in Figure 5.4(a). The rhythm tree is similar to the *phrase structure tree* that Chomsky introduced for the syntax analysis in linguistics [58] due to the structural similarities between language and music [22].

Another characteristic of the rhythm tree is the geometric progression of note durations [317]. The *duration* of notes adheres to

$$d(i) = r^i \tag{5.1}$$



(a) The *rhythm tree* [198] puts the duration of notes and rests within measures into relation.

(b) This rhythm tree shows the relationship between subsequent duration layers for *simple time signatures* only. We use two *complementary* color scales that preserve the duration order for rests and notes.

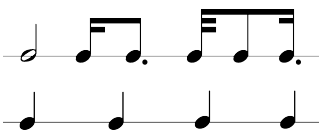
(c) A *skeleton* of the rhythmic fingerprint.

(d) *Three voices* in a measure. Notes and rests are colored according to (b).

Fig. 5.4.: We utilize (a) the hierarchy present in the rhythm tree together with (c) the clock metaphor [119] to create a space-filling concentric visualization, the *rhythmic fingerprint*. (d) Within that, we encode notes and rests in their corresponding disk. We fill the arc at the note's or rest's offset with (b) the color corresponding to their duration. This design is only suitable for simple time signatures with smaller or equal numerator than the denominator, such as $\frac{4}{4}$, $\frac{3}{4}$, or $\frac{6}{8}$.

where $i \in (-\infty, 2] \cap \mathbb{Z}$ and $r = 2$. Here, i represents the layer *index* in the rhythm tree (see Figure 5.4(b)), and r reflects the common *ratio* between successive layers (i.e., the duration length bisects at every step from top to bottom). E.g., the longest note (i.e., a whole note) has a duration of $d(2) = 4$, whereas the shortest note (i.e., a 32$^{\text{nd}}$ note) has a duration of $d(-3) = 0.125$. As discussed, geometric progression does not cover deviating. more complex rhythms, such as dotted notes and triplets.

The rhythm tree represents the complexity of rhythm in a measure as the distribution of rhythmic contents in the tree's hierarchy. This measure of complexity is similar to the Kolmogorov complexity [123]. For example, the first rhythm (1) is more complex than the second rhythm (2), which repeats a quarter note four times. Such complexity of rhythm affects the understanding and performance of musical pieces [321]. Related to this complexity is the measurement of evenness, which refers to onsets equally distributed over time [325]. In this example, rhythm (2) has a higher evenness than rhythm (1) due to identical inter-onset intervals. The *inter-onset interval* measures the temporal distance between two subsequent onsets. Considering rhythmic evenness helps to classify rhythmic patterns [325]. Maximal evenness depends on whether all inter-onset intervals are equal, as for rhythm (2). In this case, the rhythm is even mirror-symmetrical. The notion of symmetry is another relevant factor in analyzing the structure and rhythm [158, 328]. Thus, considering symmetry and evenness for visual designs supporting rhythmic analysis is essential.



5.3.2 Design Rationale

Reading a radial clock is a common task anchored in human intuition [340, 119]. Fuchs et al. show the advantage of this representation for identification and comparison tasks compared to other representations [100]. Therefore, we use a radial template (see Figure 5.4(c)) for the proposed rhythmic fingerprint design that is based on the clock metaphor [119] to reflect the temporal and repetitive aspects of rhythm. In the circular design, we lay out the content of a measure by starting at the top center and continuing in a clockwise direction. Then we map the structure of the rhythm tree to this layout (see Figure 5.4(b)). Since a whole note has the longest duration, one traversal of the circle represents a duration of $d(2) = 4$.

Our proposed design uses a concentric disk arrangement, similar to a fan chart, to radially depict the duration layers of the rhythm tree. The innermost disk represents the top layer (i.e., semibreve) of the rhythm tree (see Figure 5.4(b)). The number of layer nodes, depicted as arcs, is given by

$$a(i) = \frac{t}{d(i)} \tag{5.2}$$

where $i \in (-\infty, 2] \cap \mathbb{Z}$, $t = 4$, and d is the function given in Equation 5.1. Here, i represents the layer *index* in the rhythm tree (see Figure 5.4(b)) and t the *tactus*-level of a quarter note. Since a is strictly increasing, there are more outer than inner disk arcs, which is favorable as they have more display space devoted to them. We define the tatum to a 32$^{\text{nd}}$ note (i.e., $i \in [-3, 2] \cap \mathbb{Z}$). This general definition allows for a flexible extension to a shorter tatum such as 64$^{\text{th}}$ notes if required.

The starting point of an arc depicts a note's onset. Therefore, the starting point of each disk's first arc is located at the top center within the fingerprint's circle (see Figure 5.4(d)). For the remaining arcs, we retrieve the *starting points* according to

$$s(i) = \left\{ \frac{c}{x} \,\middle|\, x \in [1,\ a(i)] \cap \mathbb{Z} \right\} \tag{5.3}$$

where $i \in (-\infty, 2] \cap \mathbb{Z}$, $c = 360°$, and a is the function given in Equation 5.2. Here, i represents the disk's *index* c the *circumference*. The space between two starting points in a disk visually reflects the arc's duration beginning at the former starting point.

To convey the rhythmic contents of a measure, we color the arcs using the color scales illustrated in Figure 5.4(b). The outer disks use a lighter color to avoid a visual bias towards their arcs since shorter notes appear more often than longer notes. The exemplary measure in Figure 5.5 is a constructed composition of notes and rests to illustrate how three voices are encoded by the proposed rhythmic fingerprint. The third voice plays a whole note that fills the innermost disk, while the second voice comprises two half rests filling the second disk. The first voice contains multiple notes of decreasing duration encoded through arcs in different disks. This example illustrates the conflict between a note and a rest of the same duration at the same offset (i.e., a half note and a half rest at offset 0). As both would address the same arc, we prefer to show the presence of sound rather than its absence. Similarly, we aggregate multiple notes or rests of the same duration at the same offset since their multiplicity does not change the perceived rhythm [192].

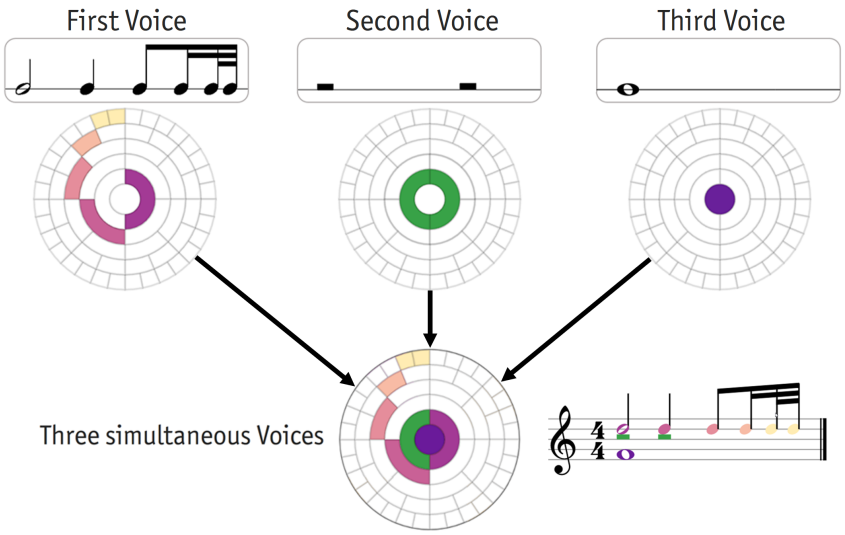


Fig. 5.5.: The rhythmic fingerprint encodes simultaneous voices. The *first* (notes) and *second* voice (rests) overlap in the *first half* segment. We favor notes over rests (here, the first half note). The third voice has no overlap with the other voices.

Apart from the durations and onsets covered by Equation 5.1 and Equation 5.3, music contains further rhythmic phenomena, such as dotted notes whose duration is given by the extension of Equation 5.1:

$$d_n(i) = \sum_{k=0}^{n} d(i-k) \tag{5.4}$$

where $i \in (-\infty, 2] \cap \mathbb{Z}$. Here, n reflects the *number* of dots that extend a note, i represents the layer *index* in the rhythm tree, and d is the function given in Equation 5.1. There are two additional mechanisms the rhythmic fingerprint utilizes to accommodate those appearances. We visualize a dotted note whose base duration belongs to disk i in disk $i+1$ to emphasize the note's extended duration. Thus, a dotted quarter will reside within the half notes disk, for example.

5.3.3 Extending the Score With Rhythmic Fingerprints

MusicXML [112] is a standard file format that is used to share digital sheet music and also to provide layout information. MusicXML has been widely adopted due to the rise of services like MuseScore [240], IMSLP [120], and their joint initiative, OpenScore [251], which distributes compositions in the MusicXML format. We leverage this advantage to extract rhythm features, including onset and duration for all notes in every voice, from music sheets using `music21` [67]. At the same time, we render the MusicXML with OpenSheetMusicDisplay[109] (OSMD), an open-source library as a Scalable Vector Graphics (SVG). By extending OSMD with D3.js [31], we can place the rhythmic fingerprints on top of the score. To ensure flawless augmentation, we enlarge the space between the musical systems to position the rhythmic fingerprints without overlapping with the CMN.

5.4 Use Cases

We present three use cases supported by the rhythmic fingerprint introduced in Section 5.3. These use cases illustrate how the rhythmic fingerprint design facilitates music analysis and interpretation tasks.

5.4.1 Interpretation of Rhythm

Rhythm exhibits many characteristics, as we outline in Section 5.2. Analysts need to examine each measure to judge a composition's complexity to extract the encoded rhythm. The rhythmic fingerprints support this task as they enable judging the complexity based on the double encoding of rhythm through color and position. Higher color diversity indicates a higher complexity, while the opposite holds for a narrow color spectrum. Along with the complexity of rhythm comes its evenness. This feature depends on onset and duration, which are difficult to measure, especially in polyrhythms. Nonetheless, evenness is vital to compare rhythmic patterns and classify music based on it. The rhythmic fingerprints support this task, reflecting evenness in their visual patterns. For performance preparations, a reader could be

interested in finding out if and where a composition is lively or slow. The color distribution of the rhythmic fingerprints aids this task, especially for longer sheet music, where most darker colors hint at slow music while lighter colors at the outer layers indicate the opposite.
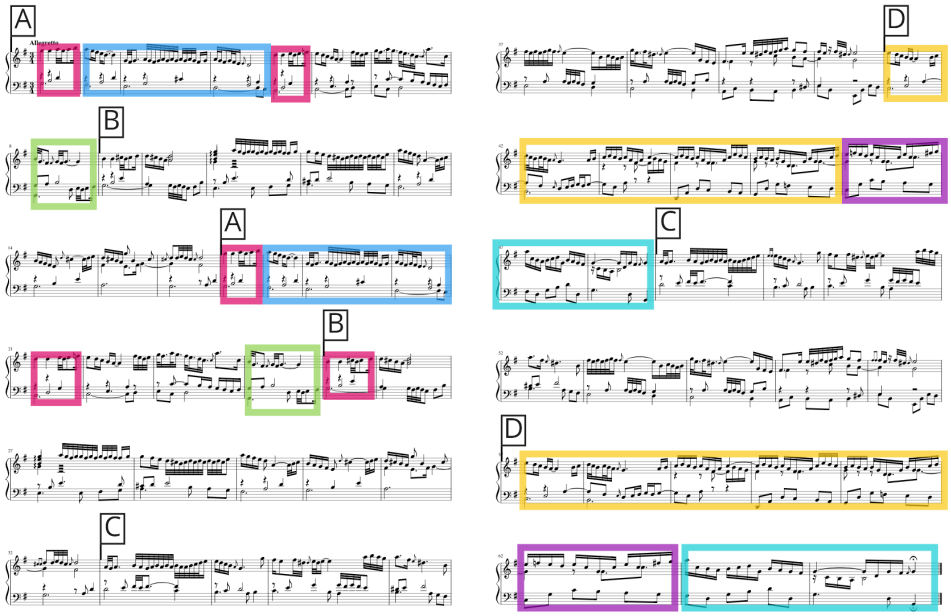
Music Structure Analysis
In Section 5.2, we discussed how rhythm affects the structure of music. The repetition of and the change in rhythmic patterns constitute section boundaries. To detect rhythmic changes and repetition, an analyst would need to closely analyze single measures by examining all notes to compare them with the content of other measures. Typically, this manual analysis process is tedious and time-consuming. The annotated rhythmic fingerprints can be of help through the visual patterns. Instead of delving into the CMN, the rhythmic fingerprints guide the user to recurrent rhythmic patterns. By keeping the connection to the CMN, analysts can verify hypotheses generated based on the fingerprints.

5.4.2 Comparative Analysis of Rhythm
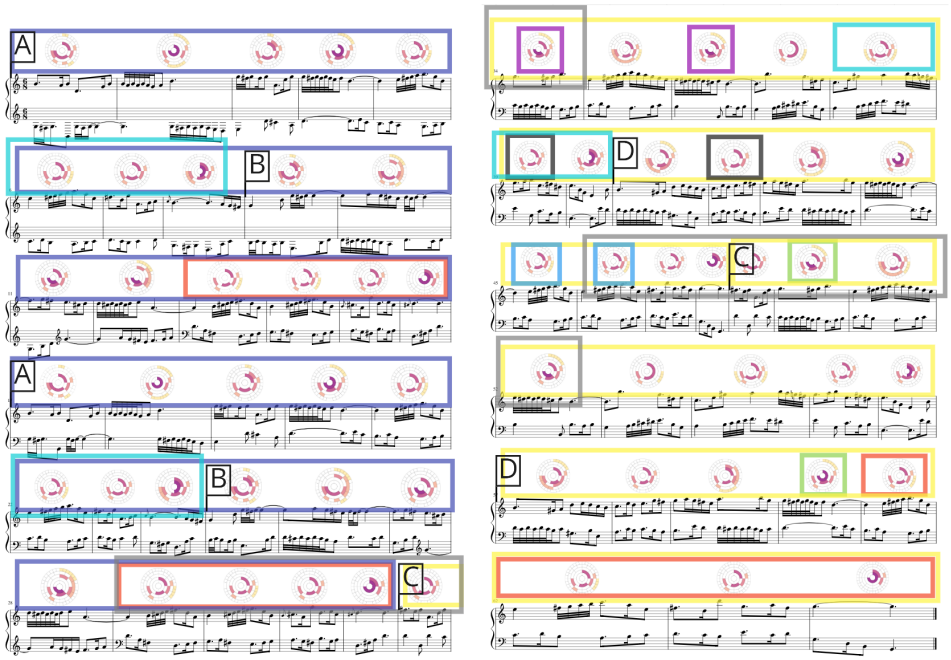
The rhythmic fingerprints facilitate the comparative analysis of multiple compositions by combining rhythm characteristics and music structure. Instead of individually comparing measures and their notes across compositions, the rhythmic fingerprints aggregate multiple voices into visual patterns. Combining their consistent structure and color encoding establishes a visual appearance of measures as single units more accessible to an analyst. They can visually match the exhibited patterns to identify rhythmic similarities, the characteristics of the present rhythms, and the music structure. The rhythmic fingerprints enable pattern matching due to their preattentive nature using color and consistent positioning. Comparing two scores with rhythmic fingerprints, a music analyst can distinguish a fast introduction from a slow opening, for example. The former features short-lasting notes in the rhythmic fingerprints' outer disks, creating bright ring-like appearances. The slow-paced opening consists of longer notes that the rhythmic fingerprints display at their center, creating a dark and compact representation. Analysts can conclude the difference between the two pieces of music at a glance, easing the comparative analysis.

5.5 Evaluation

To evaluate the rhythmic fingerprint introduced in Section 5.3, we conducted a qualitative user study similar to that presented in Chapter 4. We tasked each participant to identify patterns in two music sheets, one with the rhythmic fingerprints and the second without. This approach enables us to assess the advantages of rhythmic fingerprint augmentation and elicit qualitative user feedback. To highlight the differences between novices and experts, we divided the participants into two groups based on their expertise level regarding music theory knowledge and rhythm analysis.



(a) We asked N_2 to highlight recurring patterns in MS_1 *without* the rhythmic fingerprints. Besides identifying pattern D by three sub-patterns, he could only identify small parts (e.g., in A) or none (e.g., C) of the other patterns.



(b) We asked N_2 to highlight recurring patterns in MS_2 *with* the rhythmic fingerprints. He created the pattern C, treated AB as a single repeating pattern and found multiple nested patterns throughout larger patterns.

Fig. 5.6.: We instructed the participants to find and mark as many recurring patterns as possible by colored rectangle annotations within the presented music sheet using the CMN (a) without and (b) with the rhythmic fingerprint augmentation.

5.5.1 Study Methodology

Our qualitative assessment of the rhythmic fingerprints comprises three steps:

(1) the introduction of the design

(2) the analysis with and without the fingerprints

(3) gathering feedback during the analysis

Subsequently, we describe the used datasets, the participant characteristics, and the study procedure.

Data Sets and Ground Truth – During the study trials, the participants analyzed two music sheets, *Aria* (MS_1) [239] and *Variation VII* (MS_2) [239] from the *Goldberg Variations* by Johann Sebastian Bach. To ease the assessment of the analysis task, we exploit available ground truth [37]. Both music sheets MS_1 and MS_2 are equally long, consisting of 64 measures, and have the time signatures $\frac{3}{4}$ and $\frac{6}{8}$, respectively. We chose these compositions due to their comparable complexity and musical form. The 64 measures are divided into two parts, namely P_1 and P_2. Each part covers 32 measures and consists of a repeated pair of sections: AB for P_1 and CD for P_2 with a resulting form $ABABCDCD$ as indicated by the lettered labels in Figure 5.6. The rhythmic patterns reflect this overall structure, allowing us to compare the results between the unmodified CMN and the augmented music sheet.

Participants – We conducted our study with eight participants that we divide into two groups of different expertise levels. The first group consists of four *novices* (N_1 - N_4) who have little experience in music theory (mean score of 2) and never performed rhythm analysis before (mean score of 1). The second group includes four *experts* (E_1 - E_4) who have more advanced experience in music theory (mean score of 2.5) and are more proficient in rhythm analysis (mean score of 3.25). All participants in both groups have a university degree and are aged 29.5 ± 4.5 on average.

Procedure – The evaluation comprised three phases. In the first phase, the participants provided demographic information and their level of expertise based on a Likert scale from beginner (1) to expert (5). We familiarized the participants with the rhythmic fingerprint design during the second phase. Then, we verified their understanding of the explanations by presenting examples each participant had to solve correctly as a training step. We continued explaining the rhythmic fingerprint using an exemplary music sheet to introduce polyphonic rhythms they would encounter during the analysis. In the last study step, the participants analyzed two music sheets with and without the rhythmic fingerprints. We also randomized the condition order such that each condition applies to one novice and one expert. For every condition, we asked the participant to identify recurrent rhythmic patterns that are exact matches throughout all voices, limiting each study pass to 30 minutes. Eventually, we elicited feedback regarding the analysis and the rhythmic fingerprints, depending on the condition, during an interview. We repeated the same process for the second pass and concluded by comparing both analysis sessions in the final interview to learn about strategies for the given comparison task.

5.5.2 User Feedback Results

We elicited qualitative feedback from each participant and assessed their performance by comparing their results with the ground truth.

Pattern Identification Strategies – During the analysis, the participants followed diverse strategies to identify rhythmic patterns. Without the rhythmic fingerprints, most of the participants focused on a single stave. N_2 stated that he "*focused on the second voice because it was easier as it contained fewer notes*". Meanwhile, N_3 started differently and "*looked at the melody voice, because that's the most distinctive*". Later, both combined their findings with the other staff to match rhythmic patterns. E_3took another approach as they argued that "*16th and 32nd notes are [...] quite memorable or quarters as a starting point, they do not occur [...] often*". This idea concurs with E_4 who found that "*32nds always stand out extremely*". This strategy focuses on salient features of the visual appearance created by the CMN to form rhythmic patterns.

Given a music sheet with the rhythmic fingerprint augmentation, the participants, except for N_1 and E_1, favored the fingerprints to spot recurrent rhythmic patterns. Consequently, many users focused on the rhythmic fingerprints while neglecting the CMN. Still, E_4 "*tried to double-check [the results]*" and referred to the CMN as a means of verification. The participants preferred salient aspects of the fingerprints to identify recurrent rhythmic patterns with the rhythmic fingerprints. As examples they gave the "*dark color [which] is striking*" (N_4), "*long rhythmic runs*" (E_1), the "*yellow on the outside*" (N_2), and "*the green color [which] is [...] prominent*" (E_3).

Usefulness of the Rhythmic Fingerprints – We received valuable user feedback from our participants. All of them, except for N_1, appreciated the rhythmic fingerprints as they eased the analysis. Not only "*have [they] greatly accelerated the analysis*" (N_3) but they "*create visual[ly] clear patterns*" (N_2) such that one "*could easily find the similarities*" (N_2) and "*larger continuous patterns*" (N_3). During his first analysis with the CMN alone (see Figure 5.6(a)), N_2 did not find any ground truth patterns. By contrast, in his second analysis with the rhythmic fingerprints (see Figure 5.6(b)), he could identify the two parts, P_1 and P_2, as well as their repeating sections AB and CD, respectively. In need of the next analytical step, E_1 fittingly noticed that "*a fingerprint would be helpful for a new push because it was much easier to look for things that stand out*" while analyzing the CMN alone. E_3 even mentioned that "*someone who has never dealt with rhythm before [could] still analyze rhythm here*" and "*[wouldn't] need to know notes*".

Challenges – Before the study, no participant knew the rhythmic fingerprint and faced different challenges. At the start, many participants "*found it [...] difficult to adapt to the fingerprint*" (E_2). Besides the steep learning curve, the participants struggled with the color map. Particularly, E_3 found that "*the colors are too similar, especially the red [ones]*" which E_2 emphasized, concerned that "*the 32nds are difficult to see*" on white background. They further elaborated that "*it is more difficult to search on the outer circle compared to the inner one*" which led N_3 to " *look at them almost as closely as the notes*", weakening the rhythmic fingerprints' use. If not looked at closely enough, N_1 and E_1 feared "*a likelihood of confusion because the circles look very similar*" (N_1). N_3 and E_4 experienced a similar irritation as they "*lost the context [when only focused] on the glyphs*" (N_3) and tried "*to remember the fingerprints and to find them somewhere else*" (E_4).

5.6 Discussion

Based on the study results, we assess the pattern identification performance of the participants. Each participant had a similar or better performance with the fingerprints than with the CMN only. Since we provided an elaborate introduction to the rhythmic fingerprints before performing the analysis task, both novices and experts understood and thus profited from the rhythm visualization. This familiarity could be a reason for the users' trust in the rhythmic fingerprints. Although N_1 and E_1 primarily relied on the CMN, the remaining participants mainly worked with the visualization. Although some users reported a steep learning curve for the rhythmic fingerprints, N_1 argued that using color may inspire musicians to analyze the rhythm due to the intriguing visual design. E_1 added that the rhythmic fingerprints could give a new impulse during the rhythm analysis.

5.6.1 Lessons Learned

The main discussion points gathered during the rhythm analysis process and the user's feedback are:

(1) the fingerprint inherently encodes rhythmic complexity
(2) the fingerprint does not separate between different voices
(3) the fingerprint encodes a limited number of time signatures

Rhythm Complexity

The rhythm tree and the rhythmic fingerprints express complexity through the diversity of onsets and durations. E_2 accurately noticed that a greater variety corresponds to higher complexity. Therefore, the amount of present notes influences readability. In the final interview, E_3 reported that due to the limited space of the inner disks and the adjacent placement, it was challenging to distinguish the colored arcs. Analogously, E_1 stated that the outer disks were more challenging to compare than the inner disks due to the smaller size. While the CMN enables music readers to view multiple voices and their rhythmic progression separately, the rhythmic fingerprint aggregates the commonalities and reveals the rhythm perceived by listeners. Consequently, the fingerprints do not replace the well-established CMN, which analysts can use to understand the details of multiple voices, which is similar to the results we achieved from the study about the harmonic fingerprints in Chapter 4. We also learned that it is currently difficult to see the arcs at the outer disk due to the small lightness difference to the background. To address this or other visual impairment issues, there would be no effort to tailor the color scale to the reader's needs or maybe enable the users to adapt the representation for themselves.

Voice Separation

The current design does not support readers distinguishing separate voices in a *multi-voice setting* as in the CMN. To achieve this, readers must still examine the CMN to understand voice separation. Due to our decision to emphasize the presence of notes over rests in cases of overlap with the same offset and duration, the rhythmic

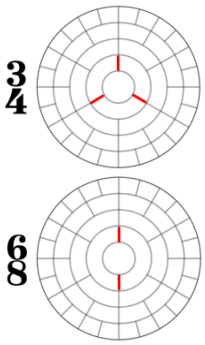
fingerprint aggregates simultaneous notes and rests. As outlined in Section 5.3, we favor the presence of sound over its absence. Consequently, the rhythmic fingerprint is not bidirectional. It visually simplifies parallel rhythmic content to extract perceived rhythm, which is not altered by multiple notes with identical onset and duration. Due to this simplification, the rhythmic fingerprint design does not support multiple voice analysis. Tailoring the design to such tasks would increase its complexity since more information would be represented.

Time Signature

One traversal of the rhythmic fingerprint represents a whole note's duration. Thus, even longer durations cannot be displayed correctly. Simple time signatures can be analyzed without issues, such as common, $\frac{3}{4}$, or cut time. E_2 noticed that the design of the rhythm visualization allowed him to determine a composition's time merely by looking at the appearance of the rhythmic fingerprint. Meanwhile, none of the other participants faced issues during their analysis due to our design choice. On the contrary, the visualization does not properly accommodate all possible time signatures, such as $\frac{5}{4}$. Since such time signatures would require more than one traversal, a more flexible design would be required.

5.6.2 Limitations and Future Work

While the study showed that the rhythmic fingerprints support the analysis of rhythmic patterns even when used by novice readers, there are still restrictions and limitations of the proposed approach. First, there are numerous time signatures of varying complexity [269]. The most common time signatures are simple (e.g., $\frac{3}{4}$) and compound (e.g., $\frac{6}{8}$) meter. The current fingerprint skeleton (see Figure 5.4(c)) only addresses simple meters. While it is theoretically possible to visualize compound meters with the current design, it does not properly accommodate the tertiary grouping, e.g., three eighth notes encompassed by a dotted quarter note. The two skeletons to the right show how $\frac{3}{4}$ and $\frac{6}{8}$ differ (i.e., red lines in the 2nd inner disk) regarding their skeletons. The main difference between the duple and the triple meter is the perception of emphasized beats in a measure [198], which is especially relevant for performers who need to know which notes to emphasize. Typically, in $\frac{6}{8}$ meter, dotted quarter notes indicate the composition's regular pulse.



There are different directions in which the proposed design could be further extended. Figure 5.7 demonstrates a different approach to consider different meter signatures that generalize towards multiple simultaneous voicings to avoid overlapping. This design would optimize voices with shorter notes to be placed on the outside of the ring, which would, in many cases, lead to scenarios where the lower voices tend to be placed in the center and voices in the middle or higher ranges on the outside of the ring. The applied colormap is slightly different from the one in Figure 5.4(b) balancing the color intensity throughout the note duration classes.

Figure 5.8 shows the same eight measures as Figure 5.1 at the start of this chapter for comparison. The direct comparison shows that some of the glyphs of the new

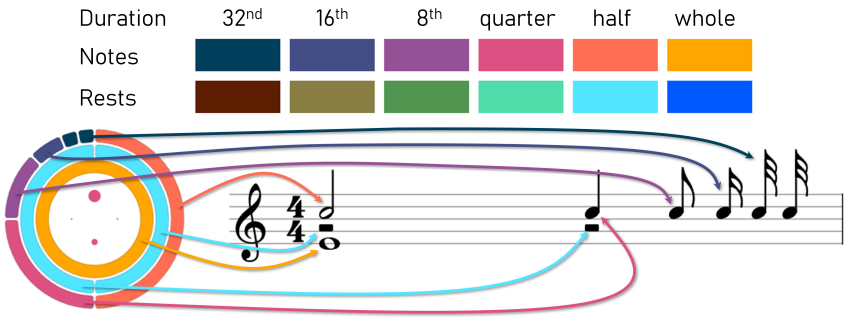


Fig. 5.7.: Alternative rhythmic fingerprint design to address the limitation of voice overlapping for simultaneous voices using a slightly different color map compared to Figure 5.4(b). This example contains three voices and the half break and half note at the start of the measure would overlap in favor of the note in the initial rhythmic fingerprint design evaluated in this chapter.

design are smaller or larger depending on the number of voices detected within single measures while trying to minimize the number of voices within a fingerprint. We also see that the space at the center is used to emphasize the meter (here $\frac{3}{4}$). An advantage of the initial design is that the fingerprint remains the same independent of the meter signature, facilitating the comparison between different signatures and keeping the same note durations in the same ring, while the newer design addresses the different voicings and the overlapping issue. Note that the color maps between the two designs are different and should not be directly compared in the illustration of Figure 5.8. The size of the latter fingerprints depends on the number of voices to be encoded, which could lead to scalability issues, especially when we try to visualize sheet music of orchestral music with dozens of instruments. In orchestral scenarios, multiple instruments often play the same rhythmic structure which then could be condensed in a transformative processing step before augmenting the sheet music.

The design illustrated in Figure 5.7 especially supports meter signatures that have a larger numerator than the denominator, such as $\frac{5}{4}$ or $\frac{12}{8}$, which would always lead to unwanted overlapping in initial rhythmic fingerprint design. For a more sophisticated argumentation, it would be required to make a comparison study of which of the designs is favorable depending on a given task, such as voice comparison or rhythmic analysis of larger music collections. This newer design only suggests how changing

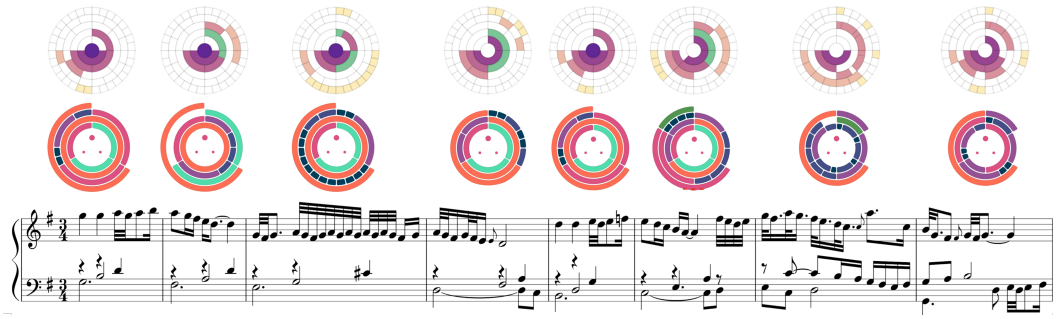


Fig. 5.8.: Direct comparison on the initial example (see Figure 5.1) between the two different rhythmic fingerprint designs. While the initial design is more consistent between varying meter signatures, the novel design improves voice separability.

the design could tackle existing limitations. We did not compare the two designs, but we encourage other researchers to pursue original research on this issue.
Besides this additional example, we consider it an open research gap to investigate how the design can be adapted to support other time signatures [269] with additional interaction techniques. Consequently, tailoring the visualization to the individual needs of music analysts to overcome the current restrictions of static rhythmic fingerprints would be worthwhile, too. We want to provide a visual music analysis interface that integrates such interaction opportunities with digital sheet music to facilitate access based on MusicXML [112]. To address this, in Chapter 7, we introduce MusicVis, which introduces this integration into an application leveraging the proposed rhythmic fingerprints. With MusicVis, we extend the temporal range for the fingerprints from single measures to larger sections of a musical composition to provide overviews of the rhythmic content. As a result, we argue that music analysis benefits from visualization research for close and distant reading, an important concept in digital humanities [145].

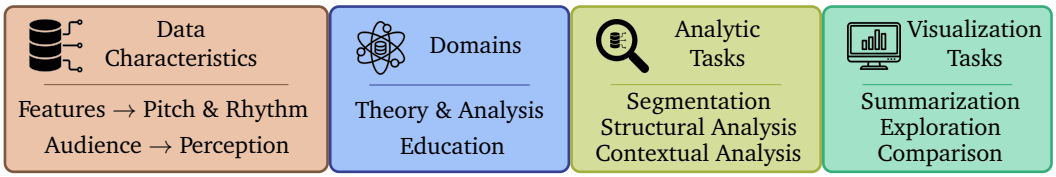
5.7 Summary

In this chapter, we proposed a visualization design for rhythm analysis by exploiting a music theory concept, the rhythm tree, to extend traditional music notation similarly as discussed in Chapter 4 concerning the harmonic fingerprints. The augmentation of sheet music with the rhythmic fingerprints combines typical close reading with the distant reading of sheet music. We conducted a qualitative user study with four novices and four experts to evaluate our approach. The user performance assessment indicates that our visualization improves the identification of rhythmic patterns. During the analysis, users could determine the music structure and characteristics of rhythm with the help of the rhythmic fingerprints. Through the participants' feedback, we also identified difficulties that analysts face while working with the proposed rhythmic fingerprints. These include the readability of complex rhythms, the separation of voices, and the display of irrational or complex time signatures. In Chapter 7, we combine the rhythmic and harmonic fingerprints as exchangeable entities augmenting digital sheet music based on MusicXML by enabling the joint analysis of harmony and rhythm using abstract and aggregated representations.

6

MELODYVIS: Visual Analytics for Melodic Patterns in Sheet Music

> *The best visualizations are those that expose something new about the underlying patterns and relationships contained within the data.*
>
> **— Martin Wattenberg**
> *American Data Visualization Scientist*



Manual melody detection is a tedious task requiring high expertise level, while automatic detection is often not expressive or powerful enough. Thus, we present MELODYVIS, a visual application designed in collaboration with musicology experts to explore melodic patterns in digital sheet music. MELODYVIS features five connected views, including a Melody Operator Graph and a Voicing Timeline. The system utilizes eight atomic operators, such as transposition and mirroring, to capture melody repetitions and variations. Users can start their analysis by manually selecting patterns in the sheet view and then identifying other patterns based on the selected samples through an interactive exploration process. We conducted a user study to investigate the effectiveness and usefulness of our approach and its integrated melodic operators, including usability and mental load questions. We compared the analysis executed by 25 participants with and without the operators. The study results indicate that the participants could identify at least twice as many patterns with activated operators. MELODYVIS allows analysts to steer the analysis process and interpret results. Our study also confirms the usefulness of MELODYVIS in supporting common analytical tasks in melodic analysis, with participants reporting improved pattern identification and interpretation. Thus, MELODYVIS addresses the limitations of fully-automated approaches, enabling music analysts to step into the analysis process and uncover and understand intricate melodic patterns and transformations in sheet music.

Chapter 6 is based on the following publication. Please refer to Section 1.4 for contribution clarifications.

[223] **Matthias Miller**, Daniel Fürst, Maximilian T. Fischer, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. **"MelodyVis: Visual Analytics for Melodic Patterns in Sheet Music"**. *To be resubmitted.*

6.1 Introduction

For a very long time, music has been an expression of personality, cultural identity, and a universal factor in religious traditions [137, 136]. Thus, it has been integral to society, human culture, and individual experiences [64]. Especially the melodic character plays a pervasive role, possessing the remarkable ability to stir emotion, create social bonds, and enhance cognitive processes when people engage with music [348]. Throughout time, different musical trends developed across many cultures [291]. The importance of music and its overall cultural evolution for humans is estimated to have started around 43 to 50 thousand years ago [220]. For a long time, music was only preserved through oral traditions [132]. Apart from physical artifacts and depictions of musicians and instruments, the oldest known musical notation, dating back to approximately 1,400 BCE, was discovered on a Sumerian clay tablet in modern-day Syria, near Ras al-Shamra, the ancient city of Ugarit [121, 84]. The tablet features engraved melodies, including Hurrian Hymn No. 6, also known as the "Seikilos Epitaph" which historians consider the earliest complete surviving musical composition in a notated form. Several music notations have been developed to pass musical knowledge on to future generations and composed hymns and melodies as part of the cultural identity. Most common to many today is the Common Western Music Notation (CWMN) [297].

The advantage of visual representation of music is the extraction of detailed musical notation from auditorial information that is subject to interpretation by a singer or musician [45]. Musical notation, in particular sheet music [86], provides a precise representation of a composition, allowing for a more accurate investigation of patterns and themes compared to audio recordings [182]. Both formats describe musical information on different semantic levels. Sheet music comprises abstract parameters such as notes, keys, measures, and other instructions in a visual form,

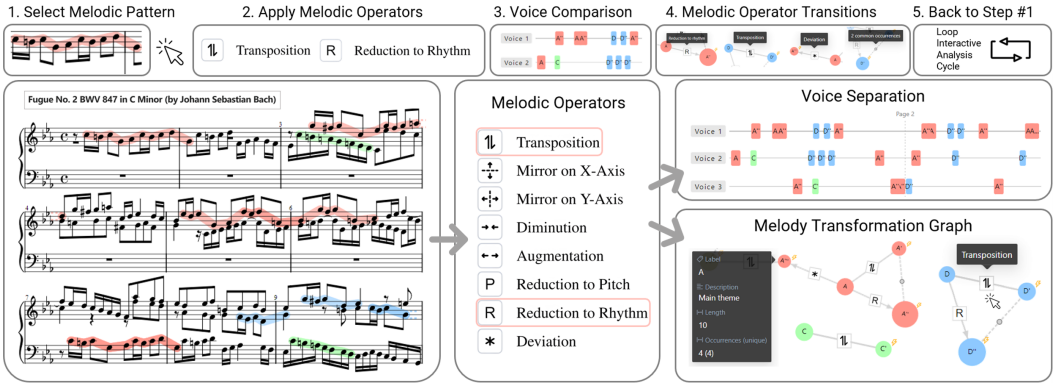


Fig. 6.1.: MELODYVIS provides an interactive workflow for analyzing melodic patterns in sheet music. Analysts can interactively select a pattern (1) that is used to apply melodic operators, (2) leading to further highlights in the *sheet view*. Then, the *Voice Separation View* (3) provides an overview and distribution of the different voices, and the *Melody Transformation Graph* (4) supports analyzing melodic variations. Finally, the analyst can continue adding further patterns by looping through the analysis process (5).

providing a detailed visual explanation of how a musical composition should be performed. In contrast, audio recordings result from music in an acoustic form comprising detailed nuances in interpretation and expression, which are generally not written down but make music "come alive" [97]. Thus, analyzing patterns in sheet music facilitates the study of a composer's intentions and the structural elements inherent in a piece while audio analysis involves the interpretation of a musician, which may introduce deviations or inaccuracies [97]. Moreover, sheet music captures essential musical elements, such as pitch, rhythm, dynamics, and articulation [225], in a visually accessible format that simplifies pattern identification and comparison across different compositions from a larger corpus [227]. Sheet music's notational unambiguity enables researchers to examine subtle variations in melodic and harmonic structures and uncover relationships between different pieces or composers that may not be readily apparent in audio recordings. For this reason, most musicologists prefer written (sheet) music to analyze the meaning of music due to its compactness and explicitness [83].

Melodic themes, characterized by reoccurring note sequence patterns and their variations in sheet music, are often a composition's most salient and memorable aspects and of specific interest to musicologists [305]. Analyzing sheet music for melodic patterns is also crucial for assessing the originality of melodic content [233], understanding stylistic and genre-specific attributes [70], and gaining insights into historical and cultural contexts. Integrating the analyst into the analysis process through Visual Analytics (VA) is crucial because it capitalizes on human expertise and intuition [229], addressing the limitations of fully automated methods that may overlook subtle, context-specific or complex patterns in the music [358]. Visual analysis of sheet music allows for examining multiple voices simultaneously, enabling a comprehensive understanding of the interplay between various elements within a composition. In contrast, the analysis of audio recordings may struggle to disentangle overlapping voices [57], making it difficult to discern individual patterns and melodic relationships. Focusing on visual analysis of sheet music data improves the preservation of a composer's intent, enhancing pattern identification accuracy and facilitating the examination of complex musical interrelationships. This approach makes sheet music more suitable for in-depth melodic analysis, thus facilitating analysts' understanding of musical compositions [224].

In music analysis, melodic transformations are crucial in understanding how melodic material is processed and developed within a composition [80]. Researchers can gain insights into techniques employed by composers by focusing on such transformations. Studying these transformations on sheet music instead of audio recordings ensures a more precise and accurate examination of the underlying melodic structures and relationships. Manually identifying these patterns and their transformations in sheet music can be challenging and require music analysts to understand music notation thoroughly [225]. Visual Analytics offers a promising solution for visually analyzing melodic patterns in sheet music. By integrating the analyst into the process, VA facilitates a more transparent and understandable analysis process compared to fully-automated methods [229]. It can support learning and interpretation tasks and enable music analysts to explore and uncover intricate patterns and relationships that might otherwise remain hidden [224].

To address the issue of how information visualization research can contribute to melody analysis, we consider the RQ: *How can we employ Visual Analytics to enable the semi-automatic investigation of melodic patterns based on sheet music?* By focusing on melodic operators, transformations that manipulate melodic material, we developed MELODYVIS. This interactive analysis interface facilitates the visual investigation of melodic content throughout a musical composition.

Contributions – This work contributes MELODYVIS, a visual interactive application encompassing multiple seamlessly connected components based on an interactive workflow for analyzing melodic patterns. We offer a comprehensive problem characterization for melodic pattern analysis, addressing visualization requirements in the context of sheet music data. We also provide an overview of potential melody analysis tasks and describe typical sheet music melody analysis users. We conducted qualitative and quantitative evaluations to assess the applicability and effectiveness of MELODYVIS, discussing its current potential and remaining limitations. Finally, we present potential future research directions to inspire interdisciplinary collaboration between the information visualization community and the field of musicology, with a special focus on melody analysis.

6.2 Background and Related Work

Music analysts investigate melodic themes in sheet music to understand musical compositions [202]. Chapter 4, Chapter 5, and Chapter 7 demonstrate the value of augmenting digital sheet music to enhance the analysis of the structure and progression of harmony and rhythm. On a more abstract level, Sapp used tonal material of compositions to enable visual analysis of key modulation providing KeyScapes [290]. The introduction of numerous digital formats for musical scores, including MIDI [249], MusicXML [112], Humdrum `**kern` [141], Lilypond [310], ABC notation [344], MEI [280], and GUIDO notation [133], and the technological progress enabled musicologists to take advantage of computational approaches that use algorithms to extract valuable information from musical pieces automatically [122].

In this work, we specifically focused on investigating melodic sequences based on digital sheet music employing Visual Analytics. Subsequently, we discuss the primary visual approaches that facilitate melodic analysis, focusing on the existing computational methods. Then, we provide an overview of published applications that assist in melodic information retrieval and the analysis of recurring patterns.

6.2.1 Sheet Music Retrieval Based on Melodic Sequences

In the related work, various approaches aim to retrieve musical compositions from a dataset based on melodic content, employing methods to search for sheet music. These methods can be categorized into pattern-matching notes, filtering by genre or duration, n-gram approaches, contour-based searches, and multimedia queries.

For example, Koláček introduced a simple note query based on the CMN, where a user can easily add a few notes to search the *Global Chant Database* [170]. He does not provide more sophisticated melodic transformations to identify variations. Instead, a few additional fields allow for query text or genre information. The melody

query input provided on Hymnary.org is similar [295]. Still, it has a few additional options to define the duration of notes and an interactive keyboard to facilitate selecting the pitch for a query. Dig That Lick provides a similar pattern similarity search interface for different Jazz music databases introduced by Henry et al. [126]. Burghardt and Lamm use n-gram matching with a specific note and a relative interval search to identify motif transpositions. Still, they do not enable further operators or the direct selection of melodic examples from a piece [39]. They also provide an option similar to Musipedia [143] where an analyst can use the Parsons Code [259], also known as pitch contour [285]. Themefinder by Kornstädt separates the contour query into a gross and refined contour query to enable slight melodic nuances in the query in addition to an absolute pitch and relative interval query [171]. Musipedia provides some flexibility with a keyboard search, rhythm search, and microphone/hum query to look for music pieces.
Analyzing the relationship of musical patterns inside a single composition is impossible, while a user can not select a query template based on a given composition [143]. Fast Melody Finder by Rho and Hwang synthesizes an audio input to create a pitch contour query string to search a database for matches addressing singing and humming search tasks [273]. A significant drawback of most approaches is that the user does not know whether there are any results before executing the query, leading to a tedious trial-and-error approach.

6.2.2 Melodic Pattern Mining and Extraction

Besides music information retrieval tasks, some approaches address melodic pattern mining based on a given data set. In this scenario, the users don't need to provide a specific search template in advance. Still, rather algorithmic approaches identify salient and frequent melodic patterns potentially interesting to a music analyst. For example, Frieler et al. used Melodia [285], a state-of-the-art melody estimation method to extract jazz solo melodies from the audio-based Weimar Jazz Database [98]. Wen et al. created a corpus of compositions from Haydn, Mozart, and Beethoven using the MusicXML format [112] to perform melodic feature extraction and then applying a k-means clustering on top of it [352]. To identify the similarity of the main melodies of cover songs of popular songs, Tsai et al. introduced a method to match the main theme based on audio data [330].
Rigaux and Travers introduced a designed a scalable search engine for melodic queries based on ELASTIC-SEARCH for collection indexing employing variant matching [275]. Their work disembogues in the approach by Zhu et al. in an extensive framework for feature extraction and a text-oriented search engine [360] to manage large collections of digitized music documents. A drawback of these approaches is the unavailability of a user to directly explore and steer the results, as they only perform an algorithmic matching of which the results are not further analyzable by a musicologist on a lower level. They don't leverage a visual interactive interface to make the search results analyzable within the context of their findings. Rather, they focus on their approach's overall accuracy in statistically measuring the quality of the identified results compared to a ground truth data set. Soum-Fontez et al. investigated how textural characteristics can help identify melodic voices in polyphonic

music, explicitly focusing on string quartets [313]. Our approach is similar to separating multiple voices into disjoint voices based on envelope extraction [117].
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 [218]. Urbano proposes the library "*MelodyShape*" that enables similarity search for melodies based on the shape [335] 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. Thus, these approaches are not directly accessible to a broad audience that requires intuitive and interactive visualization. Miller et al. presented a first visual approach to enable direct melody analysis for sheet music data, but their melody query only allows for exact matching, transpositions, and contour search [224]. While we consider this a first step in the right direction, we identified that none of the existing visual approaches allows for a voice-separated investigation of melodic transformations. Only the Verovio Humdrum Viewer by Sapp already provides a set of melodic analysis functions, including imitation, melisma, dissonance labels, composite rhythm, and meter analysis with limited opportunity for visual investigation [289]. Therefore, we aim to address this research gap by leveraging information visualization techniques to support the melodic analysis using multiple components that support low and high-level analysis tasks.

6.2.3 Applying Methodology Transfer

Concepts from the text analytics domain within the research field of digital humanities inspire this chapter. MelodyVis is a visualization system for exploring the relationships of melody patterns in a musical composition in sheet music, inspired by the methodology transfer from Poemage [210], a system designed for analyzing the sonic patterns of poems. This is similar to the methodology transfer use case we already discussed in Section 2.5 [229]. The need for close reading [145] in melody analysis is essential, as it enables the in-depth examination of melodic patterns and relationships within sheet music, closely aligning with the manual analysis process musicologists traditionally employ. As we already discussed, melody analysis examines the complex structures formed by the interaction of melodic patterns — musical phrases connected through melodic transformations, such as transpositions or inversions — throughout the composition. MelodyVis was developed as part of an ongoing, highly exploratory collaboration between data visualization experts and musicologists. This methodology transfer adapts the exploration of sonic patterns in textual data to the analysis of melodic patterns in sheet music, enabling users to efficiently explore, interpret, and understand complex melodic relationships while maintaining their domain knowledge and intuition in the analysis process.
Based on the VMG we introduced in Chapter 2, we can categorize the contribution of MelodyVis within the field of Visual Musicology [229]. Here, we use the *Structural Features* – *Pitch* and *Rhythm* – from the *Data* branch and address the *Analytical Tasks* – *Information Retrieval* and *Contextual Analysis* within the *domains* of *Theory & Analysis* and *Education*. MelodyVis considers the three *Visualization Tasks*: *Overview/Summarization*, *Navigation/Exploration*, and *Comparison*.

6.3 Structure of Melodic Patterns

To analyze melodic patterns, it is essential to understand the meaning and relationship of MELODYVIS's main building blocks:

Melodic-Related Musical Features – Melody is a compound musical feature intricately woven from pitch and time dimensions [224]. Pitch refers to the perceived frequency of musical notes and is encoded vertically on the sheet through high or low positions [347]. This element dictates the progression and contour of the notes in a melody. Time, the horizontal dimension, involves rhythm, which combines patterns of onsets and duration in a sequence of musical notes [103], contributing to the overall flow and progression of the melody. A note's length depends on the tempo and meter given for a piece. The note's *onset* defines its temporal placement in a composition [327]. In contrast to notes, *rests* are the explicit information of the absence of sound providing temporal structure information. By focusing on the interplay of pitch and rhythm, we can analyze a melodic line's characteristics, capturing the essence of this compound feature. While other elements such as harmony [222], articulation, and timbre [225] are also related to melody, they will not be the primary focus of our analysis.

We clearly distinguish between a single melody and a chord progression. By our definition, any melodic instrument, such as the recorder, can perform a melody, which is impossible with notes that must be played simultaneously. Further, we can define *motifs* and their *variations*. Melody is the most salient feature as it is usually remembered by the listener, e.g., influencing catchy tunes. Especially the aspect of the leitmotif [38] is of high interest to music analysts, which is the leading motive that influences the central theme of a musical piece.

Voicing and Melodic Texture – Combining multiple simultaneously played voices leads to diverse musical textures [140]. We consider three melodic texture types: monophony, polyphony, and homophony [42].
Monophony, also known as monody, describes the succession of single notes of only one voice comprising a melodic line without accompanying voices [347]. It is the most basic form of musical texture with little information in the vertical component, often occurring in folk and early music [274]. Melody analysis based on monophonic music focuses solely on the primary line, which forms the entire musical structure.
Polyphony comprises multiple independent melodic lines or voices played simultaneously, creating a complex and layered texture as the different melodies intersect but still fit well together [347]. Polyphony is a common feature of music from the Baroque and Classical periods and is also found in many other musical traditions worldwide. Melody analysis for polyphonic music is more complex, as it entails understanding the interplay between these lines and their relationships in its context.
Homophony features a main melody supported by harmonic accompaniment, emphasizing the relationship between melody and harmony. Homophony also describes music in which two or more voices synchronously move together in harmony [347]. Studying homophonic compositions requires analyzing both aspects. Homophony is a prevalent texture in Western classical music and popular music.

This chapter focuses on monophony and polyphony by considering independent voices [57] without addressing harmonic progressions.

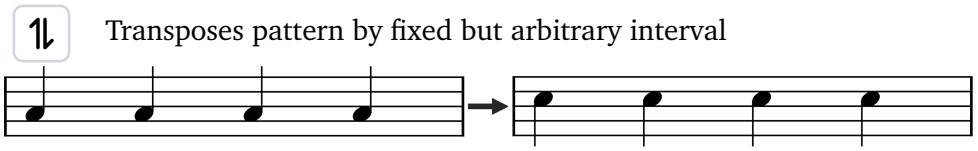
Melodic Transformation – Besides voicing and musical texture, melodic transformations such as transposition and inversion are relevant in shaping the structure of melodic patterns and their development throughout a musical composition. These techniques enable composers to craft interesting connections between different themes, leading to a broad and dynamic melodic landscape. In the context of information visualization, studying melodic transformations offers a unique opportunity to explore the underlying patterns within musical compositions, providing insights into the creative processes of composers to develop melodic material. The transitions between the different states and using the transformations as functions to adapt a pattern can be modeled as a graph. We use this visual idea for the graph visualization we implemented as a component in MelodyVis, as described in Subsection 6.4.2. Then, in a subsequent step, combining multiple transformations allows for disclosing more intricate patterns that are not easily accessible through manual analysis.

Information visualization researchers can contribute to a deeper understanding of how composers have utilized and combined these building blocks by focusing on the atomic melodic operators as the basis of these transformations. Furthermore, this interdisciplinary collaboration can lead to the development of innovative visualization techniques and tools that reveal hidden relationships, commonalities, and differences between existing melodic themes, ultimately enhancing our ability to analyze, interpret, and appreciate the complexity of musical compositions.

6.3.1 Atomic Melodic Operators

Melodic transformations in musical compositions are characteristic means composers and artists use to develop melodic material. Different transformations can be applied to a musical motif to identify repetitive melodic themes and variations in musical compositions to process the original melodic material. Subsequently, we describe eight operators we extracted from the literature that composers commonly use:

Transposition `[OP1]`

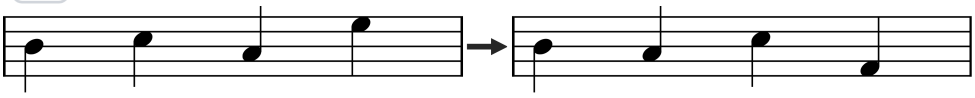


The transposition operator detects all patterns in which the pattern notes are shifted by a fixed interval while maintaining their relative positions.

We consider transposing melodic patterns as the most basic operator, as it is typically easy to identify such changes by simply listening to a song. Even detecting such occurrences of a given melodic theme in sheet music is the least difficult but requires specific musical expertise, especially for polyphonic music. In the CMN, it is also necessary for any of these operators to know how accidentals change the relative progression of a note to be able to exactly identify the application of operators. Moreover, in Chapter 7 I implemented the exact transposition feature as part of the *Melody Query*, as it is also the most often used by musical composers.

Mirror on X-Axis `[OP2]`

Mirrors pitches on x-axis through first pitch

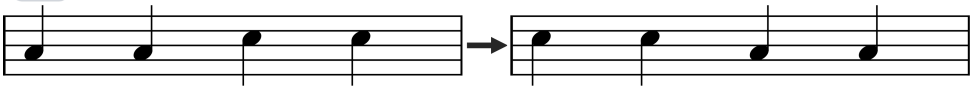


The x-axis-mirror operator identifies all patterns where the pitches of notes are reflected across the x-axis, using the first note's pitch as a reference point.

Mirroring a melodic pattern based on the x-axis is also known as *Inversion* [351]. Composers leverage this operator to create melodic variety by maintaining coherence and a sense of familiarity with the already-used theme material but presenting it in a new form. This technique is also common to establish unity within a composition but allows for further opportunities for exploring and developing a motif.

Mirror on Y-Axis `[OP3]`

Mirrors notes on y-axis of middle note(s)

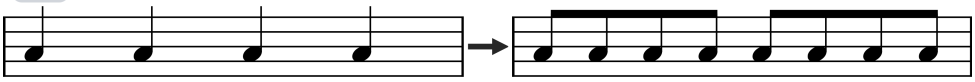


The mirror on the y-axis operator finds all patterns where the notes are reflected across the y-axis, using the middle note(s) of the pattern as a reference point.

The Y-Axis-Mirror operator is similar to *Inversion*, and musicologists also call the operator *Retrograding* [80]. This operator also leads to melodic variety and can introduce novelty, but it maintains a connection to the initial theme. Moreover, the two operators can be intentionally combined to create an even more intricate melodic landscape. For example, Arnold Schönberg extensively employed *Inversion* and *Retrograde* techniques in his twelve-tone compositions [47].

Diminution `[OP4]`

Shortens consistently the notes' length of a melody

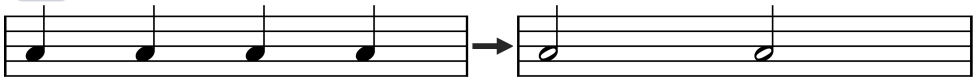


The diminution operator locates all patterns where the duration of notes is reduced by half, resulting in faster rhythmic patterns.

Diminuting melodic material also supports intensified excitement that can lead up to a climax in a composition [257]. While in the given example above, all the notes have the same duration, one can easily think of a less uniform rhythm of a melody. Then, applying *Diminution* can create more sophisticated melodic lines through the proportional shortening of the note sequence. While this technique has already been used in the Renaissance period, it continues to be a valuable composing tool for maintaining a coherent melodic core structure but introducing a sense of energy and momentum that can enhance the emotional impact or intensify urgency. The operator is also used to develop contrapuntal textures, which add depth and complexity to compositions while showcasing a composer's creativity and overall compositional skill as has been employed by Händel [207], Bach [53], and Wagner [106].

Augmentation [OP5]

↔ Proportionally elongates all notes of a melodic motif

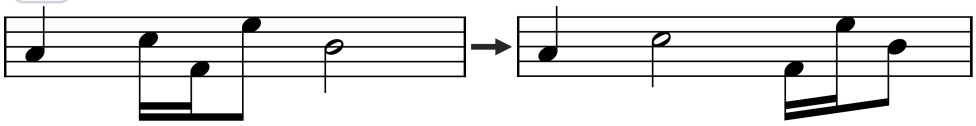


The augmentation operator uncovers all patterns where the duration of notes is increased by twice, resulting in slower rhythmic patterns.

Opposing to *Diminution*, *Augmentation* leads to a proportional elongation of melodic lines. Musicologists consider that diminution and augmentation enhance the overall structure and musical narrative by generating contrast between different sections. This operator is often applied to create accompanying melodic textures [134].

Reduction to Pitch [OP6]

P Only considers pitches of pattern

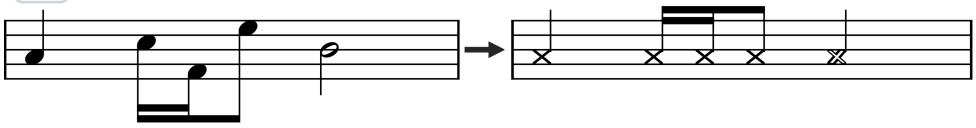


The reduction to pitch operator focuses on the pitch values of notes in the pattern, leading to various possible patterns based on pitch alone.

Only considering the pitch content allows focusing on the essential melodic structure, enabling the creation of less restricted and more digestible versions of an original theme. This operator also helps to clarify the core ideas of a more complex melody. This technique allows for developing thematic material by applying the rhythmic structure of another melody, thus combining two previously unconnected motifs.

Reduction to Rhythm [OP7]

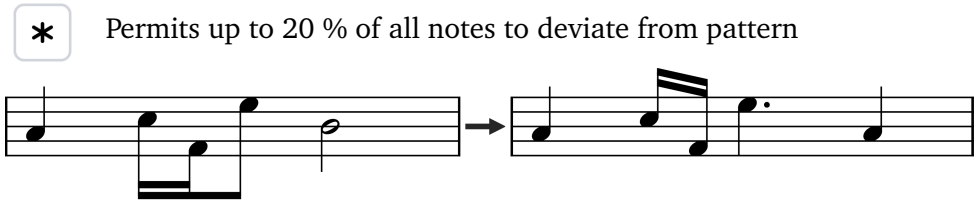
R Only considers rhythm of pattern



The reduction to rhythm operator emphasizes the duration values of notes in the pattern, leading to various possible patterns based on rhythm alone.

Similarly, taking only the rhythmic structure and content of an existing melody to connect two unrelated melodic fragments. It can be used in a larger context, such as an opera, to connect two separate acting roles musically by using their specific pitch material with the other rhythmic content opening new ways of interpretation.

Deviation `[OP8]`



Deviation allows a fixed percentage of notes (e.g., 20%) in the pattern to differ, leading to various possible patterns based on the degree of deviation.

This operator considers deviation of rhythm and pitch allowing music analysts to identify how far a composer departed from melodic themes. The *Deviation* operator is especially useful, to be more flexible in a search for similarities without the need to specify exact parameters.

6.3.2 Complex Chaining of Melodic Operators

The atomic melodic operators described in Subsection 6.3.1 provide a foundation for identifying and understanding individual melodic transformations in musical compositions. In practice, composers often apply multiple transformations simultaneously or consecutively, creating complex melodic variations. To capture such relationships and transformations, we can chain and combine the atomic operators to create complex melodic operators. Chaining operators involves applying one operator after another, where the operators' order influences the patterns identified. For instance, applying Transposition `[OP1]` followed by Mirror on X-Axis `[OP2]` would result in a different pattern than applying them in the reverse order. By chaining operators, we can uncover more subtle and sophisticated melodic variations, gaining a deeper understanding of a composer's creative process and the underlying structure of the composition. Similarly, combining operators involves applying multiple operators simultaneously. This process can reveal patterns that exhibit multiple transformations simultaneously, such as an Inversion combined with a Retrograde. Applying the same operator multiple times could sometimes serve as an identity function, which might not necessarily result in meaningful combinations. However, this flexibility allows for discovering a wider range of patterns, offering a more comprehensive analysis of melodic transformations.

There are benefits of complex operator chaining. (1) Increased flexibility: Complex operators allow analysts to discover a broader range of melodic patterns and transformations, improving understanding of a composition's structure and development. (2) Enhanced understanding of compositional techniques: Analysts can gain insights into composers' creative processes and intentions by uncovering subtle melodic variations. Likewise, there are also potential drawbacks to consider. (1) Increased complexity: As the number of chained or combined operators increases, the complexity of the analysis grows, which could make it more challenging for analysts to interpret the results. (2) Risk of overfitting: When searching for patterns using complex operators, there's a risk of identifying patterns that are not musically meaningful, especially if the analyst broadens the operators' parameter space too much.

6.3.3 Analysis Tasks and Requirements

We cooperated with domain experts and musicologists to extract requirements and tasks regarding the melodic analysis. In addition, we consider an exemplary Fugue analysis provided by the School of Music at Western Michigan University. This example is displayed in Figure 6.2, revealing how manual melodic analysis is performed through handwritten annotations, which accords with the expert input.

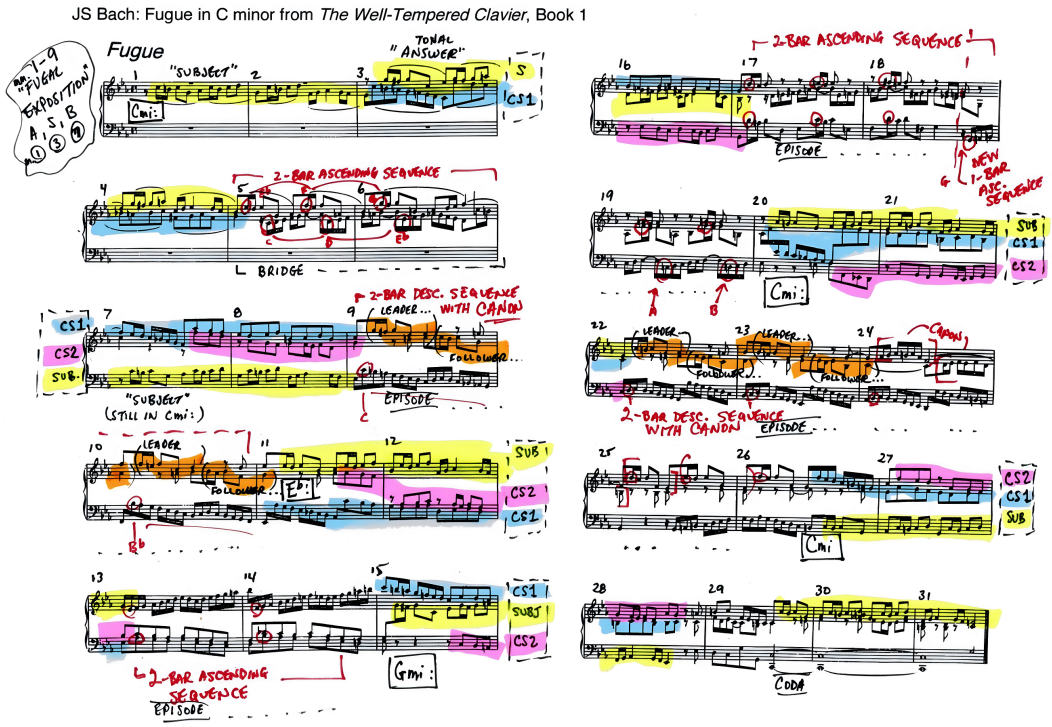


Fig. 6.2.: This manual fugue analysis example shows how a music analyst would perform a melodic pattern analysis using color and annotations. This example stems from an exam published by the School of Music from Western Michigan University [296].

Drawing from this analysis example, we identify the following four fundamental requirements for supporting music analysts in exploring melodic patterns and their transformations as intended by the composer.

R1: Identify relevant melodic sequences.
An analyst must be able to identify and select melodic motifs. It is relevant only to consider notes from the same voice. Thus, constraining patterns to remain within a single voice is crucial. This pattern identification is typically an in-situ annotation using the CMN [297], the representation analysts are most familiar with.

R2: Detect melodic pattern reoccurrences across a piece in different voices.
Identifying reoccurrences and variations of a theme typically requires analyzing multiple voices separately to determine how a composer develops a melodic topic during a composition. Thus, a visual analysis system must be capable of locating and highlighting these melodic repetitions across different voices, facilitating a comprehensive understanding of a piece's polyphonic texture.

R3: **Identify sophisticated variations and reoccurrences of a given pattern.** The handwritten analysis also reveals instances where the original melodic patterns are transformed through operators such as inversion [OP2] or transposition [OP1]. Therefore, automatically recognizing and displaying these modified patterns is relevant, enabling the analyst to track the evolution of melodic themes.

R4: **Interpret the meaning, variation type, and position of melodic patterns.** The exemplary analysis provides insights into the musical significance and context of the identified patterns and variations. MELODYVIS should integrate techniques that help understand the implications of the detected patterns, including the type of variation applied and their role in the larger composition context.

6.4 MELODYVIS – Visual Interface for Melody Analysis

MELODYVIS is a visual analytics (VA) approach to melodic pattern analysis of sheet music catering for music analysts. Below, we provide conceptual details about the interconnected components of the application and introduce the required processing steps. We also propose directions on how our approach could be further improved.

6.4.1 Data Preparation

Analyzing a music sheet in MELODYVIS requires initial pre-processing steps: First, we require any composition to be available using the MusicXML format, a widely used, general, and flexible format to encode symbolic music information [112]. We leverage the music21 Python library [67] to load the digital sheet music files continuing with the data processing. To identify coherent voices, we leverage the envelope extraction technique also used by Gray and Bunescu [117] that allows retrieving the separate voice from a composition. MELODYVIS could be easily extended to support sheet music import in addition to the predefined set of compositions.

6.4.2 Interaction Components

The visual interface for analyzing melodic patterns based on digital sheet music comprises five interconnected components reflecting the interactive analysis workflow as indicated in Figure 6.1: (1) Sheet View, (2) Melodic Operator Selection, (3) Voice Separation Overview, (4) Melody Transformation Graph, and (5) Melodic Pattern Configuration. The five separate components are updated based on the interactions that analysts typically use in the same order as we describe them subsequently.

Sheet Music View

The sheet view (SV) facilitates the selection of initial melodic patterns based on the separated voices [R1]. This intuitive interface enables analysts to configure note coloring based on voices, allowing visual differentiation as needed. The SV features in-situ placement of detected patterns by MELODYVIS, utilizing different colors to distinguish between variations [R3]. Furthermore, the component automatically adjusts measure splitting according to the analyst's screen size, optimizing readability and user experience. The SV is the starting point of the melodic analysis workflow.

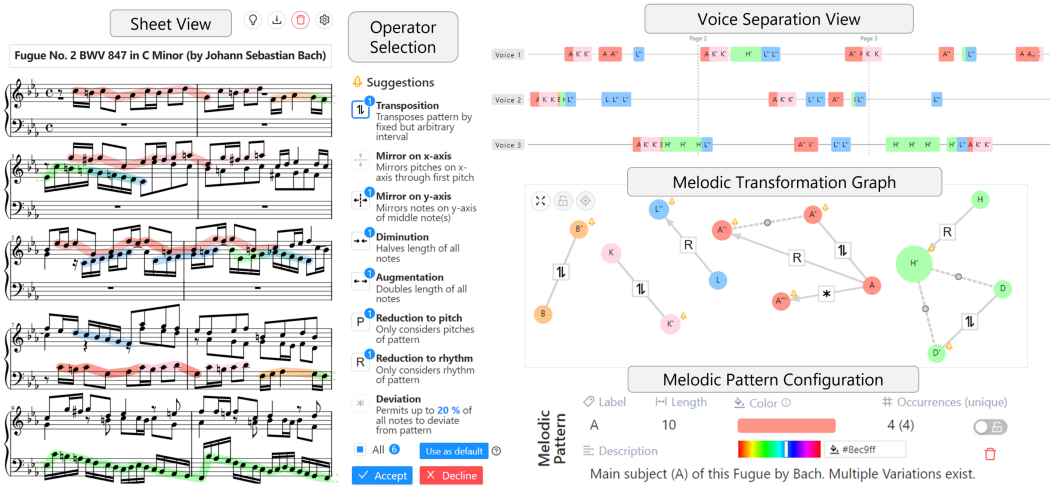


Fig. 6.3.: MELODYVIS consists of five core interaction components interconnected through linking and brushing [155]: (1) Sheet Music View, (2) Melodic Operator Selection, (3) Voice Separation View, (4) Melodic Transformation Graph, and (5) Melodic Pattern Configuration. In combination, they facilitate a comprehensive and interactive analysis workflow to explore melodic patterns in digital sheet music.

Combining these features allows the SV to provide a suitable starting point for melody analysis while highlighting identified patterns and their variations. As the first step in the analysis, the SV highlights the subsequent notes from the same voice when selecting an initial note, assisting users in completing their pattern selection. To achieve this, it is necessary to separate the correct voicing in the first place. If the number of voices varies throughout compositions, voices may not be accurately separated, leading to wrong voice assignments in the suggested selection.

Melodic Operator Selection

As shown in Figure 6.1, the introduced atomic melodic operators are applied to the initial melodic patterns selected within the sheet view. The application of these operators generates a list of pattern variations, which allows users to further explore the melodic patterns by chaining multiple operators together. This Melodic Operator Selection (MOS) enables the creation of more advanced, complex transformation chains, ultimately facilitating an efficient expansion of the search space during melodic pattern investigation and exploration tasks.

Voice Separation View

The *Voice Separation View* (VSV) provides a comprehensive overview of detected patterns throughout the sheet. Page markers as vertical lines are a navigational aid that show at which positions the page breaks occur in the SV. The mouse position in the VSV shows the corresponding page in the SV. This view effectively displays the distribution of patterns across all voices within a given sheet, using consistent coloring as in the sheet view. Users can access additional detailed information about the patterns when hovering over elements. If required, the VSV enables melody analysts to select patterns for further annotation editing. With the appropriate coloring settings applied, analysts can easily identify the first occurrence of a specific pattern within a particular voice, streamlining the analysis.

Melodic Transformation Graph
The Melodic Transformation Graph (MTG) is a supportive and more abstract visualization component designed to display the transitions of a melodic pattern selection to a set of other patterns resulting from applying a specific operator (see Figure 6.3). Nodes within the graph represent either single patterns when selected or entire pattern groups based on an applied operator, with the direction of transformation indicated by representative arrows. Operator icons on the edges between nodes facilitate identifying which operator was used to obtain the resulting pattern set.

Connected graph components reveal related patterns, while dashed lines depict similar patterns in different result sets when applying distinct operators, thus highlighting similarities between operators. This comprehensive visualization enables analysts to gain deeper insights into the complex relationships between melodic patterns and the effects of applied transformation operators. Analysts can explore the intricate connections between melodic patterns and their variations within a given musical composition through the MTG. This component also helps differentiate between user-selected melodic patterns and automatic detection corresponding to a prior user selection.

Melodic Pattern Configuration
The Melodic Pattern Configuration (MPC) is a versatile component providing additional details and customization options. Users can access melodic patterns within any other components (SV, VSV, MTG). This dedicated component shows summarized pattern information, such as unique or overlapping occurrences and pattern length (i.e., number of notes). Analysts can tailor short labels to suit their needs and modify the coloring of a single pattern or a pattern group using a simple color map. While MELODYVIS primarily uses distinct pastel colors to emphasize different patterns, users can still adjust colors to represent similarities based on their preferences. Alternatively, the coloring could be adapted based on the similarity space of two given different patterns, enabling an improved semantic understanding of the connections. The component also allows for adding a detailed description (up to 280 characters) for knowledge externalization [R4]. Also, a deletion button is provided, enabling users to remove single patterns as needed. This comprehensive configuration component facilitates a more customized and efficient analysis experience for music analysts. The manually changed annotations are linked via tooltips to the other components such that they can easily be reused during the analysis and for comparing multiple melodic patterns.

6.4.3 Visual Interaction Pipeline

MELODYVIS follows a visual interaction pipeline to facilitate the interactive analysis of melodic patterns in sheet music. As described in Figure 6.1, analysts can begin by selecting an initial melodic pattern within the SV, which is then utilized to apply melodic operators that generate further pattern highlights. Following this, the VSV provides a comprehensive overview and distribution of the different voices in the composition. To support the analysis of melodic variations, the MTG visually displays the relationships between the initial pattern and its transformed counterparts. As the analyst progresses, they can seamlessly loop through this analysis process to continue discovering and exploring additional melodic patterns within the sheet music.

6.5 Use Cases

Hereinafter, we will present three distinct use cases to demonstrate the interactive analysis workflow with MelodyVis and highlight its capabilities in analyzing digital sheet music for melodic patterns. With these use cases, we aim to showcase our prototype's practical applications and provide a solid understanding of how users can leverage MelodyVis to gain insights into musical compositions, emphasizing its potential in educational and research contexts.

6.5.1 The Four Seasons: Spring – by Antonio Vivaldi

The first use case demonstrates how a user can explore and identify melodic patterns in Vivaldi's "The Four Seasons - Spring (La Primavera)" 1st movement [342]. The composition features five distinct voices: three violin parts, one viola, and one violoncello part as depicted in the VSV at the top in Figure 6.4.

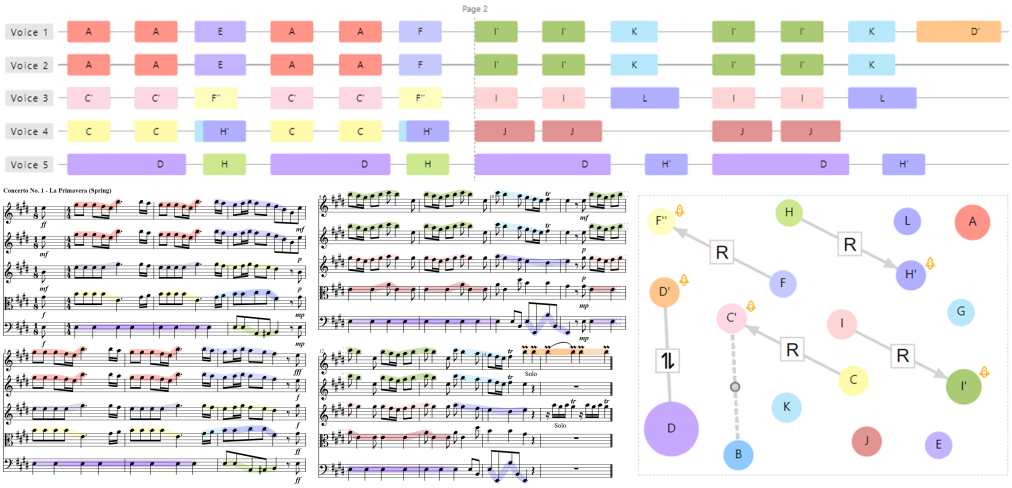


Fig. 6.4.: Visual analysis of Vivaldi's 1st movement of "The Four Seasons - Spring" [342]. (Top) VSV with color-coded highlights indicating corresponding melodic patterns across the voices. (Bottom Left) The SV displays the selected patterns with overlay annotations. (Bottom Right) The MTG illustrates the relationships between the identified patterns and their transformations within the composition, showing dashed edges with pattern sets with similar patterns.

First, the user selects the second note from the first voice (Violin 1) directly on the sheet music (SV) as the starting note of the first pattern. A second note from the same voice must be selected to complete the first pattern [R1]. The user can use this interactive exploration process to identify similar occurrences of this first pattern (light red) in the second voice [R2]. Next, the user examines the melodic variations in the 3rd and 4th voices (Violin 2 and Viola), which play along with a similar pitch [OP6]. By doing so, the user can observe the aligned similarities between these voices. In the fifth voice (Violoncello), the user can see how it lays a solid foundation by playing longer notes with lower pitches and repeating the melodic texture. The Voice Separation View proves to be particularly helpful in identifying the similarities between the corresponding voices and the regular structure of the composition. The

MTG aids the analyst in visualizing the connections [R3] between exact patterns and instances where the main themes are reduced to rhythm only [OP7]. Finally, in the MPC, the analyst adds a short description of the main theme that Vivaldi creates a vivid representation of the Spring season's essence [R4].

6.5.2 BWV 846: Prelude in C-Major – by J. S. Bach

In this second use case, we examine Bach's "Prelude in C-Major (BWV 846)" [15]. Its upward-directed melodic line and distinctive broken chord patterns characterize this piece. The analysis begins by selecting an upward arpeggio pattern in the first voice. Using the VSV and SV, the analyst can view all transposed matches of the same pattern in the same voice (see Figure 6.5).

The VSV emphasizes the simple repetitions of the same pattern in the second and third voices throughout the piece, while the most melodic changes are present in the first voice. By showcasing how MELODYVIS can effectively analyze and illustrate transposition and upward broken chord patterns (arpeggios) in a well-known composition like BWV 846, this use case demonstrates the tool's potential for enhancing the understanding of repetitive melodic structures and relationships.

The second and third voices comprise only two transposed patterns. While musicological experts can identify uniformity by examining sheet music, precise recurrences and transpositions in arpeggios are easier to discern through linking and brushing on the VSV. The contrast between the simplicity of the second and third voices and the more varied first voice showcases MELODYVIS's ability to reveal varying complexity within a single composition. Our prototype is versatile enough to handle diverse melodic structures through integrated atomic operators independent of underlying complexity. This contrast emphasizes the importance of visual techniques in providing traceable, detailed analysis of melodic patterns and relationships, ultimately enhancing music understanding.

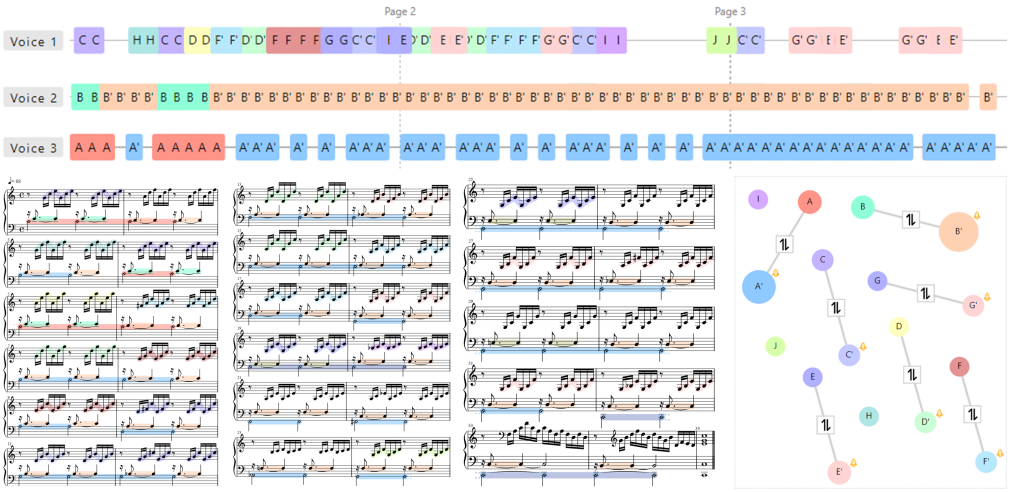


Fig. 6.5.: Analysis of J.S. Bach's "Prelude in C Major (BWV 846)", focusing on the arpeggio patterns in the first voice, where the most interesting melodic development occurs, while illustrating the simple repetitions occurring in the second and third voice.

6.5.3 Tetris Theme – by Nikolai Alexejewitsch Nekrassow

The third use case explores a single-voice version of the Tetris theme by Nikolai Alexejewitsch Nekrassow comprising an iconic melody recognized worldwide.

The analysis begins with identifying the starting theme [R1], which occurs three times (highlighted in light red in Figure 6.6). The MTG displays the theme's relationships, allowing users to visualize the abstract structure of the melody patterns and their recurring motifs [R2]. Some edges between the graph nodes are drawn using dashed lines in specific scenarios. These dashed lines identify bridges between pattern sets resulting from using the operators [R3]. Tooltips on demand indicate how many similar patterns occur on both sides of the edge. The MTG also reveals independent graph components emphasizing connected melodic pattern groups, proving all hitherto identified patterns. In contrast to the main theme, in bars 9-11, the downward half notes using minor and major thirds provide a contrasting melodic sequence, adding further, even though very simple, variety.

MELODYVIS helps gain insights into this well-known theme [R4], illustrating its applicability for analyzing various musical compositions, from classical works to more straightforward, single-voice melodies.
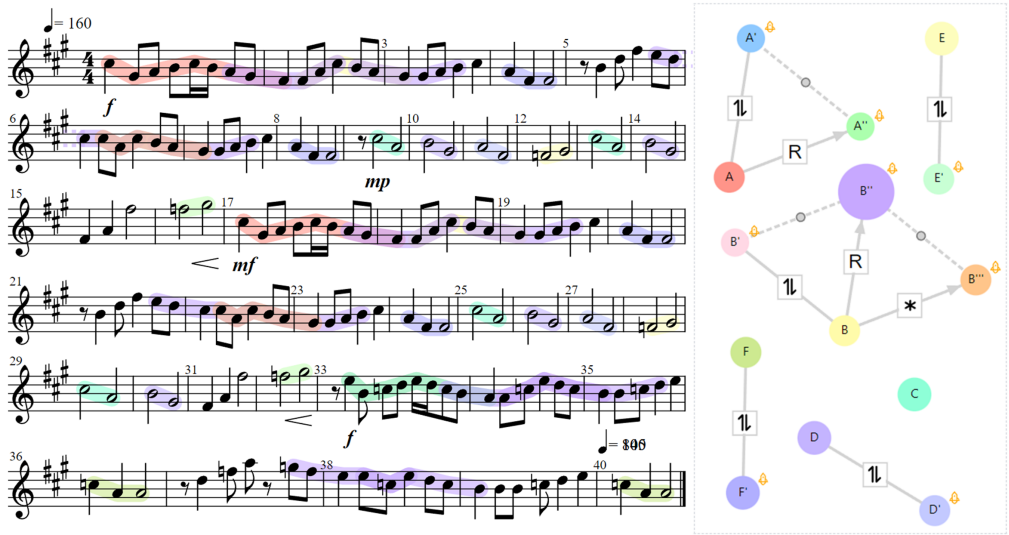


Fig. 6.6.: Tetris theme analysis: On the left, the highlighted Sheet View displays the recurring motifs and contrasting melodic sequences. The Melodic Transformation Graph on the right reveals the simple relationships yet memorable patterns used in this iconic melody.

6.6 Online User Study of MELODYVIS

We conducted an online user study to assess our MELODYVIS prototype in facilitating melody analysis tasks for participants of diverse expertise levels.

Participants – We recruited 25 participants of different expertise levels through various channels, including e-mail and online platforms such as Discord, LinkedIn, Reddit, and Twitter. Our e-mail recruitment primarily targeted mailing lists for profes-

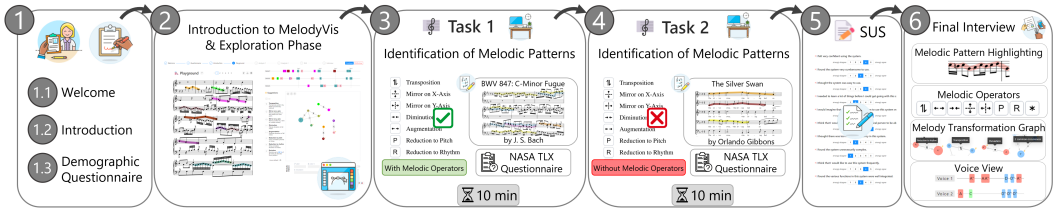


Fig. 6.7.: User Study Procedure – (1) Welcome, introduction video, demographic questionnaire, and explanation of melodic operators; (2) Familiarization with the prototype; (3) First task, followed by NASA-TLX workload evaluation; (4) Second task with randomized music piece and operator availability; (5) System Usability Scale (SUS); (6) Final survey with an interaction questionnaire about the components.

sional musicians and musicologists, music associations, members of philharmonic orchestras, and selected individuals with a keen interest in music. On Discord and Reddit, we posted calls for study participation in communities that bring together hobby musicians and musicologists. This multi-channel recruitment strategy allowed us to gather a heterogeneous group encompassing different backgrounds, experiences, and levels of expertise in music theory, melody analysis, and data visualization.

Study Tasks – The participants were tasked with identifying melodic patterns they deemed relevant in a given music sheet by highlighting a melodic pattern (i.e., selecting the first and the last note). To measure the impact of the melodic operators, the study involved **two rounds in randomized order** to minimize learning effects: analyze one of the music sheets (`SM1` and `SM2`) once *with* and once *without* the melodic operators available leading to a total of four different trial setups. For the former, we also instructed the participants to use the available melodic operators to identify melodic patterns. They were advised to focus on identifying longer melodic patterns in favor of shorter ones, particularly those that would stand out to a listener of this music sheet. Participants were expected to give meaningful labels and descriptions of their identified melodic patterns. There was no particular solution to this task; hence, the participant could complete it as soon as they considered it completed. We especially considered the amount and type of patterns Participants could expect to identify between 10 and 20 melodic patterns and had 10 minutes to complete each round of tasks.

Study Procedure – When our participants opened our web-based study prototype (`visual-musicology.com/melodyvis-study`), they were provided with a short textual introduction to the study procedure and prompted to agree to the terms of the study. Our study design to assess MelodyVis is described in detail in Figure 6.7. We used Bach's "Prelude in A minor, BWV 889" in the exploratory phase to let the participants get acquainted with the components [16].

6.6.1 Study Results

Qualitative Results

In a final interaction questionnaire survey, we gathered positive and negative feedback from our participants. Initially, we excluded all participants who did not provide patterns or complete the study.

Several participants emphasized our approach's benefits, including its ability to classify different operators and **assist** in analyzing and improving sheet music ($\mathbf{P_1}$). The system's **practicality** and intuitiveness were praised, with users finding the melody highlighting and timeline features particularly helpful ($\mathbf{P_2}$, $\mathbf{P_{19}}$, $\mathbf{P_{20}}$). The **speed and efficiency** of MELODYVIS in identifying patterns were appreciated by participants ($\mathbf{P_4}$, $\mathbf{P_{18}}$). Furthermore, the visually **seamless interface** contributed to a positive user experience ($\mathbf{P_5}$, $\mathbf{P_{16}}$). Other valued features included the ease of **marking melody patterns** ($\mathbf{P_{13}}$, $\mathbf{P_{15}}$), the **automatic recognition** of recurring motifs ($\mathbf{P_{12}}$), and the **suggestions** of melodic operators ($\mathbf{P_6}$, $\mathbf{P_7}$, $\mathbf{P_8}$). Overall, the graphical representation on sheet music was well received ($\mathbf{P_{14}}$), and the highlighting of melody components ($\mathbf{P_{20}}$).

A few participants also raised concerns regarding their experience with MELODYVIS. Some encountered **difficulties selecting notes**, as the highlight would jump to other staves ($\mathbf{P_5}$). The lack of an **auditorial reference** made it challenging for some users to analyze the melodic patterns ($\mathbf{P_{19}}$), especially without the operators in place. A few users reported that the application failed to recognize specific melodic patterns due to **chromatic variations** ($\mathbf{P_{16}}$).

Task-Solving Strategies

From the final questionnaire, we identified multiple strategies the participants employed to identify patterns during their tasks. They relied on (1) *familiarity and repetition* by recognizing recurring elements and focusing on frequently appearing motifs. Users also paid attention to (2) *musical characteristics*, such as large leaps, unusual rhythms, fugal subjects, and imitated fragments to guide their analysis. They considered the (3) *context and structure* of the music, looking for patterns around the beginnings of sections, voice entrances, lyrics, or changes in the composition's character. Lastly, participants used their (4) *intuition and prior knowledge*, imagining the melody in their heads or trying sight-singing to better understand the music and identify related patterns. By understanding these strategies, we can refine MELODYVIS to better support users' melodic analysis efforts.

Results of Selected Patterns

When comparing the selected patterns of both compositions used in the study, we identified that the users could detect at least twice as many patterns when using the melodic operators. Figure 6.8 depicts the music sheet for `SM1` indicating the

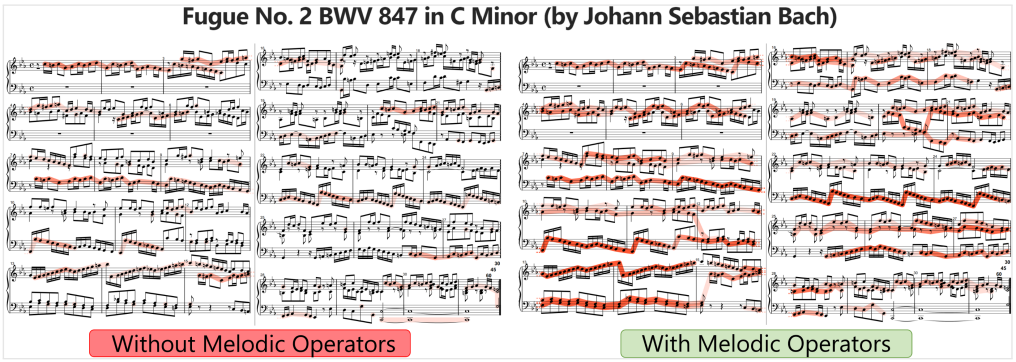


Fig. 6.8.: Resulting patterns on `SM1` for both conditions. The two music sheets reveal that with the operators (right), at least twice as many patterns could be identified compared to those without (left).

distribution of all found patterns using red highlighting. While some participants found patterns without the operators, others did not. In contrast, we found that most participants could identify the same patterns with the operators, showing the tendency of their supportive characteristics.
The four pattern identification heatmaps displayed in Figure 6.9 for both conditions and music sheets separately confirm that for both pieces, the amount of patterns with the operator is a magnitude higher than without the melodic operators available. Especially in the voices 3-5 of the “Silver Swan” the detected patterns are clearer.

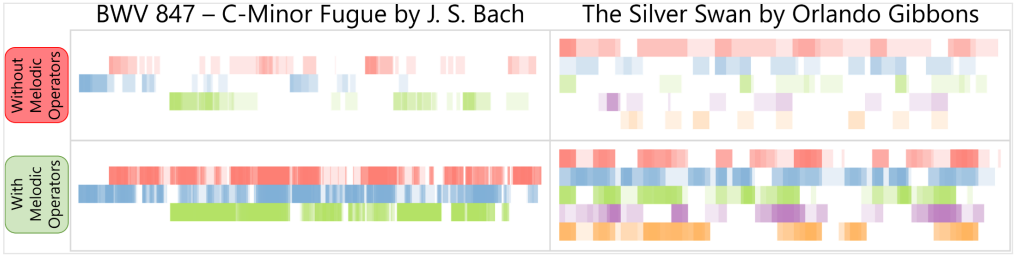


Fig. 6.9.: Direct comparison of heat maps based on selected melodic patterns from all participants without the operators (top, red) and with the operators (bottom, green) for `SM1` (left) and `SM2` (right).

NASA TLX & SUS Questionnaire
The primary purpose of the TLX questionnaire between the two conditions was to estimate whether the addition of the melodic operators would deteriorate the overall effort of the analyst since extended automation can lead to overstraining. The average task load index for both conditions is between 60 and 75 without any salient differences, indicating that the operators’ availability leads to at least double as many patterns as without. We present further details in the supplementary material of the respective publication [223].

6.7 Discussion

Subsequently, we discuss the implications of our system design, our study findings, the current limitations, and opportunities for future research.

6.7.1 Take-Home Messages

Atomic Operators and Mixed-Initiative Analysis
The concept of atomic operators and their combination into complex operators can be highly valuable for information visualization research. Our semi-automatic solution demonstrates how integrating computational techniques with human control and guidance can result in more scalable and effective pattern detection. This mixed-initiative approach retains human agency in the analysis process and can be adapted to other domains, such as text or bio-genome data analysis.

Study Insights and Scalability
Our user study showed that our designed melodic operators significantly increased the number of patterns the participants could identify with low cognitive effort. Simultaneously, the familiarity of the sheet music view offered a suitable starting and reference point during the subsequent analysis. This demonstrates the potential for scaling up complex problem-solving in information visualization while retaining human involvement and expertise in the analysis process.

Methodology Transfer to Other Domains
Our techniques can be transferred to other data analysis contexts, such as text or temporal data. Combining operators, visualizations, and interactions can provide a solid foundation for tackling intricate pattern detection tasks across domains, fostering interdisciplinary collaboration and innovation in InfoVis research.

Collaborative Research
Embracing interdisciplinary research, such as musicology and visualization, can lead to innovative techniques that support operator chaining and the simultaneous analysis of multiple musical characteristics. This allows for a deeper understanding of complex musical relationships, ultimately fostering knowledge exchange and leading to novel insights within *Visual Musicology*.

6.7.2 Limitations and Potential Extensions

Our results are limited by the number of participants, the shortness of their interactions, and the broad spectrum of participants. However, the current insights show that our system matches the initial requirements while indicating the following needs for future extensions. Adding audio playback could enhance the user experience and understanding of the analyzed music. The scalability of MelodyVis could be improved by increasing the performance for applying operators. The currently used colors, which rely on transparency and blended colors, could make interpreting the relationships between applied operators difficult. Currently, our prototype supports only a preselected number of pieces, and an option to import custom sheet music could improve its applicability. In the MTG, individual patterns resulting from the operators are currently combined into single nodes, which could be extended. Adding parameterization for the operators may provide greater flexibility in supporting additional use cases. Yet, the length of a proposed pattern is fixed, with no operator available to search for meaningful sub-sequence patterns. At this point, user interaction is mandatory in the current implementation, as automatic pattern suggestions are unavailable yet. Future versions of MelodyVis could learn from user interactions and suggest patterns for improved analysis based on previous selections.

6.7.3 Open Research Opportunities

The interdisciplinary research between musicology and visualization and data analysis at the intersection of visual musicology [229] provides compelling opportunities to uncover new insights into musical compositions, enabling scholars to explore and analyze complex melodic patterns through the fusion of traditional musicological methods and cutting-edge visual and data-driven techniques.

Melody–Lyrics Relationships and Genre Comparison – Exploring the interconnections between lyrics and melodic patterns can reveal deeper meanings and artistic intentions behind compositions. Additionally, comparing various genres to identify unique and common melodic patterns can provide insights into how composers use complex melodic combinations in different forms of music, thus shedding light on each genre's stylistic characteristics and conventions.

Temporal Evolution of Melodic Motifs – Using abstract visualization to demonstrate the temporal evolution of melodic motifs can help researchers understand how specific melodic patterns have been employed by different composers across multiple epochs and cultures. To achieve this, it involves developing visualization techniques that track the evolution and adaptation of melodic motifs over time.

Advanced Visualization Techniques – Developing novel visualization techniques for abstract melodic analysis could help identify differences between motifs across composition collections. This would enable researchers to easily navigate between abstract representations and specific melodic patterns [145] while facilitating a better understanding of the structure and development of various melodies.

Leveraging Intelligent Pattern Recognition – Utilizing state-of-the-art machine learning and AI-based techniques bears significant potential for automated pattern recognition and classification in melodic analysis. Leveraging advanced algorithms and methods could improve the accuracy and efficiency of identifying melodic patterns in single pieces and whole composition collections.

6.8 Summary

We presented MelodyVis, an interactive application for semi-automatic melodic pattern investigation in digital sheet music. Integrating musicology experts, we developed an interface with connected views like the Melody Operator Graph and Voicing Timeline for interactive pattern exploration. MelodyVis uses eight atomic operators for effective melody analysis. Our user study with 25 participants showed the tool's usefulness and effectiveness, doubling pattern identification compared to manual analysis. Users experienced improved pattern identification and interpretation. MelodyVis overcomes fully-automated approach limitations by involving analysts for transparent and accurate melodic pattern investigation, advancing interdisciplinary collaboration between information visualization and musicology.

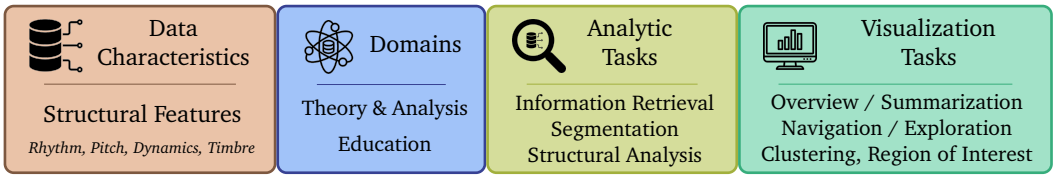
Part III

Multi-level Sheet Music Analysis Techniques

7

Bottom-Up | MusicVis Augmenting Sheet Music through Visual Analytics
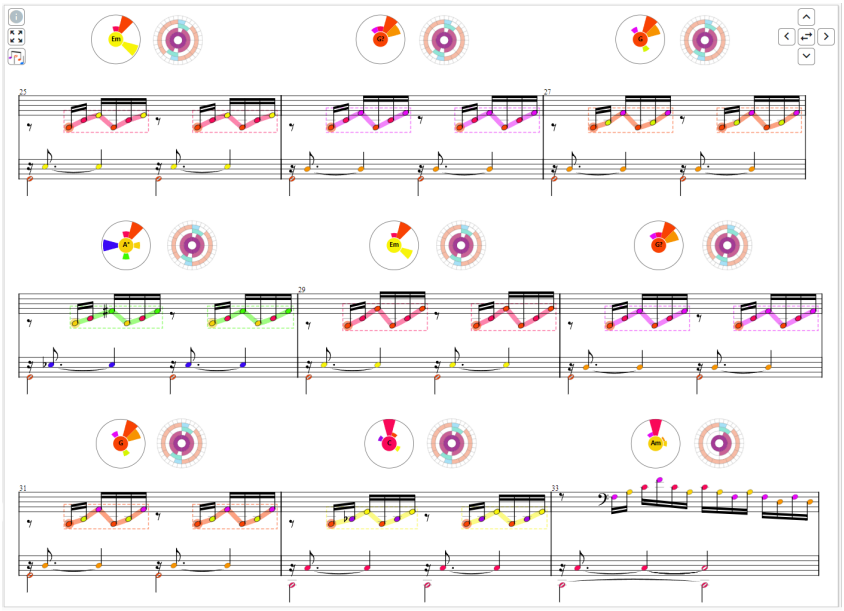
> *Visual Analytics integrates the analytic capabilities of the computer and the abilities of the human analyst.*
>
> **— Jim Thomas**
> *American Computer Scientist*



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.

Chapter 7 is based on the following publication. Please refer to Section 1.4 for contribution clarifications.

[224] **Matthias Miller**, Daniel Fürst, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. **"Augmenting Digital Sheet Music through Visual Analytics"**. *Computer Graphics Forum, Vol. 41, No. 1, 2022.*



(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. Split management facilitates the identification of a composition's structure based on semantic split criteria.

Fig. 7.1.: *MusicVis* supports composition analysis by augmenting sheet music with feature glyphs (Figure 7.1(a)). These can be interactively explored through (Figure 7.1(b)) visual *harmony, rhythm*, and *melody* queries. Applying semantic split criteria facilitates structure analysis.

7.1 Introduction

Music analysts and musicologists extract meaningful relationships from musical compositions regarding, for instance, style, epoch, or composers [163]. Besides meta-information, other features such as rhythm, dynamics, and harmony are crucial when comparing musical pieces [229]. Music analysts typically perform manual sheet music analysis 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 [190], annotation, and comparative analysis. Without the help of additional tools, these tasks are time-consuming and tedious due to the visual complexity of CMN [225]. 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 [331, 346]. Integrating the analyst into the analysis process to support knowledge generation and verification [282] is necessary since exclusively computational approaches [337] cannot model human intuition about semantic meaning. In the visualization domain, graph-based [127], matrix-based [152], and projection-based [199] techniques have been proposed to support the visual analysis at different abstraction levels. 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 chapter introduces a visual analytics technique for bridging CMN with glyph visualizations minimally intrusively. Music analysts can benefit from the CMN and abstract visualizations by addressing this gap. We demonstrate the technique's capabilities by presenting an instantiation called *MusicVis*. MusicVis uses existing visual designs such as the harmonic and rhythmic fingerprint from our previous work [222, 103] 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. These designs are interchangeable since the technique is independent of the specific visual encoding. Thus, MusicVis enables music analysts to perform close reading of the presented data, analogous to slow analytics [35]. Our technique further enables distant reading, including corpus analysis, maintaining an overview of the abstract level following Shneiderman's "Visual Information Seeking Mantra" [306].

Contributions
This chapter contains these 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 tightly connected through staged animations [124].
(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 7.1).
(4) An evaluation using a pair analytics study with 16 participants of different user groups to demonstrate its broad applicability.

7.2 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 chapter is *Theory & Analysis* and *Education* [229]. This approach focuses on the analytical tasks *Information Retrieval*, *Segmentation*, *Structure Analysis*, and *Contextual Analysis* of the *Visual Musicology Graph* [229]. We consider musical features such as structure, harmony, rhythm, and melody, which Orlovaitt describes as "middle-term" features [252]. Donnini et al. describe substantial differences and emphasize the separate consideration of symbolic music and its interpreted version [79]. Similarly, we argue that sheet music deserves more attention, especially from the perspective of visual analytics [282]. Fully automated approaches produce high-quality results but lack the opportunity to step into the process, preventing analysts from influencing the output [71].

7.2.1 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* [188]. Existing abstract visualization approaches, such as *Isochords* from Bergstrom et al. [21] and *Tonnetz* by Tymoczko et al. [331] reveal the geometry of chords. Sapp proposed a visualization to support hierarchy-based tonality analysis [290]. Mardirossian and Chew propose a music visualization tool to investigate tonal progressions of compositions at an abstract level while dropping CMN [208]. Malandrino et al. introduce a technique visualizing harmony to analyze the structure of a composition [206], which is limited to four-voice chorales and is only applicable at the sheet level. Exploiting the circle of fifths [125], we use the abstract harmonic fingerprint design that we introduced in our previous work [222] 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 [263]. Foote et al. present an analysis approach for rhythm structure grounding on self-similarity [95]. Regarding Ren et al., rhythm plays an essential role in automatic pattern extraction [272]. 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 [196]. 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 subsection 7.5.1). Similarly, Benadon exploits a circular plot to visualize and analyze rhythmic information [19]. For a more detailed description of the complexity of the musical feature *rhythm*, we refer to Toussaint, who provides different concepts for its visualization [327].

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 [218]. Urbano proposes the library "*MelodyShape*" that enables similarity search for melodies based on the shape [335] 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 [183]. 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 [143]. 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 [181].

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 [75, 248]. 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 [49]. 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 [21]. ImproViz by Syndal et al. reveals structure by the melodic landscape but omits CMN as well [312]. 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 [268]. 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 interactively segment a piece manually or use structural elements to explore structural ambiguities.

7.2.2 Methodology Transfer

Digital humanities concepts inspire this work. We apply a methodology transfer from other fields, such as text analytics, to the music domain [229]. For example, Poemage is a technique that reveals the beauty of poems by extracting the sonic topology based on textual features, enabling close reading [210]. El-Assady et al. combine close and distant reading views for text data to investigate forum discussions [10]. Keim and Oelke introduce a visualization to display textual features through literature fingerprints, enabling distant reading on text data [157]. Similarly, we use musical fingerprints for the independent musical features of pitch and timing. We combine close and distant reading by juxtaposing sheet music with glyph visualizations, creating interactive visualizations to analyze music. Jänicke et al. highlight the importance of combining close and distant reading [145]. 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 [124] enabling slow analytics [35] for musicologists.

7.3 Problem and Data Characteristics

The notational complexity of the musical dimensions [225] 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. [347] and Khulusi et al. [163] who provide a profound basis for music analysis tasks and data characteristics. While we did not consult 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.

7.3.1 Musical Features

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. Vast combinations exist, particularly when increasing the number of simultaneous notes. If rhythm is added, we can describe chord progressions and more complex harmonic patterns that require theory knowledge to understand [189]. Chord progressions influence the perceived *tonality* or *modulation*.

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

Melody represents a specific note 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 that differs from multiple simultaneously played notes. Based on

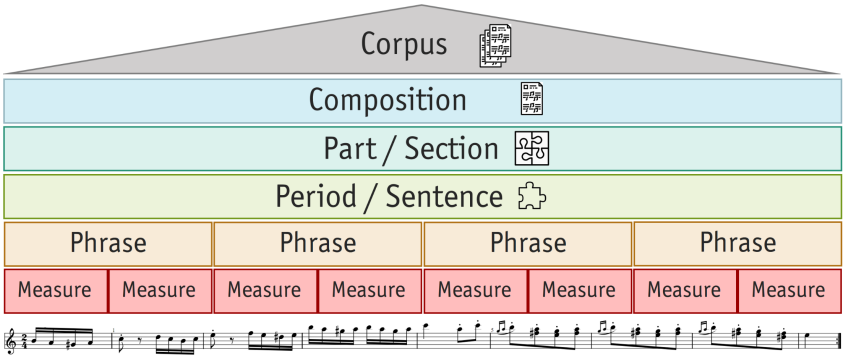


Fig. 7.2.: The hierarchical segmentation of music compositions ranges from single notes at the measure level to the corpus level with multiple layers in between comprising phrases, periods or sentences, parts or sections, and compositions.

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 the leitmotif [38] are of high interest for music analysts, which is the leading motive affecting a musical piece's central theme.

Abstraction and Aggregation
Subsection 7.3.2 introduces different structural levels. We institute statistical abstractions to analyze features at any level. While losing details, this abstraction facilitates analysis tasks such as comparison or interpretation.

7.3.2 Hierarchical Music Structure

Music contains hierarchical structures at different levels [190]. Weihs et al. introduce *Structure and Form Analysis* [**T3**] as a relevant high-level task [347]. We can identify such structures by applying segmentation on both the corpus level or the measure level. Many musical elements affect structural ambiguity. The music structure comprises a *horizontal* and a *vertical* dimension. For instance, the simultaneous performance of multiple instruments, e.g., orchestral scores, or instruments simultaneously playing multiple pitches (e.g., the piano) represent the *vertical* dimension. The vertical structure separates melodic from harmonic progressions. Melodies comprise only single notes [201]. The *horizontal* structure represents the temporal aspect. Figure 7.2 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 [347].

7.3.3 Single and Multi 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, dynamic, and further instrument-specific 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 a more complex structure at higher levels. Time-span reduction builds up a hierarchy of stability, providing a tonal center [190].

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 [81] 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.

Multi Composition

Similar to text corpora [56], sheet music collections can be grouped by characteristics such as composer, epoch, style [163]. Comparing collections requires high abstraction levels, as viewing every musical detail as low-level entities is impossible. The musical structure depends on the horizontal (temporal) dimension, which can be inferred from the compound features of rhythm or harmony. Structuring music based on these musical entities requires analysts to extract relevant patterns their salience can identify in a piece. Eventually, the reader must develop a complex mental model of the inner workings of musical relationships.

7.3.4 Analysis Tasks

Numerous analysis tasks interest music readers and analysts based on the vertical and horizontal structure in combination with the musical features [347].

The following presents an extensive task list of varying importance for different users. Based on the horizontal structure levels introduced in the previous section, we provide a concise task overview that can be performed at each level. If we look at the features, we can extract various music analysis tasks.

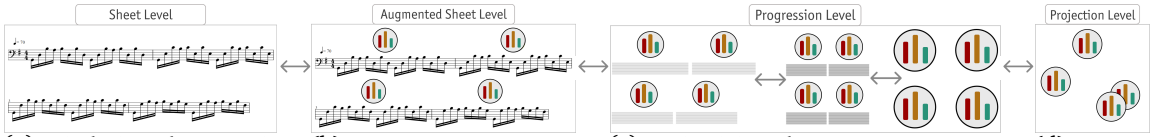
Measure-Level – At the measure level, a common harmonic task is chord detection based on simultaneous pitches. Motif extraction is a melodic task enabling the identification of repetitive note sequences. Understanding the fundamental rhythm complexity is substantial. At the measure level, only low-level 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 a composition's primary melodic content and main themes. Understanding the differences and similarities of melodic movements between different voices is another essential task that can be performed at the phrase level. E.g., this task facilitates understanding melody sequences in a fugue that is shared and repeated at different transposed stages. Besides melody, understanding chord progressions and their recurrences is required to elicit modulations from the harmonic data characteristics.

Composition-Level – A fundamental task at the piece level is the analysis of instrumentation and its usage throughout a composition. This is directly related to the interpretation of voice development. The extraction of tonalities, including identifying modulation, tonal dominance, and unexpected occurrences of tones entirely foreign to the current tonal center, is crucial. Understanding the different rhythmic patterns combined 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 conclude a composition regarding the date of creation or composer's originality. 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 and perform comparative analysis. These only reveal different information leading to other interpretations. Of course, feature summarizations can show the underlying statistics scales at any analysis level and inform structure analysis. Analysts can be interested in the frequency or scarcity of a pattern which is independent of the analysis level.



(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 [124] to omit CMN.

(d) MDS projection [63].

Fig. 7.3.: From Close to Distant Reading: Four abstraction levels connected through staged animations [124] provide different perspectives. The sheet level (a) is the familiar CMN context which is a basic requirement by typical analysts. 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 by applying cosine similarity-based MDS [63].

7.4 Technique: Augment 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 MusicVis, a visual analytics technique to add glyph visualizations as minimally intrusive augmentations to CMN. Hence, analysts can explore sheet music at multiple abstraction levels [**T2**] from the sheet to the projection level by maintaining the intermediate connection through seamless animated transitions between each designated layer [166].

7.4.1 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 [145], including music analysts (see Figure 7.3(a)). CMN does not provide extra space, requiring suitable visual modifications before augmenting it with visualizations. The needed additional space depends on the visual characteristics of a given visual design. We consider glyph visualizations suitable due to their size and the flexible application of visual attributes (color, size,

etc.) [30]. Besides, glyphs can be placed independently, which allows to unveiling of relationships between single entities [**T1**]. Glyph-based designs can communicate multivariate data while preserving the connection to the spatial context [30]. 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 visualizations. The technique positions the glyph by horizontally centering it closely to the top of the respective measure as displayed in Figure 7.3(b). The proximity to the CMN maintains 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 other sheet music perspectives. 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 voices. For instance, conductors need to understand several instruments' simultaneous progression and interplay. Encoding parallel voices in a single glyph would not be optimal in this case. Showing separate (parallel) glyphs would help to differentiate 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 must be reflected by a corresponding glyph design. If possible, using discrete bins facilitates the comparability of the 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, libraries such as jSymbolic [213] can extract e.g. melodic, rhythmic, and pitch features from symbolic sheet music representations 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.
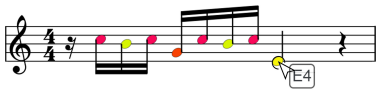
Aggregation of Adjacent Glyphs – Enabling analysts to aggregate musical information is integral to the feature analysis. Aggregating adjacent glyphs should create parent glyphs that contain the summarized content of their children. This aggregation allows for transitioning from a low (measure level) 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. Analysts can compare larger sections at aggregated stages to identify salient differences [**T4**].

Interactive Exploration and Querying – The query design should fit 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.

7.4.2 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 the analyst can trace the 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 subsection 7.3.4. 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.

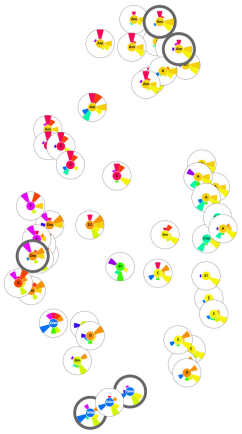
Sheet Layer – Analysts can read music as usual based on CMN at the *sheet-level* (see Figure 7.3(a)). Hitherto, no abstract information is displayed, allowing one to focus on the unaltered CMN. Optional note coloring can enhance different characteristics or tasks such as pitch class identification. Coloring notes also improves accessibility by mitigating reading difficulties [92] and introduces cognitive pleasure [179].



Augmented Sheet Layer – The second layer (see Figure 7.3(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 7.3(c)) keeps the sequential order, but omits the CMN through collapsing all instruments through *staged animations* [124] to set the focus on the glyph visualizations. Figure 7.3(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 [63] rearranges the glyphs' position based on the feature. The projection can reveal clusters and outliers. The example of the instantiation we introduce in section 7.5 applies a MDS projection based on the harmonic glyph design that we published in our previous work [222]. 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.



7.5 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 [103, 222]. Figure 7.1 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 [222] and rhythm [103]. 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.

7.5.1 Feature Visualization

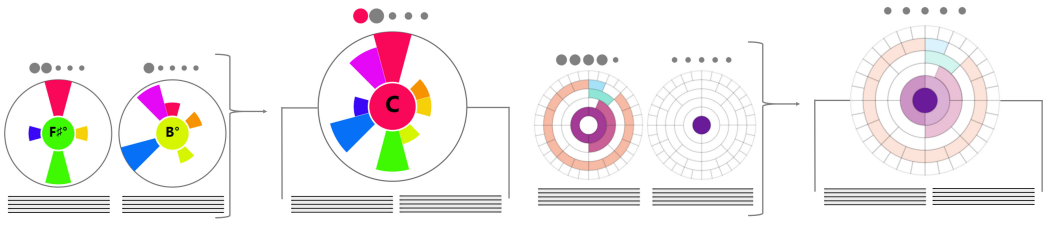
MUSICVIS uses varying visual encodings for harmonic, rhythmic, and melodic aspects. The harmonic and rhythmic glyphs cover respective problem/data aspects and analysis tasks (see Section 7.3). As discussed in Section 7.4, the glyphs are modular and interchangeable when keeping the technique's requirements. The separate feature encoding enables focusing on single aspects.

Harmony – MUSICVIS uses the harmonic fingerprint design introduced in Chapter 4 to visualize pitch-class distributions enabling analysts to compare the harmonic complexity. We refer to Chapter 4 for more details about the visual encoding based on the *circle of fifths* [125]. The proximity of arc segments indicates harmonic similarity and sound more consonantly together (e.g., C is closer to G than to B). Pitch classes 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 I discussed in Chapter 5. 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 [192]. The rhythmic fingerprint supports analysts in detecting the rhythmic similarities and differences between 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 7.7 illustrates how lines based on the respective melody notes' positions emphasize their contour. A melodic passage's dominant pitch or chord root 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.

Aggregation of Harmony and Rhythm – As required by the technique, MUSICVIS allows analysts to aggregate and split the rhythm and harmony glyphs at each close and distant reading level. The aggregated glyphs provide statistical summaries of the underlying measures. Figure 7.4 illustrates the aggregation of harmony and rhythm glyphs. The harmonic example shows how the pitch class distribution in the aggregation reflects the underlying content revealing the 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



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

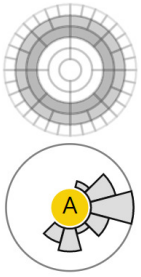
Fig. 7.4.: Summarizing neighboring harmonic and rhythmic glyphs to get a more abstract representation of the underlying features.

differences in 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.

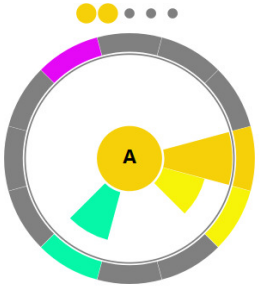
7.5.2 Visual Querying

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.

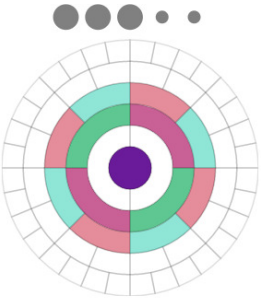
7.5.3 Querying Harmony and Rhythm



MusicVis provides visual query selectors tailored to the respective feature glyph design, enabling 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 identify rotated search template examples. The queries can reverse the match opacity scales to emphasize dissimilar glyphs, which is particularly helpful for identifying outliers.



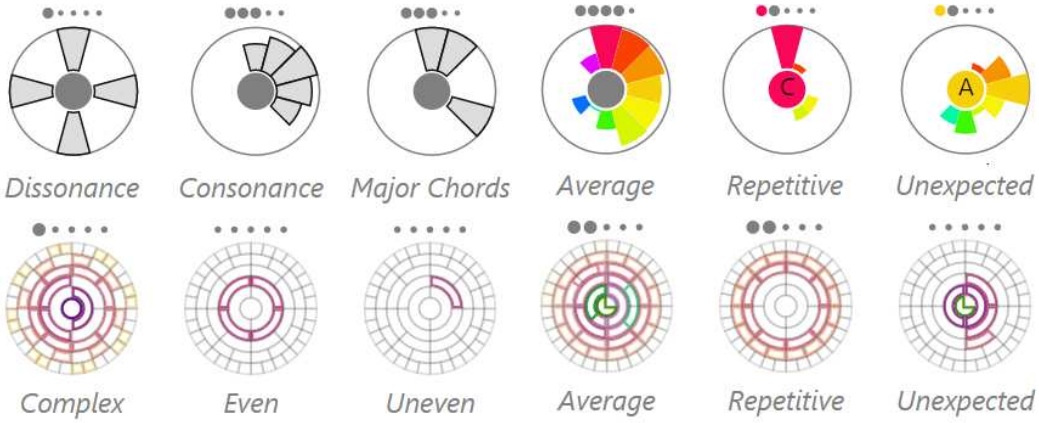
Harmony Query – This figure shows an exemplary A-Major query using the pitch classes A, C♯, and E while prohibiting the pitch class of F since the outer segment is selected, but the area is set to zero. The outer ring segments indicate required or optional pitches. Combining optional and mandatory pitches creates custom queries ("*Match Custom*"). The five dots above the query display the relevance score of the selected root pitch (*color*) and the query shape match (*gray dots*). The technique matches the rotated glyph versions if any rotation is allowed in the query. 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^{th} or 32^{th} notes or rests. The relevance score at the top indicates a relatively high shape match of the rhythmic query. 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 custom queries of *complex* or *even* patterns.

Semantic and Statistical Features

Music analysts can discover extracted glyph templates, including the *most common* harmony and rhythm glyphs (see Figure 7.5) automatically. 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 7.5(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**]. The appendix contains a detailed overview of harmonic fingerprints regarding the twelve pitch classes and their functional purpose in major scales in Section A.1. Figure A.4 contains a decent overview of many chord structures based on the harmonic fingerprint design, from single root notes to very complex Jazz chords comprising up to six. The appendices also contain a more systematic investigation of the symmetry in the rhythmic fingerprints. The details about these rhythmic fingerprints accompanied with CMN are given in Section A.2.

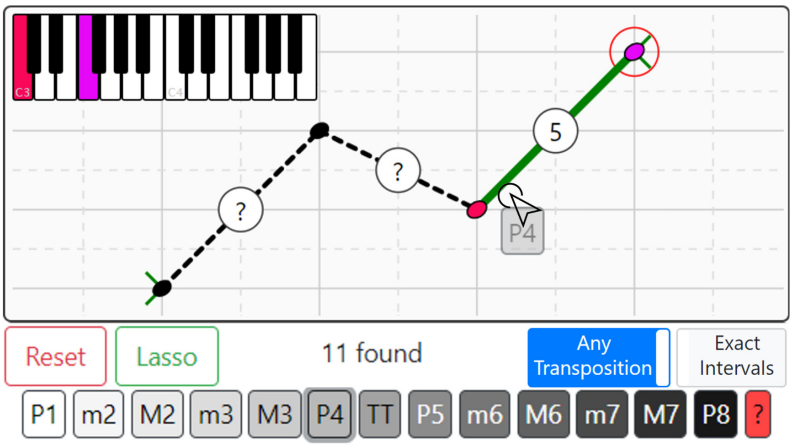


(a) Semantic Query Templates (b) Statistical Query Templates

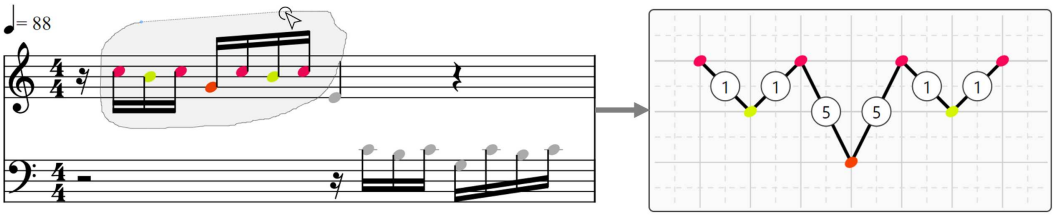
Fig. 7.5.: Querying by semantic and statistical feature templates.

7.5.4 Querying Melody

An essential contribution of MusicVis is the separate visual query for melodic motifs (see Figure 7.6). The melody query allows analysts to explore occurrences of a melodic motif. Figure 7.6(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 7.6(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.



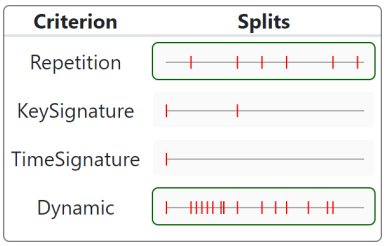
(a) Interactive melody query user interface.



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

Fig. 7.6.: Melody Query. The melody search interface (a) enables analysts to explore melodic patterns. Different parameters allow the creation of custom queries by assigning intervals or pitches. The lasso selector (b) allows for querying existing melody themes directly from the CMN.

7.5.5 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 – Based on the query results, the analyst can also use the harmony, rhythm, or melody query to apply a splitting. 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. Hence, they can explore and compare different combinations of custom and predefined split criteria without repeatedly recreating expensive manual splittings.

7.5.6 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.

7.5.7 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 [32] to create visual elements and animations. The backend is implemented using Python with Flask [118] and provides a REST API that 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 facilitate querying single melodic themes.
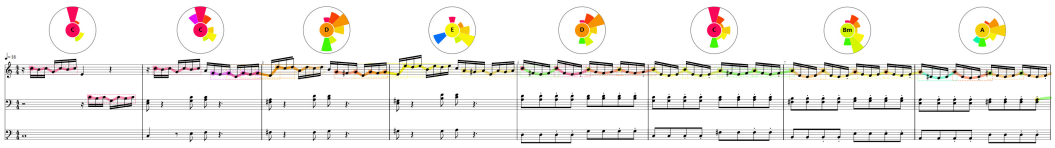


Fig. 7.7.: The first eight measures of *BWV 553* at the *augmented sheet level* reveal a modulation from C-Major to E-Major. The example contains two melodic motifs. The second voice repeats the first motif. The second motif uses *circulatio* [167].

7.6 Exemplary Analysis Workflow

We take the *Prelude in C-Major (BWV 553)* [242] by J. S. Bach (only 29 measures) to showcase *MusicVis*' components. Figure 7.7 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 7.6(b) and back to C in measures 2–6. The melody highlighting indicates how the starting motif (red) transforms to a repetitive *Circulatio* pattern [167] 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 32nd notes. The rests' shortest duration is shown at the 8th ring. In contrast, the harmony summary glyph has a high shape match score (4/5 gray dots), indicating that about 80% of the piece has the root C. This glyph also reveals that the prelude primarily comprises 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**].

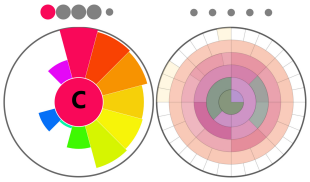
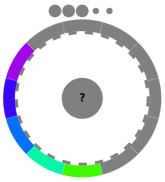




Fig. 7.8.: 65% of the harmony glyphs in Bach's BWV 553 contain notes that do not belong to the C-Major scale (F♯, C♯, G♯).

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 7.8) to realize that Bach only sticks to the C-Major scale notes if the *leitmotif* (see Figure 7.6(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.



7.7 User Study

We conducted a pair analytics study with different user types covering various application scenarios to evaluate our proposed visualization technique. We decided to use the pair analytics study method [151] 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 [314].

7.7.1 Study Design

We used online video conferencing software to share MUSICVIS via the study director's screen. We gained valuable insights into students, teachers, professionals, and hobby musicians' challenges when using sheet music. We gathered initial expectations and interactive session feedback. We inquired final retrospective feedback, revealing to what extent MUSICVIS met the participants' initial expectations.

Methodology – We performed a combined *pair analytics* and *observation* study [8] 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*" [241] 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 subsection 7.3.4). In addition, every participant selected a second piece that the other participants did not use to show the general applicability of *MusicVis*. Table 7.1 provides an overview of all selected custom musical compositions based on MusicXML selected by the study participants comprising information about composers and composition titles. Moreover, the table also includes direct references to the respective MuseScore documents for further reference in the right column [243].

Participant	Composer / Artist	Piece Title	URL
Intro #1	Thomas Hasenberger	Alle meine Entchen	https://musescore.com/score/141525
Intro #2	Joseph Haydn	Minuet in A. No. 6	https://musescore.com/score/5491192
All participants	Wolfgang Amadeus Mozart	Sonata No. 11 "Rondo Alla Turca"	https://musescore.com/score/49143
$\mathbf{P_1}$	Georg Friedrich Händel	Allegro HWV 407 for Solo Flute	https://musescore.com/score/6092784
$\mathbf{P_2}$	Georg Philipp Telemann	Sonata in F-Minor Mvt. 4	https://musescore.com/score/4861340
$\mathbf{P_3}$	Franz Schubert	Winterreise, D.911, No.24 - Der Leiermann	https://musescore.com/score/5014023
$\mathbf{P_4}$	Billie Eilish	Bad Guy	https://musescore.com/score/5512876
$\mathbf{P_5}$	Ludwig van Beethoven	Violin Sonata No.9 Op.4	https://musescore.com/score/2855611
$\mathbf{MT_1}$	Johann Sebastian Bach	Prelude BWV 846	https://musescore.com/score/719631
$\mathbf{MT_2}$	Yiruma	River Flows In You	https://musescore.com/score/5905914
$\mathbf{A_1}$	Zequinha Abreu	Tico Tico no fubá	https://musescore.com/score/4939446
$\mathbf{A_2}$	Johann Sebastian Bach	Johannes Passion - BWV 245 - Nr. 15 - "Wer hat dich so geschlagen"	https://musescore.com/score/5389251
$\mathbf{A_3}$	Johann Sebastian Bach	Johannes Passion - BWV 245 - Nr. 68 - "Ach, Herr, Laß dein lieb Engelein"	https://musescore.com/score/5396583
$\mathbf{H_1}$	Johann Sebastian Bach	Suite No. 1 in G-Major BWV 1007: Prelude	https://musescore.com/score/1031651
$\mathbf{H_2}$	Yann Tiersen	Comptine d'un autre été (l'après midi)	https://musescore.com/score/4809374
$\mathbf{H_3}$	Arctic Monkeys	Golden Trunks	https://musescore.com/score/5990418
$\mathbf{H_4}$	Ludwig van Beethoven	Sonata No. 14 Op. 27 No 2	https://musescore.com/score/880516
$\mathbf{D_1}$	Ludwig van Beethoven	Für Elise	https://musescore.com/score/5507375
$\mathbf{D_2}$	Johann Sebastian Bach	Invention No. 8 in F-Major, BWV 779	https://musescore.com/score/1015261

Tab. 7.1.: As part of the study preparation, all participants selected a specific piece that they analyzed during the study in addition to Mozart's Sonata No. 11, "Rondo Alla Turca." The first two rows in the table are the two *short* pieces used to introduce all features of MusicVis as preliminary training. We ensured that all participants selected different pieces with different numbers of instruments, styles, and clefs. This list demonstrates the general applicability of various musical compositions including even conductor's scores.

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.

7.7.2 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 investigating 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 builds on existing 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 require literacy. Before the study, they often had no idea 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 of the harmonic complexity [**T2**] which helped her confirm the present composition's belonging epoch. Overall, she explained that at the projection level, the occurrence of clusters helped her visually identify 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 the 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 of the glyph representations. Similarly, [Q13–Q16] address rhythm, [Q17, Q18] cover the melody query, and [Q19, Q20] the structure analysis. [Q21–Q23] cover the close and distant reading levels, including the staged and seamless 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 ∅ score of ∼3.65. Grouped by topic, the participants preferred the melody features (∅: ∼3.93), then harmony (∅: 3.85) and structure (∅: ∼3.83) analysis. **Q17** ("*I can identify Leitmotifs by using the melody search*") received the best average score: ∼**4.26**. **Q9** ("*The application supports understanding dominant key signatures*") received the second best average score of ∼**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 ∼**3.26**.

7.8 Discussion & Summary

The 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 and obtain a more comprehensive understanding even if already familiar with them. Providing abstract and statistical information through visual representations of basic music concepts engages analysts to further analyze 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 better interpret and understand them. The study feedback revealed that our approach can be profitable for different application contexts, including music theory education, performance preparation, and music information retrieval. Thus, we argue that music students, teachers, musicians, and musicologists will benefit from this work.

Limitations – We instantiated the visual analytics technique to support various 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, in-depth analysis is challenging on smaller displays, but music analysts would not expect to read a symphony score on a single printed page. In this work, we focus on demonstrating 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 use to perform a quantitative comparison 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 better understand the system over time and can apply it to learning musical pieces.

Lessons Learned – Based on our study, we confirmed that building from a visual data representation that users are familiar with is essential. 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 [13]. This would not be feasible on printed sheet music, usually used for teaching. Moreover, gamification [303] poses a great opportunity: students could provide information automatically checked by the system, e.g., the chords of each measure. 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 [11] 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 [308]. As Borgo et al. investigated, there are several alternative glyph designs [30] that could be applied on musical data. Our technique demonstrates how glyphs can supply aggregated information about musical characteristics. We argue that research at the interface of musicology and visualization holds great potential for interdisciplinary collaboration. Specifically, it would be interesting to invent glyph designs that address analysis tasks such as chord identification, instrument segregation, and motive detection. We plan to conduct further studies tailored to specific user groups and analysis tasks to fathom how our work could be applied in different scenarios.

Summary

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.

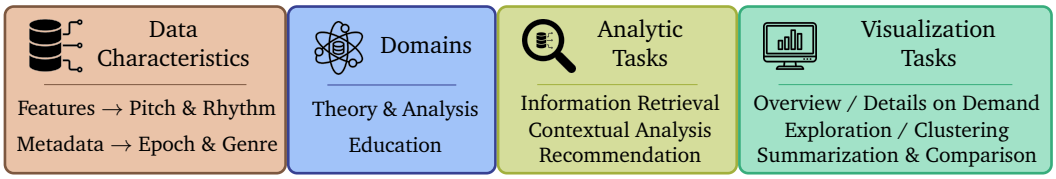
We present MusicVis, an instantiation of the technique that integrates four tightly connected close and distant reading levels using staged animations. MusicVis offers different visual queries for enhanced exploration of harmonic, rhythmic, and melodic features. A structure management view supports the interactive analysis of compositions' structural organization through splitting criteria such as repetitions, 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 feature queries enable music analysts to explore dominant and uncommon harmonic, rhythmic, and melodic patterns. We elicited further research opportunities based on the study feedback, such as applying MusicVis in collaborative teaching contexts.

8

Top-Down | CorpusVis Visual Analysis of Music Collections

> *Interactive visualization empowers you to explore, understand, and communicate complex data.*
>
> **— Hans Rosling**
> *Swedish Physician & Academic*



Manually investigating sheet music collections is challenging for music analysts due to the magnitude and complexity of underlying features, structures, and contextual information. However, applying sophisticated algorithmic methods would require advanced technical expertise that analysts do not necessarily have. Bridging this gap, we contribute CorpusVis, an interactive visual workspace, enabling scalable and multifaceted analysis. Our proposed visual analytics dashboard provides access to computational methods, generating varying perspectives on the same data. The proposed application uses metadata, including composers, type, epoch, and low-level features such as pitch, melody, and rhythm. To evaluate our approach, we conducted a pair-analytics study with nine participants. The qualitative results show that CorpusVis supports users in performing exploratory and confirmatory analyses, leading them to new insights and findings. In addition, based on three exemplary workflows, we demonstrate how to apply our approach to different tasks, such as exploring musical features or comparing composers.

Chapter 8 is based on the following publication. Please refer to Section 1.4 for contribution clarifications.

[227] **Matthias Miller**, Julius Rauscher, Daniel A. Keim, and Mennatallah El-Assady. **"CorpusVis: Visual Analysis of Digital Sheet Music Collections"**. *Computer Graphics Forum, Vol. 41, No. 3, 2022.*

8.1 Introduction

Investigating music collections is relevant for domain experts such as musicologists, music theorists, and usual music consumers [347]. Predominantly, the audio format has been the starting point for many applications, including music classification [99], personalized music recommendation [27], structure analysis [262], and the generation of music playlists [284]. On the contrary, we are convinced that the investigation of sheet music collections based on digital symbolic formats such as MusicXML [111], MIDI, Humdrum [286], and Lilypond [247]have received less attention compared to audio.

While services such as sheet music recommendations offered by, for example, OKTAV [169] leverage musical features and individual preferences to provide sheet music suggestions for piano players, it remains unclear how such recommendations are computed. To understand which musical features (e.g., structure or melody) reveal differences and similarities, computational approaches for sheet music analysis have been proposed in the community, supporting various tasks such as genre and composer classification [9]. In addition to close reading of sheet music as a typical task in musicology, distant reading is also an essential task in the digital humanities [145]. Close reading supports musicologists to investigate musical compositions on a detailed level. Abstract visualization contains the potential to provide an overview of a large set of musical compositions within their context, thus enabling distant reading. While approaches that solely rely on algorithmic or computational approaches mitigate the tedium of manual analysis of a dataset performed by human analysts, they often are not accessible to a broader user group. Still, they are limited to music analysts who have programming skills. To address this gap, tailored visualizations combined with user interaction can be used to increase access to music analysis methods for a larger user group while involving the human analyst in the analysis loop, which is a fundamental aspect of knowledge generation in visual analytics [282]. An essential aspect of the understanding process is critical thinking by the analyst about the subject, which was introduced by Bradley et al. as *slow analytics* [33].

The availability of sheet music datasets such as KernScores [288] or MuseScore [240] has the potential for analysis at different scales, benefiting from using projection techniques based on underlying features [213]. While MuseScore contains over a million compositions, the uploaded material's quality varies from content provided by users who ensure that the provided content is faithful to the original compositions to pieces that suffer from inaccuracy or incompleteness. Therefore, it is challenging to maintain a high quality of the underlying data for the analysis when considering the full dataset, requiring expensive data cleaning steps before performing analysis tasks on it. Manual curation or selection processes are required to set the focus on a particular subset that is provided by users that only provides high-quality content. Alternatively, analysts can manually view each composition that shall be part of a curated dataset, which is quite a tedious process. The research field of optical music recognition provides approaches to convert printed sheet music into symbolic formats such as MusicXML. However, there is still a need for further improvements until musicologists trust these automatic results [304]. Enabling analysts to influence the data foundation is crucial for analysis.

In addition to data quantity, increasing the data quality is crucial for effective analysis. This includes a targeted selection of representative samples. For example, a basic issue in humanities research is avoiding excluding marginalized positions, a typical risk of data colonialism. For instance, famous classical composers such as *Johannes Sebastian Bach*, *Wolfgang Amadeus Mozart*, *Joseph Haydn*, or *Ludwig van Beethoven* are often more prominent analysis subjects in musicology than less-known composers such as *Muzio Clementi*, *Domenico Scarlatti*, or *Francois-Joseph Gossec* [115]. This marginalization of composers is reaffirmed by the vast amount and duplicates of famous compositions in datasets such as MuseScore. A corpus-level overview is essential for large-scale analysis of sheet music collections, enabling analysts to audit and refine corpora and compare music sheets on numerous features. Applying abstract visualization techniques to sheet music corpora at a larger scale can provide insights about composers' work or typical differences between composition types without manually analyzing every detail of single data.

Driven by the need for a corpus-level music analysis technique, in this work, we address the research question: *How can music analysts explore, investigate, and compare sheet music collections using interactive visualization?* We aim to provide them with a multiscale, multi-perspective bird's-eye view of sheet music corpora. Therefore, we designed an interactive visual workspace with multiple tailored analysis components in an interconnected dashboard. The visual analysis is supported by computational methods, such as clustering, to aid pattern finding. Analysts can openly explore the underlying corpus, investigate pre-configured use cases, or verify hypotheses by crafting their own analysis workflow.

Contributions – This work contributes a problem characterization addressing visualization requirements concerning the analysis of sheet music corpora. We provide a list of relevant data, task characteristics, and a description of target audiences. A major contribution is the Visual Analysis Workspace for sheet music collections through a combination of multiple components that are seamlessly connected through linking and brushing. We conducted a qualitative evaluation to assess the applicability of our approach and provide details about its benefits and drawbacks. Finally, we discuss open research opportunities to inspire interdisciplinary collaboration at the interface of visualization and musicology.

8.2 Related Work

Musicology is a wide research area that covers heterogeneous research questions and challenges. The subfield *Visual Musicology* at the interface of musicology and visualization research, as introduced in Chapter 2, illustrates the vast opportunities for visualization to support domain-related scientific issues. Their framework highlights the potential of interactive visualization to perform analysis tasks, including *information retrieval*, *exploration*, and *comparison*. We use the *visual musicology graph* to classify the work presented in this paper accordingly. Specifically, we focus on *Structural Features* and *Meta-Information* of sheet music and the visualization tasks *Overview/Summarization*, *Navigation/Exploration*, *Clustering*, *Comparison*, and *Details on Demand* within the domains *Theory & Analysis* and *History*.

Visual Analysis of Music Collections – The visual investigation of musical data collections has already received attention from visualization researchers. Khulusi et al. created a web-based interface called "musiXplora" as part of a digitization project that enables humanities researchers to investigate meta-information (e.g., gender, religious denomination, profession, institutions) of musicians and instrument makers [162]. For instance, this interactive prototype facilitates information retrieval and browsing about the life and work of different composers. Similarly, Jänicke et al. designed an interactive visualization tool for interactive profiling of musicians and their relationships based on their meta-information [144]. They did not consider specific features of the compositional work but rather focus on the contextual information about the life, profession, and instruments they mastered. Pampalk used SOMs to visualize and classify the genre of music data as "Islands of Music" based on audio-extracted features such as the loudness information from audio files [255]. Chen and Putz designed an interactive user interface for browsing and organizing music collections, allowing one to listen and visually explore the similarity of musical pieces based on low-level features that reflect the genre and style of pieces from different composers [55]. Weiß et al. show how exact tonal features (pitch information) and metadata from audio data help to detect style changes over the different epochs from baroque to modern music [349]. Georges and Nguyen use a dataset of 500 classical composers to visually analyze different epochs using dendrograms and MDS projection techniques [105]. By that, they demonstrate how abstract visualization methods support the comparison of composers [34].

Feature Extraction from Music Data – Music is available and can be stored based on different formats, such as symbolic (e.g., MusicXML [111]) or audio data (e.g., mp3). Depending on this format, different features can be extracted afterward for different tasks, such as music recommendation [292]. Similar to extracting features from audio signal data (e.g., spectral information) for further processing [187], it is possible to extract low-level features from symbolic music [66]. For example, music21 [66] is a Python library that enables programmers to extract jSymbolic features from different symbolic sheet music formats, including MusicXML. McKay and Fujinaga discuss how such features can be used for MIR research tasks [213] such as composer classification as done by Verma and Thickstun leveraging convolutional networks [339]. McKay leverages audio and symbolic music data sources to perform music classification and create music information retrieval tools using his own jMIR software suite that can even consider contextual/cultural metadata [214]. Corrêa and Rodrigues published a survey about music genre classification that reveals that such tasks primarily use automatic algorithms without using advanced visualization techniques [62]. Merely computational approaches do not enable music analysts to explore low-level features, making it impossible to step into the analysis process to get a better overview of the inner workings of the applied algorithms.

Visualization of Collections in the Digital Humanities – Besides musicology, information visualization researchers have already addressed other scientific issues within digital humanities, such as the visual analysis of poems [210]. Jänicke et al. discuss many digital humanities projects that require text processing methods [146] while Kirschenbaum also confirms the necessity of text for humanities research, ranging from close via not-reading to distant reading [168]. For example, Bludeau et al. implemented a web-based prototype to visually investigate literature and hand-

written notes from *Fontane's Handbibliothek* [26]. For instance, text analysis deals with comparison tasks for plagiarism detection by identifying typical features for specific authors [107]. Keim and Oelke use textual features to create a literature fingerprinting visualization that facilitates the analysis of whole books [157]. We argue that existing visualization methods successfully applied to similar endeavors in related research areas enclose the potential for methodology transfer [229]. Consequently, we consider it useful to get inspired by existing visual techniques to address unsolved challenges, such as the visual analysis of sheet music collections.

Research Gap – While there are several projects within this research field, they are often limited to audio data only [55, 349]. Thus, we argue that the analysis of music based on symbolic formats has received less attention. We argue this shift towards audio data could be due to larger user groups of music applications such as (e.g., Spotify) which benefit from effective music recommendation methods. We see the reason for this unbalance in the smaller group size of sheet music consumers, which is much smaller. Liem et al. discuss the issue of focusing on audio signals only and argue for multimodal and user-centered analysis strategies to improve the accessibility of digital music data [194, 237]. Nevertheless, the increased use of digital symbolic music formats led to services such as OKTAV [169], which provides sheet music recommendations based on symbolic features. While automatic solutions have their application areas, employing visual methods allows analysts to step into the analysis process, helping them gain new insights or generate and confirm hypotheses, also known as human-in-the-loop approaches [282]. To our knowledge, no previous work readily supports the analysis of sheet music corpora, facilitating the visual comparison of metadata and low-level features.

8.3 Problem Description

Our primary objective is to provide an *interactive visual workspace* to support the investigation and exploration of sheet music collections. The target audience primarily comprises musicologists and music theorists. We assume that music librarians and Music Information Retrieval (MIR) researchers could benefit from our work as well. For this reason, we gathered information about the data and task requirements by conducting initial expert interviews and consulting existing literature.

Feature Characteristics – When analyzing musical compositions, musicologists need to know the relationship to the *composers*, which allows including domain knowledge in the analysis process. Usually, one can assign musical pieces to a certain *type or style*, which often correlates to a specific genre. Considering this, contextual information can help detect and compare entities from a musical score collection. As discussed in Section 8.2, there exist various frameworks that allow for extracting *statistical features* from symbolic music. We use music21 to extract jSymbolic features from digital sheet music [67] that provides the foundation for the visual analysis workspace we introduce in the following sections. Specifically, these low-level features contain rhythmic, melodic, and pitch information that provide informative attributes about single compositions or numerical distributions when investigating a score collection. In addition, we consider available contextual metadata, including composer, composition type, and temporal characteristics.

Type	Feature	Description
Metadata	Title	Title of the composition.
Metadata	Composer	Composer of the composition.
Metadata	Form	Musical form of the composition.
Metadata	Opus	Opus number of the composition.
Melodic	Amount of Arpeggiation	Fraction of horizontal intervals that are repeated notes, minor thirds, major thirds, perfect fifths, minor sevenths, major sevenths, octaves, minor tenths or major tenths.
Melodic	Average Melodic Interval	Average melodic interval (in semitones).
Melodic	Chromatic Motion	Fraction of melodic intervals corresponding to a semitone.
Melodic	Direction of Motion	Returns the fraction of melodic intervals that are rising rather than falling. Unisons are omitted.
Melodic	Distance Between Most Common Melodic Intervals	Absolute value of the difference between the most common melodic interval and the second most common melodic interval.
Melodic	Duration of Melodic Arcs	Avg. number of notes that separate melodic peaks and troughs in any part: calculated as the total number of intervals (ignore unisons) div. by the number of melody direction changes.
Melodic	Melodic Fifths	Fraction of melodic intervals that are perfect fifths.
Melodic	Melodic Octaves	Fraction of melodic intervals that are octaves.
Melodic	Melodic Thirds	Fraction of melodic intervals that are major or minor thirds.
Melodic	Melodic Tritones	Fraction of melodic intervals that are tritones.
Melodic	Most Common Melodic Interval	Melodic interval with the highest frequency.
Melodic	Most Common Melodic Interval Prevalence	Fraction of melodic intervals that belong to the most common interval.
Melodic	Number of Common Melodic Intervals	Number of melodic intervals that represent at least 9% of all melodic intervals.
Melodic	Relative Strength of Most Common Intervals	Fraction of melodic intervals that belong to the second most common interval divided by the fraction of melodic intervals belonging to the most common interval.
Melodic	Repeated Notes	Fraction of notes that are repeated melodically.
Melodic	Stepwise Motion	Fraction of melodic intervals that corresponded to a minor or major second.
Melodic	Size of Melodic Arcs	Average span (in semitones) between melodic peaks and troughs in any part. Each time the melody changes direction begins a new arc. The average size of melodic arcs is defined as the total size of melodic intervals between changes of directions - or between the start of the melody and the first change of direction - divided by the number of direction changes.
Pitch	Importance of Bass Register	Fraction of Notes between MIDI pitches 0 and 54.
Pitch	Importance of High Register	Fraction of Notes between MIDI pitches 73 and 127.
Pitch	Importance of Middle Register	Fraction of Notes between MIDI pitches 55 and 72.
Pitch	Interval Between Strongest Pitches	Absolute value of the difference between the pitches of the two most common MIDI pitches.
Pitch	Interval Between Strongest Pitch Classes	Abs. value of the diff. between the pitch classes of the two most common MIDI pitch classes.
Pitch	Most Common Pitch	Bin label of the most common pitch.
Pitch	Most Common Pitch Class	Bin label of the most common pitch class.
Pitch	Most Common Pitch Prevalence	Fraction of Note Ons corresponding to the most common pitch class.
Pitch	Most Common Pitch Class Prevalence	Fraction of Note Ons corresponding to the most common pitch.
Pitch	Number of Common Pitches	Number of pitches that account individually for at least 9% of all notes.
Pitch	Pitch Class Variety	Number of pitch classes used at least once.
Pitch	Pitch Variety	Number of pitches used at least once.
Pitch	Primary Register	Average MIDI pitch.
Pitch	Relative Strength of Top Pitch Classes	The freq. of the 2nd most common pitch cl. divided by the freq. of the most common pitch cl.
Pitch	Relative Strength of Top Pitches	The freq. of the 2nd most common pitch divided by the frequency of the most common pitch.
Rhythm	Average Note Duration	Average duration of notes in seconds.
Rhythm	Average Time Between Attacks	Average time in seconds between Note On events.
Rhythm	Changes of Meter	Set to 1 if the time signature is changed one or more times during the recording.
Rhythm	Compound or Simple Meter	Set to 1 if the initial meter is compound (numerator of time signature is greater or equal to 6 and is evenly divisible by 3) and 0 if it is simple.
Rhythm	Duration	The total duration in seconds of the music.
Rhythm	Initial Tempo	Tempo in beats per minute at the start of the recording.
Rhythm	Maximum Note Duration	Duration of the longest note (in seconds).
Rhythm	Minimum Note Duration	Duration of the shortest note (in seconds).
Rhythm	Note Density	Average number of notes per second.
Rhythm	Quintuplemeter	Set to 1 if numerator of initial time signature is 5, set to 0 otherwise.
Rhythm	Staccato Incidence	Number of notes with durations of less than a 10th of a second div. by the total number of notes.
Rhythm	Triplemeter	Set to 1 if numerator of initial time signature is 3, set to 0 otherwise.
Rhythm	Variability of Note Duration	Standard deviation of note durations in seconds.
Rhythm	Variability of Time Between Attacks	Standard deviation of the times, in seconds, between *Note On* events.

Tab. 8.1.: This table comprises all precomputed features available in the feature matrix visualization in CorpusVis. **M**elodic, **p**itch and **r**hythm features were extracted using *music21* [67]. In the future, extending this list to support further analysis aspects would be easily possible.

Task Characteristics – Content-based MIR aims at uncovering music data relevant characteristics that can be employed for similarity computations and retrieval tasks [105]. For instance, a specific retrieval task [T1] would be the identification of all compositions from *Mozart* that have a time signature of a *triple meter*. Besides performing inquiring based on exact information, music analysts could also be interested in exploring similarities of pieces within a dataset [T2] with the objective of creating new hypotheses or gaining new insights about the underlying information [163]. Additionally, the comparison of two or even a set of compositions [T3] can be relevant during music analysis to identify commonalities of the complete work of composers or different versions of the same composition [336]. Not only can comparison tasks benefit from grouping and clustering based on meta information [T4], but also the investigation of single composers based on statistical values, they provide a quick overview without the need to view all items separately. Depending on the quality of the dataset, it can be crucial for analysts to identify duplicates within a dataset [T5] to omit works from further analysis or other issues such as detecting musical plagiarism [43]. Based on our cooperation with musicologists, we focus on the following tasks that address typical analysis challenges within the scope of sheet music corpus analysis:

- [T1] Retrieve/Filter compositions based on title, composer, or type
- [T2] Exploratory analysis of a musical score collection
- [T3] Comparison of scores, composers, types, and features
- [T4] Clustering/grouping of pieces, composers, and types
- [T5] Detection of similar entities in a given corpus

Visualization Requirements – Based on the described data and task characteristics, we can identify essential requirements for the visualizations necessary to support the analysis process. For the information retrieval task [T1], we need a separate filter and sort functionality for each meta feature and the low-level characteristics. Suppose a user is looking for a specific composer or composition. In that case, we need an interactive search bar that allows for direct querying of the underlying dataset to avoid tedious searching for data items within the provided visualizations. To associate the temporal aspects of composers, we need a timeline visualization that can also be directly used to filter composers to be analyzed by the user. The type of composition is another fundamental characteristic for music analysts that can be used for exploration [T2], comparison [T3], or identification of interesting pieces. To extract similar compositions within the dataset, we can employ a dimension reduction technique such as MDS [63] to project dozens of low-level features into a two-dimensional space [T5]. Then, identifying groups based on a clustering method is required to detect multiple pieces that have related features [T4].

Used Dataset – MuseScore comprises a heterogeneous dataset with over a million compositions. Instead of using all these files, we manually selected a subset from certain users who upload curated files, e.g., ClassicMan [60]. Typically, these compositions have high community ratings. Because of this preselection process, we ensure that the compositions in the corpus are of better quality.

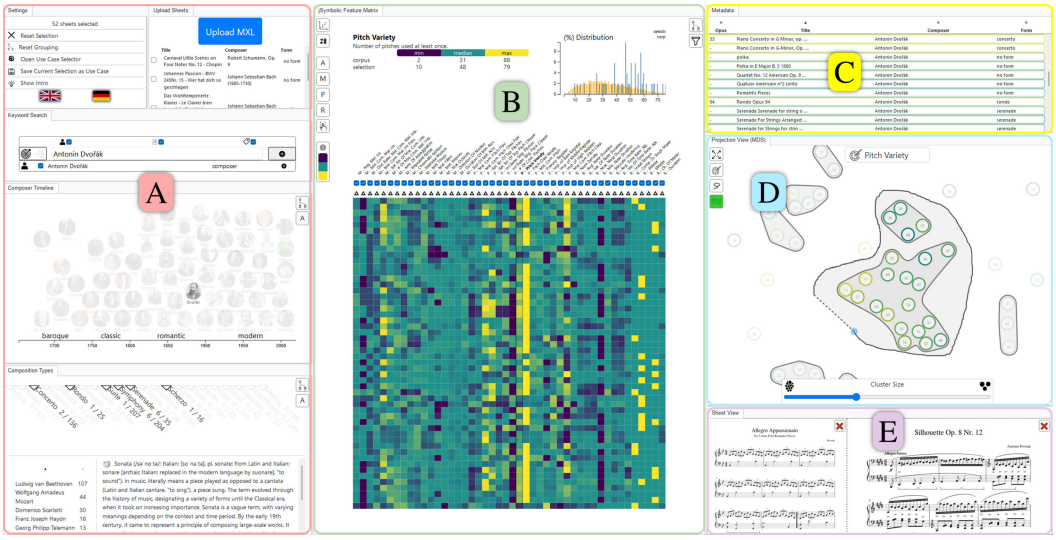


Fig. 8.1.: The visual interactive workspace supports the analysis of sheet music collections through visualizations connected through linking and brushing. After filtering a sheet music corpus A, music analysts can perform exploration, comparison, and detection tasks. The feature matrix B provides detailed information on the low-level characteristics of musical compositions. The metadata table C enables browsing through titles, composers, and composition forms. An MDS projection view D helps to identify similar compositions or clusters based on selected features. The sheet view E allows analysts to view compositions using the familiar notation. Domain experts can use CorpusVis to confirm and generate hypotheses and detect interesting patterns between composers and composition types.

8.4 Visual Interactive Analysis Workspace

This section introduces the interactive workspace and three exemplary analysis workflows to illustrate how to use it for analyzing sheet music corpora. CorpusVis is available as a web-based application via `https://visual-musicology.com/corpus`.

8.4.1 CorpusVis – User Interface

The analysis workspace consists of five components (see Figure 8.1) supporting data filtering and visual analysis following Shneiderman's Visual Information Seeking Mantra [306]. The components are connected through animated linking and brushing, showing the user relevant details during the analysis process. For instance, if the user hovers a circle (represents a composition or aggregations) in the projection view D, then the corresponding composer and forms A and matrix row(s) matrix B are highlighted. The settings widget (Figure 8.2, top left) enables users to switch between German and English and provides a guided tour for all components.

Data Import and Selection – This component enables uploading custom sheet music datasets (*.mxl* format) or filtering existing datasets by keywords (composer names, title, etc.), composition types, and via composer selection from the timeline `[T1]`. The user can save an interesting data selection as a use case or reset

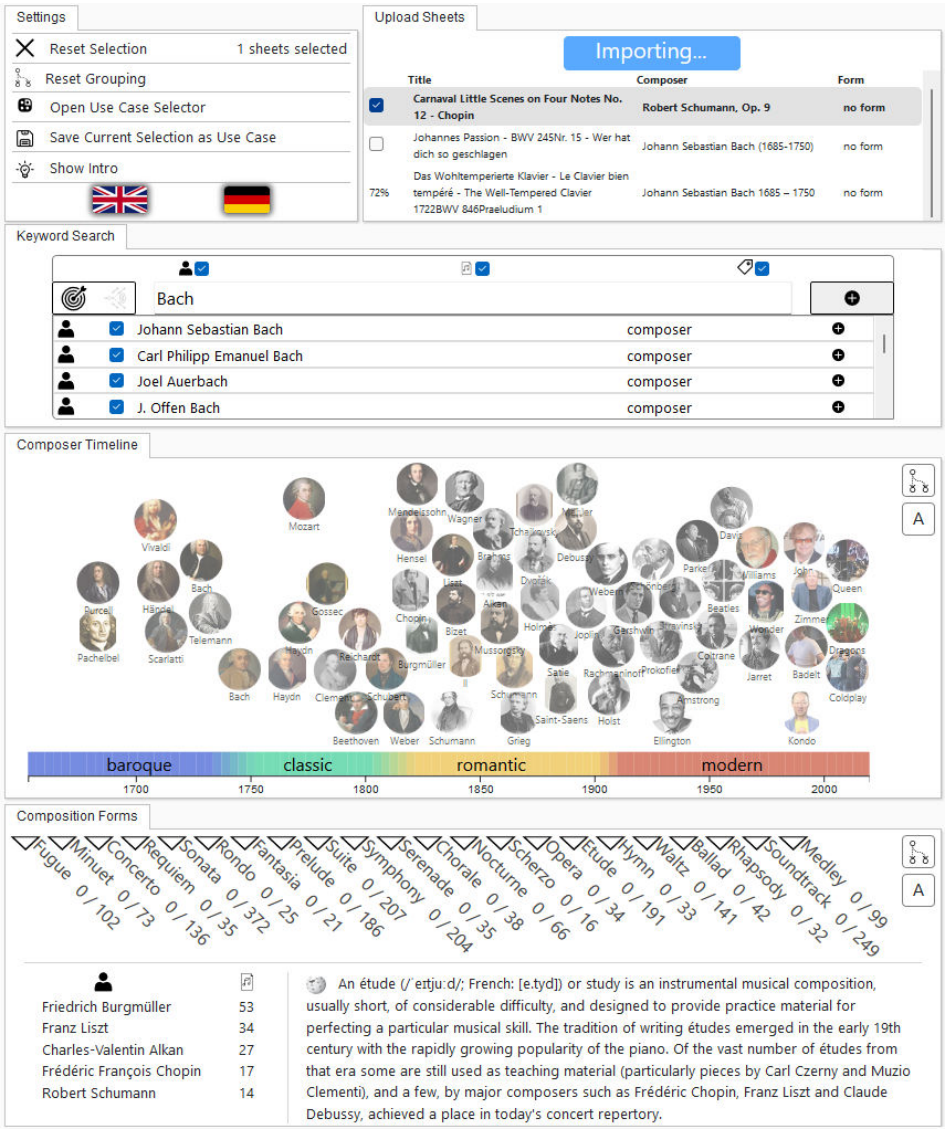


Fig. 8.2.: The keyword search allows filtering composers and composition types. Users can extend the dataset by uploading custom data. Further settings: reset the system state, save use cases, or open an interactive introductory guide through CorpusVis.

the analysis through further options provided in the settings widget. The composer timeline shows 62 composers from which we could extract their life data directly from the files. Similar to Weiß et al. [349], we use four historical periods (baroque, classic, romantic, modern) to position the composers on the timeline. We apply an animated layout that uses the vertical axis to avoid overlapping.

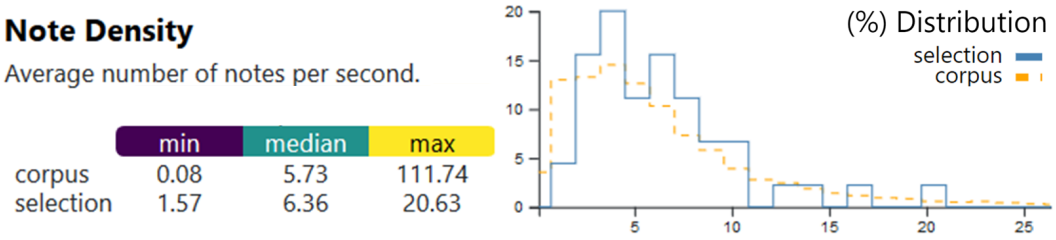


Fig. 8.3.: The feature distribution chart provides statistical characteristics (min., median, max.) of a subset and enables comparing the subset to the rest of the corpus regarding a selected feature.

jSymbolic Feature Matrix – Depending on the grouping, the feature matrix encodes statistical information extracted from single compositions or clusters at the workspace' center. User interaction allows sorting and filtering all features to focus only on subsets if needed [T3]. Analysts can also inform themselves by getting detailed information about each available feature. The features support identifying salient patterns, and the matrix can be used for comparison tasks [T2]. Overall, the feature matrix can display up to 46 different jSymbolic features extracted with music21 [66]. The feature distribution chart (see Figure 8.3) helps analysts compare the statistics of a feature to the whole corpus, including the minimum, median, and maximum value of the feature for each group [T3]. This example shows the statistical information of the feature *Note Density*, which corresponds to the average number of notes played within a second. The values show the corpus contains pieces with a note density above 100. At the same time, the selection is situated and has higher values on the left side of the distribution compared to the overall

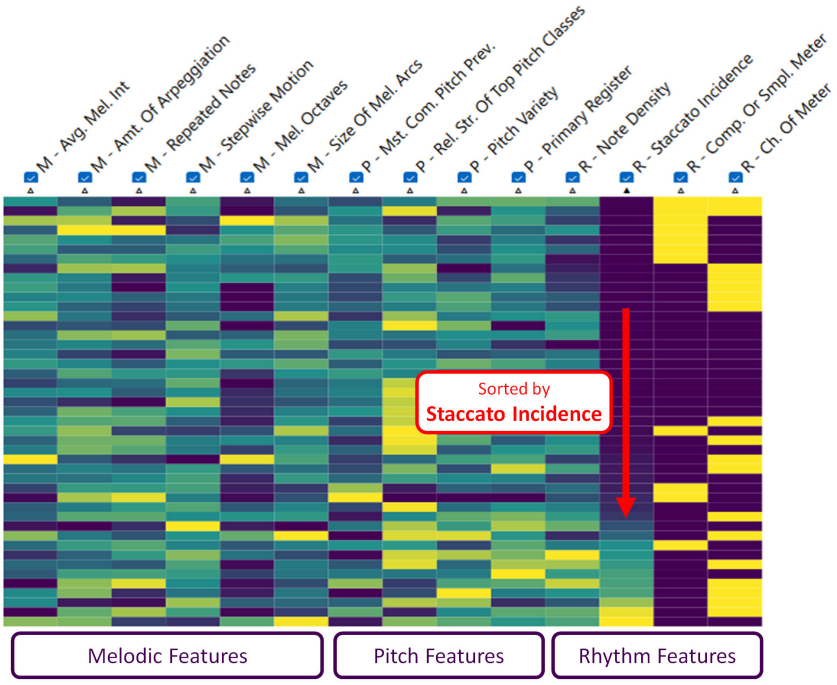


Fig. 8.4.: This exemplary matrix shows 14 features from *Clara Schumann* and *Modest Petrovich Mussorgsky*. Users can detect low (purple) and high (yellow) values.

corpus. The distribution chart illustrates that the feature regarding the selection is similar to the rest of the corpus (left-skewed). Figure 8.4 displays 14 features horizontally sorted by category (M, P, R) with the rhythmic feature *Staccato Incidence* emphasizing a few pieces with a high amount of staccato usage.

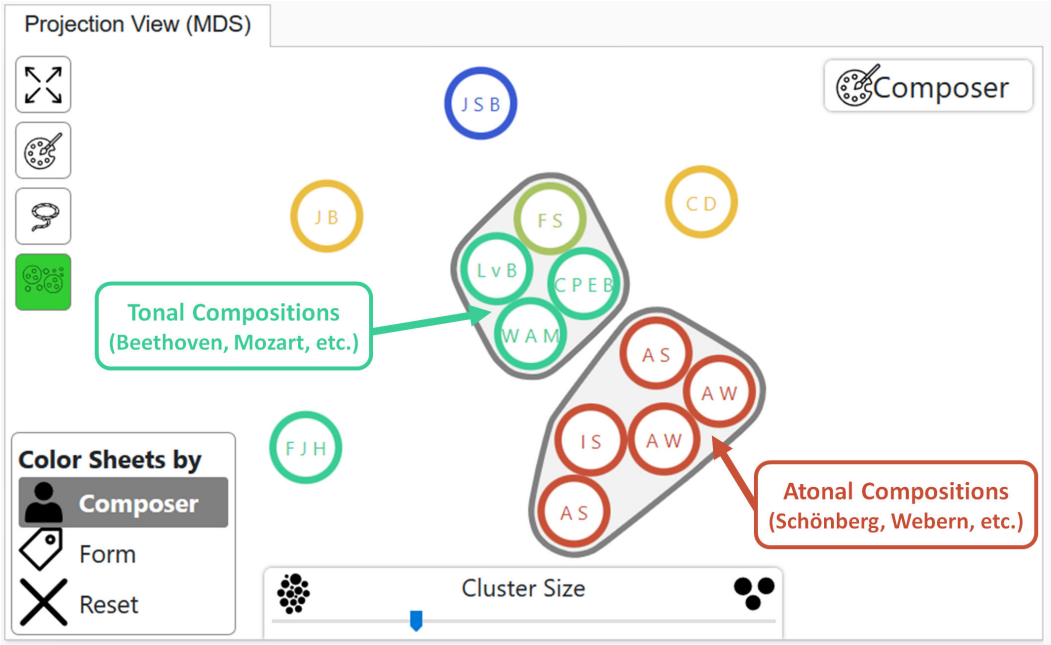


Fig. 8.5.: MDS Projection View: If no grouping is applied, each circle in the projection view represents a single composition. The color provides the information to which epoch a piece is assigned.

MDS Projection View – We compute a multidimensional scaling projection based on the selection of low-level features to show the similarity between single compositions or clusters in a separate component `[T5]`. The circle labels displayed in Figure 8.5 are the initials of the composer names (e.g., "L v B" for Ludwig van Beethoven). In this case, the color of the circles provides epoch information, and the respective color scale is shown in the composer timeline (see Figure 8.2). Analogously, a separate color scheme is used for the composition types. On-demand, the user can execute a DBSCAN clustering [130] and manipulate its *EPS* parameter via the cluster slide at the bottom to identify groups of similar entities `[T4]`. DBSCAN has the advantage that it does not require a parameter for the number of clusters and can identify outliers. We set $\mathrm{MinPTS} = 2$ to enable the identification of duplicates. The visual grouping of the clusters uses concaveman [2], a JavaScript library that builds on an algorithm based on work by Park and Oh [256], which allows the creation of concave hulls around given clusters.

Metadata Table – As a quick reference to the feature matrix, the workspace contains a metainformation table of the current selection, as displayed in Figure 8.6. Users can alphabetically sort composers, types, titles, and opus numbers. This rather simple data representation facilitates the connection to the rows of the more abstract feature matrix (see Figure 8.4) and projection view (see Figure 8.5).

Metadata

Opus	Title	Composer	Form
3	Serenade from String Quartet, ...	Franz Joseph Haydn	serenade
0	Arpeggione Sonata	Franz Schubert	sonata
9	Nocturne in E Flat Major Op. 9 ...	Frédéric François Chopin	nocturne
0	The Firebird (part 3)	Igor Stravinsky	no form
0	Minuet in G Major BWV Anhang 1 ...	Johann Sebastian Bach	minuet
0	Hungarian Dance no. 5	Johannes Brahms	no form
0	Für Elise Bagatelle No. 25 in ...	Ludwig van Beethoven	no form
439	Serenade No. 1 for three Clari ...	Wolfgang Amadeus Mozart	serenade

Fig. 8.6.: Analysts can use the metadata table to sort selected compositions and look up interesting pieces from the list.

Sheet View – Viewing sheet music using standard visual encoding is crucial for music analysts. Thus, the last component is the *sheet view,* which music analysts can use to estimate the quality of the underlying content (see Figure 8.7). Users can open the first page of compositions and relate it to the features via the context menu. Hence, the analyst can better understand abstract features by seeing the direct relation through the familiar notation. Analysts can also listen to the compositions on demand via the playback icons in the sheet view's top left corner of each item. We use the command-line version of MuseScore [240] to convert the underlying MusicXML files into their .mp3 version. These files are also cached to reduce loading times if multiple users want to play back the same compositions.

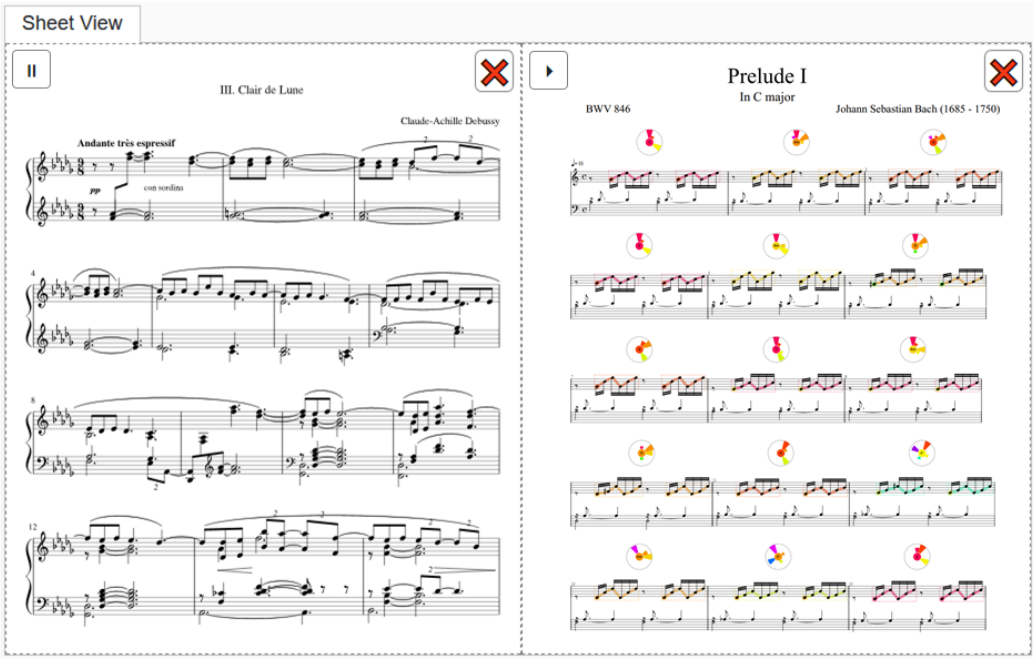


Fig. 8.7.: The sheet view allows viewing the first page of compositions. At the sheet level, analysts can perform harmonic and rhythmic analysis tasks using MusicVis [224].

8.4.2 Analysis Use Cases

CorpusVis offers multiple prepared analysis use cases consisting of preselected data and specific component configurations tailored to the analysis of the respective topic. Users can create additional scenarios by manually filtering compositions from the prepared dataset or uploading their datasets. In the remainder of this section, we will introduce three exemplary use cases. The first two represent a hypothetical user interacting with the system. At the same time, [UC3] has been provided by one of our expert users, showcasing how CorpusVis covers the introduced tasks.

Investigation of Epoch Characteristics [UC1] – In this use case, the analyst wants to explore the differences between the historical epochs [T2]. Thus, the analyst selects *Epoch Comparison* from the list of use cases. This use case considers 51 heterogeneous compositions from composers like *Pachelbel* (baroque), *Wagner* (romantic), and *Schönberg* (modern). Here, only a subset of eight features of the three feature categories (M, P, R) B such as *Repeated Notes* and *Pitch Variety* are selected that separate the four epochs within the projection view [T3]. While the modern and baroque compositions are distinguishable from all others, the pieces from the intermediate epochs, classic and romantic, have more overlap D. The analyst identifies composers between the romantic and modern epochs, such as *Stravinsky*, *Rachmaninoff*, *Schönberg*, and *Webern*, that reside within classical and romantic composer groups. Using the column of the feature *Size of Melodic Arcs* to sort the feature matrix, the analyst detects that romantic composers (yellow circles in the projection view) often used large melodic arcs compared to artists from other epochs.

Atonality versus Tonality [UC2] – An interesting aspect regarding the harmonic development between the different epochs is the discussion of tonality and atonality. Atonal music lacks a tonal center, meaning none of the twelve available pitch classes is favored over the others, as is typically the case for classical music. Moreover, atonal music comprises more dissonant intervals (minor second and tritone intervals) available in the pitch feature group B. Specifically, low values of the feature *Most Common Pitch Class Prevalence* are a good indicator for a composition to be atonal. In this example, the analyst is presented with 18 compositions from classical composers, including Mozart and Beethoven, who composed rather tonal music, and composers like Schönberg and Webern, well-known by musicologists for their atonal compositions A. After starting with the analysis, the analyst uses the projection view D to see which compositions are positioned closer together to detect similar pieces [T5] (see Figure 8.5). As described, the analyst knows tonality depends more on pitch and melodic features than rhythm. The configuration of the feature matrix B in this use case does not contain any rhythmic feature, which covers the mental model of the analyst. By viewing the columns in the feature matrix, the analyst detects as expected that the *Melodic Tritones* feature has much higher values for the atonal compositions, confirming the hypothesis that atonal compositions contain more dissonant intervals [T2]. By investigating the feature distribution chart of *Number of Common Pitches*, the analyst discovers that the compositions' subset of the configuration comprises more pieces with zero common pitches compared to the overall distribution in the corpus [T3]. Since the correlation matrix confirms a positive correlation of the *Number of Common Pitches* to *Most Common Pitch Class Prevalence*, the analyst is interested in seeing the single compositions' values. The

analyst is able to detect that the pieces by Schönberg and Webern try to avoid a tonal center, as none of the available pitches seems to be favored compared to the works by the classical composers [T5].

Composer Comparison [UC3] – In the first scenario, the domain expert compares similar pieces from two different composers [T3] and selects *"Erlkönig"* [T1] by Franz Liszt and Franz Schubert A. As a first step, the analyst manually inspects the first page using the sheet view E and detects the use of melodic octaves in the left hand of Liszt's interpretation, which appears to be a salient difference between the compositions. Based on this observation, the expert focuses on the features *Melodic Octaves* and *Average Melodic Interval*, which he assumes that these features should reflect this characteristic. The feature matrix B confirms the analyst's initial finding, as the values for the *Melodic Octaves* feature is 28.1% for Liszt, but only 2.05% for Schubert [T3]. Consequently, the analyst wants to explore whether this aspect is a typical difference between the two composers [T2] and selects all available pieces in the dataset from both composers and aggregates them based on the composer attribute [T4]. Although the distinction decreases, the trend remains (6.30% - Liszt vs. 3.8% - Schubert). While the two features selected by the user are correlated, the analyst discovers that they represent discriminating characteristics between Liszt and Schubert. As the last step, the analyst is eager to identify any unusual pieces (outliers) and ungroups the two clusters back into single compositions. By activating the clustering in the MDS projection view D, the analyst groups similar pieces into groups, finding many similarities on the feature level between Liszt and Schubert [T5]. For instance, while the compositions *"11. Écossaise D.781"* (Schubert) and *"Vive Henri IV"* (Liszt) are quite similar on the feature level. The user finds *"Ave Maria - Ellen's dritter Gesang"* by Schubert to be an interesting outlier compared to most of the other compositions. Eventually, the analyst uses the metadata table C to sort the matrix using the composer attribute to explore differences between the two collections [T2]. The analyst confirms the less surprising hypothesis that Liszt's works have a much higher *Pitch Variety*. Based on the feature *Number of Common Pitches*, he discovers that while Liszt almost always has zero pitches which individually account for at least 9% of all notes, Schubert even has compositions with five common pitches such as *"Nachtviolen (D.752)"* [T3].

8.5 Pair Analytics Study

We conducted pair analytics sessions with users of different expertise levels to evaluate CorpusVis introduced in section 8.4. Since the number of available qualified domain experts is limited, we opted for the pair analytics study method [8].

Preliminary Expert Feedback – Before the actual study, we consulted three domain experts to gather initial feedback in an informal pre-study. The experts appreciated interacting with an exploratory visualization. One of the experts liked how music information could be retrieved by simply entering keywords or mouse interactions on the composer timeline or types [T1]. The interlinked components helped them understand the connection between low-level features and the compositions [T2]. They found exploring the differences between composers and types useful by grouping [T3] or using the clustering in the projection view [T4]. They appreciated the

multiple components and the use of color to show connections of similar entities visually [T5]. Especially, they could use the feature filtering functionality to identify outliers within single feature dimensions. They also noted that the detailed description of the features in connection with the sheet view allows music students to experience new characteristics of sheet music [T2]. In their opinion, the sheet music component is a core element to verify hypotheses and investigate the quality of compositions. However, they also critiqued our color map and the lack of a feature to upload their own music sheets. Both of these issues were addressed in a refinement cycle before our study. In addition, based on our observations from the interactions in the pre-study, we added the features of interactive clustering and distribution-based grouping and filtering.

8.5.1 Study Design

We conducted the study to gather user feedback to identify the benefits and drawbacks of the introduced analysis workspace. For the evaluation, we used video conferencing software to show the interactive workspace to the participants during the study trials which enabled them to easily interact with CorpusVis ensuring equal study conditions. Based on the elicited *task characteristics* described in section 8.3, the participants carried out different analysis questions provided by the study director. The study results gave us useful insights into our work and its potential impact.

Study Participants – In sum, nine participants with different expertise levels partook in the study. Due to the required musical knowledge to use the analysis application, we focused our study on the target user group of domain experts, including musicologists and music teachers. Five participants had an experience level of at least ten years, including four university professors and a music teacher, which we further refer to as **s**ubject **m**atter **e**xperts ($\mathbf{SME_1}$–$\mathbf{SME_5}$). The remaining four participants were Ph.D. students ($\mathbf{ST_1}$–$\mathbf{ST_4}$) who are professionally involved with music on a daily basis but with less experience in music analysis compared to the SMEs.

Methodology – We applied a combined *observation* and *pair analytics* study [8] to investigate the usefulness of CorpusVis. In total, we performed nine study trials (average duration of 95 minutes) that were each led by a visual analytics expert (VAE) of our team. Each study session was structured into four sections.
The first part started with an *introduction* into the topic of sheet music collection analysis to frame the context of our work. Each participant completed a *demographic questionnaire* to provide information about their age, gender, level of education, and their level of expertise regarding music analysis. The first phase ended with an interview about the participants' *expectations* regarding an analysis system intended to support the investigation of sheet music collections.

After the general introduction, each participant received a detailed, textually guided *walkthrough* regarding all components of the study prototype (see Figure 8.8). Then, we introduced the five analysis tasks [T1]–[T5] from the analyzed requirements to each participant. Besides the analysis tasks, the introduction comprised an explanation of the relevant metadata features, including composer, type, epoch, and the low-level features extracted from the musical compositions visualized by the different components. In addition, the participants were informed about the

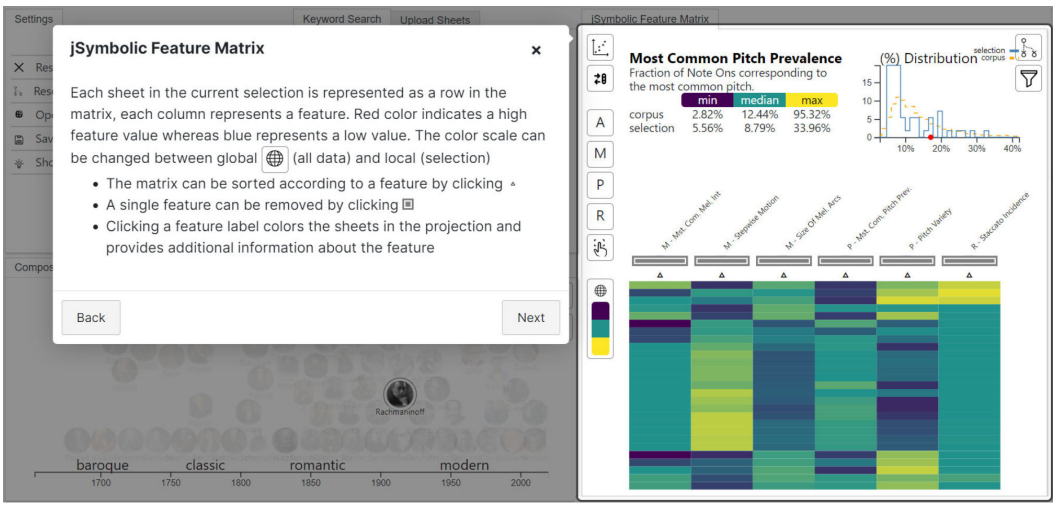


Fig. 8.8.: We used Intro.js [219] to prepare a visual explanation of all components of our prototype to ensure a consistent introduction of the interactive analysis workspace for all study participants.

origin of the used dataset (MuseScore [240]) and its content, addressing the overall amount of considered composers and musical pieces. Phase 2 closed with a simple and unguided exploration (similar to `[T2]`) of the visual interactive workspace, allowing the participants to get acquainted with the application. This introductory task contained a preselected set of ten different compositions and five features. As a quick entry point, we prepared a use case *feature explanation* that can be directly loaded from the use case selector in the interface that provides this preselection. We chose the pieces based on contrasting (extreme) feature values, such as low and high *Most Common Pitch Prevalence* values (3% and 65%). The participants were asked to explore the *Feature Matrix* (see Figure 8.4) to sort the selected sheets according to features of their interest. Then they used the *sheet view* (see Figure 8.7) to relate the feature values to the underlying common music notation.
During the main part of the study, we asked the participants to execute the analysis tasks, as described in the next section. We encouraged them to verbalize their thoughts and interactions (*think-aloud concept*). After completing the prescribed analysis tasks, we performed an interview to receive feedback about the participants' experience regarding the functionality and usability of the prototype. We closed the study trial with a final questionnaire containing nine questions answered via a five-point Likert scale to identify the benefits and drawbacks of CorpusVis.

Study Tasks – During the pair analytics session, all participants analyzed sheet music collections based on three use cases introduced in subsection 8.4.2. `[UC1]` comprised 51 compositions from various composers of all epochs. After getting acquainted with the metadata of the selected pieces, the participants selected all pieces from a single epoch of choice `[T1]` and identified at least two discriminating features `[T3]`. For this task, each participant used the feature matrix for identifying similarities between the compositions `[T5]`. In addition, the participants could add additional pieces through the keyword search to extend the selection `[T1]`. We encouraged each participant to freely explore the visualizations with individual compositions of interest `[T2]`.

A second task [UC2] addressed the comparison of *atonal* with *tonal* pieces (Tonal versus Atonal), which used contemporary and classical works. Similarly, the participants were asked to detect characteristic feature values for both groups by comparing the music sheets in the projection view and the feature matrix [T3]. The last task addressed comparing two or more composers relevant to the participant. For this, the *Composer Timeline* helped to filter the respective composers [T1], followed by a grouping of all pieces of single composers [T4]. Then, we asked the participant to visually examine the feature differences [T5] of the composers [T2].

8.5.2 Evaluation Results

We gathered feedback from the qualitative evaluation during the different study parts that we grouped thematically in the following.

Participants' Expectations – During the expectation interview, the participants stated various aspects of an application intended to support the musical analysis of sheet music corpora. $\mathbf{ST_1}$ remarked that he expected the application to group a musical repertoire into genres and composers automatically. $\mathbf{SME_1}$ expected the analysis system to use all information available in musical scores to perform musical analysis, which should present differences between musical styles, since "*classical is different from pop music*". $\mathbf{SME_2}$ mentioned that she expects an interface that allows her to provide information about her interests. Then, the "*system would automatically provide analysis results without the need for manual analysis*".

Exploratory Analysis – To familiarize themselves with the compositions of [UC1], the participants started by investigating the metadata table to see which composers are selected. Applying the color scheme of the composers' timeline to the projection view revealed the similarity of the compositions. $\mathbf{SME_1}$ mentioned the separate grouping of contemporary Jazz and Pop pieces compared to Baroque and Classical compositions, which also were separate groups but had some overlapping. Using the lasso to select the yellow-colored (modern) pieces, she identified that compared to the rest, these pieces have more note repetitions, a high proportion of consonant melodic intervals, and fewer minor/major second intervals. $\mathbf{ST_4}$ checked the quality of a sheet file based on features such as pitch variety. Typically, complex compositions such as études and symphonies have a high-pitch variety. Thus, participants could delineate custom arrangements from more original compositions.
While analyzing the configuration of [UC2], the participants were presented with tonal compositions from composers such as Mozart and Beethoven, which were compared with atonal pieces from Schönberg and Webern (see Figure 8.5). The participants used the projection view to compare similarities and appreciated the sorting functionality in the feature matrix, which helped them confirm that atonal pieces use tritone intervals more often. They identified that the feature *Number of Common Pitches* for the atonal compositions is always zero, which reflects that composer's objective is not to favor certain pitch classes. Especially $\mathbf{SME_2}$ appreciated the sheet view, which allowed her to compare the values from the feature matrix to the common music notation. The third use case started with aggregating all compositions from the composers displayed in the composer timeline. The participants found that *Changes Of Meter* is the most discriminating feature, revealing that contemporary music seldom changes the meter signature within a composition.

Comparative Analysis – The participants used the feature matrix to identify similarities between single compositions or composer groups. $\mathbf{SME_4}$ and $\mathbf{SME_5}$ found the melodic features useful to investigate the melodic content. In addition, $\mathbf{SME_3}$ and $\mathbf{ST_1}$ appreciated the features regarding the intervals, which they considered to be particularly helpful in identifying dominant melodic aspects. $\mathbf{SME_1}$ mentioned that the matrix helped confirm that contemporary music such as "Imagine Dragons and Coldplay is more consonant and has more note repetitions but fewer stepwise motions" than Schönberg (Op. 11) or Webern (Op. 4). Using CorpusVis, $\mathbf{SME_1}$ learned that minuets from Bach have much lower *Pitch Class Variety* than his fugues. $\mathbf{ST_1}$ noted that based on the available pieces, Schönberg's work has more chromatic motion (25%) than Webern's (20%). $\mathbf{ST_2}$ and $\mathbf{ST_3}$ emphasized that the pitch range and density information helped them investigate the pitch material to better understand the complexity without even considering the sheet view. The participants considered the detailed descriptions at the top of the feature matrix, combined with the distribution of the whole dataset and the current selection in the respective context, to support the exploration. A few participants mentioned the missing linking and brushing between the composition forms and the composers in the timeline on hover events, which would be easily possible to add in the future. To compare chorals and fugues, $\mathbf{SME_5}$ focused his analysis on the rhythmic features. $\mathbf{ST_3}$ could identify that Beethoven uses 20 fewer pitches compared to Liszt, who, as he stated, uses more chromatic melodic progressions. During the investigation, the participants found different matrix features to be relevant. They could easily understand the separation into the three categories of rhythm, pitch, and melody, which were of different importance depending on the task. Depending on the task at hand, different features were deemed important. For instance, the time signature is a salient feature that helps to determine certain composition types. The pitch variety helped them detect complex compositions and differentiate tonal from atonal compositions. The chromatic motion feature supported identifying twelve-tone pieces.

Closing Questionnaire – All participants completed a final questionnaire, answering nine questions to assess the usability. The questions and results based on a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree" are displayed in Figure 8.9. The questionnaire results reveal that our participants gained new insights (Q2) about sheet music collections by enabling them to perform groupings of the dataset to carry out comparison tasks based on epochs, composers, and composition types (Q7). Most participants agreed that CorpusVis is suitable for visually analyzing sheet music corpora (Q1, Q9). Q6 received the lowest scores, which illustrates that the computationally extracted features do not sufficiently address all users' aspects, such as dynamics, instrumentation, or harmonic progressions. Nevertheless, the scores of Q5 show that the computed statistical features still enable the comparison, even though some aspects the analysts would be interested in are missing. All but one of the participants stated that they would recommend the interactive workspace to a colleague or use it for their research (Q4). For most participants, the currently used dataset in CorpusVis contained the compositions they wanted to add to their analysis (Q8). Overall, the questionnaire results convey that the participants found potential for analyzing sheet music collections. They would continue using it in the future, especially if it is improved based on this study's outcomes.

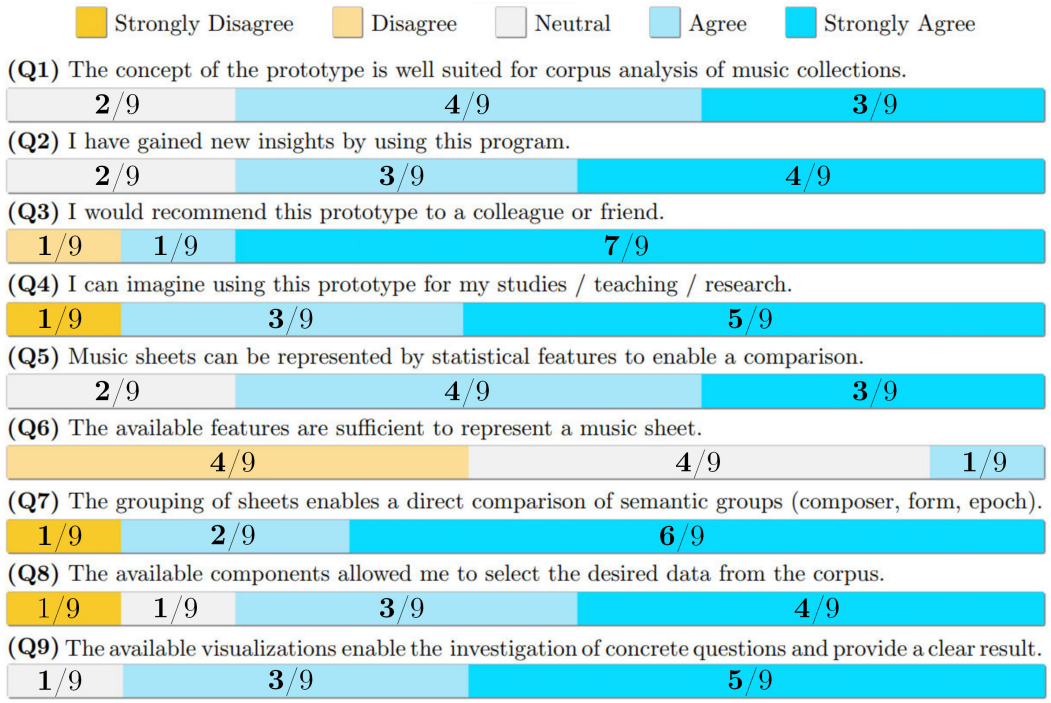


Fig. 8.9.: The final questionnaire comprised nine questions to assess the usability and functionality of CorpusVis for analyzing sheet music collections. Most questions achieved good ratings, emphasizing an overall high satisfaction level.

8.6 Discussion

We identified benefits and drawbacks from the study which showed that CorpusVis supports specific analysis tasks such as comparing composition types or even the entire work of composers. Many participants questioned the quality of the underlying MuseScore dataset, which contains volunteered sheet music. To address this issue, CorpusVis enables the upload of sheet music collections, allowing musicologists to use their curated datasets if desired. This way, we mitigate the negative influence of datasets that suffer from poor quality. Despite the underlying dataset's limited quality, CorpusVis demonstrates the positive impact that tailored visualizations have on the work of musicologists.

Musicologists controversially discuss whether and where to place clear borders between different epochs of music history. In our approach, we simplify the timeline of musical history into four major epochs *baroque*, *classic*, *romantic*, and *modern*. While we assigned individual colors to each epoch, study participants appreciated that we introduced smooth color gradients to emphasize the transition period.

During the different musical periods, the focus on specific characteristics of music has shifted. While CorpusVis already includes several features, other music styles may require extending the application. For instance, starting from the second half of the 20th century, new musical frontiers emerged, including serial, electronic, minimal, spectral, and experimental music. These contemporary music styles may sometimes be unavailable through transcribed common Western music notation. They often

rely on auditory or physical features such as timbre or spectral data. Therefore, such music compositions may only be available through audio files, not symbolic music. CORPUSVIS would either require an audio extension or a suitable symbolic notation that can be used to extract features that represent the underlying information.

We encountered several challenges during the implementation and design of the presented interactive workspace. A typical issue in this field is *copyright*. MuseScore users sometimes upload compositions from contemporary musicians or artists who are still alive so that these pieces may be deleted from MuseScore again. To avoid legal issues caused by sheet music that should not be visible in our system, we only include compositions from artists who died more than 70 years ago that are still available on MuseScore. Whenever sheets are deleted from MuseScore for any reason, as a consequence, they are no longer available in our system.

During the development of this system, we were closely collaborating with domain experts from musicology, allowing us to gather early feedback on CORPUSVIS. We quickly identified challenges in this interdisciplinary collaboration, especially in creating visualizations that were understandable by experts while allowing them to focus on aspects relevant to their intended analysis. In particular, we found that the field of music analysis is extremely broad, necessitating a flexible system with a myriad of different options. However, adding additional functionality also complicated the system's design, in turn making it more difficult for experts to understand and use. As a result, in this work, we focus on specific tasks and data characteristics to limit the scope.

8.6.1 Limitations

While the study results show that CORPUSVIS supports different music analysis tasks, it also surfaced some limitations. For instance, some composers of the 20th or 21st century heavily use timbral or textural elements. These features are often not encoded using the common Western music notation at all. As a result, CORPUSVIS cannot be of help in analyzing work from such contemporary composers. Future extensions of CORPUSVIS could integrate auditory features derived from recordings rather than only music sheets to capture the expression with which musicians played the compositions to close this gap. In its current implementation, CORPUSVIS does not capture the instrumentation of a composition. Several participants ($\mathbf{SME_2}$–$\mathbf{SME_5}$) stated that they would like to filter all compositions of a composer that contain piano or stringed instrument parts. Especially, identifying how the pitch range is distributed between different instruments would provide useful semantic information about the role of instruments and overall timbre of a piece.

While CORPUSVIS contains works from several composers, the visualization only partly scale based on the available aggregations. The projection view and feature matrix scale towards adding additional compositions from the same composer using aggregation based on the composers, as it uses equal visual space. Yet, an analyst may need to focus the analysis on a subset of musical features if other characteristics should be considered as the cognitive complexity increases with the number of features to be analyzed. Nevertheless, CORPUSVIS already uses semantic groups for melodic, rhythmic, and pitch features to enable users to focus on a subset. To improve this, CORPUSVIS could be adapted to allow users to create custom feature

groups to focus the analysis and additional features that could be easily added if required. The application's performance depends on the library music21, which loads every composition once at the beginning to prepare all features for visualization.

$\mathbf{SME_2}$ mentioned that classifying the pieces into their types may be ambiguous and that enabling analysts to re-classify compositions could increase the analyst's influence on a dataset's quality, besides the option of uploading custom datasets. $\mathbf{ST_1}$ and $\mathbf{SME_1}$ remarked that the vertical combination of notes (simultaneously played notes are referred to as harmony) is not covered by the available features, since the employed jSymbolic features are limited to pitch, melodic, and rhythm characteristics. Although CorpusVis already offers many selection and filtering mechanisms, some participants suggested additional options for filtering a dataset by key signatures, voicing, and lyrics in a future version of the application.

8.6.2 Take-Home Messages

From the work in this chapter, we distilled three major take-home messages.

(1) Corpus-level music analysis is challenging due to the magnitude and complexity of underlying features. Manual analysis of sheet music collections is a tedious process requiring profound domain knowledge. CorpusVis supports the analysis of extensive sheet music collections, which is not feasible when done manually. While abstract visualizations open up new perspectives, they should not be considered a replacement for traditional analysis methods, but rather an extension. Thus, integrating domain experts during the design phase and the requirements analysis is crucial to address their specific needs. This enables users to combine close- (sheet view) and distant reading (projection, features) and get a corpus overview.

(2) Visualization dashboards can enable multiple perspectives on the data, supporting scalable and multi-faceted analysis. Especially when musicologists want to explore, understand, and interpret music sheet corpora, visual analytics helps to address the individual needs of music analysts. Automatic computational methods do not provide visualizations to users facilitating interactive engagement with a sheet music collection to understand its underlying characteristics better. We allow various interconnected analysis workflows by tailoring the visual design of the provided components to their targeted task. The availability of multiple components covering different characteristics of the dataset enables users to set their analysis focuses on those attributes they are interested in their analysis.

(3) Tailored interactions enable a task-driven analysis of corpora using diverse metadata and statistical features. Interactive visualization allows investigating sheet music from different perspectives, supporting various aspects music analysts may be interested in. In particular, for analyzing complex questions involving multiple meta-data and features, we rely on the interactive interlinking of visual components, enabling an expressive yet visually uncomplicated analytics system. Enabling analysts to import their own curated datasets facilitates their investigation and generation and confirmation of hypotheses.

8.6.3 Future Work and Research Opportunities

We postulate that our work only represents a starting point for future cooperation between visualization and musicology researchers. We want to emphasize that it is worth "risking the drift" for such interdisciplinary collaborations [131]. So far, the classification of the compositions' metadata available in CORPUSVIS is fixed and cannot be adjusted by the analyst. In the future, enabling users to change meta information may be worthwhile to increase the dataset quality. Extending the application through labeling would further improve the analysts' ability to capture and externalize analysis results. While we have considered the temporal aspects of music, we have neglected the geographic information of where composers have lived and worked, which has already been studied by Khulusi et al. [162]. It would be interesting to investigate how the location influenced the style of specific composers. While the low-level characteristics are just statistical measures, it would be interesting to see how they correlate with the mood and emotions of listeners. So far, we have not considered the dynamics, instrumentation, or timbre of the instruments that the compositions are written for, which would be worthwhile to add support for in the future. If pianists or guitarists are looking for new pieces to play, it would be helpful to provide specific instrument filters that help them find relevant information. Moreover, as the dataset's quality plays a primary role in analysis, it would be helpful to provide classification means for analysts to identify low-quality pieces to exclude unwanted items. Eventually, CORPUSVIS focuses the analysis on a high level. In combination with MusicVis that enables analysts to investigate single compositions at the sheet level [224], which we discussed in the previous chapter, we can now employ CORPUSVIS to allow analysts to filter relevant compositions at a higher level, which can then be further analyzed at the sheet level using MusicVis.

8.7 Summary

In this chapter, we have introduced CORPUSVIS, a web-based interactive workspace for the visual analysis of sheet music collections. We discussed data and task characteristics relevant to analyzing sheet music corpora. Multiple connected components through linking and brushing support exploration, detection, and comparison tasks. Based on initial expert feedback, we identified the benefits and challenges of our approach. In addition to three exemplary use cases that demonstrate the analysis workflow to compare composers and types, we conducted a pair of analytics studies with nine participants. The qualitative study results show that music analysts benefit from interactive visualization addressing different use cases scenarios, such as comparing composers, epochs, and composition types based on low-level features, including pitch, rhythm, and melody.

9 Conclusion and Future Perspectives

> *My heart, O God, is steadfast, my heart is steadfast; I will sing and make music.*
>
> **— Psalm 57:7 by David**
> *The Old Testament, The Bible*

To conclude this doctoral thesis, in this chapter, we will have a retrospective review of implemented research prototypes, qualitative studies with users from different levels of musical expertise, and proposed visualization concepts. This review aims to put the discussed topics into a broader context, revealing auspicious research directions for the coming decades at the intersection of visual analytics and musicology.
This thesis addressed a detailed consideration of sheet music, primarily based on the digital format of MusicXML. The presented approaches introduced novel designs for the most central musical features, including harmony, rhythm, and melody. We employed foundational musical concepts for the proposed harmonic and rhythmic fingerprints to facilitate a natural connection to well-known domain-related aspects. We discussed these novel visualization concepts and demonstrated their effectiveness through qualitative user studies. While each chapter is self-contained, several cross-references emphasize how the presented approaches depend on and relate to each other. Especially the bottom-up (Chapter 7) and top-down (Chapter 8) approaches that integrate the foundational aspects and contributions of Part I and Part II reveal how visual analytics helps music analysts to explore, compare, and understand single and multiple compositions in collections. Consequently, the resulting interactive visualization techniques help present musicological data legibly and scalable, enabling comparison, exploration, and identification tasks through suitable visual abstraction methods based on underlying sheet music data.

9.1 Summary of Research Contributions

In summary, the presented thesis comprises several contributions to both musicology and information visualization research. Each chapter contributes distinctively to the advancement of information visualization, particularly in the research field of musicology. Chapter 2 introduces the pioneering concept of Visual Musicology and the Methodology Transfer Model. This chapter bridges the gap between musicology and visualization analytics and highlights the immense potential for interdisciplinary research, thereby setting a new precedent in the field [229]. It delves deeply into the diverse applications within musicology and paves the way for collaborative solutions between visualization experts and music domain professionals. We used the VMG to categorize and embed the focus and contribution of each component of this

thesis. Figure 9.1 summarizes the coverage of the VMG by emphasizing which topics this dissertation comprises through checkmarks juxtapositionally placed beside the relevant keywords. The keywords used for each of the four major branches are stated at the beginning of each chapter to highlight the VMG categorization. This final overview indicates which aspects have not been considered in this thesis.

Of course, this is only a simplification even for the highlighted leaves, as different combinations of tasks and data characteristics are open research gaps since it is impossible to explore any possible composition of tasks and data aspects in this work.

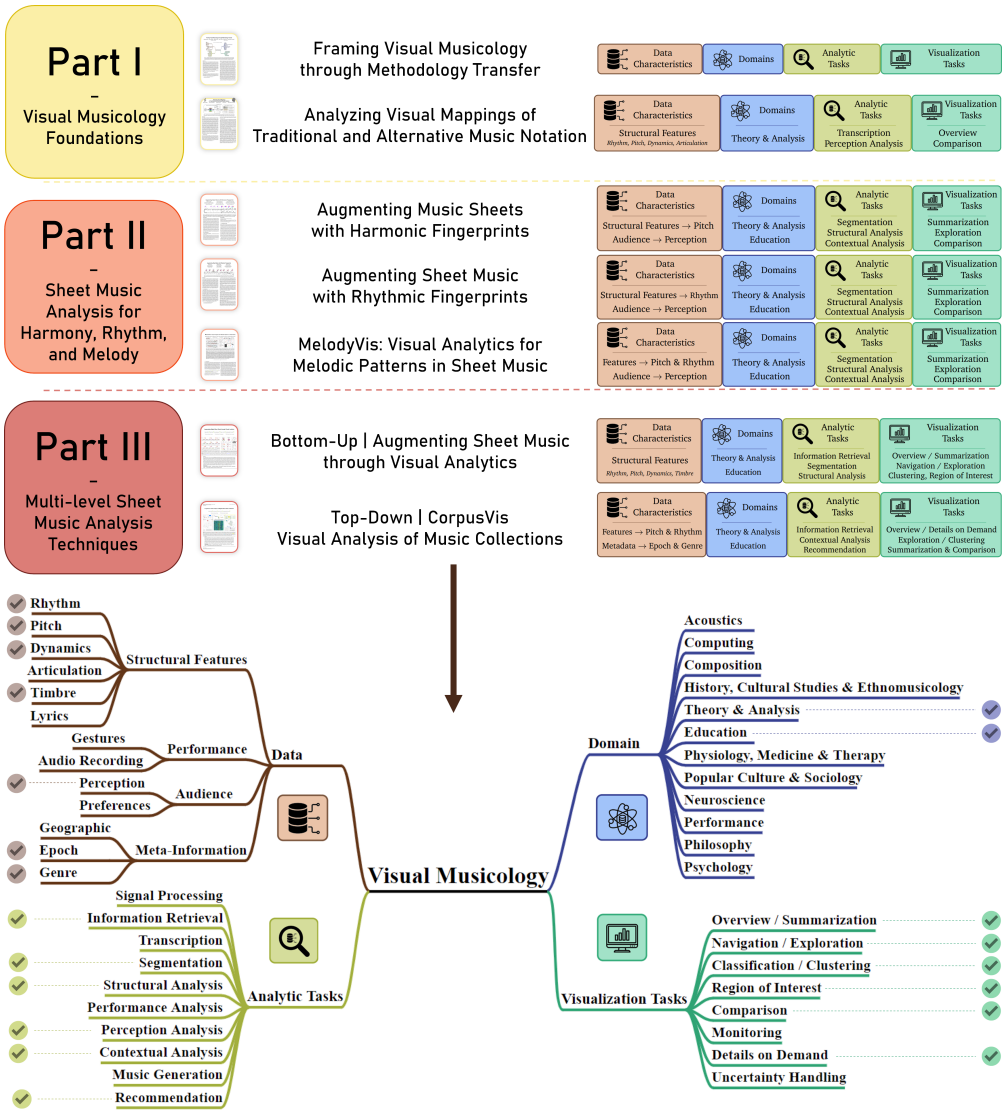


Fig. 9.1.: This thesis' chapters are segmented into three parts: foundational theory, detailed sheet music analysis, and advanced multi-level techniques. At the top, we can see the overview of the VMG topics for each chapter and part. By summarizing all chapters based on the four major VMG branches, the checkmarks placed left and right to the branches show which Visual Musicology aspects we considered in this thesis. This overview also helps identifying future research opportunities, especially the branch *Domains*, based on the vast possible combinations of the four categories: *Data Characteristics*, *Domains*, *Analytic Tasks*, and *Visualization Tasks*.

Chapter 3 takes a significant step forward by **exploring existing visual encodings in music notation**. It challenges and extends the boundaries of traditional visualization techniques, **introducing a pipeline that aids in understanding visual variable mappings and developing alternative notations** [225]. This chapter discusses visual encodings based on Gestalt Laws and Bertin's visual variables, offering innovative perspectives and identifying open research opportunities for novel music notation designs. Chapter 4 and Chapter 5 respectively introduce **harmonic and rhythmic fingerprint glyphs** [222, 103]. These chapters make necessary efforts to **enhance the interpretability of sheet music** through intuitive and familiar design concepts. The introduction of these glyphs serves as an example of new visualization techniques in musicology, making music theory more accessible and understandable to a broader audience, including novices and experts alike. Chapter 6 presents MELODYVIS, an example of interactive visual analytics applied to analyzing melodic patterns. With its multiple connected components and **atomic operators**, this application enables detailed and **interactive exploration of melodic patterns**, aiding music analysts. The interactive workflow of MELODYVIS and its support for pattern identification and interpretation emphasizes the necessity of user-driven analysis in music visualization [223]. In Chapter 7, MUSICVIS emerges as a novel method that **seamlessly integrates CMN with glyph visualizations** grounding on Chapter 4 and Chapter 5 [224]. This visual analytics technique exemplifies the potential of **combining traditional and abstract visualization methods** to enhance music analysis. The chapter's **in-depth pair analytics study** further shows the approach's intuitiveness and effectiveness, showcasing how MUSICVIS can lead to new insights and increased engagement in music analysis. Eventually, Chapter 8 introduces CORPUSVIS, pioneering a novel approach to **corpus-level music analysis**. This approach empowers analysts in **exploring, investigating, and comparing sheet music collections** through interactive visualization [227]. The chapter not only assesses the applicability of this approach but also highlights the opportunities it opens for interdisciplinary collaboration at the intersection of visualization and musicology.

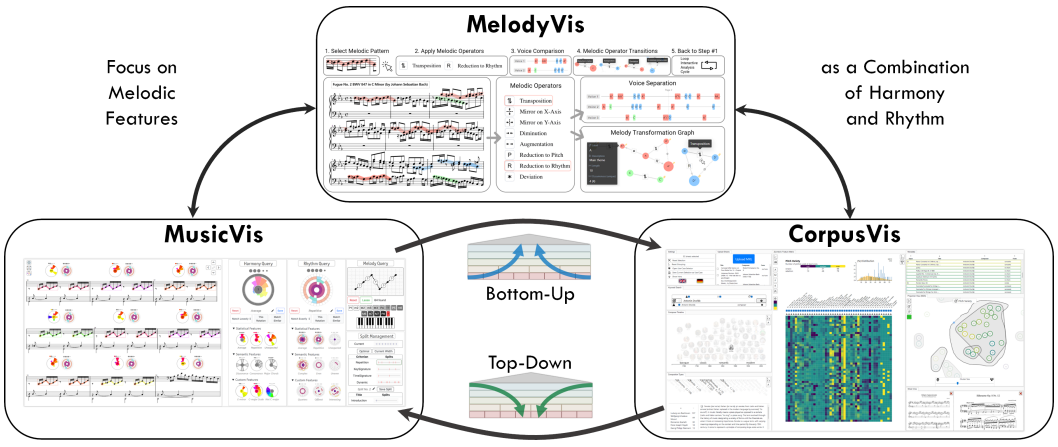


Fig. 9.2.: MUSICVIS, CORPUSVIS, and MELODYVIS build a team of three major implementations, a core contribution of this thesis. The figure illustrates the bottom-up and top-down transitions, while MELODYVIS focuses on the visual analysis of melodic relationships on a lower level, combining rhythmic with harmonic aspects.

Figure 9.2 illustrates the transition between the three applications discussed and evaluated in this thesis, illustrating the top-down and bottom-up connection between MUSICVIS and CORPUSVIS. We introduced MELODYVIS as a separate application to specifically focus on combining harmony and rhythm through using melodic operators to visually explore melodic relationships in sheet music.
Collectively, these chapters enrich the field of *Visual Sheet Music Analytics* by integrating advanced visualization techniques with musicological insights. They contribute substantially to the information visualization and musicology communities, pushing the boundaries of how sheet music is analyzed, understood, and visualized.

9.2 Lessons Learned and Experiences

We learned several lessons and encountered challenges based on the different approaches and techniques we presented in the three parts. Subsequently, we will report best practices and special guidelines on conducting research at the intersection of information visualization and musicology. Still, we argue that many also apply to digital humanities and interdisciplinary research collaborations in general.

Leverage domain-specific visual metaphors to augment instead of replace.
Research fields that have existed for centuries developed their own terminologies, concepts, methods, taxonomies, and ontologies over a long period of time. Therefore, such fields have historically grown, and while the rise of digital technologies in recent decades can drastically change the methods and techniques, it could be profitable for domain experts to leverage innovative developments, especially from visual analytics. A major challenge is that such effective visualizations have a steep learning curve which could lead to domain users unwilling to employ such techniques. A practical example from this thesis is leveraging visual metaphors such as the harmonic and rhythmic fingerprints and the melodic sheet music overlays. Instead of replacing the familiar domain-specific "visualization" concepts, we found augmenting an advisable initial step to lower the barrier for domain users to approach abstract visualization while maintaining original representations. After setting up the connection and developing an understanding of new visualization designs, it becomes easier for musicologists to transition to more abstract analysis tasks. Otherwise, they would probably not use them in practice. Combining this seamless transition with faithfully designing visualization concepts based on domain-related concepts such as the *circle of fifths* or *rhythm tree* is a second guideline that exploits standard music-theoretical concepts. This has two-fold advantages: First, domain experts familiar with the musical concepts that inspired the visualization methods' design rationale are faster in understanding the representation than more abstract visualizations which could be applied to visualization research, such as pixel or matrix visualizations. Second, musicology novices unfamiliar with such concepts can employ techniques to draw the connection between CMN and music-theoretical concepts they may need to know anyway, facilitating educational tasks by improving engagement with music theory concepts. Eventually, from our experience combining domain-inspired visual metaphors following the augmentation strategy instead of merely replacing them is a recommendable guideline for more abstract visual representations, especially for *slow analytics* [34] tasks often conducted in digital humanities research.

Prioritize scalability in visualization and data wrangling.
When starting interdisciplinary research projects, we had to make several decisions regarding implementing techniques while identifying suitable data sources for preparing use cases that demonstrate the value of novel visualization methods. Thanks to the sheet music data source provided by MuseScore [243], such as the MusicXML files carefully curated by ClassicMan, we were able to develop the sheet music analysis applications. Instead of implementing our own CMN library for displaying sheet music, we employed the *OpenSheetMusicDisplay Library (OSMD)* [109]. We decided to use *SVG* instead of *Canvas Graphics* [338] because of its support within the d3.js library that highly simplifies the process of implementing interactive visualizations with seamless animations [31]. While OSMD is a powerful library, the SVG representation was not intended to support scalability since typical use cases involve displaying single or few pages of sheet music. Therefore, our initial decision to opt for SVG led to scalability issues making it difficult for us to display larger compositions, especially for orchestra or long piano sonatas. Hence, we recommend taking the extra effort to start with HTML Canvas Graphics, in contrast to SVG, might facilitate the later analysis process in dealing with large datasets without needing to prepare a separate application such as we did with CorpusVis. Integrating these strategies would especially benefit direct comparison and analysis tasks across compositions, which are crucial for both musicologists and for music students and teachers.

Balance high-level abstractions with low-level details.
While we started our work by designing the harmonic and rhythmic fingerprint visualization designs in Part II, we had already planned to continue employing our proposed visualization techniques in the early stages. This enabled visual analytics tasks on a higher, more abstract level as presented in Part III. The prepared seamless transitions enabled our domain experts to understand the abstract concepts and develop hypotheses about the underlying based on, e.g., the projection layer. The annotations they could assign on each layer allowed them to trace and balance the multi-staged analysis tasks back to CMN, but also in the opposite direction. Therefore, we learned that carefully blending different design concepts to find a balance between them can be crucial in providing suitable analytics methods.

Pursue mutual benefits through interdisciplinary research collaborations.
Domain-specific issues such as sheet music exploration, comparison, and analysis tasks discussed in these chapters hold the potential for valuable advancements in two initially independent research fields. While the humanities and other fields deal with complex issues such as social network dynamics, they may benefit from innovative or even existing visualization techniques that have not yet been applied to these problems. While we do not argue that every task should replaced with visualization, we postulate that any field could benefit from cooperating with visualization researchers. On the other hand, domain-related issues could provide new analytical challenges for information visualization research that may require specific tailoring for suitable solutions, which is why we introduced the *Methodology Transfer Model* in Chapter 2 (see Figure 2.2). This methodological approach will foster interdisciplinary research, especially within the *Visual Musicology*. We consider that our work introduced in this thesis only touches the surface of possible research endeavors, being a starting position for abundant future interdisciplinary research projects at the interface of musicology and information visualization and even other research fields.

Use ground truth data for quantitative evaluations.
A major lesson I learned during our research projects was that quantitative analysis depends on reliable ground truth data. As previously mentioned, we used sheet music data from MuseScore, containing mostly volunteered-provided data of very heterogeneous quality [243]. To automatically use this information on a larger scale, it becomes extremely difficult to make reliable hypotheses. Moreover, the content of this thesis deliberately focuses on the qualitative side of evaluation, considering users of different expertise levels to identify the usefulness of the proposed methods. Eventually, there are two sides to every coin. This means that conducting both quantitative and qualitative research is viable and desired. Therefore, we leave it to future *Visual Musicology* research projects to apply visualization metrics specifically designed for musicological tasks. Many publications originating from the *International Society for Music Information Retrieval (ISMIR)*[1] are focusing on the quantitative evaluation of music data, especially based on music in audio format[2]. Various computational challenges still need much better performance which will also have a positive influence on the work we conducted. For example, better *Optical Music Recognition (OMR)* [304] methods will lead to larger high-quality datasets, which could then again be imported to our system or similar applications depending on sheet music using symbolic formats such as MusicXML [111]. Working on an integrated pipeline that provides a service to extract exact CMN from given audio data is an open challenge. Solving this issue holds the potential to be groundbreaking for sheet music visualization research, as audio data has achieved significantly more research attention by companies such as Deezer and Spotify [12]. We assume that improving these OMR systems will increase the number of services such as OKTAV to facilitate instrument learning for musicians [169]. Also, such pipelines will increase the performance of quantitative methods with the rise of larger, *high-quality* sheet music datasets.

Optimize granularity for detailed analysis.
Composers who use CMN to notate their compositions determine which notes are arranged in each measure. Based on this, we decided to take single measures as the lowest level for aggregation of musical content in the rhythmic and harmonic fingerprints. Also, we required the harmonic and rhythmic fingerprints to be interchangeable as described in Chapter 7. Breaking down the harmonic fingerprints on the quarter note level would be required to balance the granularity further and improve detailed analysis. Yet, this would conflict with the rhythm tree design of the rhythmic fingerprints, which depend on the radial structure of the meter signatures such as $\frac{4}{4}$. While our decision was required to manage a similar level of abstraction between the two features, we could have mentioned it in Chapter 4, which is a limitation of our work. Our approach underpins the critical nature of how design decisions influence what tasks are supported. For example, in baroque music, chords often changed on the more fine-grained, quarter note level [211]. Tailoring toward this level would help harmony analysis tasks for such types of music. In contrast, modern pop music is often more simple and doesn't have chord changes on such a low level, such as *Perfect* by Ed Sheeran[3]. In such scenarios, our design rationale would suffice and cover the requirements for harmonic analysis of these pieces.

[1] https://ismir.net/
[2] https://github.com/ismir
[3] `https://musescore.com/user/8203336/scores/4634731`

Explore music notation beyond Western sheet music.
Another lesson of our approaches is that conducting visualization design research that generalizes for music notations originating from different cultures and music styles is challenging. While our proposed methods are biased toward Western music encoding, we want to emphasize that designing music visualization to bridge notations of different styles and cultures is crucial to understanding the connection and comparing varying notations. The pentatonic relationship between notes, which originates from the frequency relationship of notes on a physical level, is a common denominator between music styles of different cultures. For example, Killick proposed a cross-cultural more global notation to bridge this gap by informing his design in keeping the pitch information on the y-axis and the temporal dimension on the x-axis [165]. Using his design to extend with scalable visualization might improve a more integrated analysis of sheet music analysis using varying notations. For example, he encodes musical instruments using the visual channel *shape*. Eventually, while minimizing bias in such scenarios could improve understanding of the connection between different cultures, it will still be difficult to convince CMN users to switch after overcoming the steep learning curve. Because of this, CMN could never be completely replaced by a newer design but was rather adapted to meet different needs, such as the Braille music notation for blind people [87] or instrument-specific notation such as fretboard visualizations for guitarists [129].

On the shoulders of giants.
At some point, we have to emphasize that the work presented in this thesis would not have been possible without the preceding work of other parties. As with most other research projects, ours also depends on existing research publications and previous technical advancements. We want to highlight these foundations by providing the deserved references in this paragraph. First, for all frontend visual implementations, such as the harmonic and rhythmic fingerprints and the CMN overlay visualizations, we employed the JavaScript library d3.js by Bostock et al. [31]. To extract musical jSymbolic features, we applied the Python library `music21` [67], which can be directly applied to symbolic sheet music based on the MusicXML format [112]. Other data sources use other symbolic formats, such as Humdrums `**kern` format by Huron [141], Lilypond by Nienhuys [247], or MEI by Roland [280]. There are even more formats; working on a solution that bridges all these formats would lead to larger datasets of consistent formatting. Originally, the jSymbolic features were proposed by McKay in his master thesis [212]. Together with Fujinaga, he distilled 160 musical features using the Java programming language [215], which were then adapted by `music21`. Although `music21` only implemented a subset of these features, it facilitates the reading and iterating through the music-specific elements (e.g., notes, measures, instruments), which helped us to focus on the visualization concepts. Eventually, the whole research work presented in this thesis depends on CMN, which has been developed and influenced by a myriad of people. Before that, Neumes emerged from European monasteries as a new kind of notation for the need to notate Gregorian chorals [318]. Two of the earliest people who formalized and influenced CMN are Guido d'Arezzo [343] and the monk Hucbald of Saint-Amand [88]. In sum, we emphasize that building on top of the work of others should be considered a privilege. It also means that searching for the potential of work that others have published will help to continue research projects that may become relevant later.

9.3 Future Directions and Outlook

This thesis presented and evaluated several approaches to augment digital sheet music with advanced visual design concepts to support different tasks, including analysis, exploration, understanding, and externalization. We want to highlight specific next steps where we see great potential for further research. With these considerations, we aim to stimulate future research projects derived from our research, addressing the intersection of Visual Musicology with diverse topics.

Integration of Big Data in Musicology – The research work pursued in this thesis barely touches on the immense possibilities of utilizing extensive digital sheet music collections for Visual Musicology [1]. Dominating platforms such as MuseScore are pivotal, offering a wealth of musical compositions [243]. Unfortunately, the reliance on community-contributed content leads to varying quality, posing a challenge for consistent, reliable research data. In addition, access limitations, while necessary to prevent misuse, further restrict research potential. These barriers narrowed the scope of our research and highlighted a broader issue in the field. Collaborations between corporate entities and researchers in musicology, focusing on information visualization, are essential to address these challenges. Using established standardized formats for sheet music data, such as MusicXML, across platforms would significantly streamline research processes and data compatibility. Advocating for open access and flexible licensing models is crucial to balance the rights of creators with the expansive potential of academic research. Creating a dedicated collaborative research community within Visual Musicology can encourage invaluable knowledge sharing and joint ventures between academia and industry. Tackling these issues would expand the research scope and enhance the interdisciplinary essence of Visual Musicology. Moving towards integrating big data more effectively in musicology provides ample chances for research and discovery, changing how we approach this evolving discipline.

Advancements in AI and Machine Learning for Musicology – The improvements of advanced artificial intelligence (AI) and machine learning methods herald a transformative era in music analysis [28]. These methods promise to automate and refine the understanding of complex musical structures and patterns, going beyond the capabilities of traditional analysis. By harnessing AI, researchers and musicians will probably uncover nuanced insights into composition techniques, performance styles, and musical evolution. This streamlines the analytical process for discovering new connections within music genres, styles, and historical periods. Integrating AI in music analysis and generation tasks will redefine our comprehension and appreciation of music, blending precision and creativity that was previously impossible. A model example of AI advancement in music generation is Google's Magenta project[113], which employs machine learning to produce novel audio compositions, showcasing AI's growing capability in (semi-)automatically creating musical pieces [135]. This approach augments the creative process and pushes the boundaries of how we perceive and interact with music. Similarly, generating MusicXML with AI would mark a significant leap on the composition side. AI-driven MusicXML generation aids composers by automating aspects of notation, enabling more efficient and flexible composition workflows [78]. These developments shed light on a future where AI enhances our music understanding and actively contributes to the creative process, combining traditional musicology with state-of-the-art technological progress.

Extending Visual Musicology through Virtual, Augmented, and Mixed Reality The intersection of Visual Musicology with Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) could enhance music education, collaboration, and cultural engagement. VR opens new dimensions in interactive music theory education, allowing immersive environments where learners can explore music theory concepts. This technology could also enable virtual composition workshops, providing a novel platform for experimenting with music creation in a simulated real-world environment. Moreover, AR could further transform instrument learning by overlaying interactive instructional elements onto physical instruments. This could be combined with sheet music, where AR-enhanced music scores come alive with interactive elements. This approach would improve comprehension and engagement for learners of diverse skill levels by directly connecting their instrument with the musical score through catered visualization techniques. For music collaboration, VR offers a unique platform where musicians and students from any location can learn, practice, and perform together in virtual spaces. Concerning health applications, this technology would also extend its benefits to therapeutic applications, such as experimental VR-based music therapy programs contributing to cognitive development and stress relief. Mixed Reality (MR) in musicology allows interactive exhibits and cultural experiences. MR can be utilized to craft immersive displays that reveal music's evolution and cultural heritage, offering interactive content that can be widely disseminated to educational and cultural institutions. The ubiquity of smartphones further enhances the potential of these advanced musicological tools. With their widespread availability and advanced capabilities, smartphones can serve as a gateway to these immersive experiences, making them accessible to a broader audience and democratizing the music experience. By leveraging smartphones, the benefits of VR, AR, and MR in musicology can be extended beyond specialized settings, reaching people everywhere and enriching their engagement with music in its many forms.

Visual Tools and Techniques for the Future of Music Education – My doctoral thesis lays the groundwork for advanced visual techniques to enhance sheet music's understanding, analysis, and exploration [229]. This foundation allows for various innovative applications in music education. Integrating interactive sheet music and music theory applications presents a transformative opportunity. Students can gain a more intuitive and engaging understanding of complex musical concepts by combining dynamic, interactive sheet music with applications that visually teach scales and music theory. This synergy enables a deeper, more interactive learning experience where theory and practice blend seamlessly. The gamification of music learning platforms is another avenue with immense potential. Such platforms could significantly increase student motivation and engagement by embedding game mechanics into music education. This approach can transform mundane practice sessions into exciting and rewarding challenges, enhancing the learning curve. Furthermore, music analytics for performance improvement stands to revolutionize music education. Software that provides visual feedback on pitch, rhythm, and dynamics can help students quickly identify and correct their mistakes, leading to more efficient and effective practice sessions. Lastly, data visualization in music history, including cultural heritage, can enrich students' understanding of music's evolution and context. By visually mapping music history and trends, students can gain a more comprehensive understanding of how music and composers have been influenced by and

have influenced other cultures and music creators. Building upon the content of this thesis, these advancements promise to make music education more engaging and effective and pave the way for a deeper, more nuanced understanding and appreciation of music as an art form.

To sum up, in synthesizing the future direction of Visual Musicology, it's essential to recognize the multifaceted nature of this field, as encapsulated in the VMG. This graph illustrates the various data types and highlights the diverse domains encompassed by musicology, from psychological and physiological to physical and societal. The potential of visualization tasks to aid musicologists in addressing complex domain problems through computational methods is profitable. Visual Musicology underscores the interdisciplinary essence of this field. It bridges musicology and information visualization, fostering a fertile ground for innovative research collaborations. This cross-disciplinary approach is vital for addressing the myriad challenges presented by the different subdomains of musicology, which include the analysis of sheet music and the psychological aspects like the emotions of audiences and the gestures of musicians, as already discussed in Chapter 2. The future direction of Visual Musicology lies in leveraging this comprehensive framework to address the multifarious needs of the musicological domain. While visualization cannot solve and address all music-related challenges, the VMG serves as a reminder of the unexplored potential in this field. It calls for a joint effort to develop tailored solutions that cater to the specific requirements of each subdomain and user, acknowledging that some areas may benefit more from analytical techniques than others. This diversity and complexity promise ample cross-disciplinary research and innovative solutions as we look to the future. Integrating big data, AI, and machine learning into music analysis, coupled with the exploration of VR, AR, and MR, offers a pathway to revolutionize how we understand and interact with music and educate and engage with it on a societal level. By utilizing interdisciplinary research and leveraging the VMG, we can prepare the way for a future where musicology and information visualization research synergistically create a more profound understanding and appreciation of music.

9.4 The Value of Visual Musicology

We believe that there are ample opportunities for research at the intersection of information visualization and musicology, which we framed as a new research area called *Visual Musicology*. We argue that ongoing research will provide interesting challenges for researchers on both sides, bridging two disciplines by considering their reciprocal potential. This thesis especially focused on exploring and investigating digital sheet music. We discussed how interactive visual analysis can make advanced visualization techniques more accessible to visualization laypeople within musicology. One advantage of our approach is that people from the musicology domain get systematic instructions on using more abstract visualization for understanding, interpreting, and inspecting sheet music data. First, this relevance of visualization becomes more apparent for musicological tasks. This thesis emphasizes the need for information visualization experts to tackle musicological challenges. It advocates for liaisons who, like the author of this thesis, possess a background in musicology and expertise in visualization design and programming.

Visualization can be an effective instrument to build on intuitive domain-specific concepts to enable more profound insights into compositions, performances, and music theory. For example, in music educational contexts, we can use musical concepts to reveal the importance of visualizations as illustrative means in contrast to leveraging tailored visualizations to understand complex relationships in music data structures. Second, musical data can be highly systematic when based on symbolic formats. At the same time, the fundamentals of music processing deal with audio data, which contains far more nuanced details and poses its own difficulties for analysis tasks. We consider a specialized investigation of the links between the symbolic and audio format based on big data sources will bring further mutual progress for both fields, such as audio-to-score alignment [309].

We argue that Visual Musicology displays its strengths in cross-pollinating collaborations, transforming the role of modern musicology to enrich it with data science techniques, hence empowering music educators to improve music education through visual applications, for instance. In addition, combining visualization techniques for musicological tasks will lead to increasingly democratized access to useful visual applications in the musicology domain, enhancing education, performance, and music exploration with visual tools. While recent research endeavors such as Khulusi et al.'s survey about visualizations for musical data exist [163], we postulate that providing an encompassing categorization framework will help future research projects to continue following a systematic strategy, which we discussed in this thesis. Since the methodical approaches presented in this thesis are a starting point, there are strong indicators that modern technologies will change how musicological tasks will be approached. To summarize, even if technological progress has the potential for groundbreaking paradigm shifts, the respective user communities first must approach each other and accept or even pursue new methods and the application of new technologies. The omnipresence of technical devices and the increasing usage of digital technologies hint toward practitioners and experts using experimental research and industrial products for many different purposes.

9.5 Concluding Reflections and Final Remarks

In retrospect, the research conducted and discussed in this thesis started from the overarching research question *"How can we employ visual analytics to support the understanding and investigation of sheet music?"*

We contributed to developing and designing novel visualization techniques and applications, which we evaluated through several pair-analytics user studies to carry *Visual Musicology* research forward by setting a specific focus on digital sheet music. Our study outcomes provide evidence that exploiting familiar concepts such as the circle of fifths or the rhythm tree are very effective as baselines for interaction and visualization methods, enabling the interpretation, exploration, and overall understanding of harmonic, rhythmic, and melodic connections extracted from sheet music. Having an overview of the intersection between musicology and information visualization regarding data, task, and domain characteristics facilitates the identification of relevant challenges, which we showcased for education and analysis. Following the systematic strategy of building the previous parts and chapters on one another demonstrates how initial designs can progress toward more complex interaction scenarios, increasing the level of interaction and abstraction, which remains accessible through consistent design choices. In this thesis, the initial harmonic and rhythmic fingerprint design was static, which we extended for interactive visual querying in MusicVis, further enhancing music analysis support. Finally, we introduced CorpusVis as an opposing but complementing technique, facilitating the investigation of complex patterns and relationships between composers, musical epochs, composing styles, and musical features based on sheet music collections.

Based on the Visual Musicology Graph and the Methodology Transfer Model, we illustrated various ways to continue research on top of the contributions and findings presented in this thesis. We are looking forward to the forthcoming research, especially at the intersection of musicology and visual analytics, which will provide even more effective interaction designs, visualization techniques, and applications addressing music analysis tasks following completely different strategies.

Appendix

A

This chapter provides additional visual material to facilitate further insights into musical features, especially the geometric relationship between chords, leveraging the harmonic fingerprint design. The illustrations are carefully designed based on the prototypes presented in Part III. The figures prepared for the appendices were built using MuseScore to notate the sheet music based on the CMN, which was then loaded into MusicVis to create and extract the fingerprints from the carefully prepared sheet music. The respective musicxml files for the figures in this appendix are provided here: `https://osf.io/q57kh/`. These sheet music files can be downloaded and adapted for further reference or used to continue on this work.

Figure A.1 gives a helpful overview of chords that comprise one to three different pitch classes. The top row contains the respective root notes based on the order of the circle of fifths that is explained in detail in Subsection 4.3.1. The left column reveals the general template that is rotational invariant for each row since it can be rotated to match any instance. The juxtaposition to the sheet music facilitates the mapping between the harmonic fingerprints and CMN. Overall, the chords are constructed by 30° clockwise rotations between each chord step from left to right. This means that the chords in the first column would result from such a rotation when applied to the fingerprints on the most-right column (here F).

Figure A.2 has the same structure but contains more complex chords starting from harmonic compositions of four different pitch classes, such as *Major* 7^{th} chords up to six pitch classes in the last two rows, such as *Half-Diminished* 13^{th}. These chords are typically used in genres with a higher harmonic complexity, such as Jazz music. For example, the *Fully-Diminished* 7^{th} chord is especially remarkable due to the four pitch classes on the four opposite sides leading to a “wheel”-shape. The opposite sides reveal rather dissonant chord compositions, which most listeners perceive as less pleasant sounds. This chord is also exceptional compared to most other chords since it repeats itself already after three iterations, such that the chords based on the root notes C, A, F♯, and E♭ have identical harmonic fingerprints. This noticeable geometric anomaly can be identified in other chords, revealing dissonant simultaneous sounding of notes.

Moreover, the appendix (see Section A.2) provides examples of rhythmic fingerprints that investigate rhythmic patterns and compare varying meter/time signatures.

A.1 Systematic Investigation of the Harmonic Geometry

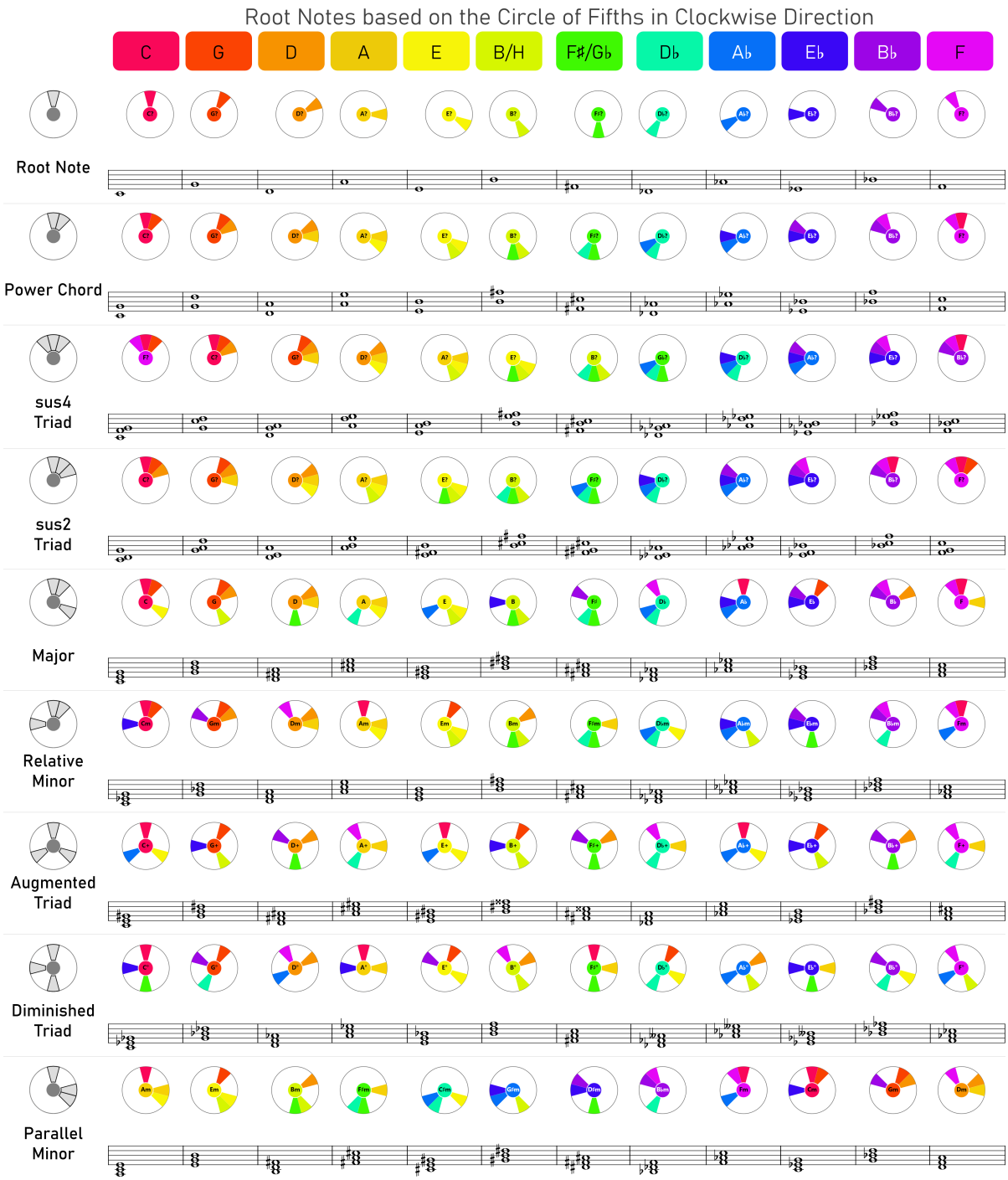


Fig. A.1.: This figure provides an overview of harmonic fingerprints for simple chords starting from *Root Notes* (only one pitch class) over *Power Chords* (two pitch classes with a distance of seven semitones or quint interval), up to three note *Triads* such as *sus2/sus4*, *Major*, *Relative and Parallel Minor*, *Augmented*, *Diminished* chords. The left column shows the chord templates for each row, while the top column represents the root notes for the respective columns.

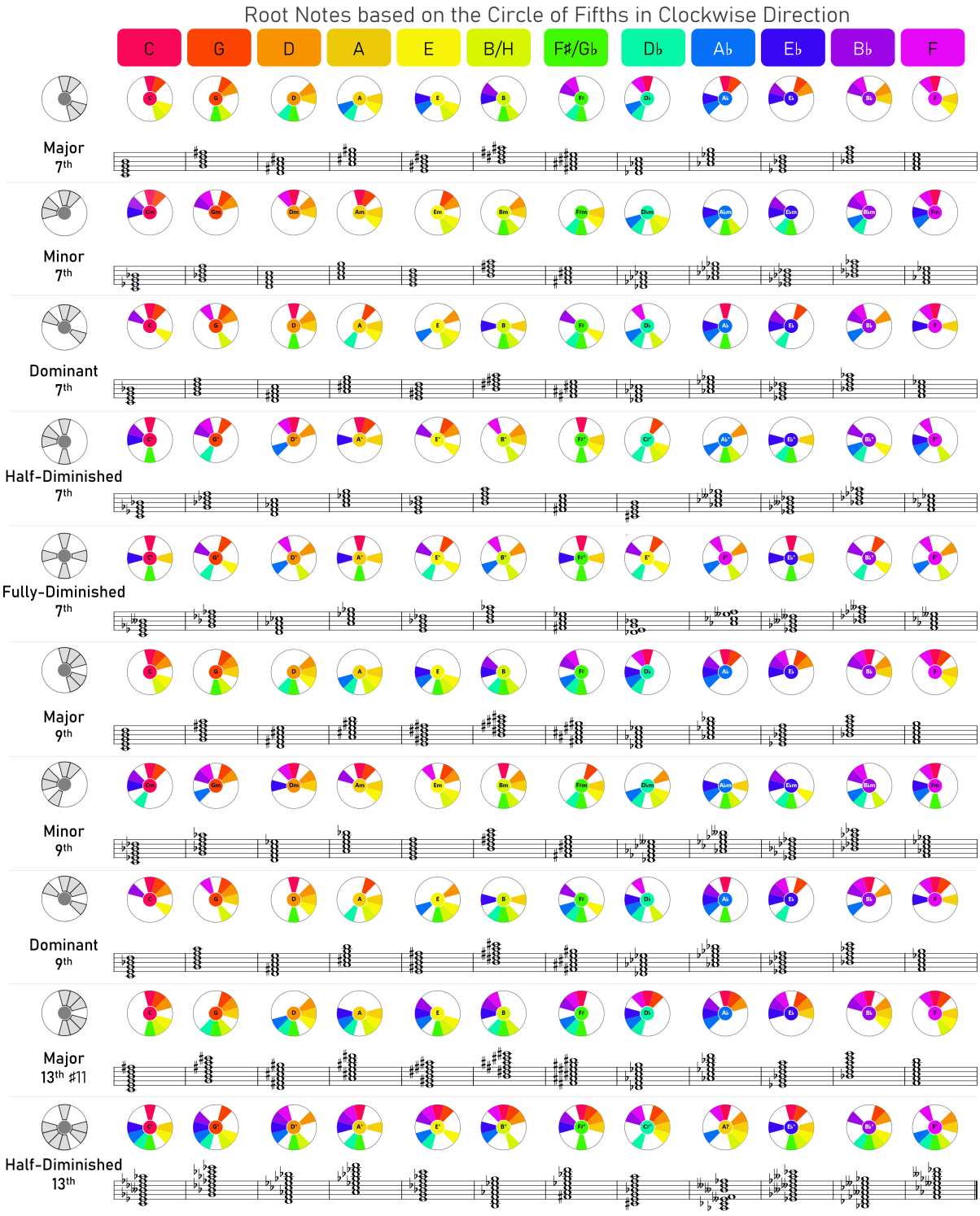


Fig. A.2.: Extension of the listed chords from Figure A.1 by adding chord compositions comprising four to six different pitch classes. Again, the top row displays the *root notes* while the first entries in each row represent the general chord templates that build up the rows when considering the respective root notes. In general, the chords result from stacked third intervals such as the *Minor* 9^{th}. Applying 30°clockwise rotations connects neighboring fingerprints from left to right.

A.1.1 Compact Harmonic Fingerprints Summary Overview

Starting with Figure A.1 and Figure A.2, we removed the CMN (sheet music) details to focus only on the harmonic fingerprints by merging them into a single overview which is displayed in Figure A.4. The overall structure remains the same, while the figure provides a decent summary at a glance. While the text within the centered circles is small, we emphasize that the same root note for almost every row is given in the header (first row still shows the root notes).
Based on the subsequent figure, we can identify exciting interrelationships between the chords by combining fingerprints from different rows and templates. For instance, the added arrows in Figure A.3 illustrate how the *Power Chords* in the second is constructed by merging two *root notes* from the first row, such as B and F♯ into the B -*Power Chord*. Figure A.3 also shows the rotational representation, since the F -*Power Chord* is a combination of the root notes F and C.

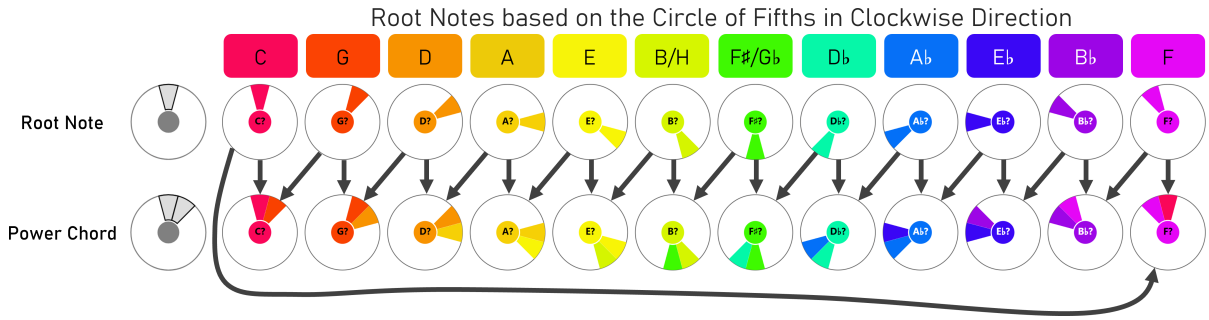


Fig. A.3.: *Power Chords* are compositions by single root notes. The arrows indicate how *Root Notes* need to be combined (↓ , ↙) to create power chords. We can also see that the F -Power Chord is created together with C, which is in the first (or better second) column, revealing the circular structure originating from the standard design of the COF (see subsection 4.3.1).

Note that this structural type of combining two fingerprints becomes more complex in the later rows, which cannot be generated by simply merging two neighboring fingerprints into the one below, such as A -Minor 7^{th}. For example, the *Fully-Diminshed* 7^{th} chords in row 14 are a composition of *Diminished Triads* (row 8) and the root note that is three columns to the right (e.g., D → B).
We could even add an additional consideration of basic intervals (e.g., second, third, etc.), creating a tabular structure on how combining two notes of intervals would lead to shapes. The quint interval is already in the second row (e.g., power chord).

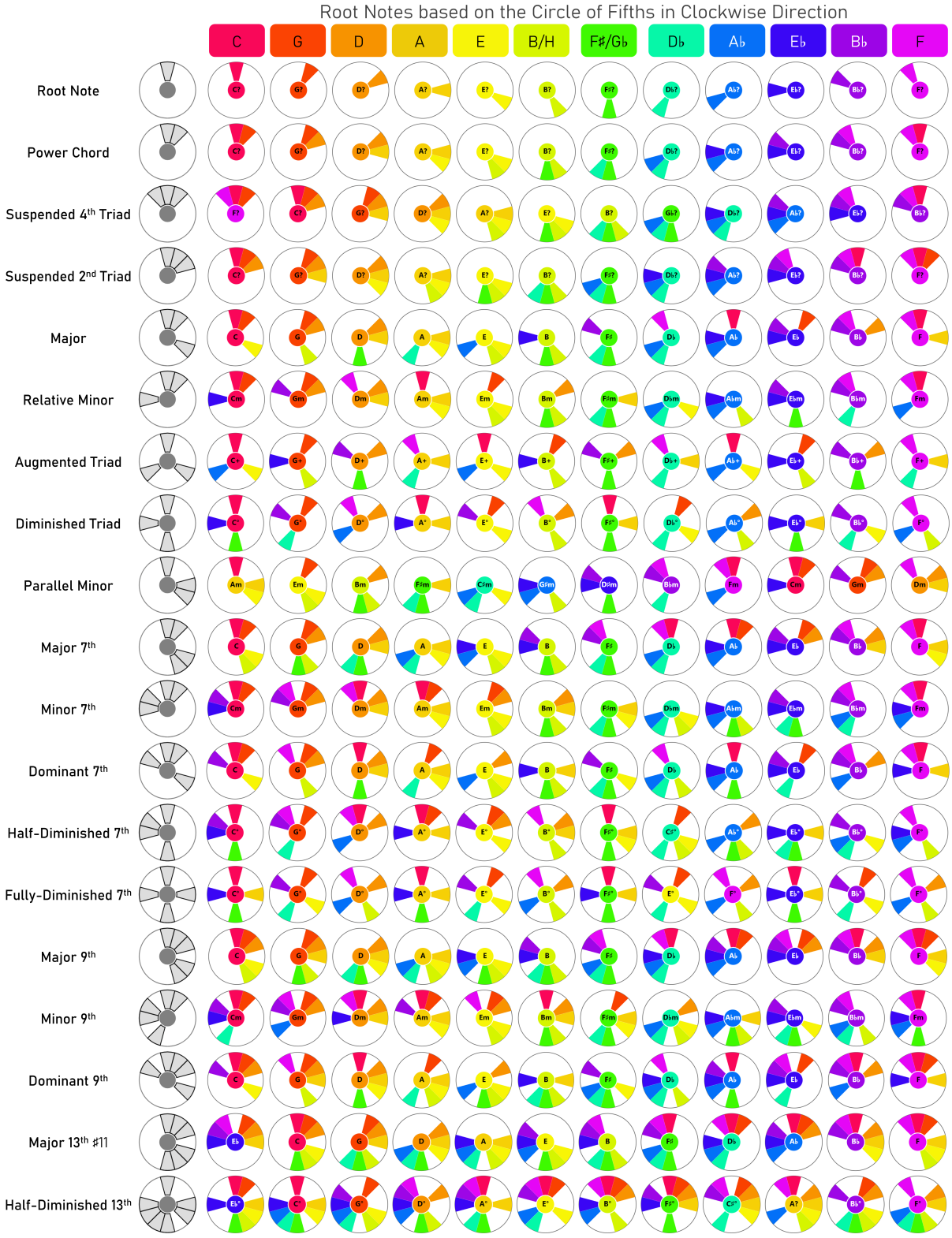


Fig. A.4.: Merged overview of Figure A.1 and Figure A.2 by deliberately omitting the CMN to achieve a better, more abstract overview, similar to the progression layer of MusicVis (please refer to Figure 7.4.2 for more details). The color encoding based on the root notes helps identify differences and commonalities between the chord structures and layers. The color legend is displayed in the first row.

A.1.2 Harmonic Fingerprints for the Major Scale Degree Chords

We want to consider how the harmonic fingerprints can be used when focusing on the specific chords typically constructed on top of notes within a scale, such as the major scale. In the C-Major scale, e.g. , the chord on the third degree (E -Minor chord) is E G B as illustrated in Table A.1, which also indicates how three different scales are connected to each other based on the scale degrees (`I`, `ii`, etc.). Capital and lowercase letters refer primarily to major and minor chords, respectively.
Figure A.5 and Figure A.6 are an extension of Table A.1 showing the same chords but using two different orders. Figure A.5 arranges the twelve scale root notes based on the COF, while Figure A.6 shows the same fingerprints but based on chromatic (semitones) upward motions, basically "jumping" through the COF. In both figures, we can see that the chords in the major scale on the degrees `I`, `IV`, and `V` have the same template structure since it is invariant with respect to rotation. The same applies to the minor chords on the degrees `ii`, `iii`, and `vi`. This scale's degree $\texttt{VII}^0$ is an exception, as it contains a diminished chord. The table also shows that `i`, `iv`, and `v` can exist as minor chords in case of minor scales such as the A -Minor scale in the right column of Table A.1. The original table provided by Krumhansl et al. does not differentiate between uppercase and lowercase lettering to disambiguate between the chord qualities (Major, Minor) [176]. In addition, we included the bracketed chords to display how the `Roman Numerals` can be used together with ♭ and ♯ accidentals to notate the respective chords. Note that the chord quality is consistent in each row. We could extend this table for all pitch classes similar to Figure A.4, but this example should be illustrative enough.

Chord	Quality	#	Components	C-Major	G-Major	A-Minor
C	Major	C	C E G	`I`	`IV`	(`III`)
D	Minor	d	D F A	`ii`	(`v`)	`iv`
E	Minor	e	E G B	`iii`	`vi`	`v`
F	Major	F	F A C	`IV`	♭`VII`	`VI`
G	Major	G	G B D	`V`	`I`	(♭`VII`)
A	Minor	a	A C E	`vi`	`ii`	`i`
B♭	Major	B♭	B♭ D F	(♭`VII`)	(♭`III`)	(♭`II`)
B	Diminished	$\texttt{b}^0$	B D F	$\texttt{VII}^0$	—	`ii`
D	Major	D	D F♯ A	(`II`)	`V`	(`IV`)
F♯	Diminished	$\texttt{f}\sharp^0$	F♯ A C	(♯`IV` ^{0})	$\texttt{VII}^0$	(♯`VI` ^{0})
B	Minor	b	B D F♯	(`vii`)	`iii`	(`ii`)
E	Major	E	E G♯ B	(`III`)	(`VI`)	`V`
G♯	Diminished	g♯°	G♯ B D	(♯`V` ^{0})	(♯`I` ^{0})	$\texttt{VII}^0$
C	Augmented	$\texttt{C}^+$	C E G♯	(`I` $^+$)	(`IV` $^+$)	`III` $^+$

Tab. A.1.: Chord note component positions in C-Major, G-Major, and A-Minor. This table, inspired by Krumhansl et al. [176], helps to identify which chords are built on top of which degree note (chord root notes). The scale degrees relate to the harmonic fingerprints in the first column of Figure A.5.

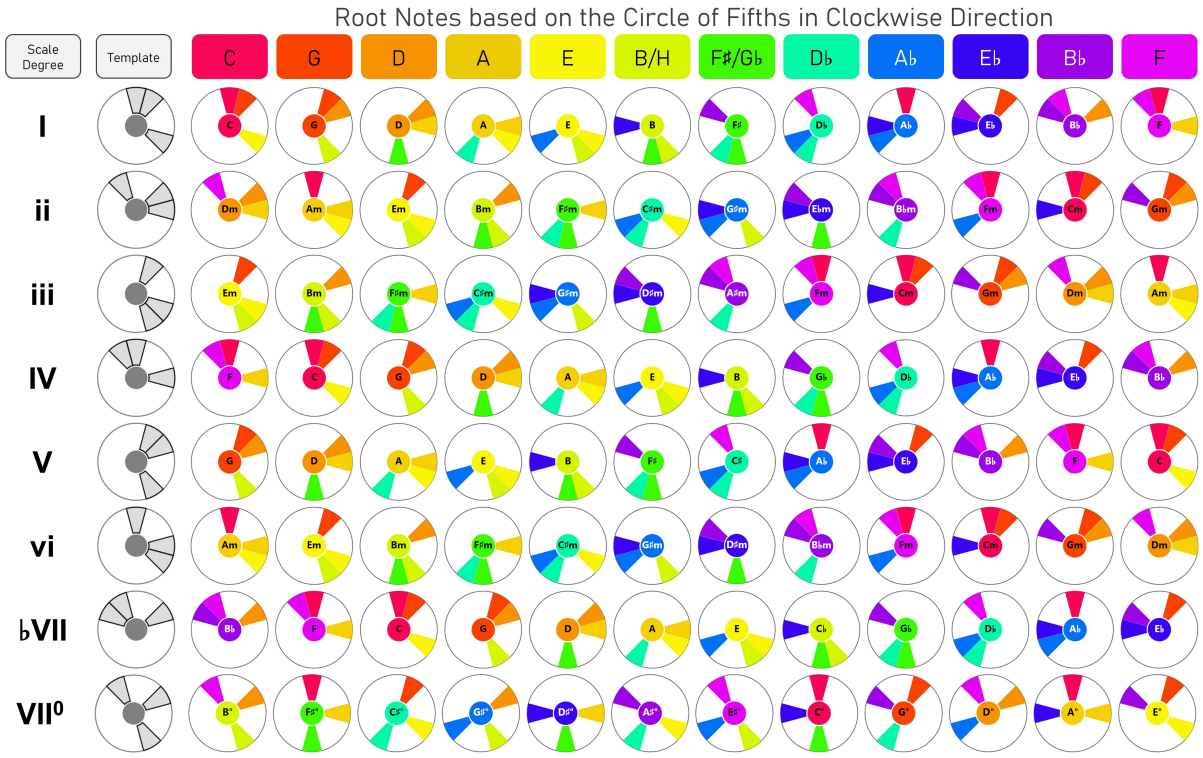


Fig. A.5.: The left column shows the scale degrees introduced in Table A.1. The remaining columns show the templates for each row and the harmonic fingerprints for each scale root, such as C in the third column. The column arrangement uses the circle of fifths instead of chromatic/semitones as in Figure A.6. The templates show the same fingerprint template for `I`, `IV`, `V`, and `♭VII`.

Figure A.5 and Figure A.6 arrange the harmonic fingerprints based on the scale degree in major scales for each of the existing twelve pitch classes. Figure A.5 uses the same column order as the previous figures to maintain the connection. The rows indicate the same scale degrees that are discussed in more detail in Table A.1. We can see the same template for the degrees `I`, `IV`, `V`, and `♭VII` representing major chords. The `♭VII` represents an exception since the B♭ is not a typical note in the C-major scale. Nevertheless, it is often used as a transition chord in chord progressions, due to its proximity to the pitch class F in the COF (neighbors). Similarly, `ii`, `iii`, and `vi` represent minor chords which are typically consistent in major scales. Figure A.6 contains the same harmonic fingerprints, but the columns are ordered based on a chromatic upward motion, which is achieved simply by applying the column swaps as indicated by the bidirectional arrows in the figure. Based on this upward semitones representation, the harmonic fingerprints are not rotating in the clockwise direction by 30°steps anymore which is demonstrates clearly the functionality of the circle of fifths on the visual layout.

An interesting characteristic of Figure A.6 is that none of each two horizontally neighboring glyphs even share a single segment. This finding emphasizes the dissonant relationship between the chords of the keys based on root notes that have a distance of only a minor second interval. The last row in both figures represents the diminished chord that is constructed on the last note of each major scale, which is the penultimate measure and fingerprint in Figure A.7. We repeat the first chord in each scale at the end to show the repetitiveness of the scales.

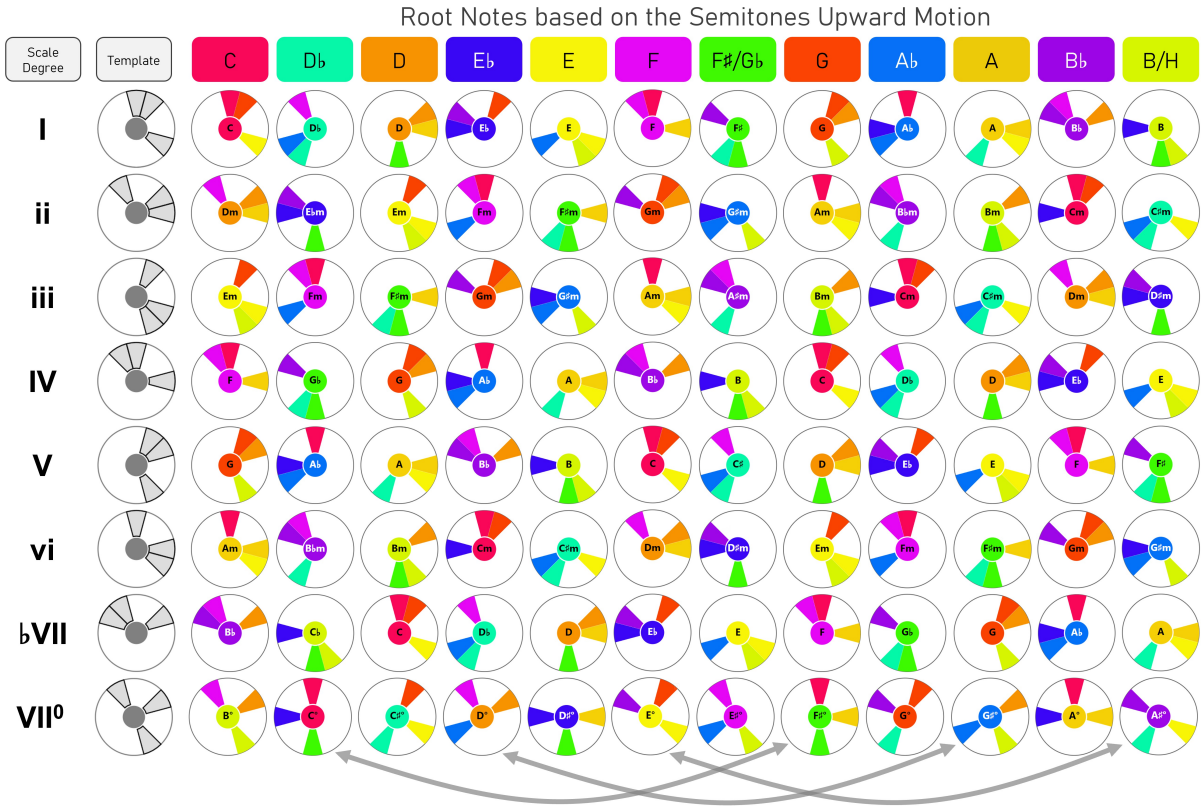


Fig. A.6.: A similar arrangement of the scale degrees as in Figure A.5. The bidirectional arrows at the bottom show how the fingerprint column needs to be swapped to achieve a chromatic or semitone upward order of the root notes, as indicated in the header of the figure. The chord templates (wrt. rotation invariance) are not changed, and the scale degrees' order remains the same.

Figure A.6 and Figure A.7 show the same harmonic fingerprints, but the latter provides the exact details relating to the underlying sheet music. As in typical major scales, there are no flat seventh chords, so we omitted them in Figure A.7. Moreover, we transposed the order between the two figures such that it fits on a single page.

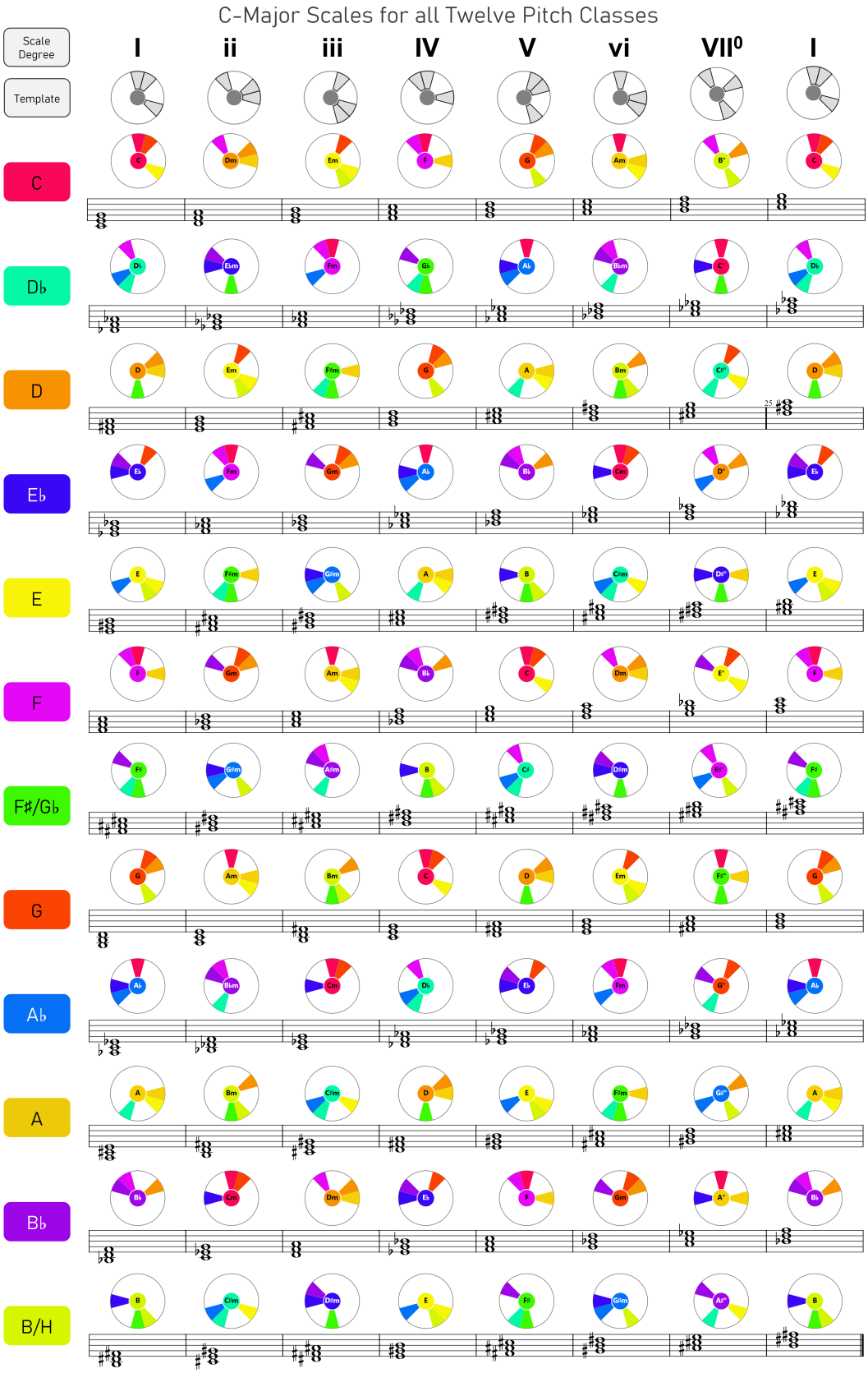


Fig. A.7.: The harmonic fingerprints for each scale degree in all keys. The ordering is based on a chromatic increase through all notes similar to Figure A.6. Similarly, we could set up figures for other scales, such as Minor, Mixolydian, or Lydian.

A.2 Systematic Exploration of the Rhythmic Fingerprints

Besides the harmonic fingerprint, we do not want to neglect the rhythmic fingerprint design. Hence, in this section, we consider patterns based on the design of the rhythmic fingerprint, introduced in Figure 5.4, in more detail. Figure A.8 shows regular rhythms maintaining the same length of all notes within single measures to summarize how the notes are distributed for such regular rhythmic fingerprints.

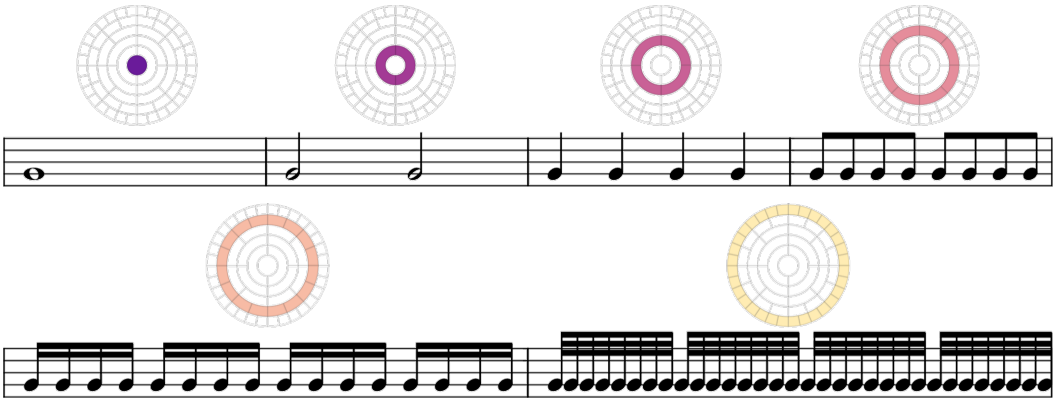


Fig. A.8.: Decreasing note duration throughout six measures from a whole note to 32nd notes. While the darker color in the inner part represents longer notes, the outer ring has a lighter hue for the 32nd notes. Here, we can see that the overall size of the fingerprint remains the same, even if more space is needed for CMN.

Figure A.9 shows how the moving the offset of a rhythmic pattern in a bar rotates a rhythmic fingerprint. The first row shows how changing the offset rotates the rhythmic fingerprints' segments. The second row contains a more complex rhythm, which leads to visual changes (here, the 7th measure) because of the quarter note, which is shifted by an 8th note. Here we have a $\frac{4}{4}$ meter signature. By subdivisions of two, we can change the angle from 90° to 45° (for 8th notes) and so on.
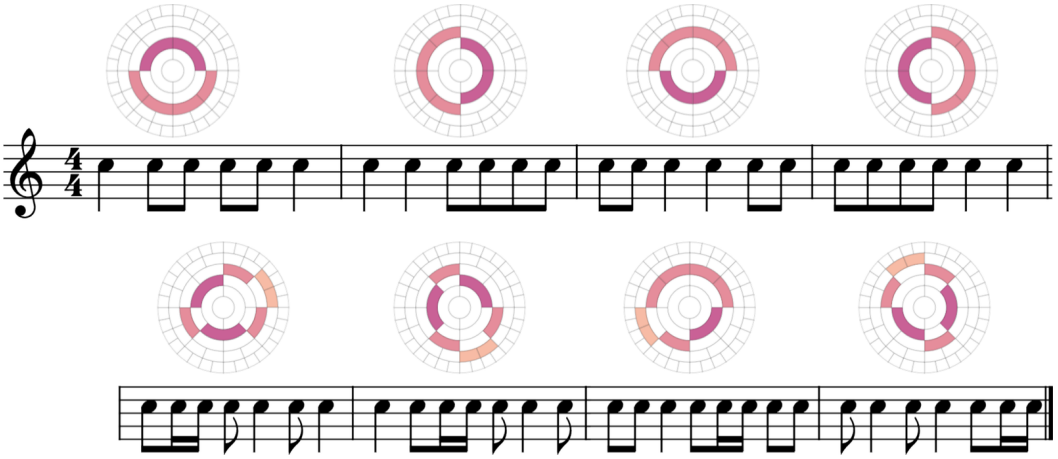


Fig. A.9.: Two examples of rotating rhythmic patterns. Since we only represent a single voice, we can see how only one segment from the inner to outer direction is filled.
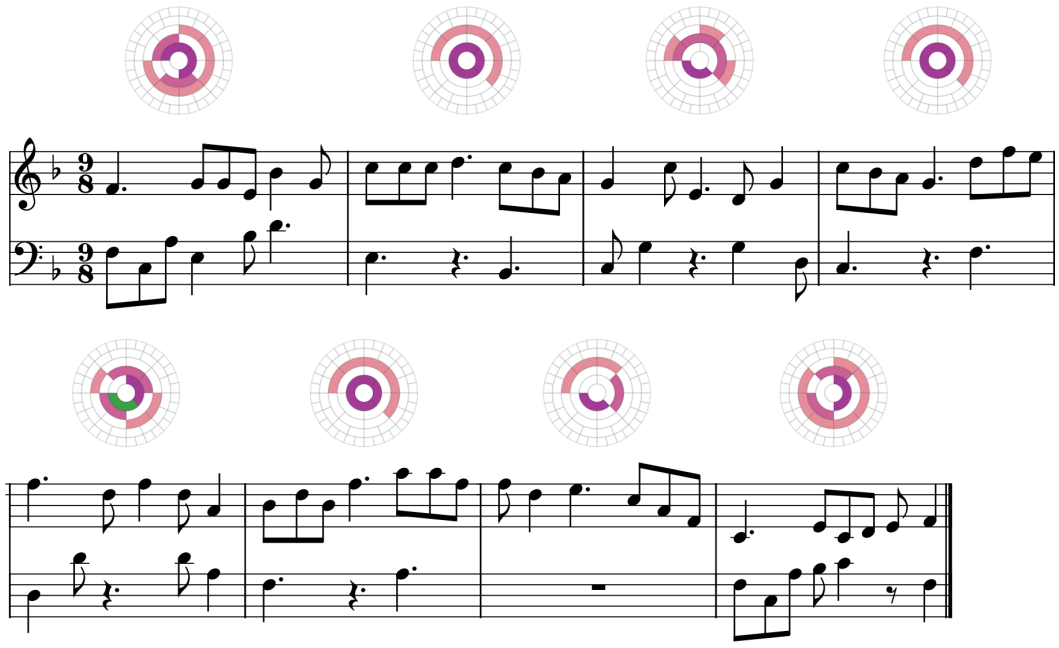


Fig. A.10.: Eight measures show the limitation of the basic rhythmic fingerprint representing a piece with $\frac{9}{8}$ meter signature. The last 8^{th} note in every bar is missing in the rhythmic fingerprint.

The rhythmic complexity in Figure A.10 is increased while maintaining a consistent time or meter signature of $\frac{9}{8}$. Our presented rhythmic fingerprint design is limited to a maximum of a whole note. Hence, the rhythmic fingerprints in measure one and two Figure A.11 look the same, although the time signatures are not the same ($\frac{2}{2} \neq \frac{3}{2}$). The same happens for $\frac{4}{4}$, $\frac{5}{4}$, and $\frac{6}{4}$, as well as $\frac{9}{8}$ and $\frac{12}{8}$. Figure A.10 illustrates that the fingerprint design facilitates the identification of different voices that do not share the same note durations and offset, such as in measures one and two. The sixth 8^{th} note in measure stands for itself, indicating that either both voices are having an 8^{th} note at this offset or one of the voices is having a break. The alternative rhythmic fingerprint design discussed in Figure 5.7 addresses the voice segmentation issue. While losing some scalability features, the alternative design mitigates both the limitation described in Figure A.11 and the simultaneous voicings.

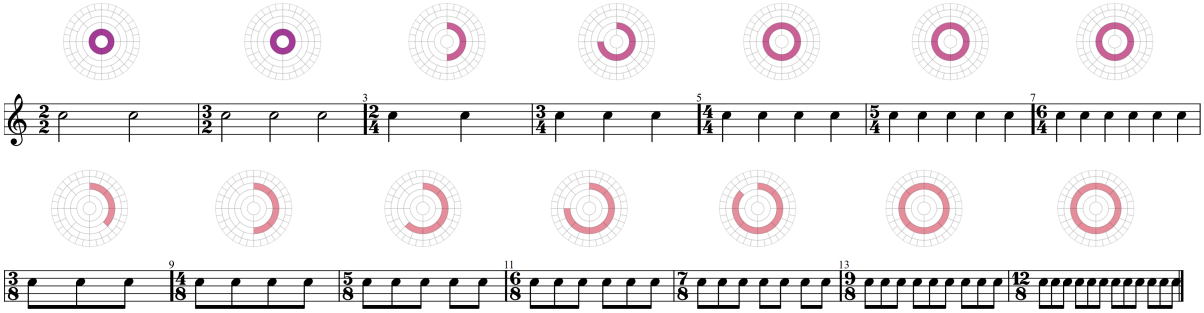


Fig. A.11.: Overview of 14 different time signatures with the same note length as the denominator of the respective time signature. The rhythmic fingerprints over the length of a whole note (e.g., $\frac{4}{4}$) can not be distinguished.

A.3 Comparing Aggregated Feature Glyphs

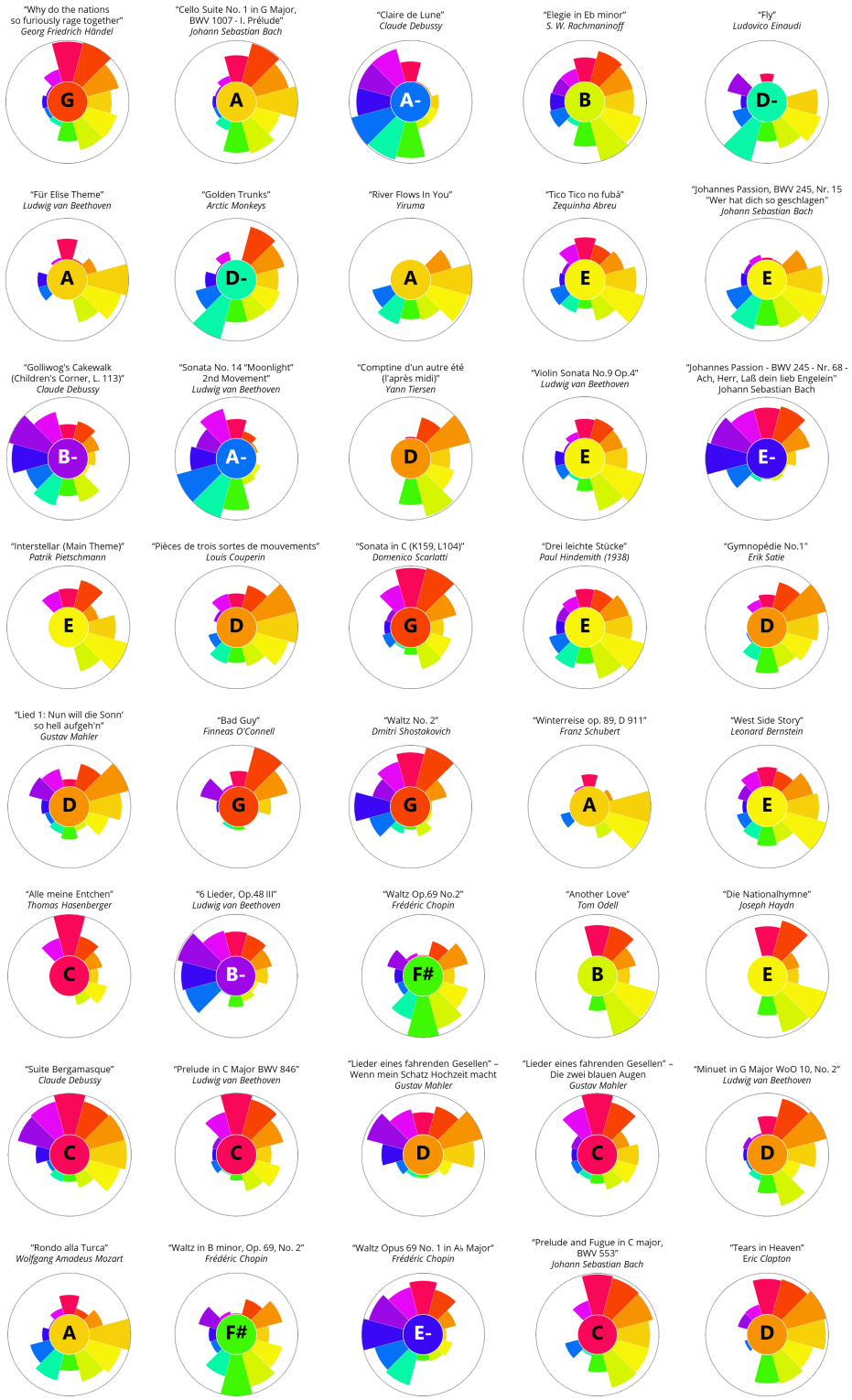


Fig. A.12.: Aggregated harmonic glyphs of 40 varying compositions.

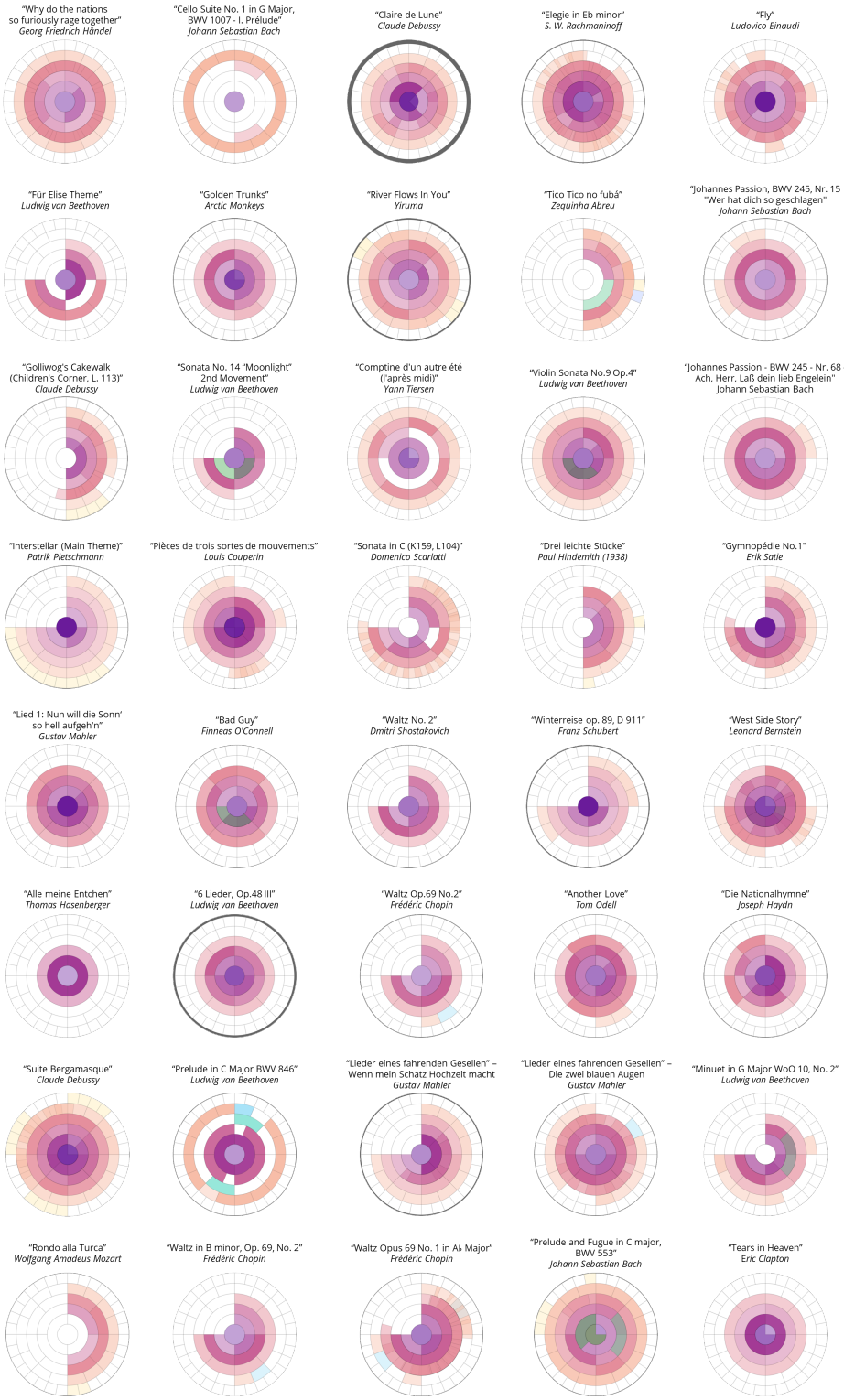


Fig. A.13.: Aggregated rhythmic glyphs of 40 varying compositions.

Figure A.12 and Figure A.13 provide an overview of 40 different composers such as Frédéric Chopin, Erik Satie, Johann Sebastian Bach, Ludwig van Beethoven, Paul Hindemith, Yann Tiersen, Eric Clapton, Claude Debussy, Domenico Scarlatti, Georg Friedrich Händel, Ludovico Einaudi, and even more of different epochs and composing styles. The position in the overview grid is the same for the harmonic and rhythmic fingerprints to enable direct comparison between the features and between the compositions.

For example, Tiersen's 'Comptine d'un autre été (l'après midi)' (Row 3, Column 3 or 3,3) is a harmonically more simple composition, as it only comprises seven different pitch classes, while many others, including Debussy's 'Suite Bergamesque' (7,1), contain all pitch classes, even if there is still a shift towards a side in the harmonic fingerprint. On the other hand, the viewer must be aware that identical harmonic fingerprints don't mean that the compositions are harmonically identical but only contain almost the same amount of notes for each resp. pitch class, independent of their location in the composition. Therefore, the analyst must not draw simple conclusions from this representation but can only get a glimpse of what might lie below in the respective composition.

Likewise, the aggregated glyphs of the rhythmic feature must be viewed with caution. While some compositions seem to be rhythmically extremely regular, such as Bach's 'Cello Suite No. 1 in G Major, BWV 1007 - I. Prélude' (1,2), which can be verified by the underlying composition. An advantage of the rhythmic fingerprint is that typical notes at a certain position in all the measures can be identified, while single notes that might only occur once or twice might be even almost not visible to to the applied transparency. Overall, an analyst can create hypotheses from the rhythmic fingerprint about the potential of largo sections based on darker areas at the glyphs' center, such as Bach's 'Johannes Passion, BWV 245, Nr. 15 "Wer hat dich so geschlagen"' (2,5). Analysts can identify potential vividness in compositions by focusing on the lighter small segments in the outer two rings (32^{nd} and 16^{th} notes), such as Mozart's 'Rondo alla Turca' (8,1), Debussy's 'Golliwog's Cakewalk' (3,1), or Abreu's 'Tico Tico no fubá' (2,4).

Of course, considering the exact glyph sequence at the measure level and identifying similarities between compositions there would be more accurate while doing this comparison only on this abstract level can lead to false conclusions. Nevertheless, the aggregated fingerprints can still emphasize strict rhythmic patterns, as in Beethoven's 'Prélude in C Major BWV 846' (7,2) and 'Für Elise' (2,1). In contrast, Rachmaninoff's 'Elegie in E♭Minor' (4,1) comprises more diverging rhythmic artifacts testifying of higher rhythmic complexity than many of the other compositions, which stem from many quintuplets and triplets, among other rhythmic elements. Looking back into the original compositions is always a good idea to confirm hypotheses that arise at such an abstract level as these summary fingerprints.

List of Figures

1.1 Thesis Outline – Visual Musicology Foundations, Analysis of Harmony, Rhythm, and Melody, and Visual Sheet Music Analytics Techniques . . 4
1.2 Framework Classification Using the Visual Musicology Graph 5

2.1 Visual Musicology: Integrating Diverse Data Types and Subdomains . . 21
2.2 Methodology Transfer Model: Bridging Domains for Solution Transfer . 23
2.3 Three Methodology Transfer Use Cases in Visual Musicology 24

3.1 Visual Encoding of Music Notations: Segmenting Musical Features and Exploiting Gestalt Laws for Creating Novel Notations 30
3.2 Multi-Step Music Notation Visualization Pipeline 33

4.1 Analyzing Recurring Harmonic Patterns With Glyph Fingerprints 43
4.2 Schroer: Illustrating Harmonic Relationships in Musical Epochs 44
4.3 Exploring Enharmonic Key Signatures and Notational Differences . . . 46
4.4 Circle of Fifths (COF): Depicting Key Relationships in Western Music . 47
4.5 Using the COF to Encode Note Distribution and Pitch Classes 50
4.6 Analyzing Themes and Rhythmic Variations in 'Grande Valse Brillante' 51
4.7 Participants Identify and Annotate Harmonic Patterns and Themes . . . 55
4.8 Using the Circle of Fifths to Map Chords to Harmonic Fingerprints . . . 59

5.1 Analyzing Recurring Rhythmic Patterns Using Rhythmic Fingerprints . 64
5.2 Drum Tabulature Illustrating Rhythmic Guidance 66
5.3 Visualizing Rhythms: Standard, Timecycle, and Necklace Notation . . . 67
5.4 Clock-Inspired 'Rhythmic Fingerprint' Design for Encoding Note and Rest Durations Based on Radial Rhythm Tree 68
5.5 Encoding Simultaneous Voices in the Rhythmic Fingerprint 70
5.6 User Study: Rhythmic Annotation for Standard and Augmented CMN . 73
5.7 Alternative Rhythmic Fingerprint Design for Simultaneous Voice Clarity 78
5.8 A Comparison of Two Rhythmic Fingerprint Designs 78

6.1 Multi-Component Interaction in MelodyVis for Melody Exploration . . 82
6.2 Example of Melodic Pattern Analysis Based on a Fugue 92
6.3 MelodyVis' Core Components for Melody Exploration 94
6.4 Melodic Pattern and Transformation Analysis of Vivaldi's Spring 96
6.5 Examining Melodic Patterns in Bach's "Prelude in C-Major" 97
6.6 Tetris Theme: From Sheet View to Melodic Transformation Graph . . . 98
6.7 User Study Design: Evaluation of MelodyVis and Melodic Operators . 99
6.8 Comparing Identified Patterns With and Without Operators 100
6.9 Contrasting Heat Maps of Melodic Patterns During the Evaluation . . . 101

7.1 MusicVis: Augmented CMN Analysis With Interactive Feature Glyphs 108
7.2 Hierarchical Segmentation of Music Compositions 113
7.3 Close to Distant Reading: Abstraction Levels and Seamless Transitions 115
7.4 Fingerprint Summarization for Enhanced Abstract Representation . . . 119
7.5 Semantic and Statistical Feature Template Queries 121
7.6 Melody Query Interface and Lasso Selection 122
7.7 BWV 553 Analysis: Harmonic Modulation from C-Major to E-Major . . 124
7.8 Harmony Glyphs in BWV 553: Exploration of Non-diatonic Notes . . . 124

8.1 CorpusVis – Visual-Interactive Analysis of Sheet Music Collections . . 138
8.2 Keyword Search, User Data Import, and Interactive Introduction Guide 139
8.3 Statistical Comparison of Music Features – Distribution Chart for Corpus Subset Analysis: Example for Note Density 140
8.4 Compositional Feature Matrix for Schumann and Mussorgsky 140
8.5 Visualizing Compositions and Epochs in MDS Projection View 141
8.6 Utilizing Metadata Table for Composition Exploration and Sorting . . . 142
8.7 Sheet View: Utilizing MusicVis for Harmonic and Rhythmic Analysis . . 142
8.8 Intro.js: Introducing CorpusVis' Interactive Workspace Components . 146
8.9 CorpusVis Evaluation: Feedback Results and User Experience 149

9.1 Thesis Structure Synopsis through Visual Musicology Graph Topics . . 156
9.2 The Trio of MusicVis, CorpusVis, and MelodyVis 157

A.1 Chord Spectrum for Root, Power, and Triad Harmonic Templates 170
A.2 Extended Harmonic Fingerprints from 4 to 6 Pitch Class Combinations 171
A.3 Visualizing the Relationship Between Power Chords and Root Notes . . 172
A.4 Merged Harmonic Fingerprint Overview Omitting CMN 173
A.5 Analyzing the Mapping Between Scale Degree Chords and Harmonic Fingerprints Based on Major Scales Using the COF Ordering 175
A.6 Scale Degree Arrangement for Chromatic/Semitones Order 176
A.7 Harmonic Fingerprints for Each Scale Degree in Major Scales 177
A.8 Whole Notes to 32nd Notes in the Rhythmic Fingerprints 178
A.9 Rotating Rhythmic Fingerprints by Changing the Offset 178
A.10 Limitation of Rhythmic Fingerprints Longer Than Whole Notes 179
A.11 14 Different Time Signatures With Matching Note Lengths 179
A.12 Aggregated Harmonic Glyphs of 40 Varying Compositions. 180
A.13 Aggregated Rhythmic Glyphs of 40 Varying Compositions. 181

List of Tables

3.1 Investigating the Visual Encoding of Music Notations 34

7.1 Participant-Selected Pieces for Study Preparation Demonstrating the Generalizability of MusicVis 126

8.1 Extracting jSymbolic Musical Features using `music21`. 136

A.1 Deconstructing Chords Into Notes Based on Their Scale Degree 174

Nomenclature and Acronyms

ACM	Association for Computing Machinery
AI	Artificial Intelligence
BMLO	Bayerisches Musiker-Lexikon Online
BWV	Bach-Werke-Verzeichnis
C(W)MN	Common (Western) Music Notation
CGF	Computer Graphics Forum
COF	Circle of Fifths
D^3	Data-Driven Documents
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DH	Digital Humanities
GTTM	Generative Theory of Tonal Music
IMSLP	International Music Score Library Project
ISMIR	International Society for Music Information Retrieval
InfoVis	Information Visualization
MDS	Multidimensional Scaling
MEI	Music Encoding Initiative
MIDI	Musical Instrument Digital Interface
MIR	Music Information Retrieval
MOS	Melodic Operator Selection
MPC	Melodic Pattern Configuration
MTG	Melody Transformation Graph
MTM	Methodology Transfer Model
MXL	MusicXML
NASA-TLX	NASA Task Load Index
OMR	Optical Music Recognition
OP	(Melodic) Operator
OSMD	Open Sheet Music Display
RQ	Research Question
SME	Subject Matter Expert
SV	Sheet View
SVG	Scalable Vector Graphics
TENOR	Int. Conf. on Technologies for Music Notation and Representation
VA	Visual Analytics
VAE	Visual Analytics Expert
VIS4DH	Visualization for the Digital Humanities
VM	Visual Musicology
VMG	Visual Musicology Graph
VSV	Voice Separation View
XML	Extensible Markup Language

Bibliography

[1] Samer A. Abdallah, Emmanouil Benetos, Nicolas Gold, Steven Hargreaves, Tillman Weyde, and Daniel Wolff. “The Digital Music Lab: A Big Data Infrastructure for Digital Musicology”. In: *ACM Journal on Computing and Cultural Heritage* 10.1 (2017). DOI: 10.1145/2983918.

[2] Vladimir Agafonkin. *A very fast 2D concave hull algorithm in JavaScript*. Retrieved March 9, 2022, from https://github.com/mapbox/concaveman. 2017.

[3] Kofi Agawu. “Analyzing Music under the New Musicological Regime”. In: *The Journal of Musicology* 15.3 (July 1997). DOI: 10.2307/763911.

[4] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of Time-Oriented Data*. Springer, 2011. DOI: 10.1007/978-0-85729-079-3.

[5] Mohammad Alharbi and Robert S. Laramee. “SoS TextVis: An Extended Survey of Surveys on Text Visualization”. In: *Computers* 8.1 (2019). DOI: 10.3390/computers8010017.

[6] Shinta Widiastuty Anggerainy, Dessie Wanda, and Nani Nurhaeni. “Music Therapy and Story Telling: Nursing Interventions to Improve Sleep in Hospitalized Children”. In: *Comprehensive Child and Adolescent Nursing* 42 (2019).

[7] Pavlos Antoniadis and Frédéric Bevilacqua. “Processing of Symbolic Music Notation via Multimodal Performance Data: Brian Ferneyhough’s Lemma-Icon-Epigram for Solo Piano, Phase 1”. In: *Proc. of TENOR*. 2016.

[8] Richard Arias-Hernández, Linda T. Kaastra, Tera Marie Green, and Brian D. Fisher. “Pair Analytics: Capturing Reasoning Processes in Collaborative Visual Analytics”. In: *Conf. on Syst. Sc.* 2011. DOI: 10.1109/HICSS.2011.339.

[9] Marcelo Gabriel Armentano, Walter A. De Noni, and Hernán F. Cardoso. “Genre classification of symbolic pieces of music”. In: *J. Intell. Inf. Syst.* 48.3 (2017). DOI: 10.1007/s10844-016-0430-7.

[10] Mennatallah El-Assady, Rita Sevastjanova, Daniel A. Keim, and Christopher Collins. “ThreadReconstructor: Modeling Reply-Chains to Untangle Conversational Text through Visual Analytics”. In: *Comput. Graph. Forum (CGF)* 37.3 (2018). DOI: 10.1111/cgf.13425.

[11] Mennatallah El-Assady, Rita Sevastjanova, Fabian Sperrle, Daniel A. Keim, and Christopher Collins. “Progressive Learning of Topic Modeling Parameters: A Visual Analytics Framework”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 24.1 (2018). DOI: 10.1109/TVCG.2017.2745080.

[12] Willian Garcias de Assunção and Luciana Aparecida Martinez Zaina. “Evaluating User Experience in Music Discovery on Deezer and Spotify”. In: *Proc. Brazilian Symp. on Human Factors in Comput. Syst.* ACM, 2022. DOI: 10.1145/3554364.3560901.

[13] Hagit Attiya, Sebastian Burckhardt, Alexey Gotsman, Adam Morrison, Hongseok Yang, and Marek Zawirski. “Specification and Complexity of Collaborative Text Editing”. In: *Proc. on Principles of Distributed Computing*. ACM, 2016. DOI: 10.1145/2933057.2933090.

[14] Jean-Julien Aucouturier and Francois Pachet. “Representing Musical Genre: A State of the Art”. In: *Journal of New Music Research (JNMR)* 32.1 (Mar. 2003). DOI: 10.1076/jnmr.32.1.83.16801.

[15] Johann Sebastian Bach. *BWV 846 – Prelude in C-Major*. Retrieved March 31, 2023, from https://musescore.com/user/101554/scores/117279. MuseScore.

[16] Johann Sebastian Bach. *Fugue No. 20 BWV 889 in A-Minor*. Retrieved March 30, 2023, from https://musescore.com/classicman/scores/1777656. MuseScore, 2016.

[17] Stephen A. Barney, W. J. Lewis, J. A. Beach, and Oliver Berghof. *The Etymologies of Isidore of Seville*. Cambridge University Press, 2006. DOI: 10.1017/CBO9780511482113.

[18] David Beard and Kenneth Gloag. *Musicology: The Key Concepts*. Routledge, May 2005.

[19] Fernando Benadon. “A Circular Plot for Rhythm Visualization and Analysis”. In: *Music Theory Online* 13.3 (2007). DOI: 10.30535/mto.13.3.1.

[20] Emmanouil Benetos, Simon Dixon, Dimitrios Giannoulis, Holger Kirchhoff, and Anssi Klapuri. “Automatic Music Transcription: Challenges and Future Directions”. In: *J. of Intell. Inf. Sys.* 41.3 (2013). DOI: 10.1007/s10844-013-0258-3.

[21] Tony Bergstrom, Karrie Karahalios, and John C. Hart. “Isochords: Visualizing Structure in Music”. In: *Proc. Graph. Interf.* ACM, 2007. DOI: 10.1145/1268517.1268565.

[22] Tatiana Bershadskaya. “Analogies and Parallels in the Structure of Music and Verbal Languages”. In: *Lang., Mus., and Comput.* 2015. DOI: 10.1007/978-3-319-27498-0_1.

[23] Jacques Bertin. *Semiology of Graphics*. University of Wisconsin Press, 1983.

[24] Thierry Bertin-Mahieux, Daniel P. W. Ellis, Brian Whitman, and Paul Lamere. “The Million Song Dataset”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf.* 2011.

[25] Louis Bigo, Jérémie Garcia, Antoine Spicher, and Wendy E. Mackay. “PaperTonnetz: Music Composition with Interactive Paper”. In: *Sound and Music Computing*. Copenhague, Denmark, 2012. DOI: 10.1145/2468356.2479608.

[26] Mark-Jan Bludau, Viktoria Brüggemann, Anna Busch, and Marian Dörk. “Reading Traces: Scalable Exploration in Elastic Visualizations of Cultural Heritage Data”. In: *Comput. Graph. Forum (CGF)* 39.3 (2020).

[27] Dmitry Bogdanov, Martín Haro, Ferdinand Fuhrmann, Anna Xambó, Emilia Gómez, and Perfecto Herrera. “Semantic audio content-based music recommendation and visualization based on user preference examples”. In: *Inf. Process. Manag.* 49.1 (2013). DOI: 10.1016/j.ipm.2012.06.004.

[28] Alexandra Bonnici, Roger B. Dannenberg, Steven T. Kemper, and Kenneth P. Camilleri. “Editorial: Music and AI”. In: *Frontiers Artif. Intell.* 4 (2021). DOI: 10.3389/FRAI.2021.651446.

[29] Rita Borgo, Johannes Kehrer, David H. S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. “Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications”. In: *Eurographics 2013 - State of the Art Reports*. The Eurographics Association, 2013.

[30] Rita Borgo, Johannes Kehrer, David H. S. Chung, Eamonn Maguire, Robert S. Laramee, Helwig Hauser, Matthew Ward, and Min Chen. “Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications”. In: *STAR*. Eurographics Association, 2013. DOI: 10.2312/conf/EG2013/stars/039-063.

[31] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. "D^3 Data-Driven Documents". In: *IEEE Trans. Vis. Comput. Graph.* 17.12 (2011). DOI: 10.1109/TVCG.2011.185.

[32] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. "D^3 Data-Driven Documents". In: *IEEE Trans. Vis. Comput. Graph.* 17.12 (2011). DOI: 10.1109/TVCG.2011.185.

[33] Adam James Bradley, Mennatallah El-Assady, Katharine Coles, Eric Alexander, Min Chen, Christopher Collins, Stefan Jänicke, and David Joseph Wrisley. "Visualization and the Digital Humanities: Moving Toward Stronger Collaborations". In: *IEEE Comput. Graph. and Appl.* 38.6 (2018). DOI: 10.1109/MCG.2018.2878900.

[34] Adam James Bradley, Hrim Mehta, Mark Hancock, and Christopher Collins. "Visualization, Digital Humanities, and the Problem of Instrumentalism". In: *Proc. VIS4DH at IEEE VIS* (2016).

[35] Adam James Bradley, Victor Sawal, and Christopher Collins. "Approaching Humanities Questions Using Slow Visual Search Interfaces". In: *Workshop on VIS4DH at IEEE VIS*. 2019.

[36] Lee Brasseur. "Florence Nightingale's Visual Rhetoric in the Rose Diagrams". In: *Technical Communication Quarterly* 14.2 (2005). DOI: 10.1207/s15427625tcq1402_3.

[37] Werner Breig. "Bachs Goldberg-Variationen als zyklisches Werk". In: *Archiv für Musikwissenschaft* 32.4 (1975). DOI: 10.2307/930762.

[38] Matthew Bribitzer-Stull. *Understanding the Leitmotif: From Wagner to Hollywood Film Music*. Cambridge University Press, 2015. DOI: 10.1017/CBO9781316161678.

[39] Manuel Burghardt and Lukas Lamm. "Entwicklung eines Music Information Retrieval-Tools zur Melodic Similarity-Analyse deutschsprachiger Volkslieder". In: *Inf.* (2017).

[40] Remo Aslak Burkhard. "Learning from Architects: The Difference between Knowledge Visualization and Information Visualization". In: *Int. Conf. on Information Visualisation*. 2004. DOI: 10.1109/IV.2004.1320194.

[41] Eren Cakmak, Udo Schlegel, **Matthias Miller**, Juri Buchmüller, Wolfgang Jentner, and Daniel A. Keim. "Interactive Classification Using Spectrograms and Audio Glyphs". In: *IEEE Conf. on Visual Analytics Science and Technology (VAST Challenge MC1)*. 2018.

[42] Emilios Cambouropoulos. "'Voice' Separation: Theoretical, Perceptual and Computational Perspectives". In: *Conf. on Music Perception and Cognition*. Citeseer. 2006.

[43] Samuel Cameron. "What Is Plagiarism and What Is Musical Plagiarism?" In: *An Econ. Appr. to the Plagiar. of Music*. Springer, 2020. DOI: 10.1007/978-3-030-42109-0_1.

[44] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think*. CA, USA: Morgan Kaufmann Publishers Inc., 1999.

[45] Cornelius Cardew. "Notation—Interpretation, Etc." In: *Tempo* 58 (1961). DOI: 10.1017/S0040298200045873.

[46] M. Sheelagh T. Carpendale. "Considering Visual Variables as a Basis for Information Visualisation". In: (2003). DOI: 10.5072/PRISM/30495.

[47] Patricia Carpenter. "The Piano Music of Arnold Schoenberg". In: *Theory and Practice* 30 (2005).

[48] Giusy Caruso, Esther Coorevits, Luc Nijs, and Marc Leman. "Gestures in Contemporary Music Performance: A Method to Assist the Performer's Artistic Process". In: *Contemporary Music Review* 35.4-5 (Sept. 2016). DOI: 10.1080/07494467.2016.1257292.

[49] Wing-Yi Chan, Huamin Qu, and Wai-Ho Mak. "Visualizing the Semantic Structure in Classical Music Works". In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 16.1 (2010). DOI: 10.1109/TVCG.2009.63.

[50] Winnie Wing-Yi Chan. *A Report on Musical Structure Visualization*. 2007.

[51] Michael Chanan. *Repeated Takes: A Short History of Recording and its Effects on Music*. Verso, 1995.

[52] Mona Lisa Chanda and Daniel J. Levitin. "The neurochemistry of music". In: *Trends in Cognitive Sciences* 17.4 (2013). DOI: 10.1016/j.tics.2013.02.007.

[53] Terence Charlston. "Patterns of play: Orlando Gibbons, Simon Lohet, and JS Bach's Fugue in E Major BWV 878/2. Part II". In: *Clavichord International* 24.1 (2020).

[54] Muhammad Faisal Cheema, Stefan Jänicke, and Gerik Scheuermann. "Enhancing Close Reading". In: *Digital Humanities 2016, Conf. Abstracts, Jagiellonian University & Pedagogical University*. 2016.

[55] Ya-Xi Chen and Andreas Butz. "MusicSim: Integrating Audio Analysis and User Feedback in an Interactive Music Browsing UI". In: *Proc. on Intell. User Interf.* ACM, 2009.

[56] Yanhua Chen, Lijun Wang, Ming Dong, and Jing Hua. "Exemplar-based Visualization of Large Document Corpus". In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 15.6 (2009). DOI: 10.1109/TVCG.2009.140.

[57] Elaine Chew and Xiaodan Wu. "Separating Voices in Polyphonic Music: A Contig Mapping Approach". In: *Computer Music Modeling and Retrieval*. Springer, 2005. DOI: https://doi.org/10.1007/978-3-540-31807-1_1.

[58] Noam Chomsky. *Syntactic Structures*. 14th ed. Janua Linguarum. Series minor 4. De Gruyter Mouton, 1985.

[59] Peter Ciuha, Bojan Klemenc, and Franc Solina. "Visualization of Concurrent Tones in Music with Colours ". In: *Proc. of the International Conference on Multimedia*. ACM Press, 2010. DOI: 10.1145/1873951.1874320.

[60] ClassicMan. *Sheet Music Collection of MuseScore user ClassicMan*. Retrieved July 29, 2021, from https://musescore.com/user/19710/sheetmusic. 2021.

[61] Kate Cook. *Musical Style and Genre: History and Modernity*. Routledge, 2013. DOI: 10.4324/9780203059722.

[62] Débora C. Corrêa and Francisco Aparecido Rodrigues. "A survey on symbolic data-based music genre classification". In: *Expert Syst. Appl.* 60 (2016). DOI: 10.1016/j.eswa.2016.04.008.

[63] Michael A. A. Cox and Trevor F. Cox. "Multidimensional Scaling". In: *Handbook of Data Visualization*. Springer, 2008. DOI: 10.1007/978-3-540-33037-0_14.

[64] Ian Cross. "Music, Cognition, Culture, and Evolution". In: *Annals NY Acad. of Sci.* 930.1 (2001). DOI: 10.1111/j.1749-6632.2001.tb05723.x.

[65] Leandro Cruz, Vitor Rolla, Juliano Kestenberg, and Luiz Velho. "Visual Representations for Music Understanding Improvement". In: *Proc. of the Int. Symp. on CMMR*. 2017.

[66] Michael Scott Cuthbert. *music21.features.jSymbolic*. Retrieved July 8, 2021, from https://web.mit.edu/music21/doc/moduleReference/moduleFeaturesJSymbolic.html. 2021.

[67] Michael Scott Cuthbert and Christopher Ariza. "Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data". In: *Proc. ISMIR Conf.* 2010.

[68] Bruce Dalby. “Toward an Effective Pedagogy for Teaching Rhythm: Gordon and Beyond”. In: *Music Educators Journal* 92.1 (2005). DOI: 10.2307/3400228.

[69] Roger B. Dannenberg. “Music Representation Issues, Techniques, and Systems”. In: *Computer Music J.* 17.3 (1993).

[70] Roger B. Dannenberg. “Style in Music”. In: *The Structure of Style: Algorithmic Approaches to Understanding Manner and Meaning* (2010).

[71] Roger B. Dannenberg and Masataka Goto. “Music Structure Analysis from Acoustic Signals”. In: *Handbook of Signal Processing in Acoustics*. Springer, 2008.

[72] Maria-Iuliana Dascălu, Melania Coman, Roxana Postelnicu, and Cosmin Nichifor. “Learning to Play a Musical Instrument in Adulthood–Challenges and Computer-mediated Solutions”. In: *Procedia-Social and Behavioral Sciences* 142 (2014). DOI: 10.1016/j.sbspro.2014.07.639.

[73] W. Bas De Haas, José Pedro Magalhães, Frans Wiering, and Remco C. Veltkamp. “Automatic Functional Harmonic Analysis”. In: *Computer Music J.* 37.4 (2013). DOI: 10.1162/COMJ_a_00209.

[74] Roberto De Prisco, Delfina Malandrino, Donato Pirozzi, Gianluca Zaccagnino, and Rocco Zaccagnino. “Evaluation Study of Visualisations for Harmonic Analysis of 4-part Music”. In: *Int. Conf. Inf. Vis. (IV)*. IEEE, 2018.

[75] Roberto De Prisco, Delfina Malandrino, Donato Pirozzi, Gianluca Zaccagnino, and Rocco Zaccagnino. “Understanding the Structure of Musical Compositions: Is Visualization an Effective Approach?” In: *Information Visualization* 16.2 (Apr. 1, 2017). DOI: 10.1177/1473871616655468.

[76] Frederik L. Dennig, **Matthias Miller**, Daniel A. Keim, and Mennatallah El-Assady. “FS/DS: A Theoretical Framework for the Dual Analysis of Feature Space and Data Space”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* PP (2023). DOI: 10.1109/TVCG.2023.3288356.

[77] Frank Desmet, Luc Nijs, Michiel Demey, Micheline Lesaffre, Jean-Pierre Martens, and Marc Leman. “Assessing a Clarinet Player’s Performer Gestures in Relation to Locally Intended Musical Targets”. In: *Journal of New Music Research* 41.1 (Mar. 2012). DOI: 10.1080/09298215.2011.649769.

[78] Hao-Wen Dong, Ke Chen, Julian J. McAuley, and Taylor Berg-Kirkpatrick. “MusPy: A Toolkit for Symbolic Music Generation”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2020.

[79] Roberto Donnini. “The Visualization of Music: Symmetry and Asymmetry”. In: *Comput. & Math. with Applic.* 12.1-2 (1986).

[80] Walter Jay Dowling. “Recognition of Melodic Transformations: Inversion, Retrograde, and Retrograde Inversion”. In: *Perception & Psychophysics* 12 (1972).

[81] Kenneth Drake. *The Beethoven Sonatas and the Creative Experience*. Indiana Univ. Press, 2000.

[82] Geoffrey M. Draper, Yarden Livnat, and Richard F. Riesenfeld. “A Survey of Radial Methods for Information Visualization”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)*. 2009. DOI: 10.1109/TVCG.2009.23.

[83] Norton Dudeque. *Music Theory and Analysis in the Writings of Arnold Schoenberg (1874-1951)*. Ashgate Publishing, Ltd., 2005.

[84] Richard J. Dumbrill. *The Archaeomusicology of the Ancient Near East*. Trafford Publishing, 2005.

[85] Calvin Elliker. “Toward a Definition of Sheet Music”. In: *Notes* 55.4 (1999).

[86] Calvin Elliker. “Toward a Definition of Sheet Music”. In: *Notes* 55.4 (1999).

[87] Benoit Encelle, Nadine Jessel, Josiane Mothe, and Bachelin Ralalason. “BMML: Braille Music Markup Language”. In: *Proc. Int. Conf. on Internet Computing*. CSREA Press, 2008.

[88] Raymond Erickson and Claude V Palisca. *Musica Enchiriadis: And, Scolica Enchiriadis*. Yale University Press, 1995.

[89] Don P. Ester, John W. Scheib, and Kimberly J. Inks. “Takadimi: A Rhythm System for All Ages”. In: *Music Educators Journal* (2006). DOI: 10.2307/3878473.

[90] Maximilian T. Fischer, Simon D. Hirsbrunner, Wolfgang Jentner, **Matthias Miller**, Daniel A. Keim, and Paula Helm. “Promoting Ethical Awareness in Communication Analysis: Investigating Potentials and Limits of Visual Analytics for Intelligence Applications”. In: *Proc. of FAcct ’22: ACM Conf. on Fairness, Accountab., and Transp.* Assoc. for Comput. Machinery, 2022. DOI: 10.1145/3531146.3533151.

[91] Maximilian T. Fischer, Yannick Metz, Lucas Joos, Matthias Miller, and Daniel A. Keim. “MULTICASE: A Transformer-based Ethics-aware Multimodal Investigative Intelligence Framework”. In: (2024). DOI: 10.48550/arXiv.2401.01955.

[92] Nanke Flach, Anneke Timmermans, and Hanke Korpershoek. “Effects of the design of written music on the readability for children with dyslexia”. In: *Int. J. of Music Education* 34.2 (2016). DOI: 10.1177/0255761414546245.

[93] Josef Focht. *Bayerisches Musiker-Lexikon Online*. Retrieved Nov. 28, 2023, from http://www.bmlo.lmu.de/. 2006.

[94] Joyce Horn Fonteles, Maria Andréia Formico Rodrigues, and Victor Emanuel Dias Basso. “Creating and evaluating a particle system for music visualization”. In: *J. of Visual Languages & Computing* 24.6 (2013). DOI: 10.1016/j.jvlc.2013.10.002.

[95] Jonathan Foote and Matthew L. Cooper. “Visualizing Musical Structure and Rhythm via Self-Similarity”. In: *Proc. of Int. Comput. Music Conf.* Int. Comp. Music Assoc., 2001.

[96] Jason Freeman. “Extreme Sight-Reading, Mediated Expression, and Audience Participation: Real-Time Music Notation in Live Performance”. In: *Comput. Music J.* 32.3 (2008). DOI: 10.1162/comj.2008.32.3.25.

[97] Christian Fremerey, Meinard Müller, and Michael Clausen. “Towards Bridging the Gap between Sheet Music and Audio”. In: *Knowledge Representation for Intelligent Music Processing*. Vol. 09051. Dagstuhl Seminar Proceedings. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 2009.

[98] Klaus Frieler, Dogac Basaran, Frank Höger, Hélène-Camille Crayencour, Geoffroy Peeters, and Simon Dixon. “Don’t hide in the frames: Note-and pattern-based evaluation of automated melody extraction algorithms”. In: *Int. Conf. on Digital Libraries for Musicology*. 2019.

[99] Zhouyu Fu, Guojun Lu, Kai Ming Ting, and Dengsheng Zhang. “A Survey of Audio-Based Music Classification and Annotation”. In: *IEEE Trans. Multim.* 13.2 (2011). DOI: 10.1109/TMM.2010.2098858.

[100] Johannes Fuchs, Fabian Fischer, Florian Mansmann, Enrico Bertini, and Petra Isenberg. “Evaluation of Alternative Glyph Designs for Time Series Data in a Small Multiple Setting”. In: *Conf. on Human Factors in Computing Systems*. 2013. DOI: 10.1145/2470654.2466443.

[101] Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, **Matthias Miller**, and Daniel A. Keim. “EduClust - A Visualization Application for Teaching Clustering Algorithms”. In: *Eurographics - Education Papers*. The Eurographics Association, 2019. DOI: 10.2312/eged.20191023.

[102] Johannes Fuchs, Petra Isenberg, Anastasia Bezerianos, **Matthias Miller**, and Daniel A. Keim. "Teaching Clustering Algorithms With EduClust: Experience Report and Future Directions". In: *IEEE Comput. Graph. and Appl.* 40.2 (2020). DOI: 10.1109/MCG.2020.2970560.

[103] Daniel Fürst, **Matthias Miller**, Daniel A. Keim, Alexandra Bonnici, Hanna Schäfer, and Mennatallah El-Assady. "Augmenting Sheet Music with Rhythmic Fingerprints". In: *Workshop on VIS4DH at IEEE VIS* (2020). DOI: 10.1109/VIS4DH51463.2020.00007.

[104] Mark Gaare. "Alternatives to Traditional Notation". In: *Music Educators J.* 83.5 (1997). DOI: 10.2307/3399003.

[105] Patrick Georges and Ngoc Nguyen. "Visualizing Music Similarity: Clustering and Mapping 500 Classical Music Composers". In: *Scientometrics* 120.3 (2019).

[106] Henry F. Gilbert. "Humor in Music". In: *The Mus. Quart.* 12.1 (1926).

[107] Bela Gipp, Norman Meuschke, Corinna Breitinger, Jim Pitman, and Andreas Nürnberger. "Web-based Demonstration of Semantic Similarity Detection Using Citation Pattern Visualization for a Cross Language Plagiarism Case". In: *Proc. of the Int. Conf. on Enterprise Information Systems*. Vol. 2. SciTePress, 2014. DOI: 10.5220/0004985406770683.

[108] Cory Gledhill. *"Pitch Bracket Notation"*. Retrieved July 2, 2018, from http://pitchbracket.com/. 2018.

[109] PhonicScore GmbH. *OpenSheetMusicDisplay*. Retrieved October 12, 2023, from https://opensheetmusicdisplay.org/. 2023.

[110] Alan Goldman. "The Value of Music". In: *The Journal of Aesthetics and Art Criticism* 50.1 (1992). DOI: 10.2307/431065.

[111] Michael Good. "MusicXML for Notation and Analysis". In: *The Virtual Score: Representation, Retrieval, Restoration* 12.113-124 (2001).

[112] Michael Good. "MusicXML: An Internet-Friendly Format for Sheet Music". In: *XML Conf. and Expo*. 2001.

[113] Google AI. *Magenta: Make Music and Art Using Machine Learning*. Retrieved November 20, 2023, from https://magenta.tensorflow.org/. 2023.

[114] Masataka Goto, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano. "Songle: A Web Service for Active Music Listening Improved by User Contributions". In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2011.

[115] Phil G. Goulding. *Classical Music: The 50 Greatest Composers and Their 1,000 Greatest Works*. Ballantine Books, 2011.

[116] Fabien Gouyon and Simon Dixon. "A Review of Automatic Rhythm Description Systems". In: *Computer Music J.* 29.1 (2005). DOI: 10.1162/comj.2005.29.1.34.

[117] Patrick Gray and Razvan C. Bunescu. "From Note-Level to Chord-Level Neural Network Models for Voice Separation in Symbolic Music". In: *CoRR* abs/2011.03028 (2020).

[118] Miguel Grinberg. *Flask Web Development - Developing Web Appl. with Python*. O'Reilly, 2014.

[119] Chen Guo, Shuang Wei, Mingran Li, Zhenyu Cheryl Qian, and Yingjie Victor Chen. "Comparison of Circle and Dodecagon Clock Designs for Visualizing 24-Hour Cyclical Data". In: *Lecture Notes in Computer Science*. 2017. DOI: 10.1007/978-3-319-58640-3_5.

[120] Edward Guo. *International Music Score Library Project / Petrucci Music Library*. 2006.

[121] Hans G. Güterbock. “Musical Notation in Ugarit”. In: *Revue d’Assyriologie et d’archéologie orientale* 64.1 (1970).

[122] Knut Hartmann, Daniel Büchner, Axel Berndt, Andreas Nürnberger, and Carsten Lange. “Interactive Data Mining and Machine Learning Techniques for Musicology”. In: *Conf. on Interdisciplinary Musicology*. Vol. 15. Citeseer. 2007.

[123] Vahid Hashemi and Sara Ramezani. *Inducing Meter for a Rhythm and Measures for Rhythm Complexity*. 2008.

[124] Jeffrey Heer and George Robertson. “Animated Transitions in Statistical Data Graphics”. In: *IEEE Trans. Vis. Comput. Graph.* 13.6 (2007). DOI: 10.1109/TVCG.2007.70539.

[125] Johann David Heinichen. *Der General-Bass in der Composition (Dresden, 1728)*. Georg Olms Verlag, 1969.

[126] Lucas Henry, Klaus Frieler, Gabriel Solis, Martin Pfleiderer, Simon Dixon, Frank Höger, Tillman Weyde, and Helene-Camille Crayencour. “Dig That Lick: Exploring Patterns in Jazz with Computational Methods”. In: *Jazzforschung/Jazz Research* 50 (2021).

[127] Ivan Herman, Guy Melançon, and M. Scott Marshall. “Graph Visualization and Navigation in Information Visualization: A Survey”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 6.1 (2000). DOI: 10.1109/2945.841119.

[128] Marcia Herndon and Norma McLeod. *Music as Culture*. Norwood Eds., 1981.

[129] Frank Heyen, Alejandro Gabino Diaz Mendoza, Quynh Quang Ngo, and Michael Sedlmair. “Visual Guitar Tab Comparison”. In: *CoRR* abs/2311.14726 (2023). DOI: 10.48550/ARXIV.2311.14726.

[130] Alexander Hinneburg and Daniel A. Keim. “An Efficient Approach to Clustering in Large Multimedia Databases with Noise”. In: *Proc. of the Int. Conf. on Knowledge Discovery and Data Mining*. AAAI Press, 1998.

[131] Uta Hinrichs, Mennatallah El-Assady, Adam James Bradley, Stefania Forlini, and Christopher Collins. “Risk the Drift! Stretching Disciplinary Boundaries through Critical Collaborations between the Humanities and Visualization”. In: *Workshop on VIS4DH at IEEE VIS*. 2017.

[132] Mantle Hood. “The Reliability of Oral Tradition”. In: *Journal of the American Musicological Society* 12.2/3 (1959).

[133] Holger H. Hoos, Keith Hamel, Kai Renz, and Jürgen Kilian. “The GUIDO Notation Format: A Novel Approach for Adequately Representing Score-Level Music”. In: *Int. Conf. on Mathematics and Computing*. 1998.

[134] Yo-Wei Hsiao and Li Su. “Learning Note-to-Note Affinity for Voice Segregation and Melody Line Identification of Symbolic Music Data”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf.* 2021.

[135] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Ian Simon, et al. “Music Transformer: Generating Music with Long-Term Structure”. In: *Int. Conf. on Learning Representations*. OpenReview.net, 2019.

[136] Jessica Hudson. *Church and Worship Music in the US: A Research and Information Guide*. 2017.

[137] Ray Hudson. “Regions and Place: Music, Identity and Place”. In: *Progress in Human Geography* 30.5 (2006).

[138] Cecilia Hultberg. “Approaches to Music Notation: The printed score as a mediator of meaning in Western tonal tradition”. In: *Music Education Research* 4.2 (2002). DOI: 10.1080/1461380022000011902.

[139] Samuel Hunt, Tom Mitchell, and Chris Nash. “How can music visualisation techniques reveal different perspectives on musical structure?” In: *Int. Conf. on TENOR*. 2017.

[140] David Huron. *Characterizing Musical Textures*. Ann Arbor, MI: Michigan Publishing, University of Michigan Library, 1989.

[141] David Huron. “Music Information Processing Using the Humdrum Toolkit: Concepts, Examples, and Lessons”. In: *Comput. Music J.* 26.2 (2002).

[142] IMSLP. *Grande Valse Brillante, Op.18 (Chopin, Frédéric)*. Retrieved Nov. 25, 2023, from https://imslp.org/wiki/Grande_valse_brillante%2C_Op.18_(Chopin%2C_Fr%C3%A9d%C3%A9ric). 2019.

[143] Ken Irwin. “Musipedia: The Open Music Encyclopedia”. In: *Reference Reviews* 22.4 (2008). DOI: 10.1108/09504120810872337.

[144] Stefan Jänicke, Josef Focht, and Gerik Scheuermann. “Interactive Visual Profiling of Musicians”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 22.1 (2016). DOI: 10.1109/TVCG.2015.2467620.

[145] Stefan Jänicke, Greta Franzini, Muhammad Faisal Cheema, and Gerik Scheuermann. “On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges”. In: *Eurograph. Conf. on Visualization*. The Eurographics Association, 2015. DOI: 10.2312/eurovisstar.20151113.

[146] Stefan Jänicke, Greta Franzini, Muhammad Faisal Cheema, and Gerik Scheuermann. “Visual Text Analysis in Digital Humanities”. In: *Comput. Graph. Forum (CGF)* 36.6 (2017). DOI: 10.1111/cgf.12873.

[147] Kristoffer Jensen, Jieping Xu, and Martin Zachariasen. “Rhythm-Based Segmentation of Popular Chinese Music”. In: *Proc. of ISMIR*. 2005.

[148] Wolfgang Jentner, Fabian Sperrle, Daniel Seebacher, Matthias Kraus, et al. “Visualisierung der COVID-19-Inzidenzen und Behandlungskapazitäten mit CoronaVis”. In: *Resilienz und Pandemie: Handlungsempfehlungen anhand von Erfahrungen mit COVID-19*. Kohlhammer, 2022.

[149] Brian Johnson and Ben Shneiderman. “Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures”. In: *Proc. of the Conf. on Vis.* 1991. DOI: 10.1109/visual.1991.175815.

[150] Wilke Jan Kaasjager. *Klavarskribo*. Jan. 2001. DOI: 10.1093/gmo/9781561592630.article.15102.

[151] Linda T. Kaastra and Brian D. Fisher. “Field experiment methodology for pair analytics”. In: *Proc. of the Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*. ACM, 2014. DOI: 10.1145/2669557.2669572.

[152] Florian Kaiser and Thomas Sikora. “Music Structure Discovery in Popular Music using Non-negative Matrix Factorization”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. Ed. by J. Stephen Downie and Remco C. Veltkamp. 2010.

[153] Marius Kaminskas and Francesco Ricci. “Contextual music information retrieval and recommendation: State of the art and challenges”. In: *Computer Science Review* 6.2 (2012). DOI: 10.1016/j.cosrev.2012.04.002.

[154] Miika Karttunen and Atte Tenkanen. “Three-Dimensional Score: Seeing Music, Hearing Sculpture”. In: *Proc. of Bridges 2016: Mathematics, Music, Art, Architecture, Education, Culture*. Tessellations Publishing, 2016.

[155] Daniel A. Keim. “Information Visualization and Visual Data Mining”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 8.1 (2002). DOI: 10.1109/2945.981847.

[156] Daniel A. Keim, Gennady L. Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. “Visual Analytics: Definition, Process, and Challenges”. In: *Information Visualization - Human-Centered Issues and Perspectives*. Springer, 2008. DOI: 10.1007/978-3-540-70956-5_7.

[157] Daniel A. Keim and Daniela Oelke. “Literature Fingerprinting: A New Method for Visual Literary Analysis”. In: *Proc. of IEEE Conf. on Visual Analytics Science and Technology (VAST)*. IEEE, Oct. 2007. DOI: 10.1109/VAST.2007.4389004.

[158] Davorin Kempf. “What Is Symmetry in Music?” In: *Int. Rev. of the Aesth. and Soc. of Music* 27.2 (1996). DOI: 10.2307/3108344.

[159] Roger A. Kendall and Edward C. Carterette. “The Communication of Musical Expression”. In: *Music Perception: An Interdiscipl. J.* 8.2 (1990). DOI: 10.2307/40285493.

[160] Joseph Kerman. *Contemplating Music: Challenges to Musicology*. Harvard University Press, 2009.

[161] David Kestenbaum and Victor Mair Hyman. *Colored music notation system and method of colorizing music notation*. US Patent 7,148,414. 2006.

[162] Richard Khulusi, Jakob Kusnick, Josef Focht, and Stefan Jänicke. “musiXplora: Visual Analysis of a Musicological Encyclopedia”. In: *Proc. of the Int. Joint Conf. on Computer Vision, Imaging and Comput. Graph. Theory and Appl.* 2020. DOI: 10.5220/0008977100760087.

[163] Richard Khulusi, Jakob Kusnick, Christofer Meinecke, Christina Gillmann, Josef Focht, and Stefan Jänicke. “A Survey on Visualizations for Musical Data”. In: *Comput. Graph. Forum* 39.6 (2020). DOI: 10.1111/CGF.13905.

[164] Janice N. Killian. “The Effect of Audio, Visual and Audio-Visual Performance on Perception of Musical Content”. In: *Bulletin of the Council for Research in Music Education* (2001).

[165] Andrew Killick. “Global Notation as a Tool for Cross-Cultural and Comparative Music Analysis”. In: *Analytical Approaches to World Music* 8.2 (2020).

[166] Younghoon Kim, Michael Correll, and Jeffrey Heer. “Designing Animated Transitions to Convey Aggregate Operations”. In: *Comput. Graph. Forum (CGF)* 38.3 (2019). DOI: 10.1111/cgf.13709.

[167] Warren Kirkendale. “"Circulatio"-Tradition, "Maria Lactans", and Josquin as Musical Orator”. In: *Acta Musicologica* 56.Fasc. 1 (1984). DOI: 10.2307/932617.

[168] Matthew G. Kirschenbaum. “The Remaking of Reading: Data Mining and the Digital Humanities”. In: *The National Sci. Found. Symp. on Next Gen. of Data Mining and Cyber-Enabled Disc. for Innovation*. Vol. 134. 2007.

[169] David Kitzmüller and Toni Luong. *OKTAV*. Retrieved July 8, 2021, from https://www.oktav.com/. 2021.

[170] Jan Koláček. *Global Chant Database: Online Searchable Database of Gregorian Chant and Medieval Music*. Retr. March 27, 2023, from http://globalchant.org/search.php. 2009.

[171] Andreas Kornstädt. “Themefinder: A Web-based Melodic Search Tool”. In: *Computing in Musicology: A Directory of Research* 11 (1998).

[172] Menno-Jan Kraak. *Cartography: Visualization of Spatial Data*. Routledge, Nov. 2013. DOI: 10.4324/9781315847184.

[173] Matthias Kraus, **Matthias Miller**, Juri Buchmüller, Manuel Stein, Niklas Weiler, Daniel Keim, and Mennatallah El-Assady. “Breaking the Curse of Visual Analytics: Accommodating Virtual Reality in the Visualization Pipeline”. In: *Comput. Vision, Imaging Comput. Graph. Theory Appl.* Springer International Publishing, 2020. DOI: 10.1007/978-3-030-41590-7_11.

[174] Matthias Kraus, Thomas Pollok, **Matthias Miller**, Timon Kilian, et al. “Toward Mass Video Data Analysis: Interactive and Immersive 4D Scene Reconstruction”. In: *Sensors; Sel. Papers from the Int. Conf. on Imaging for Crime Detect. and Prev.* 1 (2020). DOI: 10.3390/s20185426.

[175] Carol L. Krumhansl. “Rhythm and Pitch in Music Cognition”. In: *Psychol. Bulletin* 126.1 (2000). DOI: 10.1037/0033-2909.126.1.159.

[176] Carol L. Krumhansl, Jamshed J. Bharucha, and Edward J. Kessler. “Perceived Harmonic Structure of Chords in Three Related Musical Keys”. In: *J. of Experimental Psychology: Human Perception and Performance* 8.1 (1982).

[177] Joseph B. Kruskal and James M. Landwehr. “Icicle Plots: Better Displays for Hierarchical Clustering”. In: *The American Statistician* 37.2 (1983). DOI: 10.2307/2685881.

[178] Kostiantyn Kucher and Andreas Kerren. “Text Visualization Techniques: Taxonomy, Visual Survey, and Community Insights”. In: *IEEE Pacific Visualization Symp.* 2015. DOI: 10.1109/PACIFICVIS.2015.7156366.

[179] Hannu Kukka, Heidi Oja, Vassilis Kostakos, Jorge Gonçalves, and Timo Ojala. “What Makes You Click: Exploring Visual Signals to Entice Interaction on Public Displays”. In: *Conf. on Human Factors in Comp. Sys*. ACM, 2013. DOI: 10.1145/2470654.2466225.

[180] Yi-Ting Kuo and Ming-Chuen Chuang. “A Proposal of a Color Music Notation System on a Single Melody for Music Beginners”. In: *Int. J. of Music Education* 31.4 (2013). DOI: 10.1177/0255761413489082.

[181] Olivier Lartillot. “Efficient Extraction of Closed Motivic Patterns in Multi-Dimensional Symbolic Representations of Music”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf.* 2005.

[182] Kai Lassfolk. “Music Notation as Objects: An Object-Oriented Analysis of the Common Western Music Notation System”. PhD thesis. University of Helsinki, 2004.

[183] Mikael Laurson, Mika Kuuskankare, and Kimmo Kuitunen. “The Visualization of Computer-Assisted Music Analysis Information in PWGL”. In: *J. of New Music Research* 37.1 (2008). DOI: 10.1080/09298210802259595.

[184] Marc Leman. *Embodied Music Cognition and Mediation Technology*. MIT Press, 2008.

[185] Marc Leman. *The Expressive Moment: How Interaction (with Music) Shapes Human Empowerment*. Cambridge, MA: MIT Press, 2016.

[186] Werner Lemberg. “A survey of the history of musical notation”. In: *TUGboat* 37.3 (2016).

[187] Raphael Lenain, Jack Weston, Abhishek Shivkumar, and Emil Fristed. “Surfboard: Audio Feature Extraction for Modern Machine Learning”. In: *Interspeech, Conf. of the Int. Speech Comm. Assoc.* ISCA, 2020. DOI: 10.21437/Interspeech.2020-2879.

[188] Fred Lerdahl. *Tonal Pitch Space*. Oxford University Press, 2004.

[189] Fred Lerdahl and Ray Jackendoff. “A Generative Theory of Tonal Music”. In: *MIT Press, Cambridge, MA* (1983).

[190] Fred Lerdahl and Ray Jackendoff. “An Overview of Hierarchical Structure in Music”. In: *Music Perception: An Interdisciplinary Journal* 1.2 (1983). DOI: 10.2307/40285257.

[191] Micheline Lesaffre, Pieter-Jan Maes, and Marc Leman, eds. *The Routledge Companion to Embodied Music Interaction*. Taylor & Francis Group, 2017.

[192] F. Levé, R. Groult, G. Arnaud, C. Séguin, R. Gaymay, and M. Giraud. “Rhythm Extraction from Polyphonic Symbolic Music”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2011.

[193] Jia Li. “Music Analysis Through Visualization”. In: *Proc. of the Int. Conf. on Technologies for Music Notation and Representation* (Cambridge, UK). Anglia Ruskin University, 2016. DOI: 10.5281/zenodo.924063.

[194] Cynthia C. S. Liem, Meinard Müller, Douglas Eck, George Tzanetakis, and Alan Hanjalic. “The Need for Music Information Retrieval with User-Centered and Multimodal Strategies”. In: *Proc. Int. ACM Workshop on Music Inf. Retrieval with User-Centered and Multimodal Strategies*. ACM Press, 2011. DOI: 10.1145/2072529.2072531.

[195] Shusen Liu, Dan Maljovec, Bei Wang, Peer-Timo Bremer, and Valerio Pascucci. “Visualizing High-Dimensional Data: Advances in the Past Decade”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 23.3 (2016).

[196] Yang Liu and Godfried T. Toussaint. “Mathematical Notation, Representation, and Visualization of Musical Rhythm: A Comparative Perspective”. In: *Int. J. of Machine Learning and Computing* 2.3 (2012). DOI: 10.7763/IJMLC.2012.V2.127.

[197] Justin London. *Rhythm*. Grove Music Online, Oxford Music Online. Retrieved October 12, 2023, from 10.1093/gmo/9781561592630.article.45963. 2001.

[198] H. C. Longuet-Higgins. “The Perception of Music”. In: *Interdiscipl. Sci. Reviews* 3.2 (1978). DOI: 10.1179/030801878791926065.

[199] Omar Lopez-Rincon and Oleg Starostenko. “Music Visualization Based on Spherical Projection With Adjustable Metrics”. In: *IEEE Access* 7 (2019). DOI: 10.1109/ACCESS.2019.2944083.

[200] Valerio Lorenzoni, Pieter Van den Berghe, Pieter-Jan Maes, Tijl De Bie, Dirk De Clercq, and Marc Leman. “Design and Validation of an Auditory Biofeedback System for Modification of Running Parameters”. In: *J. on Multimodal User Interfaces* (2018). DOI: 10.1007/s12193-018-0283-1.

[201] Psyche Loui. “Learning and Liking of Melody and Harmony: Further Studies in Artificial Grammar Learning”. In: *Top. in Cogn. Sci.* 4.4 (2012). DOI: 10.1111/j.1756-8765.2012.01208.x.

[202] Namunu Chinthaka Maddage, Changsheng Xu, Mohan S. Kankanhalli, and Xi Shao. “Content-based Music Structure Analysis with Applications to Music Semantics Understanding”. In: *Proc. of the ACM Int. Conf. on Multimedia*. ACM, 2004. DOI: 10.1145/1027527.1027549.

[203] Pieter-Jan Maes, Valerio Lorenzoni, Bart Moens, Joren Six, Federica Bressan, Ivan Schepers, and Marc Leman. “Embodied, Participatory Sense-Making in Digitally-Augmented Music Practices: Theoretical Principles and the Artistic Case “SoundBikes””. In: *Critical Arts* 32.3 (May 2018).

[204] José Pedro Magalhães and Hendrik Vincent Koops. “Functional Generation of Harmony and Melody”. In: *Proc. of the ACM SIGPLAN Int. Workshop on Functional Art, Music, Modeling & Design*. 2014.

[205] Delfina Malandrino, Donato Pirozzi, Gianluca Zaccagnino, and Rocco Zaccagnino. “A Color-Based Visualization Approach to Understand Harmonic Structures of Musical Compositions”. In: *Int. Conf. on Inf. Vis.* IEEE Computer Society, 2015. DOI: 10.1109/iV.2015.21.

[206] Delfina Malandrino, Donato Pirozzi, and Rocco Zaccagnino. “Visualization and Music Harmony: Design, Implementation, and Evaluation”. In: *Int. Conf. Inf. Vis.* 2018. DOI: 10.1109/iV.2018.00092.

[207] Alfred Mann. “Händels Fugenlehre–ein unveröffentlichtes Manuskript”. In: *Bericht über den musikwiss. Kongreß*. Bärenreiter, 1962.

[208] Arpi Mardirossian and Elaine Chew. “Visualizing Music: Tonal Progressions and Distributions.” In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2007.

[209] Siobhan McAndrew and Martin Everett. “Music as Collective Invention: A Social Network Analysis of Composers”. In: *Cultural Sociology* 9.1 (Mar. 2015). DOI: 10.1177/1749975514542486.

[210] Nina McCurdy, Julie Lein, Katherine Coles, and Miriah D. Meyer. “Poemage: Visualizing the Sonic Topology of a Poem”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 22.1 (2016). DOI: 10.1109/TVCG.2015.2467811.

[211] Ryan A. McIntyre. “Bach in a Box: The Evolution of Four Part Baroque Harmony Using the Genetic Algorithm”. In: *Proc. IEEE Conf. on Evolutionary Computation*. IEEE, 1994. DOI: 10.1109/ICEC.1994.349943.

[212] Cory McKay. *Automatic Genre Classification of MIDI Recordings*. Master’s thesis. Available at https://escholarship.mcgill.ca/downloads/h415p991b. 2004.

[213] Cory McKay, Julie Cumming, and Ichiro Fujinaga. “JSYMBOLIC 2.2: Extracting Features from Symbolic Music for use in Musicological and MIR Research”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2018.

[214] Cory McKay and Ichiro Fujinaga. “jMIR: Tools for Automatic Music Classification”. In: *Proc. of the Int. Comp. Music Conf.* Mich. Publ., 2009.

[215] Cory McKay and Ichiro Fujinaga. “jSymbolic: A Feature Extractor for MIDI Files”. In: *Proc. of the Int. Comp. Music Conf.* 2006.

[216] Liam McNabb. *SoS Literature Browser*. Retr. April 4, 2023, from http://sos.swansea.ac.uk/.

[217] Liam McNabb and Robert S. Laramee. “Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization”. In: *Computer Graphics Forum* 36.3 (2017). DOI: 10.1111/cgf.13212.

[218] Colin Meek and William P. Birmingham. “Automatic Thematic Extractor”. In: *J. Intell. Inf. Syst.* 21.1 (2003). DOI: 10.1023/A:1023549700206.

[219] Afshin Mehrabani. *Intro.js: Better Introductions for Websites and Features with a Step-by-Step Guide for Your Projects.* Retrieved November 29, 2021, from https://introjs.com/. 2016.

[220] Björn Merker. “Synchronous Chorusing and the Origins of Music”. In: *Musicae Scientiae* 3 (1999). DOI: 10.1177/10298649000030S105.

[221] Silvia Miksch and Wolfgang Aigner. “A Matter of Time: Applying a Data-Users-Tasks Design Triangle to Visual Analytics of Time-Oriented Data”. In: *Computers & Graphics* 38 (2014). DOI: 10.1016/j.cag.2013.11.002.

[222] **Matthias Miller**, Alexandra Bonnici, and Mennatallah El-Assady. "Augmenting Music Sheets with Harmonic Fingerprints". In: *Proc. of ACM Symp. on Doc. Eng.* 2019. DOI: 10.1145/3342558.3345395.

[223] **Matthias Miller**, Daniel Fürst, Maximilian T. Fischer, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. "MelodyVis: Visual Analytics for Melodic Patterns in Sheet Music". In: *arXiv preprint arXiv:2407.05427* (2024). Originally submitted to IEEE VIS 2023, Revision.

[224] **Matthias Miller**, Daniel Fürst, Hanna Hauptmann, Daniel A. Keim, and Mennatallah El-Assady. "Augmenting Digital Sheet Music through Visual Analytics". In: *Comput. Graph. Forum (CGF)* 41.1 (2022). DOI: 10.1111/cgf.14436.

[225] **Matthias Miller**, Johannes Häußler, Matthias Kraus, Daniel Keim, and Mennatallah El-Assady. "Analyzing Visual Mappings of Traditional and Alternative Music Notation". In: *Workshop on VIS4DH at IEEE VIS* (2018).

[226] **Matthias Miller** and Daniel Miller. "GraphStateVis: Interactive Visual Analysis of Qubit Graph States and their Stabilizer Groups". In: *IEEE International Conference on Quantum Computing and Engineering (QCE)*. 2021. DOI: 10.1109/QCE52317.2021.00057.

[227] **Matthias Miller**, Julius Rauscher, Daniel A. Keim, and Mennatallah El-Assady. "CorpusVis: Visual Analysis of Digital Sheet Music Collections". In: *Comput. Graph. Forum (CGF)* 41.3 (2022). DOI: 10.1111/cgf.14540.

[228] **Matthias Miller**, Hanna Schäfer, Alexandra Bonnici, and Mennatallah El-Assady. "Visual Pattern Analysis using Digital Sheet Music". In: *Late Breaking/Demos of the Int. Conf. on Music Information Retrieval*. 2019.

[229] **Matthias Miller**, Hanna Schäfer, Matthias Kraus, Marc Leman, Daniel A. Keim, and Mennatallah El-Assady. "Framing Visual Musicology through Methodology Transfer". In: *Workshop on VIS4DH at IEEE VIS* (2019).

[230] **Matthias Miller**, Xuan Zhang, Johannes Fuchs, and Michael Blumenschein. "Evaluating Ordering Strategies of Star Glyph Axes". In: *Proc. of the IEEE Visualization Conference (VIS)*. 2019. DOI: 10.1109/VISUAL.2019.8933656.

[231] Reiko Miyazaki, Issei Fujishiro, and Rumi Hiraga. "Exploring MIDI Datasets". In: *Sketches & Applications*. SIGGRAPH '03. ACM, 2003. DOI: 10.1145/965400.965453.

[232] Lousin Moumdjian, Jeska Buhmann, Iris Willems, Peter Feys, and Marc Leman. "Entrainment and Synchronization to Auditory Stimuli During Walking in Healthy and Neurological Populations: A Methodological Systematic Review". In: *Frontiers in Human Neuroscience* 12 (June 2018). DOI: 10.3389/fnhum.2018.00263.

[233] Daniel Müllensiefen and Marc Pendzich. "Court Decisions on Music Plagiarism and the Predictive Value of Similarity Algorithms". In: *Musicae Scientiae* 13 (2009).

[234] Meinard Müller. "Music Structure Analysis". In: *Fundamentals of Music Proc.: Audio, Analysis, Algorithms, Appl.* Springer, 2015. DOI: 10.1007/978-3-319-21945-5_4.

[235] Meinard Müller. "Tempo and Beat Tracking". In: *Fundamentals of Music Proc.: Audio, Analysis, Algorithms, Appl.* Springer, 2015. DOI: 10.1007/978-3-319-21945-5_6.

[236] Meinard Müller, Elaine Chew, and Juan Pablo Bello. "Computational Music Structure Analysis". In: *Dagstuhl Reports* 6.2 (2016). DOI: 10.4230/DagRep.6.2.147.

[237] Meinard Müller, Masataka Goto, and Simon Dixon. "Multimodal Music Processing (Dagstuhl Seminar 11041)". In: *Dagstuhl Reports* 1.1 (2011). DOI: 10.4230/DagRep.1.1.68.

[238] Tamara Munzner. *Visualization Analysis & Design*. CRC Press, 2014. DOI: 10.1201/b17511.

[239] MuseScore. *Goldberg Variations*. Retrieved October 12, 2023, from https://musescore.com/user/13172/scores/1671046. 2016.

[240] MuseScore. *MuseScore.com*. Retrieved Nov. 14, 2023, from https://musescore.com/. 2023.

[241] MuseScore. *Piano Sonata No. 11 K. 331 3rd Movement, "Rondo alla Turca"*. Retrieved Nov. 20, 2020, from https://musescore.com/classicman/scores/49143. 2020.

[242] MuseScore. *Prelude in C-Major (BWV 553)*. Retrieved Nov. 20, 2020, from https://musescore.com/user/28657557/scores/5390410. 2020.

[243] MuseScore BVBA. *"Sheet music | MuseScore"*. Retrieved November 20, 2023, from https://musescore.com/sheetmusic/. 2023.

[244] Arpege Music. *Pizzicato Music Notation*. Retrieved April 9, 2019, from http://www.arpegemusic.com/alternative-music-notation.htm. 2019.

[245] Tomoyasu Nakano, Masataka Goto, Jun Ogata, and Yuzuru Hiraga. "Voice Drummer: A Music Notation Interface of Drum Sounds Using Voice Percussion Input". In: *Proc. of the ACM Symp. on User Interface Software and Technology*. 2005.

[246] Delphine Naquin, Yves d'Aubenton-Carafa, Claude Thermes, and Maud Silvain. "CIRCUS: A Package for Circos Display of Structural Genome Variations from Paired-End and Mate-Pair Sequencing Data". In: *BMC Bioinformatics* 15.1 (2014). DOI: 10.1186/1471-2105-15-198.

[247] Han-Wen Nienhuys and Jan Nieuwenhuizen. "LilyPond, a System for Automated Music Engraving". In: *Proc. of the Coll. on Musical Informatics*. Vol. 1. Citeseer. 2003.

[248] Oriol Nieto and Morwaread Farbood. "Identifying Polyphonic Musical Patterns From Audio Recordings Using Music Segmentation Techniques". In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2014.

[249] Hélio Magalhães de Oliveira and R. C. de Oliveira. "Understanding MIDI: A Painless Tutorial on Midi Format". In: *CoRR* abs/1705.05322 (2017).

[250] Joe Onyekwelu. "Popular Music: A Satirical Weapon for the Fight against Social Ills". In: *Nsukka of Musical Arts Research* 1 (2012).

[251] OpenScore. *OpenScore - Liberating Sheet Music*. Retrieved October 12, 2023, from https://openscore.cc/. 2023.

[252] Ernesta Orlovaitė. "Visual Fingerprints: Identifying, Summarizing and Comparing Music". In: *Sum. and Compar. Mus.* (2013). DOI: 10.2139/ssrn.2373133.

[253] Berthold Over and Torsten Roeder. "MUSICI and MusMig." In: *Musicians' Mobilities and Music Migrations in Early Modern Europe*. Ed. by Berthold Over and Gesa zur Nieden. Biographical Patterns and Cultural Exchanges. Transcript Verlag, 2016.

[254] Stephen E. Palmer. "Gestalt Theory". In: *Theories and Applications of Counseling and Psychotherapy: Relevance Across Cultures and Settings* (2019).

[255] Elias Pampalk. *Islands of Music Analysis, Organization, and Visualization of Music Archives*. CiteSeer, 2001.

[256] Jin-Seo Park and Se-Jong Oh. "A New Concave Hull Algorithm and Concaveness Measure for n-dimensional Datasets". In: *J. Inf. Sci. Eng.* 28.3 (2012).

[257] Jinha Park. “Piano Music of Elisenda Fábregas: A Stylistic Analysis”. PhD thesis. University of South Carolina, 2013.

[258] Richard Parncutt, Amanda Bayley, and Ali C. Gedik. *“Editorial Board - Journal of Interdisciplinary Music Studies”*. Retrieved April 4, 2023, from `https://musicstudies.org/editorial-board/`. 2019.

[259] Denys Parsons. “The Dictionary of Tunes and Musical Themes”. In: *Spencer Brown* (1975).

[260] Jim Paterson. *“History of Music Notation - Evolution, Printing, and Computers”*. Retrieved July 8, 2023, from `https://www.mfiles.co.uk/music-notation-history.htm`. 2018.

[261] Kailash Patil, Daniel Pressnitzer, Shihab A. Shamma, and Mounya Elhilali. “Music in Our Ears: The Biological Bases of Musical Timbre Perception”. In: *PLoS Comput. Biol.* 8.11 (2012). DOI: 10.1371/JOURNAL.PCBI.1002759.

[262] Jouni Paulus, Meinard Müller, and Anssi Klapuri. “State of the Art Report: Audio-Based Music Structure Analysis”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2010.

[263] Matevž Pesek, Aleš Leonardis, and Matija Marolt. “An Analysis of Rhythmic Patterns with Unsupervised Learning”. In: *Applied Sciences* 10.1 (2020). DOI: 10.3390/app10010178.

[264] Martin Pichl, Eva Zangerle, and Günther Specht. “Combining Spotify and Twitter Data for Generating a Recent and Public Dataset for Music Recommendation”. In: *Proc. of the 26th GI-Workshop Grundlagen von Datenbanken*. 2014.

[265] Michael Poast. “Color Music: Visual Color Notation for Musical Expression”. In: *Leonardo Music J.* 33.3 (2000). DOI: 10.1162/002409400552531.

[266] Thomas Pollok, Matthias Kraus, Chengchao Qu, **Matthias Miller**, Tobias Moritz, Timon Kilian, Daniel A. Keim, and Wolfgang Jentner. “Computer Vision Meets Visual Analytics: Enabling 4D Crime Scene Investigation from Image and Video Data”. In: *Int. Conf. on Imaging for Crime Detection and Prevention* (2019).

[267] Roberto De Prisco, Delfina Malandrino, Donato Pirozzi, Gianluca Zaccagnino, and Rocco Zaccagnino. “Understanding the structure of musical compositions: Is visualization an effective approach?” In: *Information Visualization* 16.2 (2017). DOI: 10.1177/1473871616655468.

[268] Brigitte Rafael, Stefan Oertl, Michael Affenzeller, and Stefan Wagner. “Using Heuristic Optimization for Segmentation of Symbolic Music”. In: *Computer Aided Systems Theory*. Vol. 5717. Lecture Notes in Computer Science. Springer, 2009. DOI: 10.1007/978-3-642-04772-5_83.

[269] Sabbir M. Rashid, David De Roure, and Deborah L. McGuinness. “A Music Theory Ontology”. In: *Proc. of Semantic Appl. for Audio and Music*. ACM, 2018. DOI: 10.1145/3243907.3243913.

[270] Julius Rauscher, Raphael Buchmüller, Daniel A. Keim, and Matthias Miller. “SkiVis: Visual Exploration and Route Planning in Ski Resorts”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* (2023). DOI: 10.1109/TVCG.2023.3326940.

[271] Ana Rebelo, Ichiro Fujinaga, Filipe Paszkiewicz, André R. S. Marçal, Carlos Guedes, and Jaime S. Cardoso. “Optical music recognition: state-of-the-art and open issues”. In: *Int. Journal of Multimedia Information Retrieval* 1.3 (Oct. 2012). DOI: 10.1007/s13735-012-0004-6.

[272] Iris Yuping Ren, Anja Volk, Wouter Swierstra, and Remco C. Veltkamp. “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. (ISMIR)*. 2018.

[273] Seung-Min Rho and Een-Jun Hwang. “FMF(Fast Melody Finder): A Web-Based Music Retrieval System”. In: *Computer Music Modeling and Retrieval*. Ed. by Uffe Kock Wiil. Springer, 2004.

[274] Pál Richter. “Monophony in Multipart Instrumental Hungarian Folk Music”. In: *Multipart Music. Personalities and Educated Musicians in Trad. Practices.* MTA BTK Zenetudományi Intézet, 2015.

[275] Philippe Rigaux and Nicolas Travers. “Scalable Searching and Ranking for Melodic Pattern Queries”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2019.

[276] Devra L. Robledo. *Method of Representing Rhythm in Music Notation and Display Therefore*. US Patent 7,763,790. 2010.

[277] Jacques-Daniel Rochat. *Dodeka Music Notation*. Retrieved April 9, 2019, from `https://www.dodekamusic.com/how-to-learn-music-faster/`. 2019.

[278] George L. Rogers. “Effect of Colored Rhythmic Notation on Music-Reading Skills of Elementary Students”. In: *J. of Research in Music Education* 44.1 (1996). DOI: 10.2307/3345410.

[279] Katja Rogers, Amrei Röhlig, Matthias Weing, Jan Gugenheimer, et al. “P.I.A.N.O.: Faster Piano Learning with Interactive Projection”. In: *Proc. ACM Int. Conf. on Interactive Tabletops and Surfaces*. ACM Press, 2014. DOI: 10.1145/2669485.2669514.

[280] Perry Roland. “The Music Encoding Initiative (MEI)”. In: *Proc. of the Int. Conf. on Musical Applications Using XML*. Vol. 1060. Citeseer. 2002.

[281] Perry Roland, Andrew Hankinson, and Laurent Pugin. “Early Music and the Music Encoding Initiative”. In: *Early Music* 42.4 (2014).

[282] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey P. Ellis, and Daniel A. Keim. “Knowledge Generation Model for Visual Analytics”. In: *IEEE Trans. Vis. Comput. Graph. (TVCG)* 20.12 (Dec. 2014). DOI: 10.1109/TVCG.2014.2346481.

[283] Curt Sachs. *The History of Musical Instruments*. Courier Corp., 2012.

[284] Keigo Sakurai, Ren Togo, Takahiro Ogawa, and Miki Haseyama. “Music Playlist Generation Based on Graph Exploration Using Reinforcement Learning”. In: *IEEE Global Conf. on Life Sciences and Technologies*. IEEE, 2021. DOI: 10.1109/LifeTech52111.2021.9391870.

[285] Justin Salamon and Emilia Gómez. “Melody Extraction From Polyphonic Music Signals Using Pitch Contour Characteristics”. In: *IEEE Trans. Speech Audio Process.* 20.6 (2012). DOI: 10.1109/TASL.2012.2188515.

[286] Craig Stuart Sapp. “Graphic to Symbolic Representations of Musical Notation”. In: *Proc. of the Int. Conf. on Technologies for Music Notation and Representation*. 2015.

[287] Craig Stuart Sapp. “Harmonic Visualizations of Tonal Music”. In: *Proc. of the Int. Computer Music Conf.* Vol. 1. 2001.

[288] Craig Stuart Sapp. “Online Database of Scores in the Humdrum File Format”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2005.

[289] Craig Stuart Sapp. “Verovio Humdrum Viewer”. In: *Music Encoding Conf.* 2017.

[290] Craig Stuart Sapp. “Visual Hierarchical Key Analysis”. In: *Computers in Entertainment* 3.4 (Oct. 2005). DOI: 10.1145/1095534.1095544.

[291] Patrick Savage and Steven Brown. “Toward a New Comparative Musicology”. In: *Analytical Approaches to World Music* 2 (2013).

[292] Markus Schedl. “Deep Learning in Music Recommendation Systems”. In: *Frontiers in Applied Mathematics and Statistics* 5 (2019).

[293] Nina Schelehoff. “Optical Music Recognition: Overview, Challenges, and Possibilities”. In: (2020).

[294] Victor Schetinger, Sara Di Bartolomeo, Mennatallah El-Assady, Andrew McNutt, **Matthias Miller**, João Paulo Apolinário Passos, and Jane Lydia Adams. “Doom or Deliciousness: Challenges and Opportunities for Visualization in the Age of Generative Models”. In: *Comput. Graph. Forum* 43.3 (2023). DOI: 10.31219/osf.io/3jrcm.

[295] Tina Schneider and Harry Plantinga. “Hymnary.org: The Use of a New Database in Hymnology”. In: *Advances in the Study of Inf. and Religion* 2.1 (2012).

[296] Western Michigan University School of Music. *Form exam, Section 2: 18th-century Fugal Analysis*. Retrieved March 29, 2023, from https://wmich.edu/musicgradexamprep/FugueReview.pdf. 2019.

[297] Bill Schottstaedt. “Beyond MIDI: The Handbook of Musical Codes”. In: ed. by Eleanor Selfridge-Field. MIT Press, 1997. Chap. CMN.

[298] Kelly Schroer. *“Visualizing Harmony”*. Retrieved July 9, 2019, from http://kellyschroer.com/harmony.html. 2019.

[299] Hans Jörg Schulz, Zabedul Akbar, and Frank Maurer. “A Generative Layout Approach for Rooted Tree Drawings”. In: *IEEE Pacific Visualization Symp.* 2013. DOI: 10.1109/PacificVis.2013.6596149.

[300] Véronique Sébastien, Didier Sébastien, and Noël Conruyt. “Annotating Works for Music Education: Propositions for a Musical Forms and Structures Ontology and a Musical Performance Ontology”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2013.

[301] Daniel Seebacher, **Matthias Miller**, Tom Polk, Johannes Fuchs, and Daniel A. Keim. “Visual Analytics of Volunteered Geographic Information: Detection and Investigation of Urban Heat Islands”. In: *IEEE Computer Graphics and Applications* 39.5 (2019). DOI: 10.1109/MCG.2019.2926242.

[302] William A. Sethares. *Rhythm and Transforms*. Springer, 2007.

[303] Rita Sevastjanova, Hanna Schäfer, Jürgen Bernard, Daniel A. Keim, and Mennatallah El-Assady. “Shall we play? – Extending the Visual Analytics Design Space through Gameful Design Concepts”. In: *MLUI for Visualization and Analytics, IEEE VIS Workshop*. 2019.

[304] Elona Shatri and György Fazekas. “Optical Music Recognition: State of the Art and Major Challenges”. In: *Proc. of the Int. Conf. on Technologies for Music Notation and Representation*. Hamburg University for Music and Theater, 2020.

[305] Hsuan-Huei Shih, S.S. Narayanan, and C.-C.J. Kuo. “A Dictionary Approach To Repetitive Pattern Finding in Music”. In: *IEEE Int. Conf. on Multimedia and Expo (ICME)*. 2001. DOI: 10.1109/ICME.2001.1237711.

[306] Ben Shneiderman. “The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations”. In: *Proc. of the IEEE Symp. on Visual Languages*. IEEE Comput. Soc. Press, 1996. DOI: 10.1109/VL.1996.545307.

[307] Svenja Simon, Sebastian Mittelstädt, Daniel A. Keim, and Michael Sedlmair. “Bridging the Gap of Domain and Visualization Experts with a Liaison”. In: *Eurographics Conf. on Visualization*. 2015. DOI: 10.2312/eurovisshort.20151137.

[308] Federico Simonetta, Filippo Carnovalini, Nicola Orio, and Antonio Rodà. “Symbolic Music Similarity through a Graph-Based Representation”. In: *Proc. of the Audio Mostly on Sound in Immersion and Emotion*. ACM, 2018. DOI: 10.1145/3243274.3243301.

[309] Federico Simonetta, Stavros Ntalampiras, and Federico Avanzini. “Audio-to-Score Alignment Using Deep Automatic Music Transcription”. In: *Int. Workshop on Multimedia Signal Processing*. IEEE, 2021. DOI: 10.1109/MMSP53017.2021.9733531.

[310] Stephen Sinclair, Michael Droettboom, and Ichiro Fujinaga. “Lilypond for PyScore: Approaching a Universal Translator for Music Notation”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2006.

[311] Sean M. Smith and Glen N. Williams. “A Visualization of Music”. In: *IEEE Visualization ’97*. IEEE, 1997. DOI: 10.1109/VISUAL.1997.663931.

[312] Jon Snydal and Marti A. Hearst. “ImproViz: Visual Explorations of Jazz Improvisations”. In: *Ext. Abstr. Proc. of the Conf. CHI*. ACM, 2005. DOI: 10.1145/1056808.1057027.

[313] Louis Soum-Fontez, Mathieu Giraud, Nicolas Guiomard-Kagan, and Florence Levé. “Symbolic Textural Features and Melody/Accompaniment Detection in String Quartets”. In: *Int. Symp. on Comput. Music Multidisc. Res.* 2021.

[314] Fabian Sperrle, Mennatallah El-Assady, Grace Guo, Rita Borgo, Duen Horng Chau, Alex Endert, and Daniel Keim. “A Survey of Human-Centered Evaluations in Human-Centered Machine Learning”. In: *Comput. Graph. Forum (CGF)* (2021).

[315] John T. Stasko and Eugene Zhang. “Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations”. In: *IEEE Symp. on Inf. Vis.* IEEE Computer Society, 2000. DOI: 10.1109/INFVIS.2000.885091.

[316] Mark J. Steedman. “A Generative Grammar for Jazz Chord Sequences”. In: *Music Perception* 2.1 (1984).

[317] Kurt Stone. “Problems and Methods of Notation”. In: *Perspectives of New Music* 1.2 (1963).

[318] Hope Strayer. “From Neumes to Notes: The Evolution of Music Notation”. In: *Mus. Offerings* 4.1 (2013). DOI: 10.15385/jmo.2013.4.1.1.

[319] K. Swanwick. *Musical Knowledge: Intuition, Analysis and Music Education*. Routledge, 1994.

[320] Eric Robert Taylor. *The AB Guide to Music Theory*. Retrieved April 9, 2019, from https://books.google.com.mt/books?id=h7hZSAAACAAJ. Assoc. Board Royal Schools of Music, 1989.

[321] Eric Thul and Godfried T. Toussaint. “Rhythm Complexity Measures: A Comparison of Mathematical Models of Human Perception and Performance”. In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. 2008.

[322] Renee Timmers and Henkjan Honing. “On music performance, theories, measurement and diversity”. In: *Cognitive Processing* 1.2 (2002).

[323] Petri Toiviainen. “Visualization of Tonal Content with Self-Organizing Maps and Self-Similarity Matrices”. In: *Comput. Entertain.* 3.4 (2005). DOI: 10.1145/1095534.1095543.

[324] Christian Tominski and Wolfgang Aigner. *“TimeViz Browser”*. Retrieved April 4, 2023, from https://browser.timeviz.net/. 2015.

[325] Godfried T. Toussaint. “Computational geometric aspects of rhythm, melody, and voice-leading”. In: *Computational Geometry: Theory and Applications*. 2010. DOI: 10.1016/j.comgeo.2007.01.003.

[326] Godfried T. Toussaint. “The Geometry of Musical Rhythm”. In: *Discrete and Computational Geometry*. 2005. DOI: 10.1007/11589440_20.

[327] Godfried T. Toussaint. *The Geometry of Musical Rhythm: What Makes a "Good" Rhythm Good?* CRC Press, 2019. DOI: 10.1201/9781351247771.

[328] Godfried T. Toussaint, Luke Matthews, Malcolm Campbell, and Naor Brown. "Measuring Musical Rhythm Similarity: Transformation versus Feature-based Methods". In: *Journal of Interdisciplinary Music Studies* 6.1 (2012). DOI: 10.4407/jims.2012.12.002.

[329] Leo Treitler. "Oral, Written, and Literate Process in the Transmission of Medieval Music". In: *Speculum* 56.3 (1981).

[330] Wei-Ho Tsai, Hung-Ming Yu, and Hsin-Min Wang. "Using the Similarity of Main Melodies to Identify Cover Versions of Popular Songs for Music Document Retrieval". In: *Journal of Information Science and Engineering* 24.6 (2008).

[331] Dmitri Tymoczko. "The Generalized Tonnetz". In: *J. of Music Theory* (2012). DOI: 10.1215/00222909-1546958.

[332] Dmitri Tymoczko. "The Geometry of Musical Chords". In: *Science* 313.5783 (July 2006). DOI: 10.1126/science.1126287.

[333] George Tzanetakis, Ajay Kapur, W. Andrew Schloss, and Mathew Wright. "Computational Ethnomusicology". In: *Journal of Interdisciplinary Music Studies* 1 (Jan. 2007).

[334] Simon den Uijl, Henk Jan de Vries, and Deniz Bayramoglu. "The Rise of MP3 as the Market Standard: How Compressed Audio Files Became the Dominant Music Format". In: *Int. J. IT Stand. Stand. Res.* 11.1 (2013). DOI: 10.4018/JITSR.2013010101.

[335] Julián Urbano. *MelodyShape: a Library and Tool for Symbolic Melodic Similarity based on Shape Similarity*. Retrieved September 16, 2020 from https://github.com/julian-urbano/MelodyShape/. 2015.

[336] Julián Urbano, Juan Lloréns, Jorge Morato, and Sonia Sánchez-Cuadrado. "Melodic Similarity through Shape Similarity". In: *Exploring Music Contents - Int. Symp., CMMR 2010*. Vol. 6684. Lecture Notes in Computer Science. Springer, 2010. DOI: 10.1007/978-3-642-23126-1_21.

[337] Michelle Urberg. "Pasts and Futures of Digital Humanities in Musicology: Moving Towards a "Bigger Tent"". In: *Music Ref. Services* 20.3-4 (2017). DOI: 10.1080/10588167.2017.1404301.

[338] Steven J. Vaughan-Nichols. "Will HTML 5 Restandardize the Web?" In: *Computer* 43.4 (2010). DOI: 10.1109/MC.2010.119.

[339] Harsh Verma and John Thickstun. "Convolutional Composer Classification". In: *Proc. Int. Soc. of Music Inf. Retrieval Conf. (ISMIR)*. Ed. by Arthur Flexer, Geoffroy Peeters, Julián Urbano, and Anja Volk. 2019.

[340] Iris Vessey. "Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature". In: *Decision Sciences* 22.2 (1991). DOI: 10.1111/j.1540-5915.1991.tb00344.x.

[341] Tim Vets, Luc Nijs, Micheline Lesaffre, Bart Moens, et al. "Gamified Music Improvisation with BilliArT: A Multimodal Installation with Balls". In: *J. on Multimodal User Interfaces* 11.1 (Mar. 2017). DOI: 10.1007/s12193-016-0224-9.

[342] Antonio Vivaldi. *The Four Seasons Concerto No. 1 - La Primavera (Spring)*. Retrieved March 31, 2023, from https://musescore.com/musikmann/scores/135805. MuseScore.

[343] Joseph Smits van Waesberghe. "The Musical Notation of Guido of Arezzo". In: *Musica Disciplina* 5 (1951).

[344] Chris Walshaw. *ABC Notation – Symbolic Music Notation*. Retrieved March 28, 2023, from https://abcnotation.com/. 1991.

[345] Maike Watson. "MuseScore". In: *Journal of the Musical Arts in Africa* 15.1-2 (July 2018). DOI: 10.2989/18121004.2018.1534342.

[346] Martin Wattenberg. "Arc Diagrams: Visualizing Structure in Strings". In: *IEEE Symp. Information Visualization (InfoVis)*. IEEE Comput. Soc., 2002. DOI: 10.1109/INFVIS.2002.1173155.

[347] Claus Weihs, Dietmar Jannach, Igor Vatolkin, and Günter Rudolph. *Music Data Analysis: Foundations and Applications*. CRC Press/Taylor & Francis, 2017. DOI: 10.1111/anzs.12244.

[348] Anne Weisgerber, Marie Bayot, Eric Constant, and Nicolas Vermeulen. "Music and Cognitive Processing of Emotions: Individual Differences and Psychopathology". In: *Journal of Communications Research* 5.3 (2013).

[349] Christof Weiß, Matthias Mauch, Simon Dixon, and Meinard Müller. "Investigating Style Evolution of Western Classical Music: A Computational Approach". In: *Musicae Scientiae* 23 (2018). DOI: 10.1177/1029864918757595.

[350] Christof Weiß and Meinard Müller. "Quantifying and Visualizing Tonal Complexity". In: *Proceedings of the Conference on Interdisciplinary Musicology*. 2014.

[351] Robert L. Welker. "Abstraction of Themes from Melodic Variations". In: *J. of Experimental Psychology: Human Perception and Performance* 8.3 (1982).

[352] Ru Wen, Zheng Xie, Kai Chen, Ruoxuan Guo, Kuan Xu, Wenmin Huang, Jiyuan Tian, and Jiang Wu. "Music Style Analysis among Haydn, Mozart, and Beethoven: an Unsupervised Machine Learning Approach". In: *Proc. Int. Comput. Music Conf. ICMC*. Michigan Publishing, 2017.

[353] Blake West. *Hummingbird*. Retr. Nov. 13, 2019, from http://www.hummingbirdnotation.com/. 2019.

[354] Huub van de Wetering, Nico Klaassen, and Michael Burch. "Space-Reclaiming Icicle Plots". In: *IEEE Pac. Vis. Symp.* IEEE, 2020. DOI: 10.1109/PacificVis48177.2020.4908.

[355] Sheila Whiteley, Andy Bennett, and Stan Hawkins. *Music, Space and Place: Popular Music and Cultural Identity*. Ashgate Popular and Folk Music Series. Ashgate, 2005.

[356] Gerhard Widmer and Werner Goebl. "Computational Models of Expressive Music Performance: The State of the Art". In: *Journal of New Music Research* 33.3 (Sept. 2004). DOI: 10.1080/0929821042000317804.

[357] Jacek Wolkowicz, Stephen Brooks, and Vlado Keselj. "MIDIVIS: Visualizing Music Structure via Similarity Matrices". In: *Proc. of the International Computer Music Conference*. 2009.

[358] Jacek Michal Wolkowicz. "Application of Text-based Methods of Analysis to Symbolic Music". PhD thesis. Dalhousie University, Halifax, 2013.

[359] Mitchell Wong and Marcel Danesi. "Color, Shape, and Sound: A Proposed System of Music Notation". In: *Semiotica* 2015.204 (Jan. 2015). DOI: 10.1515/sem-2014-0086.

[360] Tiange Zhu, Raphaël Fournier-S'niehotta, Philippe Rigaux, and Nicolas Travers. "A Framework for Content-Based Search in Large Music Collections". In: *Big Data and Cognitive Computing* 6.1 (2022). DOI: 10.3390/bdcc6010023.