Exploring Unexploited Compositional Space in Intercultural, Cross-level, and Concurrence Features of Music

Permalink
https://escholarship.org/uc/item/0jg9d84b

Author
Lin, Hsin-Ming

Publication Date
2017

Peer reviewed|Thesis/dissertation
UNIVERSITY OF CALIFORNIA, SAN DIEGO

Exploring Unexploited Compositional Space in Intercultural, Cross-level, and Concurrence Features of Music

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in

Music

by

Hsin-Ming Lin

Committee in charge:

Professor Shlomo Dubnov, Chair
Professor Natacha D. Diels
Professor Lei Liang
Professor Miller S. Puckette
Professor Akos Rona-Tas

2017
The dissertation of Hsin-Ming Lin is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

________________________________________________

________________________________________________

________________________________________________

________________________________________________

Chair

University of California, San Diego

2017
# Table of Contents

Signature Page .......................................................................................................................... iii

Table of Contents ....................................................................................................................... iv

List of Abbreviations ................................................................................................................... viii

List of Figures ................................................................................................................................. x

List of Tables ................................................................................................................................. xiv

Vita ...................................................................................................................................................... xvi

Abstract of the Dissertation .......................................................................................................... xviii

1. Introduction ............................................................................................................................... 1

1.1 Preliminary ................................................................................................................................ 2

1.1.1 Data Science ....................................................................................................................... 2

1.1.2 Computational Musicology ................................................................................................. 6

1.2 Rationale ................................................................................................................................... 8

1.2.1 Personal Background ......................................................................................................... 8

1.2.2 Exploitation ....................................................................................................................... 10

1.2.3 Mutualism .......................................................................................................................... 12

1.3 Levels of Feature ...................................................................................................................... 16

1.3.1 Low .................................................................................................................................... 17

1.3.2 Middle ............................................................................................................................... 19

1.3.3 High .................................................................................................................................... 20

1.3.4 Meta .................................................................................................................................... 22

1.4 Levels of Representation ......................................................................................................... 22
3. Exploration
3.1 Observation
   3.1.1 Dataset by Dataset
   3.1.2 Sample by Sample
   3.1.3 Adding Tessitura and Mobility
3.2 Perspectives
   3.2.1 External
   3.2.2 Internal
3.3 Evaluation
   3.3.1 Balanced Dataset
   3.3.2 Feature Sets
   3.3.3 Classification
   3.3.4 Discussion
3.4 Composition
   3.4.1 Fundamentals
   3.4.2 Outliers
   3.4.3 Experiments
4. Conclusion
   4.1 Contribution
   4.2 Application
   4.3 Limitation
   4.4 Expectation
List of Abbreviations

AC: algorithmic composition
ANN: artificial neural network
ANOVA: analysis of variance
CAAC: computer-aided algorithmic composition
CAC: computer assisted composition
CMA: computational music analysis
CV: coefficient of variation
EFC: Essen folksong collection
JRP: Josquin research project
K-NN: k-nearest neighbors
LLOCV: leave-one-out cross-validation
MIDI: musical instrument digital interface
MIR: music information retrieval
PCC: Pearson correlation coefficient
RBF: radial basis function
SD: standard deviation
SQL: structured query language
SVM: support vector machine
Va: viola
Vc: violoncello
Vn-1: first violin
Vn-2: second violin

WJazzD: Weimar jazz database

WTC: well-tempered clavier
List of Figures

Figure 1. Data Science Process [2].................................................................3
Figure 2. Data Scientist’s Involvement [2]..................................................3
Figure 3. Example of Prediction [2]. ...............................................................4
Figure 4. Examples of Statistical Learning Algorithms [2] (originally from [3]). ....5
Figure 5. Example of Classification [2]..........................................................6
Figure 6. Evolution of Musical Topics in the Billboard Hot 100 [7].................7
Figure 7. Audio Features Clustering for Style Analysis of Classical Music [8]. ....7
Figure 8. Expectancy Based Complexity and Originality [21]............................11
Figure 9. Features from 250 Kiki Recordings and 250 Bouba Recordings [22]. ....11
Figure 10. Babbitt’s Trinity of Representational Domains with Wiggins’s
Transformations [25].....................................................................................13
Figure 11. Fundamental Realities of Music [23].............................................14
Figure 12. Stages of Communication in Music [23]........................................14
Figure 13. Positions in Musical Semiotics [23].............................................14
Figure 14. Taxonomy for Content-based MIR [29]......................................16
Figure 15. Characterization of Audio Features [31]......................................17
Figure 16. Pitch Features [33]......................................................................18
Figure 17. High-level Features [35]................................................................21
Figure 18. Universal Relationships between 32 Musical Features [36].............22
Figure 19. Comparison of Music Representation Systems [38]......................24
Figure 20. Representation Schemes [39]......................................................25
Figure 21. Symbolic Music Representations [33].........................................26
Figure 22. Example of LilyPond [40]............................................................27
Figure 23. Example of Humdrum "**kern" [41]................................................................. 28
Figure 24. Example Score [42]................................................................................................. 29
Figure 25. Example of MusicXML Source Code [42] of the First Chord in the Piano Part of Figure 20................................................................................................................... 29
Figure 26. Middle C (262 Hz) Played on a Piano and a Violin [49].................................. 31
Figure 27. Stages in Automatic Melodic Transcription [34]................................................... 31
Figure 28. Melodic Pitch Contour Extraction [51].................................................................. 33
Figure 29. Pitch Contours [51].................................................................................................. 33
Figure 30. Melody Selection [51].............................................................................................. 34
Figure 31. Vocal Jazz Excerpt [51].......................................................................................... 34
Figure 32. Roles of Manual Segmentation and Correction [32].......................................... 35
Figure 33. MedleyDB Dataset [52]........................................................................................... 36
Figure 34. Transcription of a San (Bushman) Song with Musical Bow Accompaniment by Four Ethnomusicologists at the SEM Symposium on Transcription and Analysis in 1964 [29]................................................................................................................... 37
Figure 35. The Yale-Classical Archives Corpus [57]................................................................ 39
Figure 36. One of the Scorings of Turkish Folk Song [61]..................................................... 42
Figure 37. A Cantus Firmus [83]................................................................................................ 49
Figure 38. Melodic Pitch Contour (middle line) of Figure 37 [17].................................. 50
Figure 39. A German Folk Song [14]........................................................................................ 51
Figure 40. Pitch Histogram of the Melody in Figure 39 [14]................................................. 51
Figure 41. Index of the Mobility of the Melody in Figure 39 [14]......................................... 52
Figure 42. Texture Space [66]................................................................................................... 54
Figure 43. Concurrence and Non-concurrence between Pitch and Intensity Curves in Syllables of Bird Calls (Babblers) [64]................................................................ 56
Figure 44. A Guqin Tone with Rhythmic Vibrato and Fluctuation Patterns [29]....... 59
Figure 45. Distribution of Mean “Average Melodic Interval” and Mean “Most Common Melodic Interval” from All Datasets in Table 11. ........................................ 64

Figure 46. Distribution of Mean “Register Width” and Mean “Susceptibility” from All Datasets in Table 12. .......................................................................................... 65

Figure 47. Distribution of Register Width and Susceptibility from All (except 4) Samples in Table 12. .......................................................................................... 66

Figure 48. Distribution of Register Width and Index of Tessitura from Samples in Table 12 except the MedleyDB Dataset. ................................................................. 68

Figure 49. Distribution of Index of Tessitura and Susceptibility from Samples in Table 12 except the MedleyDB Dataset. ................................................................. 69

Figure 50. Distribution of Index of Tessitura and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset. ................................................................. 70

Figure 51. Distribution of Register Width and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset. ................................................................. 70

Figure 52. Distribution of Susceptibility and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset. ................................................................. 71

Figure 53. Ranges and Means of Register Width. .................................................................................. 74

Figure 54. Ranges and Means of Susceptibility. .................................................................................. 75

Figure 55. Ranges and Means of PCC between Local Susceptibilities and Register Widths by Cumulative-size Windowing. ................................................................. 77

Figure 56. Ranges and Means of PCC between Local Susceptibilities and Register Widths by Fixed-size Windowing. ................................................................. 77

Figure 57. Melodic Profile of the 1st Movement of Mozart’s Piano Sonata No. 16 in C major, K. 545. .......................................................................................... 78

Figure 58. Melodic Profile of the 1st Movement of Beethoven’s Piano Sonata No. 21 in C major, Op.53. .......................................................................................... 79

Figure 59. Distribution and Path of Local Features from 1st Violin Part in the 1st Movement of Mozart’s String Quartet No. 15 in D minor, K. 421................................. 80

Figure 60. Distribution and Path of Local Features from Violoncello Part in the 1st Movement of Mozart’s String Quartet No. 15 in D minor, K. 421................................. 81
Figure 61. Human Classification Accuracy of Haydn’s and Mozart’s String Quartet Movements [91]........................................................................................................................................ 83

Figure 62. Accuracy on 1st Violin......................................................................................................................... 87

Figure 63. Accuracy on 2nd Violin.......................................................................................................................... 88

Figure 64. Accuracy on Viola................................................................................................................................. 88

Figure 65. Accuracy on Violoncello......................................................................................................................... 89

Figure 66. 1st Violin Distribution of M19 and M17 in Table 24. ................................................................. 94

Figure 67. Violoncello Distribution of P2+P10 and M6+M18 in Table 24. ......................................... 94

Figure 68. Ascending Diatonic Scale. ..................................................................................................................... 96

Figure 69. Melodic Pitch Contour (middle line) of Figure 68.............................................................. 96

Figure 70. Descending Diatonic Scale................................................................................................................... 96

Figure 71. Melodic Pitch Contour (middle line) of Figure 70.............................................................. 97

Figure 72. Ascending Chromatic Scale. .............................................................................................................. 97

Figure 73. Descending Chromatic Scale. .............................................................................................................. 97

Figure 74. Ascending Skips with Returning Steps. .............................................................................................. 98

Figure 75. Descending Skips and Returning Steps.............................................................................................. 98

Figure 76. Score of the “han0261.krn” File [53] in Table 16......................................................................... 99

Figure 77. Score of the “han1078.krn” File [53] in Table 16.......................................................................... 99

Figure 78. Experimental Composition 1........................................................................................................... 100

Figure 79. Distribution of Register Width and Susceptibility from All Samples in Figure 47 and Four Unprecedented Samples.................................................................................. 100

Figure 80. Experimental Arrangement................................................................................................................ 101

Figure 81. Experimental Composition 2......................................................................................................... 101

Figure 82. Experimental Composition 3......................................................................................................... 102

Figure 83. Screenshot of My Old Program [11]. .............................................................................................. 103
List of Tables

Table 1. Examples of MIR Tasks and Their Specificities [9]................................. 8
Table 2. Time Scale and Characteristics of the Different Dimensions of Music [30]. 17
Table 3. Common Low-level Features Used in Music Classification [31]. .......... 19
Table 4. Musicological Characteristics [32].......................................................... 19
Table 5. Features for MIR Tasks [9]........................................................................ 20
Table 6. Text vs. Music [37].................................................................................... 23
Table 7. Basic Representations of Music [37].......................................................... 23
Table 8. Types of Standard MIDI File (SMF).......................................................... 26
Table 9. Contents of the Meertens Tune Collections [59].................................. 41
Table 10. SymbTr Representation of the Score in Figure 36 [61]....................... 42
Table 11. Retrieved Quantities by music21 Feature Extractors. ....................... 45
Table 12. Retrieved Quantities by My Feature Extractors................................ 45
Table 13. Adopted Packages from Anaconda Python ........................................ 46
Table 14. Retrieval Information from Figure 37 [17].......................................... 50
Table 15. Example of Data in Audio Frames from MedleyDB [52]. ............... 59
Table 16. The 30 Outliers in Figure 47.................................................................. 67
Table 17. Types of Window Size and Example of Each Window’s Address........ 72
Table 18. Global Feature-to-feature PCCs in Instrumental Samples.................... 73
Table 19. Global Feature-to-feature PCCs in Vocal Samples............................... 73
Table 20. Correlations between Susceptibilities and Register Widths in Each Part of String Quartets [17].................................................................. 75
Table 21. Correlations between Susceptibilities and Register Width in the First Part of Piano Sonatas [17].................................................................. 75
Table 22. Mean Classification Accuracy across All Results of Different Numbers of Feature After Feature Selection................................................................. 86

Table 23. Mean Feature Scores........................................................................................................... 90

Table 24. Top 20 Features with the Best Mean Scores. ................................................................. 91

Table 25. The Most Significant Internal Correlations. .............................................................. 93

Table 26. Susceptibilities of the Datasets (mean ≥ 6)................................................................. 111

Table 27. Susceptibilities of the Datasets (5 ≤ mean < 6). ....................................................... 112

Table 28. Susceptibilities of the Datasets (mean < 5)................................................................. 113
Vita

2017  Doctor of Philosophy (Computer Music), Department of Music, UC San Diego

2017  Teaching Assistant, Department of Music, UC San Diego

2016  Associate Instructor, Department of Music, UC San Diego

2012–2013  Research Assistant, Music and Audio Computing Laboratory, Research Center for Information Technology Innovation, Academia Sinica

2011–2012  Engineer, Department of Audiovisual Technology, Zuei Ing Co., Ltd.

2011  Master of Arts (Music Technology), Institute of Music, National Chiao Tung University

2008–2010  Teaching Assistant, Center for General Education, National Chiao Tung University

2006–2007  Teacher, Department of Music, Chung-Hwa School of Arts


2005  Bachelor of Fine Arts (Theory and Composition), Department of Music, National Sun Yat-sen University

Publications


Abstract of the Dissertation

Exploring Unexploited Compositional Space in Intercultural, Cross-level, and Concurrence Features of Music

by

Hsin-Ming Lin

Doctor of Philosophy in Music

University of California, San Diego, 2017

Professor Shlomo Dubnov, Chair

Pioneer composers always try to explore beyond frontiers. Nowadays, they are allowed to listen, read, and play music with machines. Music information retrieval (MIR) technology has brought revolutionary computational music analysis (CMA) in recent years. It provides novel ways to dissect music. On the other hand, algorithmic
composition (AC) is able to generate countless pieces with explicit ground truth for MIR experiments. While most data scientists have been seeking the best features which are perfect for discrimination between samples in dissimilar classes, little attention was paid to unclaimed territories in any dimensional feature space. If somewhere there is a tiny ratio of outliers from diverse classes, the unexploited parts might be worth to explore.

To investigate, specific intercultural and cross-level features are implemented to extract from 18009 symbolic and audio samples in 91 sub-datasets which come from various genres in five eras across three continents. Next, the indices of tessitura and mobility are also extracted from all symbolic samples. Besides, 32 jSymbolic features from fewer samples are selected to realize the (non-)concurrence features based on internal correlations between local base features. Final results confirm that vocal styles broadly have larger susceptibilities and narrower register widths than instrumental styles in average. Distributions reveal the unexploited areas in two-dimensional space. Evaluations illustrate (non-)concurrence features’ performance and improvement on the balanced composer classification of Haydn and Mozart’s separated string quartet parts. Manual compositions demonstrate several successes to penetrate the frontiers.

Although modern artificial neural network may automatically learn features to detect unexploited compositional space, those features are not necessary perceivable by human or even not tractable through AC. Contemporary composers have to, however, select controllable features and devise appropriate algorithms and
parameters unless they do not want to compose by their own. This research initiates innovative praxis. Its ultimate goal is to evolve toward the mutualism. Composers learn from the dimensional feature space and distributions through MIR and CMA with intent to devise better algorithms and parameters to manipulate in AC. Then AC has the capability to promote better techniques and features for MIR and CMA, which again stimulate composer's imagination and creativity.
1. Introduction

Ultimately, disciplines are defined not by their methods but by the questions they ask. The development of new methods, however, can often make it easier to pursue certain questions. Conscientious scholars focus on the questions, and then acquire whatever tools best allow them to address those questions.

David Brian Huron [1]
Following the advancement of computation and storage technology, data sciences become more and more influential. Typical purposes include analyzing existing data to predict unknown or future data, whereas I find it could also point a direction of an innovative compositional strategy to intentionally generate novel data. In this chapter, I will firstly provide some preliminary knowledge. Next, I will elaborate my rationale. Finally I will review different levels of feature, levels of representation, and suggest some datasets before my own implementation in next chapter.

1.1 Preliminary

Data sciences have ubiquitously influence our daily life even though we are unaware. Companies are collecting customer’s information and analyzing behavior in many ways. I will introduce from their general to specific applications so that I can properly explain my rational in next section.

1.1.1 Data Science

Experts have been employing tools and techniques from data sciences on purpose to help them make decisions. At first, they have questions in mind. Next, they collect or record data. Third, the raw data has to be processed, cleaned, and analyzed through various statistical algorithms. Fourth, the analytical result usually has to be visualized for people to understand. Finally, they decide whether it is enough to answer the original question, or there are more new questions, as illustrated in Figure 1 and Figure 2.
Figure 1. Data Science Process [2].

Figure 2. Data Scientist’s Involvement [2].
Prediction tasks are the most common applications of data science. Data scientists analyze user behavior to know the time users spent on a social networking site. They find its relationship with the number of new friends. Therefore, they can find a prediction line via linear regression as demonstrated in Figure 3. One presumable hypothesis is that the more new friends a user has, the more time the user spends on the site.

Figure 3. Example of Prediction [2]. The best prediction line has minimal errors (vertical distances).

The above-mentioned prediction values refer to another axis, which is continuous data; if they refer to discrete data, e.g. labels, this kind of application is a classification task. For instance, credit rating may result from a function of age and income. Through statistical learning algorithms such as linear regression and k-nearest neighbors (KNN), new customers might be classified in light of their data location in the bidimensional space, cf. Figure 4. If a customer’s data is right in the middle of one
cluster, for instance, age 55 and income 65000, it is probable to be rated as “high credit”, equal to the other samples in the cluster as scattered in Figure 5. In reality, scientists often add more variables to construct the higher-dimensional space, which is much more difficult to image in human mind. Thereby, before modern people have techniques from data sciences and tools from information technology, it was never easy to find or address questions in a higher-dimensional space.

**Figure 4.** Examples of Statistical Learning Algorithms [2] (originally from [3]).
left: linear regression; middle: 15-NN; right: 1-NN, with an overfitting issue not discussed here. Any new data in the orange (upper left) regions will be labeled the same class as other orange circles. Any new data in the blue (lower right) regions will be labeled the same class as other blue circles.
1.1.2 Computational Musicology

The music industry uses analytics for many reasons, such as:

1. to determine how tracks are performing
2. to decide who gets paid for airplay or purchases, and funnel that money to the right creators and licensors
3. to best spend marketing dollars
4. to plan out tour dates and timing
5. to understand and embrace growing music trends
6. to set the price of advertising and sponsorship

Alistair Croll [4]

In music technology, data sciences are versatile, too. Researchers analyze not only user behavior (context-based) but also music itself (content-based). The “music information retrieval” [5] (MIR) technology has brought revolutionary “computational music analysis” [6] (CMA) in recent years. It has the capability to disclose historical trends as delineated in Figure 6 and to discriminate typical styles as illuminated in
Figure 7. Numerous MIR applications as sorted in Table 1 are now indispensable to commercial music industry.

Figure 6. Evolution of Musical Topics in the Billboard Hot 100 [7].

Figure 7. Audio Features Clustering for Style Analysis of Classical Music [8].
Table 1. Examples of MIR Tasks and Their Specificities [9].

<table>
<thead>
<tr>
<th>Use Case</th>
<th>Specificity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music Identification</td>
<td>H</td>
<td>Identify a compact disk, provide metadata about an unknown track, mobile music information retrieval: e.g. shazam.com</td>
</tr>
<tr>
<td>Plagiarism detection</td>
<td>H</td>
<td>Identify mis-attribution of musical performances, mis-appropriation of music intellectual property.</td>
</tr>
<tr>
<td>Copyright monitoring</td>
<td>H</td>
<td>Monitor music broadcast for copyright infringement or royalty collection</td>
</tr>
<tr>
<td>Versions</td>
<td>H/M</td>
<td>Remixes, live vs. studio recordings, cover songs. Used for database normalization and near-duplicate results elimination</td>
</tr>
<tr>
<td>Melody</td>
<td>H/M</td>
<td>Find works containing a melodic fragment</td>
</tr>
<tr>
<td>Identical Work / Title</td>
<td>M</td>
<td>Retrieve performances of same opus number or song title</td>
</tr>
<tr>
<td>Performer</td>
<td>M</td>
<td>Find music by a specific artist</td>
</tr>
<tr>
<td>Sounds like</td>
<td>M</td>
<td>Find music that sounds like a given recording</td>
</tr>
<tr>
<td>Performance Alignment</td>
<td>M</td>
<td>Mapping one performance onto another independent of tempo and repetition structure</td>
</tr>
<tr>
<td>Composer</td>
<td>M</td>
<td>Find works by one composer</td>
</tr>
<tr>
<td>Recommendation</td>
<td>M/L</td>
<td>Find music that matches the user’s personal profile</td>
</tr>
<tr>
<td>Mood</td>
<td>L</td>
<td>Find music using emotional concepts: Joy, Energetic, Melancholy, Relaxing</td>
</tr>
<tr>
<td>Style / Genre</td>
<td>L</td>
<td>Find music that belongs to a generic category: Jazz, Funk, Female Vocal</td>
</tr>
<tr>
<td>Instrument(s)</td>
<td>L</td>
<td>Find works with same instrumentation</td>
</tr>
<tr>
<td>Music-Speech</td>
<td>L</td>
<td>Radio broadcast segmentation, Music archives cataloguing</td>
</tr>
</tbody>
</table>

1.2 Rationale

In this sub-chapter, I will elaborate my reason that I am interested in the exploitation in dimensional feature space of music and my motivation to learn something from it to mutually benefit (algorithmic) composers and (music) data scientists.

1.2.1 Personal Background

During years 2010 and 2011, I created an algorithmic composition (AC) model to generate monophonic melody [10]. It borrowed “the law of conservation of energy”
from physics to mimic and to manipulate the natural melodic pitch contours [11].

After my master’s degree, I had two-year work experience in a karaoke machine company and then in Academia Sinica, the Taiwan’s national central research institute. Those jobs brought me perspectives on MIR and CMA. After entering University of California, San Diego in 2013, I contemplated transforming my compositional algorithm into an analytical program which enables me to observe abundant pre-existing music pieces from my own perspective. In 2014, I started learning Python programming language [12] to adopt the music21 library [13] since other retrieval tools for symbolic music data were outdated, immature, or inconvenient for me. Besides revising my model and gathering more datasets, I tested other musical features, too. At first, I concentrated on melodic characteristics, e.g. “index of mobility” and “index of tessitura” [14]; later, I considered more types of high-level features, e.g. [15]-[16]. In 2015, I looked trends and relationships. I also began to care about intercultural and cross-level potentialities owing to the advantages in my original model [17].

Furthermore, in my past full-time jobs from 2011 to 2013, I realized that many, if not most, users always want to tag their own compositions as “mixed” or “hybrid” in terms of either genre or style. They refuse to accept typical tags like classical, pop, rock, jazz, and so on, even if their compositions are very close to one or some of the tags (no matter the similarity is judged objectively by classification programs or subjectively by music experts). If people really want to compose mixed or hybrid ones,
they have to make the music more ambiguous to classify. An intriguing path is to borrow data sciences to explore the possibility.

1.2.2 Exploitation

Algorithmic composition (AC) is usually being applied in two ways: imitation and innovation, in other words, “generating music imitating a corpus of compositions or a specific style” and “automating composition tasks to varying degrees, from designing mere tools for human composers, to generating compositions without human intervention.” [18] By Nierhaus’s definition, to use “algorithmic procedures as supplementary compositional tools” is “computer assisted composition (CAC) in a narrower sense of the expression.” [19] When the degree of automation is low, Fernández and Vico also conceded AC to be a “computer-aided algorithmic composition (CAAC)” [18]. They, however, argued that “the divide between both ends of the spectrum of automation [...] is not clear, because any method that automates the generation of creative works can be used as a tool to aid composers, and systems with higher degrees of automation can be custom-built on top of many CAAC frameworks.” [18] Therefore, in this dissertation, I simply mention AC to include and imply any possible CAC or CAAC.

For the imitation by AC, researchers have applied data sciences to measure its success as compared in Figure 8; for the innovation by AC, it is able to generate limitless free pieces with explicit ground truth for MIR experiments as the two artificial genres which are analyzed in Figure 9. Combining both, AC with MIR is able to mix multiple composers’ styles together to generate a hybrid, e.g. [20].
Diverging from the applications of data sciences in Chapter 1.1, I feel that to search unexploited compositional space is very interesting. A musical feature is a dimension which human or machine can measure music pieces; multiple features create a higher-dimensional space. If somewhere there are sparse distributions of
music samples, it could be a unique or a barely acceptable space to compose. If somewhere there is, however, entirely empty distribution, it may be either a whole new or an absolutely unacceptable space, or may result from the inadequate or unbalanced representative data. I will give concrete examples in following chapters. Above all, my approach is somewhat contradictory to common applications of data sciences. Scientists usually select powerful features to facilitate predication and classification. Conversely, I prefer features which may be not so useful for traditional tasks but have the capability to unveil the space which was unexploited by our predecessors.

1.2.3 Mutualism

Mazzola, Park, and Thalmann have claimed that “music is an ever-evolving field of artistic and scientific expression” [23]. In previous ages, people learn composing through listening to music and reading scores. In the current age of data science, a person is allowed to listen with machine [24] and read with machine. Instead of inspiration itself, listening and reading are approaches to inspire composers. My motivation is to find unique perspectives on compositional strategy. Composers are allowed to follow or diverge from our predecessors in the space constructed by musical features (or dimensions).

Wiggins has described music as “an abstract and intangible cultural construct, which does not have physical existence in itself, but which is described by all three of these representational domains” [25] as connected in Figure 10. Nowadays, there are few datasets in the auditory (or perceived) domain but numerous datasets in the other
two domains. Hence, I focus on the graphemic (or notated) domain and the acoustic (or physical) domain.

![Diagram](image)

**Figure 10.** Babbitt’s Trinity of Representational Domains with Wiggins’s Transformations [25]. also cf. Figure 11 on page 14.

Data sciences have no role as long as there is no data. Fortunately, symbolic and audio datasets at least offer us the other two aspects of music as indicated in Figure 10 above and Figure 11 below. Their scores and recordings are composers’ expression in the middle of communication to listeners as shown in Figure 12. From the expression, listeners have to identify its significations in order to understand its content as displayed in Figure 13. A listener could be either a human audience or a computer analyzer. Nevertheless, Marsden advised that “the objective of computational music analysis should probably not be to generate ‘an analysis’ but rather, like forensic science, to answer specific music-analytical questions with a degree of complexity, speed and accuracy which is impossible by other means” [26].
On the other hand, most musical artistic effects are contextual, which often contradict the essence of statistical “global” features (i.e. one value from one piece).
For this reason, I propose to look into the internal variation, distribution, correlation, and relationship between “local” (or “sectional”) features [17] (i.e. multiple values from one piece). I will elaborate in Chapter 3.2 and Chapter 3.3.

Moreover, Sturm has warned that “what can be measured precisely or reliably does not mean it is relevant” [27]. Some, if not many, musical or audio features are neither perceivable nor accessible by human mind. Even though some are perceivable, audiences are prone to misunderstand and misinterpret. For instance, Huron discovered the differences between “how melodies are notationally structured” and “how experienced listeners think they are structured” [28].

Nevertheless, through investigating those features, composers could learn how to create new perceivable ones or how to handle old ones in another way. Although modern artificial neural network (ANN) may automatically learn features to detect unexploited compositional space, those features are not necessary perceivable by human or even not tractable in AC. Contemporary composers need to manually select. A feature is extractable from music but may not directly controllable; only what I call a parameter is the governable one. To manipulate a feature may require a composer to simultaneously adjust different parameters. On purpose to facilitate the progress, scientists had better select features which are more musicologically meaningful so that they are allowed to be deployed during AC procedures. Next, experts should consider AC parameters which might be effective to fulfill some CMA tasks, too. Finally, if composers know more about MIR features, they may devise more innovative
algorithms and parameters to compose or to analyze. As a result, it becomes a mutualism [17].

1.3 Levels of Feature

Features have various levels of abstraction as categorized in Figure 14 and varied terms of time scale as listed in Table 2. All of them may overlap each other to a certain extent because the categorization is a relative degree rather than an absolute value. Considering both aspects above, there are diverse features over those levels and terms as divided in Figure 15. A semantic gap intervenes between top-level and the other levels. Nonetheless, Wiggins has argued that “whether music can really be said to have ‘semantics’ is an ongoing debate” [25].

![Figure 14. Taxonomy for Content-based MIR][29]
Table 2. Time Scale and Characteristics of the Different Dimensions of Music [30].

<table>
<thead>
<tr>
<th>Time scale</th>
<th>Dimension</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short term</td>
<td>Timbre</td>
<td>Quality of the produced sound</td>
</tr>
<tr>
<td></td>
<td>Orchestration</td>
<td>Sources of sound production</td>
</tr>
<tr>
<td></td>
<td>Acoustics</td>
<td>Quality of the recorded sound</td>
</tr>
<tr>
<td>Middle term</td>
<td>Rhythm</td>
<td>Patterns of sound onsets</td>
</tr>
<tr>
<td></td>
<td>Melody</td>
<td>Sequences of notes</td>
</tr>
<tr>
<td></td>
<td>Harmony</td>
<td>Sequences of chords</td>
</tr>
<tr>
<td>Long term</td>
<td>Structure</td>
<td>Organization of the musical work</td>
</tr>
</tbody>
</table>

Figure 15. Characterization of Audio Features [31].

1.3.1 Low

Low-level features are the fundamental information about pitch, duration, loudness, and timbre extracted from the audio signal as highlighted in Figure 16 and Table 3. Some of them are able to reflect musicological characteristics as juxtaposed in Table 4. Interdisciplinary researchers, however, “have not yet agreed on a
methodology from an MIR point of view for approaching descriptions such as ‘sweet’ (tian), ‘mellow’ (run), ‘fragile’ (cui), ‘round’ (yuan), or ‘wide’ (kuan)” [32].

**Figure 16.** Pitch Features [33].
(a) sheet music (b) pitch-based log-frequency spectrogram (c) chromagram.
Table 3. Common Low-level Features Used in Music Classification [31].

<table>
<thead>
<tr>
<th>Class</th>
<th>Feature Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timbre</td>
<td>Zero Crossing Rate (ZCR)</td>
</tr>
<tr>
<td></td>
<td>Spectral Centroid (SC)</td>
</tr>
<tr>
<td></td>
<td>Spectral Rollloff (SR)</td>
</tr>
<tr>
<td></td>
<td>Spectral Flux (SF)</td>
</tr>
<tr>
<td></td>
<td>Spectral Bandwidth (SB)</td>
</tr>
<tr>
<td></td>
<td>Spectral Flateness Measure (SFM)</td>
</tr>
<tr>
<td></td>
<td>Spectral Crest Factor (SCF)</td>
</tr>
<tr>
<td></td>
<td>Amplitude Spectrum Envelope (ASE)</td>
</tr>
<tr>
<td></td>
<td>Octave based Spectral Contrast (OSC)</td>
</tr>
<tr>
<td></td>
<td>Daubechies Wavelet Coef Histogram (DWCH)</td>
</tr>
<tr>
<td></td>
<td>Mel-frequency Cepstrum Coefficient (MFCC)</td>
</tr>
<tr>
<td></td>
<td>Fourier Cepstrum Coefficient</td>
</tr>
<tr>
<td></td>
<td>Linear Predictive Cepstrum Coefficient (LPCC)</td>
</tr>
<tr>
<td></td>
<td>Stereo Panning Spectrum Features (SPSF)</td>
</tr>
<tr>
<td>Temporal</td>
<td>Statistical Moments (SM)</td>
</tr>
<tr>
<td></td>
<td>Amplitude Modulation (AM)</td>
</tr>
<tr>
<td></td>
<td>Auto-Regressive Modeling (ARM)</td>
</tr>
</tbody>
</table>

Table 4. Musicological Characteristics [32].

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Audio features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch register</td>
<td>Pitch histogram (1st degree)</td>
</tr>
<tr>
<td>Vibrato variability</td>
<td>Vibrato rate (SD)</td>
</tr>
<tr>
<td>Volume variability</td>
<td>Loudness (SD)</td>
</tr>
<tr>
<td>Brightness</td>
<td>Spectral centroid (mean) LTAS Tristimulus</td>
</tr>
<tr>
<td>Timbre variability</td>
<td>Spectral flux (mean)</td>
</tr>
</tbody>
</table>

1.3.2 Middle

Features in the next level are close to general terms of music theory. Those kinds of information are essential to many MIR tasks as listed in Table 5. Several of them are extractable from both audio and symbolic data. In some research, those features are already in a relative high level.
1.3.3 High

Again, a level is a relative degree, not an absolute value. Equivalent features may exist in two or more levels. In an even higher level, there are more textural, structural, and contextual features as enumerated in Figure 17. Several of them like “dance accompaniment” and “sex segregation” as reported in Figure 18 are merely extractable from metadata. There are also special features which solely applicable to a specific musical repertoire such as “a cappella flamenco singing, more specifically in debla and martinete styles” [34].
Figure 17. High-level Features [35].
(a) the musical hierarchy (b) 26 structural characters.
Figure 18. Universal Relationships between 32 Musical Features [36].

1.3.4 Meta

The top level belongs to features which include but are not limited to genre, style, mood, emotion perception, and emotion induction. They are mostly related to human perception in the auditory domain which is summarized in the beginning of Chapter 1.3. Sometimes, meta-level features are barely extractable from music and are very exceptionable seeing that “ground truth’ in music genre recognition, music emotion recognition and music autotagging, for instance, has notorious ambiguity” [27].

1.4 Levels of Representation

In addition to features, levels can also denote classes of data representation as exemplified in Table 6. Generally speaking, lower-level features are easier to be extracted from lower-level representation, and vice versa. With respect to conversion,
transforming from higher level to lower level of representation is relatively simpler than from lower to higher as compared in Table 7.

**Table 6.** Text vs. Music [37].

<table>
<thead>
<tr>
<th>Explicit structure</th>
<th>minimum</th>
<th>medium</th>
<th>maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Music representation (and examples)</td>
<td>Audio (CD, MP3)</td>
<td>Events (Standard MIDI File)</td>
<td>Music Notation (sheet music)</td>
</tr>
<tr>
<td>Text representation (and examples)</td>
<td>Audio (speech)</td>
<td>ordinary text</td>
<td>text with markup (HTML)</td>
</tr>
</tbody>
</table>

**Table 7.** Basic Representations of Music [37].

<table>
<thead>
<tr>
<th>Representation</th>
<th>Audio</th>
<th>Time-stamped Events</th>
<th>Music Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common examples</td>
<td>CD, MP3 file</td>
<td>Standard MIDI File</td>
<td>sheet music</td>
</tr>
<tr>
<td>Unit</td>
<td>sample</td>
<td>event</td>
<td>note, clef, lyric, etc.</td>
</tr>
<tr>
<td>Explicit structure</td>
<td>none</td>
<td>little (partial voicing information)</td>
<td>much (complete voicing information)</td>
</tr>
<tr>
<td>Avg. rel. storage</td>
<td>2000</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Convert to left</td>
<td>-</td>
<td>easy</td>
<td>OK job: easy</td>
</tr>
<tr>
<td>Convert to right</td>
<td>1 note/time: pretty easy; 2 notes/time: hard; other: very hard</td>
<td>OK job: fairly hard</td>
<td>-</td>
</tr>
<tr>
<td>Ideal for</td>
<td>music bird/animal sounds sound effects speech</td>
<td>music</td>
<td>music</td>
</tr>
</tbody>
</table>

Wiggins et al. has compared several representation systems with regard to range of raw data and range of manageable structures information, namely expressive completeness and structural generality as depicted in Figure 19. Selfridge-Field and Sapp have sorted more systems into three schemes as enumerated in Figure 20. More research can be found in papers from the Music Encoding Conference, which is hold
by Music Encoding Initiative annually. Below I select four of the most prevalent examples.

Figure 19. Comparison of Music Representation Systems [38].
4.1 Notation Codes
- 4.1.1 Monophonic Codes: EsAC and Plaine and Easie Code
  - 4.1.1.1 EsAC
  - 4.1.1.2 Plaine and Easie Code
- 4.1.2 Typesetting Codes: CMN, MusiXTeX, Lilypond, Guido, ABCplus
- 4.1.3 Polyphonic Codes: DARMS and SCORE
  - 4.1.3.1 DARMS
  - 4.1.3.2 SCORE
- 4.1.4 Shareware and Open-Source Codes
  - 4.1.4.1 Lilypond
  - 4.1.4.2 MuseScore

4.2 Sound-Related Codes
- 4.2.1 MIDI
  - 4.2.1.1 Main parameters and global-variable equivalents
  - 4.2.1.2 SMF Sections
  - 4.2.1.3 General MIDI Instrument Specifications
  - 4.2.1.4 Efforts to Refine and Extend the SMF
- 4.2.2 Music V and Csound
- 4.2.3 Conducting Cues and Expressive Overlays

4.3 Sychronic Approaches to Music Representation
- 4.3.1 MuseData
- 4.3.2 Humdrum (kern)

Figure 20. Representation Schemes [39].

1.4.1 MIDI

To the general public, the MIDI (Musical Instrument Digital Interface) binary code format is the most popular standard since 1980s. It can be interchanged through a wide variety of computer programs and electronic instruments. Despite the fact that it records pitch, duration, and loudness quite close to performance, the absence of information like time signatures and enharmonic spellings is infamous. Thereupon, higher-level features are sometimes troublesome to be correctly extracted from MIDI files as exhibited in Figure 21 and Table 8.
Figure 21. Symbolic Music Representations [33].
(a) sheet music (b) MIDI type 0 in a simplified tabular form (c) piano roll.

Table 8. Types of Standard MIDI File (SMF).
Tracks in type 2 are not necessarily being played simultaneously.

<table>
<thead>
<tr>
<th>type</th>
<th>track(s)</th>
<th>channel(s)</th>
<th>popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>≧ 1</td>
<td>popular</td>
</tr>
<tr>
<td>1</td>
<td>≧ 1 (parallel)</td>
<td>≧ 1</td>
<td>popular</td>
</tr>
<tr>
<td>2</td>
<td>≧ 1 (serial)</td>
<td>≧ 1</td>
<td>rare</td>
</tr>
</tbody>
</table>

1.4.2 LilyPond

LilyPond was derived from MusiXTeX, a music typesetting program. The structure of symbolic information as highlighted in Figure 22 is closer to sheet music than MIDI. Time signature and enharmonic spellings are precise. What is more, users are allowed to define their customary pitch names and accidentals to implement microtonal music, which is burdensome in MIDI to implement via crude “pitch bend change message”.
1.4.3 Humdrum

Sapp wrote that “Humdrum file format is analogous to XML: organizing structure for data without concern for content.” [41] Users may define any numbers of representations within it. The “**kern” is one of the representations to carry core information in Western music. It displays concurrent events horizontally as demonstrated in Figure 23. Microtonal music can be implemented through temperaments setting or individual notes’ altered tunings.
1.4.4 MusicXML

MusicXML is the most predominantly supported scheme for data interchange between commercial music notation software. It is a markup language like HTML (HyperText Markup Language), which is designed to create web pages. Its XML-based file format is explicit, flexible, and expandable. Altered tunings can be assigned to individual notes or staff lines. As exemplified in Figure 24 and Figure 25, it is capable of unbounded amount of organized metadata if regardless of file size. Nonetheless, it is an arduous task to write the music by directly typing texts into the file; users should edit it through graphical interfaces.
Figure 24. Example Score [42].

```
<part id="p2">
  <measure number="1">
    <attributes>
      <divisions>2</divisions>
    </attributes>
    <key>
      <fifths>-3</fifths>
      <mode>minor</mode>
    </key>
    <time>
      <beats>3</beats>
      <beat-type>4</beat-type>
    </time>
    <staves>2</staves>
    <clef number="1">
      <sign>
        <line>2</line>
      </sign>
      <clef number="2">
        <sign>
          <line>4</line>
        </sign>
    </clef>
  </attributes>
  <direction placement="below">
    <dynamics default-x="329" default-y="-75">
      <pp/>
    </dynamics>
  </direction-type>
  <staffs>1</staffs>
  <sound dynamic="40"/>
</direction>
</part>
```

Figure 25. Example of MusicXML Source Code [42] of the First Chord in the Piano Part of Figure 20.
1.5 Melodic Contour

Although there are few datasets in the auditory domain as I pointed out in Chapter 1.3, people have been extracting features from the acoustic domain and the graphemic domain with intent to estimate the psychological information. The melodic contour is such a feature which approximates the auditory domain. Moreover, it is a typical example of cross-level features. The cross-level characteristic here implies that the feature is extractable from multiple levels of representation (cf. Chapter 1.4), rather than that multiple features are respectively selected from different levels of feature (cf. Chapter 1.3).

A melodic contour is the line of movement along melody notes in the bidimensional pitch-time space. Nevertheless, researchers treat pitch contour [43]-[44] and duration contour [45] separately to compare symbolic melodic similarity. Although Marsden proved that “the perceived similarity between two melodies can be influenced by the presence of a third melody” [46], the melodic pitch contour is one of the practical solutions which are adopted by online melody search engines. For instance, Musipedia [47] allows users simply typing a short series of melodic pitch contour codes (U = up, D = down, R = repeat) to search melody.

1.5.1 Audio

Basically, there are three levels of difficulty to extract melodic pitch contour from an audio file. First of all, a monophonic melody is the most straightforward case. Fundamental frequencies ($f_0$) are often the most apparent and stable signals in sound spectrogram as shown in Figure 26 except some special or electronic instruments.
Sophisticated models also consider psychoacoustical effects, e.g. “equal-loudness contours” [48]. If the prior information about target instrument (or voice) and its acoustic properties are both known, the fundamental frequencies will be more evident.

![Image of Middle C (262 Hz) Played on a Piano and a Violin](image)

**Figure 26.** Middle C (262 Hz) Played on a Piano and a Violin [49]. The top pane shows the waveform, with the spectrogram below. Zoomed-in regions shown above the waveform reveal the 3.8-ms fundamental period of both notes.

Here the fundamental frequencies are represented frame by frame. As a consequence, there is no higher-level information about duration or rhythm. One must firstly segment the melodic pitch contour to retrieve more note-level information. Once the segmentation is accomplished, the melodic duration contour is extractable. Beyond melodic contour extraction, if the task is melodic transcription, one has to link some music theory with the object of note labelling as presented in Figure 27.

**Figure 27.** Stages in Automatic Melodic Transcription [34].
The second level of difficulty is to extract melodic pitch contour from an audio recording of a polyphonic instrument or identical instruments. As highlighted in Figure 16 on page 18, polyphonic fundamental frequencies are more complicated for sure. Now that any audio mixing file records no voicing information at all, a real problem is to draw a line or several lines as the melodic pitch contour(s). The MIR community’s common definition of melody is “the single (monophonic) pitch sequence that a listener might reproduce if asked to whistle or hum a piece of polyphonic music, and that a listener would recognize as being the ‘essence’ of that music when heard in comparison” [50]. It is obviously incomplete to represent contrapuntal music. Without score, the final result is often arguable even if multiple melodic contours are extracted. Thereby, manual corrections are often required, which are even more important in next level.

The third level of difficulty is to extract melodic pitch contour from an audio recording of dissimilar instruments. This is the realest scenario and its mature development is still in progress. The melodic pitch contour is picked from all pitch contours as illustrated in Figure 28 and Figure 29. During the melody selection, many pitch contours are removed; the smoothed melody pitch mean is refined again and again as delineated in Figure 30 and Figure 31. The final result is usually, if not always, slightly different from the ground truth. Hence, manual corrections are necessary to achieve perfection especially if there are more tasks as indicated in Figure 32 will depend upon the result.
Figure 28. Melodic Pitch Contour Extraction [51].

![Diagram of Melodic Pitch Contour Extraction]

Figure 29. Pitch Contours [51].
(a) vocal jazz (b) opera (c) pop (d) instrumental jazz.
Figure 30. Melody Selection [51].
(a) smoothing (b) removing octave duplicate (c) smoothing and removing pitch outliers (d) smoothing.

Figure 31. Vocal Jazz Excerpt [51].
(a) all pitch contours created (b) contours after filtering and melody pitch mean (thick red line) (c) final extracted melody (black) and ground truth (thick red, shifted down one octave for clarity).
1.5.2 Symbolic

To extract melodic contour from a symbolic file is relatively effortless thanks to its more hierarchical information except type-0 MIDI file, cf. Figure 21 and Table 8 on page 26. As I implied in Chapter 1.4, most symbolic representations store polyphonic music in multiple parts or tracks. A melodic contour could be extracted from every part; multiple melodic contours could be extracted from a polyphonic piece. Nonetheless, there might be chords or multiple voices in individual parts. One has to decide which notes to be included or excluded. Additionally, people sometimes intentionally dividing notes into parts which are actually played via a single instrument. It is reasonable in a MIDI file whenever they want to control its subtle difference between performance parameters. Users must be cautious about this kind of data.
1.6 Datasets

Researchers have built and released numerous datasets. They accompany with so-called ground truth of features which are generated via machine extraction and human annotation as well as correction. Thus a low-level audio dataset may accompany with high-level ground truth of features. Depending upon the levels of data and ground truth, each of the datasets may target specific MIR classification tasks like mood, genre, form, harmony, melody, and so forth.

1.6.1 Audio

There are a fast growing number of audio datasets. Most of them focus on Western or even only commercial music on account of the revenue and interest from music industry. Some datasets merely contain annotations and audio snips, e.g. 30-second recording of every song; others comprise generous data. For instance, MedleyDB is a multitrack dataset which aggregates 122 songs [52]. Isolated multiple tracks are precious with intent to evaluate automatic instrument recognition, source separation, automatic mixing, and polyphonic melodic contours. Furthermore, it has 105 full length songs and 108 melody annotations across nine genres as charted in Figure 33.

![Figure 33. MedleyDB Dataset [52].](image-url)
1.6.2 Symbolic

Anthropologist and ethnomusicologists have been collecting huge amount of music by conducting filed research even before the audio recording technology. Thus many samples are notated via varied approaches as contrasted in Figure 34. Cornelis et al. pointed out that “the differences in approach illustrate the difficulty of creating a uniform symbolic representation of music that does not use the Western concepts and categorizations” [29]. Seeing that mainstream symbolic representations aim at Western music, most of the modern symbolic digital datasets still have a bias in favor of Western music.

Figure 34. Transcription of a San (Bushman) Song with Musical Bow Accompaniment by Four Ethnomusicologists at the SEM Symposium on Transcription and Analysis in 1964 [29].
First of all, KernScores [53] is the longest dataset in the Humdrum representation, cf. Chapter 1.4.3. It collects more than 10'000 monophonic songs, harmonized songs, and instrumental compositions. Vocal samples cover folk melodies from three continents (Europe, Asia, and North America), Bach chorales, and early Renaissance genres, e.g. mass, motet, and secular song. Instrumental samples consist of Bach well-tempered clavier (WTC), Classical and Romantic keyboard works, and Classical string quartets.

Second, TAVERN is a dataset which “consists of 27 complete sets of theme and variations for piano composed between 1765 and 1810 by Mozart and Beethoven” in Humdrum representation, too [54]. They are divided into phrases and annotated with harmony like Roman numerals and function labels. Therefore, this dataset is ideal for the purpose of evaluating computational music analysis, e.g. chord labeling and phrase parsing.

Third, the Yale-Classical Archives Corpus sorts extracted features and metadata from 8980 MIDI files across 505 composers, cf. Figure 35. It does not contain the original MIDI files, which are available on the Classical Archives web site [55].
Figure 35. The Yale-Classical Archives Corpus [57].
Fourth, the music21 library [13] includes a corpus of freely distributable music. Its files are saved in Humdrum, MusicXML, and other representations. Most pieces are composed by classical music composers. Now that the library is still under active maintenance and upgrade, it may amass more pieces, composers, and genres.

Fifth, Nottingham Dataset gathers 1037 folk tunes in 9 categories as follows: Ashover, Christmas, Hornpipes, Jigs, Morris, Playford, Reels, Slip Jigs, and Waltzes. Jukedeck Research released a cleaned version of 1034 samples on March 7th, 2017 [56]. It provides MIDI files in three editions: melody, chords, and mixture.

Last but not least, the remaining one is The Aria Database [58]. It preserves almost two hundred opera aria MIDI files. I manually select 177 samples of vocal melodic part for my research.

1.6.3 Hybrid

Some datasets incorporate multiple levels of data. The audio and symbolic data is usually complementary to each other. What is more, those datasets are also beneficial to more musicologists inasmuch as the rich information is versatile beyond MIR tasks.

First, Meertens Tune Collections compiles 4’120 vocal folk song strophes “from both twentieth-century field recordings and printed books that cover roughly the same repertoire” as well as 2’368 instrumental tunes “from eighteenth-century, mostly monophonic Dutch sources, both prints and manuscripts.” [59]. In addition to MIDI, LilyPond, and Humdrum, it also encompasses other types of text and image representations as reported in Table 9.
Table 9. Contents of the Meertens Tune Collections [59].

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>File types</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTC-OGLAUDIO</td>
<td>Collection <em>Onder de groene linde</em>: 7,178 audio recordings collected by Dutch field workers during the 1950s-1980s.</td>
<td>mp3</td>
<td>20 GB</td>
</tr>
<tr>
<td>MTC-OGLSCANS</td>
<td>Scans of 3,754 transcriptions of recordings from <em>Onder de groene linde</em> as made during the 1950s-1980s. The music is hand-written, the lyrics are mostly typed.</td>
<td>jpg</td>
<td>1.6 GB</td>
</tr>
<tr>
<td>MTC-FS</td>
<td>4,120 digitally encoded strophes from vocal folk songs both from <em>Onder de groene linde</em> (2,503) and from various related written sources (1,617).</td>
<td>**kern, midi, LilyPond, png, pdf, txt</td>
<td>250 MB</td>
</tr>
<tr>
<td>MTC-INST</td>
<td>2,368 digitally encoded instrumental tunes from 18th-century Dutch manuscripts and printed scores.</td>
<td>**kern, midi, LilyPond, png, pdf</td>
<td>130 MB</td>
</tr>
<tr>
<td>MTC-ANN</td>
<td>Annotated Corpus: 360 melodies used in various publications.</td>
<td>**kern, midi, png, pdf</td>
<td>1.5 MB</td>
</tr>
<tr>
<td>MTC-LC</td>
<td>Large Corpus: 4,830 melodies used in various publications.</td>
<td>**kern, midi</td>
<td>20 MB</td>
</tr>
</tbody>
</table>

Next, Weimar Jazz Database (WJazzD) [60] upgraded to version 2.0 on March 31st, 2017. It accumulates 456 instrumental jazz solo transcriptions in Structured Query Language (SQL) database. Respective unquantized MIDI files are downloadable from its website, too.

Finally, SymbTr is a distinct non-Western music dataset. Its 1’700 pieces of Turkish art and folk music range from 155 “makams” (melodic patterns) and 100 “usuls” (rhythmic structures) to 48 forms. Its special representation is compatible with both the uneven 17-tone classical Turkish music and uneven 24-tone Turkish folk song as demonstrated in Figure 36 and Table 10. One can foresee that “data structures and algorithms developed for other musical traditions are not directly applicable to Turkish makam music” [61].
Figure 36. One of the Scorings of Turkish Folk Song [61].

Table 10. SymbTr Representation of the Score in Figure 36 [61].

<table>
<thead>
<tr>
<th>Code</th>
<th>Note53</th>
<th>Comma53</th>
<th>NoteAE</th>
<th>CommaAE</th>
<th>Num.</th>
<th>Denom.</th>
<th>Ns</th>
<th>LNS</th>
<th>VelOn</th>
<th>Syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Do5</td>
<td>318</td>
<td>C5</td>
<td>318</td>
<td>1</td>
<td>4</td>
<td>667</td>
<td>95</td>
<td>96</td>
<td>Bir</td>
</tr>
<tr>
<td>9</td>
<td>Re5b3</td>
<td>324</td>
<td>Db4</td>
<td>325</td>
<td>1</td>
<td>8</td>
<td>333</td>
<td>99</td>
<td>108</td>
<td>dal</td>
</tr>
<tr>
<td>9</td>
<td>Re5b3</td>
<td>324</td>
<td>Db4</td>
<td>325</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>99</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Do5</td>
<td>318</td>
<td>C5</td>
<td>318</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>95</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Si4b2</td>
<td>312</td>
<td>B4b1</td>
<td>313</td>
<td>1</td>
<td>4</td>
<td>667</td>
<td>95</td>
<td>72</td>
<td>da</td>
</tr>
<tr>
<td>12</td>
<td>Do5</td>
<td>318</td>
<td>C5</td>
<td>318</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>333</td>
<td>99</td>
<td>i</td>
</tr>
<tr>
<td>9</td>
<td>Do5</td>
<td>318</td>
<td>C5</td>
<td>318</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>99</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Si4b2</td>
<td>312</td>
<td>B4b1</td>
<td>313</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>95</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>La4</td>
<td>305</td>
<td>A4</td>
<td>305</td>
<td>1</td>
<td>8</td>
<td>333</td>
<td>99</td>
<td>96</td>
<td>ki</td>
</tr>
<tr>
<td>8</td>
<td>Si4b2</td>
<td>312</td>
<td>B4b1</td>
<td>313</td>
<td>1</td>
<td>8</td>
<td>42</td>
<td>99</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>La4</td>
<td>305</td>
<td>A4</td>
<td>305</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>99</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Sol4</td>
<td>296</td>
<td>G#4</td>
<td>296</td>
<td>1</td>
<td>16</td>
<td>167</td>
<td>95</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>La4</td>
<td>305</td>
<td>A4</td>
<td>305</td>
<td>1</td>
<td>8</td>
<td>333</td>
<td>99</td>
<td>96</td>
<td>ki</td>
</tr>
<tr>
<td>9</td>
<td>Si4b2</td>
<td>312</td>
<td>B4b1</td>
<td>313</td>
<td>1</td>
<td>8</td>
<td>333</td>
<td>95</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Do5</td>
<td>318</td>
<td>C5</td>
<td>318</td>
<td>1</td>
<td>4</td>
<td>1334</td>
<td>45</td>
<td>84</td>
<td>raz</td>
</tr>
</tbody>
</table>
A melody could hardly include unmelodious elements; the concept of the melodious is intimately related to the concept of singableness. The nature and technique of the primordial musical instrument, the voice, determines what is singable. The concept of the melodious in instrumental melody has developed as a free adaptation from the vocal model.

Arnold Franz Walter Schönberg [62]
With intent to implement my foregoing exploration of the compositional space, I prepare proper datasets and tools. I will elaborate the features which are created by myself and will also introduce those I borrowed from other research. In the end of this chapter, I will discuss their intercultural and cross-level capabilities.

2.1 Preparation

With an eye to spot the unexploited compositional space as I manifested in Chapter 1.2.2, I deploy feature extractors so that I can extract features from selected datasets. Since not all samples in every dataset fit any tasks, I have to clean the datasets before my final retrieval.

2.1.1 Data Cleaning

In my initial retrieval, I noticed that several datasets and samples have illogical values such as extremely wide register width. There were also problematic files which were incomplete, separate one-part notes into two parts, or mix multiple parts’ notes in a single part or MIDI track. I scrutinized every suspicious file and discard all unacceptable ones. Furthermore, I removed some symbolic files which the music21 library [13] cannot parse.

2.1.2 Samples Quantity

I employ the feature extractors which are already implemented in the music21 library [13] with the aim of extracting features in the jSymbolic feature set [15]-[16] which was specified in Chapter 2.2.3. In the end, 15173 samples from 66 datasets are retrieved as counted in Table 11.
Table 11. Retrieved Quantities by music21 Feature Extractors.

<table>
<thead>
<tr>
<th>source</th>
<th>instrumental datasets</th>
<th>vocal datasets</th>
<th>total datasets</th>
<th>instrumental samples</th>
<th>vocal samples</th>
<th>total samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aria [58]</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>177</td>
<td>177</td>
</tr>
<tr>
<td>KernScores [53]</td>
<td>12</td>
<td>49</td>
<td>61</td>
<td>792</td>
<td>10697</td>
<td>11489</td>
</tr>
<tr>
<td>Nottingham [56]</td>
<td>0</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>1034</td>
<td>1034</td>
</tr>
<tr>
<td>SymbTr [61]</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>187</td>
<td>1833</td>
<td>2020</td>
</tr>
<tr>
<td>WJazzD [60]</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>453</td>
<td>0</td>
<td>453</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>14</strong></td>
<td><strong>60</strong></td>
<td><strong>74</strong></td>
<td><strong>1432</strong></td>
<td><strong>13741</strong></td>
<td><strong>15173</strong></td>
</tr>
</tbody>
</table>

In parallel, I implement my own feature extractors on purpose to extract other features. In this case, I divide each of the vocal parts of Bach’s Chorale and each of the instrumental parts of String Quartet. Additionally, I add the MedleyDB [46] dataset with the object of testing the cross-level features, cf. Chapter 2.3. As a consequence, 18009 samples from 83 datasets are retrieved as calculated in Table 12.


<table>
<thead>
<tr>
<th>source</th>
<th>instrumental datasets</th>
<th>vocal datasets</th>
<th>total datasets</th>
<th>instrumental samples</th>
<th>vocal samples</th>
<th>total samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aria [58]</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>177</td>
<td>177</td>
</tr>
<tr>
<td>KernScores [53]</td>
<td>21</td>
<td>55</td>
<td>76</td>
<td>1857</td>
<td>12360</td>
<td>14217</td>
</tr>
<tr>
<td>MedleyDB [52]</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>47</td>
<td>61</td>
<td>108</td>
</tr>
<tr>
<td>Nottingham [56]</td>
<td>0</td>
<td>9</td>
<td>9</td>
<td>0</td>
<td>1034</td>
<td>1034</td>
</tr>
<tr>
<td>SymbTr [61]</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>187</td>
<td>1833</td>
<td>2020</td>
</tr>
<tr>
<td>WJazzD [60]</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>453</td>
<td>0</td>
<td>453</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>24</strong></td>
<td><strong>67</strong></td>
<td><strong>91</strong></td>
<td><strong>2544</strong></td>
<td><strong>15465</strong></td>
<td><strong>18009</strong></td>
</tr>
</tbody>
</table>

2.1.3 Environment

I write computer programs which depend upon the music21 [13] 3.1.0 library and several packages in Anaconda 4.3.0 distribution (64-bit) of Python 3.6.0 as described in Table 13. My programs extract features from above-mentioned datasets and calculate the statistics. I aim for musical features which are potentially applicable
to AC regardless of levels of representation. They may or may not be capable of traditional MIR tasks.

Table 13. Adopted Packages from Anaconda Python

<table>
<thead>
<tr>
<th>name</th>
<th>version</th>
<th>my purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>matplotlib</td>
<td>2.0.0</td>
<td>2D plotting</td>
</tr>
<tr>
<td>numpy</td>
<td>1.11.3</td>
<td>basic statistical information</td>
</tr>
<tr>
<td>scikit-learn</td>
<td>0.18.1</td>
<td>machine learning and data mining</td>
</tr>
<tr>
<td>scipy</td>
<td>0.18.1</td>
<td>Pearson correlation coefficient (PCC) and two-tailed p-values</td>
</tr>
</tbody>
</table>

When it comes to real classification tasks, I operate the normal linear-kernel SVM (support vector machine) classifier from the scikit-learn library [63] in Table 13 above. I adopt one of its univariate feature selections which are refined in light of its ANOVA (analysis of variance) f-value between label and feature.

2.2 Features

I derive features from my former degree thesis [11] and combine von Hippel’s features [14] with intent to investigate the difference between vocal and instrumental melodic contours. In addition, I review McKay’s feature set [15]-[16], which looks beyond melodic information. Finally, I implement concurrence and non-concurrence features [64] based on internal correlations between local features.

2.2.1 Susceptibility

I customize several musical features from my own old model [11], in which I conceived a particular parameter and coined its name “susceptibility”, a nod to the magnetic susceptibility in electromagnetism. The parameter regulated a generative monophonic melody’s sensitivity to make a return after successive pitch intervals [10]-[11]. It might inherently be capable of not only the control over melodic pitch
contour but also the “effect of tessitura” [81] and the “low-skip bias” [82] even without any prior restriction on pitch range. Nevertheless, the original model was invented for AC. Some parameters were allowed to be adjusted by a user. For this reason, I have to modify it in order to extract features from pre-existing samples [17].

First of all, the central frequency is now automatically calculated from the pitch range of a given melody. The audio frequency or its equivalent value of a symbolic pitch according to any suitable tuning system is $f$. The highest frequency in the melody is $f_{\text{max}}$; the lowest frequency in the melody is $f_{\text{min}}$. The total note numbers of the monophonic melody is $n$. The register width of the melody is defined as

$$w = \log_2 \left( \frac{f_{\text{max}}}{f_{\text{min}}} \right)$$

(1)

The central frequency of the melody is

$$f_0 = f_{\text{min}} \times \sqrt{\frac{f_{\text{max}}}{f_{\text{min}}}}$$

(2)

Second, the part of energy ratio interval is unchanged. The frequency ratio of each successive interval is

$$r_n = \frac{f_n}{f_0}; \quad n \geq 0$$

(3)

Its energy ratio is

$$e_n = (r_n)^2; \quad n \geq 0$$

(4)

The energy ratio interval is

$$i_n = e_n - e_{(n-1)}; \quad n \geq 1$$

(5)
Third, I design a constant $k$ to enlarge the initial tolerable energy ratio maximum and minimum. I assign the same tentative value ($k = 4$) to all experiments in my research. The initial maximum and minimum are no longer set in advance by the user. Conversely, the initial tolerable energy ratio maximum is given by

$$e_{\text{max},1} = \max\{e_j\} \times k; \quad j = 1, 2, 3, \ldots, n$$

while the reciprocal of initial tolerable energy ratio minimum can be written as

$$\frac{1}{e_{\text{min},1}} = \frac{1}{\min\{e_j\} \times k}; \quad j = 1, 2, 3, \ldots, n$$

Next, $s$ (susceptibility) is the largest possible value which is approached by means of my retrieval program. The following part is similar to my original thesis.

The tolerable energy ratio maximum is defined as

$$e_{\text{max},n} = e_{\text{max},n-1} - (i_{n-1} \times s); \quad n \geq 2$$

while the reciprocal of tolerable energy ratio minimum is

$$\frac{1}{e_{\text{min},n}} = \frac{1}{e_{\text{min},n-1}} + (i_{n-1} \times s); \quad n \geq 2$$

The tolerable frequency ratio maximum is

$$r_{\text{max},n} = \sqrt{e_{\text{max},n}}; \quad n \geq 1$$

The reciprocal of tolerable frequency ratio minimum is

$$\frac{1}{r_{\text{min},n}} = \sqrt{\frac{1}{e_{\text{min},n}}}; \quad n \geq 1$$

i.e.

$$r_{\text{min},n} = \sqrt{e_{\text{min},n}}; \quad n \geq 1$$
Finally, I revise the last part for better visualization. The logarithmic frequency ratio is defined as

\[ l_n = \log_2 (r_n); \ n \geq 0 \]  

(13)

The logarithmic tolerable frequency ratio maximum is

\[ l_{\max_n} = \log_2 (r_{\max_n}); \ n \geq 1 \]  

(14)

The logarithmic tolerable frequency ratio minimum is

\[ l_{\min_n} = \log_2 (r_{\min_n}); \ n \geq 1 \]  

(15)

On purpose to exemplify, I select a short melody as notated in Figure 37 from a counterpoint textbook [83] to process. The melody is notated in symbolic pitch, so I have to assign an audio frequency to every note. For the purpose of simple values and straightforward calculation, I avoid adopting the popular twelve-tone equal temperament. Nonetheless, users of this model are allowed to refer to any suitable tuning system for conversion from pitches into frequencies.

\[ \text{Figure 37. A Cantus Firmus [83].} \]

susceptibility = 10.31, see the paragraph below.

In this example, the frequency of the first pitch \((f_1)\) is 240 Hz. The highest frequency \((f_{\max})\) is 360 Hz; the lowest frequency \((f_{\min})\) is 240 Hz. As a result, the central frequency \((f_0)\) is about 294 Hz. The susceptibility is approached from zero to the largest possible value using an increment of 0.01. As soon as the susceptibility reaches 10.32, the \(e_3\) will exceed \(e_{\max_3}\). In consequence, the final retrieved susceptibility is 10.31, cf. Table 14 and Figure 38.
Table 14. Retrieval Information from Figure 37 [17].

<table>
<thead>
<tr>
<th>pitch</th>
<th>la 3</th>
<th>mi 4</th>
<th>si 3</th>
<th>re 4</th>
<th>do 4</th>
<th>si 3</th>
<th>la 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>f ( (f_0=294) )</td>
<td>240</td>
<td>360</td>
<td>270</td>
<td>320</td>
<td>288</td>
<td>270</td>
<td>240</td>
</tr>
<tr>
<td>r ( (r_0=1) )</td>
<td>0.816</td>
<td>1.225</td>
<td>0.919</td>
<td>1.089</td>
<td>0.980</td>
<td>0.919</td>
<td>0.816</td>
</tr>
<tr>
<td>e ( (e_0=1) )</td>
<td>0.667</td>
<td>1.500</td>
<td>0.844</td>
<td>1.185</td>
<td>0.960</td>
<td>0.844</td>
<td>0.667</td>
</tr>
<tr>
<td>i</td>
<td>-0.333</td>
<td>0.833</td>
<td>-0.656</td>
<td>0.341</td>
<td>-0.225</td>
<td>-0.116</td>
<td>-0.177</td>
</tr>
<tr>
<td>e-max</td>
<td>6.000</td>
<td>9.437</td>
<td>0.845</td>
<td>7.611</td>
<td>4.091</td>
<td>6.412</td>
<td>7.611</td>
</tr>
<tr>
<td>l/e-min</td>
<td>6.000</td>
<td>2.563</td>
<td>11.155</td>
<td>4.389</td>
<td>7.909</td>
<td>5.588</td>
<td>4.389</td>
</tr>
<tr>
<td>e-min</td>
<td>0.167</td>
<td>0.390</td>
<td>0.090</td>
<td>0.228</td>
<td>0.126</td>
<td>0.179</td>
<td>0.228</td>
</tr>
<tr>
<td>r-max</td>
<td>2.449</td>
<td>3.072</td>
<td>0.919</td>
<td>2.759</td>
<td>2.023</td>
<td>2.532</td>
<td>2.759</td>
</tr>
<tr>
<td>l/r-min</td>
<td>2.449</td>
<td>1.601</td>
<td>3.340</td>
<td>2.095</td>
<td>2.812</td>
<td>2.364</td>
<td>2.095</td>
</tr>
<tr>
<td>r-min</td>
<td>0.408</td>
<td>0.625</td>
<td>0.299</td>
<td>0.477</td>
<td>0.356</td>
<td>0.423</td>
<td>0.477</td>
</tr>
<tr>
<td>l ( (l_0=0) )</td>
<td>-0.292</td>
<td>0.292</td>
<td>-0.123</td>
<td>0.123</td>
<td>-0.029</td>
<td>-0.123</td>
<td>-0.292</td>
</tr>
<tr>
<td>l-max</td>
<td>1.292</td>
<td>1.619</td>
<td>-0.121</td>
<td>1.464</td>
<td>1.016</td>
<td>1.340</td>
<td>1.464</td>
</tr>
<tr>
<td>l-min</td>
<td>-1.292</td>
<td>-0.679</td>
<td>-1.740</td>
<td>-1.067</td>
<td>-1.492</td>
<td>-1.241</td>
<td>-1.067</td>
</tr>
</tbody>
</table>

Figure 38. Melodic Pitch Contour (middle line) of Figure 37 [17].
The first point of middle line denotes the original melody's central frequency.
2.2.2 Indices of Tessitura and Mobility

By von Hippel’s definition, in a melody as notated in Figure 39, the “index of
tessitura (or pitch distribution)” is the standard deviation of pitch distribution as
charted in Figure 40; the “index of mobility (or freedom of motion)” is the lag-one
autocorrelation between successive pitch values as illustrated in Figure 41. He
described the “pitch proximity” by those two features and claimed that they “offers a
more precise and detailed description of melodic structure” [14]. Thereupon, I include
them in my research.

![Figure 39. A German Folk Song [14].](image)

![Figure 40. Pitch Histogram of the Melody in Figure 39 [14].](image)

index of tessitura = 3.5.
2.2.3 jSymbolic Feature Set

I inspect a bunch of McKay’s features [15]-[16] which are implemented in the music21 library [13]. In his research, the MIR task was the automatic genre classification of MIDI files. For that reason, he considered “as wide a range of features as possible in order to characterize as full a spectrum of music as possible”, except some criteria as follows:

1. no matter whether features are perceptible to the average human or not;
2. avoid features which are difficult to objectively measure or even extract automatically at all;
3. avoid sophisticated theoretical analyses which are computationally expensive;
4. prefer features which are less tied to specific music types or limiting assumptions.

Although those features may be substantial with regard to his classification task, I prefer features which are more applicable to AC. Moreover, many of his...
features are only available or accessible in MIDI representation. They do not exist in other symbolic file formats. As a result, I select 32 features from the jSymbolic feature set as identified in Appendix A1

2.2.4 Concurrence and Non-concurrence

In the history of music, composers have developed various musical textures, e.g. monophony, biphony, heterophony, homophony, and polyphony. “In discussions of texture a distinction is generally made between homophony [...] and polyphonic (or contrapuntal) treatment.” [65] Unfortunately, in some MIR research papers, notably when dealing with audio data, authors often describe all multi-pitch music as “polyphonic”, e.g. [9], [30], [31], [33], and [49]. In fact, true polyphonic music has highly independent parts integrating together. Listeners contextually appreciate not only the vertical sonority but also the horizontal interaction between separate parts. Therefore, musical texture is an important element of style as plotted in Figure 42. Much information could be retrieved from the interrelationship between parts.
Composers sometimes interchange and overlap homophonic and polyphonic textures within one piece. For instance, in the Finale of W. A. Mozart’s String Quintet No.5 in D major, K.593, he begins the movement with a homophonic theme and introduces another theme with 3-voice polyphonic texture cycling through all 5 parts since measure 54. The third theme starts from measure 132 with 5-voice polyphonic
texture, which is immediately followed by a 4-voice stretto across all parts except the second viola.

A more magnificent example is the Finale of his Symphony No. 41 in C major, K. 551. While most sections may sound like homophonic, there are actually many contrapuntal melodies played by orchestral instruments. The coda section culminates in a 5-voice invertible counterpoint, which incorporates the 5 themes previously introduced in this movement.

After all, Lloyd and Boyle argued that “the melodic line was everything: harmony was in the making, it was being formed by writing concurrent melody.” [67] One should always remember the melodic essence even in homophonic section or piece because composers are still very likely to manipulate the degree of independence for each vocal or instrumental part. They take care of both longitudinal and latitudinal connections.

Researchers have extracted features from part-to-part interrelationship in pieces with the aim of composer classification [68]-[70]. Furthermore, in addition to composer classification from isolated movements of string quartets [71]-[72], separate parts are classifiable, too [73]-[74]. Inevitably, there is less texture information in a single part than multiple ones.

Nevertheless, a single part has its inner states of concurrence and non-concurrence, which “can be viewed as another characteristic of the texture” [64] as Cohen and Dubnov stated. Two values in a substantial concurrence have a positive correlation; in a notable non-concurrence they have a negative correlation (or inverse
concurrency) as contrasted in Figure 43. This state may serve as another form of MIR feature. I coin its name “concurrence feature” for the purpose of differentiating it from global [75], local [75]-[76], contextual [76], and event [77]-[78] features.

![Figure 43. Concurrence and Non-concurrence between Pitch and Intensity Curves in Syllables of Bird Calls (Babblers) [64].](image)

(a) in calm situations; (b) in tense situations.

A concurrence or non-concurrence feature extracted from time series data is a statistical value which looks like a global feature. It does not preserve all the changes along time, which are recorded in its local base features. It reflects, however, the interaction between two series of local base features. Even though the sample is an individual vocal or instrumental part from a piece, it is able to treat multiple series of any accessible local base feature from the single part as multiple lines. Hence, there is considerably more texture information after adding concurrence and non-concurrence features.

To the best of my knowledge, concurrence and non-concurrence features have not been quantitatively examined. In Chapter 3.3, I will demonstrate their discrimination capabilities on purpose to enhance composer classification of separate string quartet parts and discuss the implications.
2.3 Intercultural and Cross-level

Although culture-aware and domain-specific approaches, e.g. [79]-[80], are ideal for MIR, CMA, and computational musicology, they also bring innate constraints from the original style. Those constraints may limit the application to AC unless it is aiming for the exact stylistic composition. My motivation as I manifested in Chapter 1.2.2 is to unveil the space which was unexploited by our predecessors. Therefore, I prefer inclusive musical features which are extractable from various datasets. The more datasets I add in my research, they can ensure me the larger number of wide-ranging samples.

Volk and Honingh said that “to achieve a meta-theory of music, we need to rise above the study of particular examples of music tying all musical cultures together through time” [84]. It is quite challenging to find computationally extractable, objective, and universal high-level features. For instance, a so-called melody may not exist in some music. Nonetheless, it is feasible to seek intercultural ones. By my personal definition, a universal feature is extractable from (almost) all types of music; an intercultural feature is from many distinct cultures; a cross-cultural feature is from at least two.

Because of the biased cultural specificity of existing datasets and the unavailability of audio recordings of ancient music as I presented in Chapter 1.6.2, I prefer the cross-level features. Again, the cross-level characteristic implies that the feature is extractable from multiple levels of representation (cf. Chapter 1.4), rather
than that multiple features are respectively selected from different levels of feature (cf. Chapter 1.3).

Needless to say, lower-level features are more ubiquitous in all music, e.g. Table 3 on page 19. They are, however, not musicologically meaningful to be deployed during composition procedures. Besides, Volk and Honingh also argued that “it is a failure to study the musical products without the musical processes” [84]. Thus I proposed the mutualism between (algorithmic) composers and MIR scientists in Chapter 1.2.3.

The melodic pitch contour is an excellent cross-level material in both symbolic datasets and, at least monophonic, audio datasets as I explained in Chapter 1.5. It is not “bridging the audio-symbolic gap” [85]; it simply circumvents the gap by avoiding any segmentation, beat-tracking, quantization, pitch spelling, and dubious “approximate units of meaning.” [37] For instance, it is painful to extract incontrovertible note-level information from a typical guqin tone with vibrato. Conversely, the melodic contour is evident as delineated in Figure 44.
In those foregoing customized features from my original AC model, the compositional algorithm is independent of any tuning system because it takes audio frequencies in place of symbolic pitches into consideration, cf. Chapter 2.2.1. Thus my new analytical program only requires information in the level of melodic pitch contour to extract those musical features. It successfully extracts features from not only high-level data in symbolic files but also low-level data in audio frames as recorded in Table 15.

**Table 15.** Example of Data in Audio Frames from MedleyDB [52].

<table>
<thead>
<tr>
<th>timestamp</th>
<th>audio frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.510839</td>
<td>195.649</td>
</tr>
<tr>
<td>0.516644</td>
<td>196.23</td>
</tr>
<tr>
<td>0.522449</td>
<td>196.861</td>
</tr>
<tr>
<td>0.528254</td>
<td>197.167</td>
</tr>
<tr>
<td>0.534059</td>
<td>196.824</td>
</tr>
<tr>
<td>0.539864</td>
<td>196.384</td>
</tr>
<tr>
<td>0.545669</td>
<td>195.414</td>
</tr>
</tbody>
</table>
By contrast, the tessitura, mobility, and some jSymbolic features rely upon note-level information. They are not extractable from the MedleyDB dataset. It does not have symbolic music files but has three types of melody annotation according to different definitions [52]. I choose the first one in which the fundamental frequency curve of the predominant melodic line is drew from a single source.

What is more, a concurrence or non-concurrence feature itself is also cross-level, intercultural, and even universal. Caution should be exercised that it is extracted from local base features. They may be, however, culture-aware or level-specific.

In summary, cross-level features have more potential for intercultural exploration and application. They are more extractable from diverse, if not most, music and multiple levels of representations. On the other hand, they are also more convenient for composers to manipulate than lower-level features.
3. Exploration

In recent decades, the ‘space’ of that which might be explored — the possible musical vocabularies, syntaxes and strategies that may attract composers — has expanded at a seemingly undamped and possibly undampable rate.

Roger Reynolds [86]
My explorations begin in my observations on different scales. In feature space, investigation can be accomplished dataset by dataset, genre by genre, sample by sample, and even window by window, and so on and so forth. In addition, I will deviate to evaluate the concurrence and non-concurrence features, namely the internal correlations, in terms of composer classification. Finally, I will demonstrate how to compose in the specific unexploited space.

3.1 Observation

I explore the compositional space on various scales of observation. I inspect the dataset-to-dataset relationships and the sample-to-sample distributions. I consider all the datasets which were listed in Chapter 2.1.2 and all the features which were explained in Chapter 2.2.

3.1.1 Dataset by Dataset

My experiments start from the comparison between datasets. For example, I choose two jSymbolic features, “most common melodic interval” and “average melodic interval”, cf. Chapter 2.2.3. Comparing both of them dataset by dataset, the distribution of their mean values reveal the distances between each of the datasets as displayed in Figure 45. The only three Native Americans datasets and the only three quartet datasets are respectively close.

On the equal observation scale, the ranges and mean values of “susceptibility” (cf. Chapter 2.2.1) discriminate vocal style and instrumental style as sorted in Appendix A2. The ranges of susceptibilities in vocal datasets are broadly higher and wider while they are mostly lower and narrower in instrumental datasets. All 66
datasets with larger mean susceptibilities than “Aria” are vocal styles except “Keyboard\Mazurka”; all 24 datasets with smaller mean susceptibilities are instrumental styles except “MedleyDB\Vocal” [17].

In addition to susceptibility, the distances between each of the datasets are again displayed upon including the mean values of “register width” (cf. Chapter 2.2.1). Most of the vocal datasets aggregate in the middle zone; most of the instrumental datasets spread in the upper left region as shown in Figure 46. Caution should be used that the distribution of mean values is informative for neither traditional classification tasks nor my exploration of unexploited compositional space. Instead, they need the distribution of all individual retrieval values, like Figure 5 and Figure 7. Thereby, I have to compare those two features sample by sample.
Figure 45. Distribution of Mean “Average Melodic Interval” and Mean “Most Common Melodic Interval” from All Datasets in Table 11.
Figure 46. Distribution of Mean “Register Width” and Mean “Susceptibility” from All Datasets in Table 12.
3.1.2 Sample by Sample

As scattered in Figure 47, the distribution of vocal datasets delineates a cluster with some outliers; the other distribution of instrumental datasets depicts another cluster with several outliers. The two clusters virtually overlap each other in a large area. If merely extracting those two features, the clustering is unable to accurately distinguish vocal style and instrumental style.

Figure 47. Distribution of Register Width and Susceptibility from All (except 4) Samples in Table 12. samples out of scope: two at (18.85, 0.25); one at (20.15, 0.25); one at (29.38, 0.17).

Nevertheless, the distribution illuminates a whole new perspective. As a matter of fact, the overlapping clusters draw a thick but solid nonlinear trend with less than 0.17% outliers (30/18009). These outliers come from various composers, genres, eras, and geographic regions as Table 16. Outsides the clusters, the sparse population of music samples signify that it could be a unique or barely acceptable space to compose. After all, it is a hardly exploited area. Composers are allowed to follow, deviate, or
diverge from our predecessors’ trend. By adopting this strategy, one can add more innovative musical features to construct a higher-dimensional space. In addition to listening and reading with machine as I presented in Chapter 1.2.3, it inspires composers to search unknown dimensions and “‘space’ of that which might be explored” [86] as I proposed in Chapter 1.2.2.

**Table 16.** The 30 Outliers in Figure 47.

<table>
<thead>
<tr>
<th>dataset</th>
<th>filename</th>
<th>length</th>
<th>register width</th>
<th>susceptibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aria</td>
<td>trav3_brindisi_milano.mid</td>
<td>209</td>
<td>1.42</td>
<td>9.02</td>
</tr>
<tr>
<td>Harmo\Bach Chorale\371\Tenor</td>
<td>chor253.krn</td>
<td>64</td>
<td>1.08</td>
<td>9.32</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>mazurka50-1.krn</td>
<td>343</td>
<td>2.58</td>
<td>16.30</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>mazurka-51.krn</td>
<td>474</td>
<td>3.00</td>
<td>6.93</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>mazurka56-2.krn</td>
<td>297</td>
<td>2.00</td>
<td>7.97</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>mazurka59-2.krn</td>
<td>416</td>
<td>3.00</td>
<td>13.79</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>mazurka67-1.krn</td>
<td>259</td>
<td>3.00</td>
<td>13.55</td>
</tr>
<tr>
<td>Mono\EFC\China\han</td>
<td>han0261.krn</td>
<td>29</td>
<td>1.42</td>
<td>14.67</td>
</tr>
<tr>
<td>Mono\EFC\China\han</td>
<td>han1078.krn</td>
<td>103</td>
<td>1.42</td>
<td>12.80</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\boehme</td>
<td>deut2742.krn</td>
<td>36</td>
<td>1.00</td>
<td>13.48</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\erk</td>
<td>deut1434.krn</td>
<td>22</td>
<td>0.67</td>
<td>11.49</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\erk</td>
<td>deut1576.krn</td>
<td>76</td>
<td>1.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\erk</td>
<td>deut1913.krn</td>
<td>33</td>
<td>1.00</td>
<td>11.21</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\erk</td>
<td>deut2111.krn</td>
<td>28</td>
<td>0.83</td>
<td>13.96</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\fink</td>
<td>deut317.krn</td>
<td>66</td>
<td>1.42</td>
<td>10.34</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr090.krn</td>
<td>37</td>
<td>0.67</td>
<td>12.86</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr092.krn</td>
<td>40</td>
<td>0.67</td>
<td>12.86</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr093.krn</td>
<td>25</td>
<td>0.67</td>
<td>14.12</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr102.krn</td>
<td>39</td>
<td>1.00</td>
<td>14.21</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr184.krn</td>
<td>28</td>
<td>0.67</td>
<td>12.86</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>kindr195.krn</td>
<td>38</td>
<td>1.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Mono\EFC\Europa\misc</td>
<td>emmenth1.krn</td>
<td>46</td>
<td>1.42</td>
<td>10.52</td>
</tr>
<tr>
<td>Mono\EFC\Europa\oesterrh</td>
<td>oestri002.krn</td>
<td>86</td>
<td>1.67</td>
<td>8.76</td>
</tr>
<tr>
<td>Mono\EFC\Europa\oesterrh</td>
<td>oestri024.krn</td>
<td>86</td>
<td>1.67</td>
<td>8.76</td>
</tr>
<tr>
<td>Mono\EFC\Europa\schweiz</td>
<td>suisse22.krn</td>
<td>47</td>
<td>1.42</td>
<td>10.52</td>
</tr>
<tr>
<td>Mono\EFC\Europa\schweiz</td>
<td>suisse55.krn</td>
<td>35</td>
<td>1.42</td>
<td>9.90</td>
</tr>
<tr>
<td>Mono\Friuli (Italy)</td>
<td>friuli046.krn</td>
<td>50</td>
<td>0.67</td>
<td>11.49</td>
</tr>
<tr>
<td>Mono\Friuli (Italy)</td>
<td>friuli049.krn</td>
<td>42</td>
<td>0.67</td>
<td>11.49</td>
</tr>
<tr>
<td>Mono\Friuli (Italy)</td>
<td>friuli082.krn</td>
<td>51</td>
<td>0.67</td>
<td>11.49</td>
</tr>
<tr>
<td>String Quartet\Beethoven\Vn-2</td>
<td>quartet14-3.krn</td>
<td>39</td>
<td>1.67</td>
<td>14.96</td>
</tr>
</tbody>
</table>
3.1.3 Adding Tessitura and Mobility

In search of other feature space, I include the index of tessitura and the index of mobility as I declared in Chapter 2.2.2. They are not extractable from the MedleyDB dataset as I explained in Chapter 2.3. Hence, I must exclude it temporarily.

First, the register width and the index of tessitura are linearly correlated with each other, cf. Figure 48. There are, however, not so many outliers very far from the trend. The completely empty areas may be either a whole new or an unacceptable space as I warned in Chapter 1.2.2. In fact, the register width naturally restricts the longest range of pitch deviation. Meanwhile, the index of tessitura is the standard deviation of pitch distribution as I informed in Chapter 2.2.2. Thereby, some areas are even unachievable due to their inherent relationship. This situation is also evident in the bottom left empty area of Figure 47 on page 66.

![Figure 48. Distribution of Register Width and Index of Tessitura from Samples in Table 12 except the MedleyDB Dataset.](image-url)
Next, the index of tessitura and the susceptibility have a similar correlation between the register width and the susceptibility, cf. Figure 47 on page 66 and Figure 49 below. It is reasonable as a result of the aforementioned linear correlation between the register width and the index of tessitura. Nonetheless, samples spread a bit more loosely in Figure 49. The trend is still clear but is not so strong. Some distant outliers exist, too.

![Figure 49](image)

**Figure 49.** Distribution of Index of Tessitura and Susceptibility from Samples in Table 12 except the MedleyDB Dataset.

Last but not least, the index of mobility does not discriminate between vocal samples and instrumental samples although low-mobility samples mostly come from vocal datasets, cf. Figure 50, Figure 51, and Figure 52. With regard to exploration, their trends are relatively vague in all the three distributions; their dispersions are slightly more even, particularly in Figure 52.
Figure 50. Distribution of Index of Tessitura and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset.

Figure 51. Distribution of Register Width and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset.
Figure 52. Distribution of Susceptibility and Index of Mobility from Samples in Table 12 except the MedleyDB Dataset.

To sum up, the most unexploited area exists in the bidimensional feature space of the register width and the susceptibility while a very small minority of the samples from yet diverse datasets stands beyond the frontier of the cluster’s territory as illustrated in Figure 47 on page 66. It goes without saying that one may consider more than two features at once to explore a higher-dimensional space as I reviewed in Chapter 1.2.2. Nevertheless, caution should be exercised that all distributions unavoidably become sparser in the higher space due to the “curse of dimensionality” [87], especially whenever the number of samples is insufficient.

3.2 Perspectives

As I indicated in Chapter 1.2.3, a statistical global feature is not capable of illustration of contextual effects. For that reason, I propose to observe features through windowing a piece. The sequential windows may overlap each other; the observation
scale could be small or large. That is to say, a feature can be extracted from audio frames, notes, measures, phrases, passages, or finally a piece. The variation, distribution, correlation, and relationship between features on each observation scale will reveal how artistic attributes differentiate from each other and hopefully what compositional tactics or analytical features could be conceived.

As a similar concept in signal processing, the length of each section is the window size; the length of each overlap is the hop size. The sizes determine the resolution of measurement and affect its accuracy as well as its precision. All sizes on each observation scale’s unit could be fixed, cumulative, or dynamic as exemplified in Table 17. A fixed size in larger unit is often made of varied sizes in smaller unit. For instance, every piece may comprise measures; every measure may comprise notes; every note may comprise frames (if in audio format); and so forth. In this sub-chapter, I compare the divergence between the external perspective (outside a piece) and the internal perspective (inside a piece).

<table>
<thead>
<tr>
<th>type</th>
<th>1st window</th>
<th>2nd window</th>
<th>3rd window</th>
</tr>
</thead>
<tbody>
<tr>
<td>fixed</td>
<td>0–4</td>
<td>1–5</td>
<td>2–6</td>
</tr>
<tr>
<td>cumulative</td>
<td>0–4</td>
<td>0–5</td>
<td>0–6</td>
</tr>
<tr>
<td>dynamic</td>
<td>depend</td>
<td>varied</td>
<td>varied</td>
</tr>
</tbody>
</table>

Table 17. Types of Window Size and Example of Each Window’s Address.

window size = 5; hop size = 1.

3.2.1 External

The correlation between global features is valuable information as connected in Figure 18 on page 22. Although Hall argued that “good feature subsets contain features highly correlated with the class, yet uncorrelated with each other” [88]-[89],
the feature-to-feature correlations might be relevant to an intercultural relationship, a stylistic tendency, or a composer’s habit. Thereupon, I calculate the global feature-to-feature Pearson correlation coefficients (PCC) through the scipy library which was specified in Table 13. This observation as recorded in Table 18 and Table 19 also leads me to the external versus internal perspectives on string quartet datasets.

Table 18. Global Feature-to-feature PCCs in Instrumental Samples.
cf. Chapter 2.2.1 and Chapter 2.2.2.

<table>
<thead>
<tr>
<th>feature</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>B: highest frequency</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>C: lowest frequency</td>
<td>0.01</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>D: central frequency</td>
<td>0.03</td>
<td>0.85</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>E: register width</td>
<td>0.07</td>
<td>0.54</td>
<td>-0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>F: largest energy ratio interval</td>
<td>0.07</td>
<td>0.47</td>
<td>-0.31</td>
<td>0.03</td>
<td>0.84</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G: index of tessitura</td>
<td>0.34</td>
<td>0.36</td>
<td>-0.44</td>
<td>-0.07</td>
<td>0.83</td>
<td>0.69</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H: index of mobility</td>
<td>0.23</td>
<td>0.30</td>
<td>0.12</td>
<td>0.25</td>
<td>0.25</td>
<td>0.10</td>
<td>0.28</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>I: susceptibility</td>
<td>-0.07</td>
<td>-0.20</td>
<td>0.31</td>
<td>0.06</td>
<td>-0.54</td>
<td>-0.52</td>
<td>-0.26</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

Table 19. Global Feature-to-feature PCCs in Vocal Samples.
cf. Chapter 2.2.1 and Chapter 2.2.2.

<table>
<thead>
<tr>
<th>feature</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: length</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B: highest frequency</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>C: lowest frequency</td>
<td>-0.06</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>D: central frequency</td>
<td>-0.02</td>
<td>0.94</td>
<td>0.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>E: register width</td>
<td>0.17</td>
<td>0.37</td>
<td>-0.29</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>F: largest energy ratio interval</td>
<td>0.13</td>
<td>0.19</td>
<td>-0.27</td>
<td>-0.03</td>
<td>0.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G: index of tessitura</td>
<td>0.18</td>
<td>0.29</td>
<td>-0.25</td>
<td>0.04</td>
<td>0.82</td>
<td>0.65</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>H: index of mobility</td>
<td>0.32</td>
<td>0.26</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.41</td>
<td>0.05</td>
<td>0.43</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>I: susceptibility</td>
<td>-0.12</td>
<td>-0.28</td>
<td>0.29</td>
<td>-0.02</td>
<td>-0.85</td>
<td>-0.57</td>
<td>-0.73</td>
<td>-0.47</td>
<td>1.00</td>
</tr>
</tbody>
</table>

One of the several features extracted from a wealth of datasets is the “register width” as I clarified in Chapter 2.2.1. The ranges and mean values of this feature correspond with my, if not everybody’s, general impression on each part of classical
string quartet, namely first violin, second violin, viola, and violoncello, cf. Figure 53. In average, each composer’s two outer parts (first violin and violoncello) have wider register width than the two inner parts (second violin and viola). In this genre, the dominant melodic role is usually solely played by first violin. As a result, it has the largest mean register width.

![Figure 53. Ranges and Means of Register Width.](image)

\( B = \text{Beethoven}; \ H = \text{Haydn}; \ M = \text{Mozart}. \) 1~4 denote part numbers of string quartet.

The mean values of another feature, the “susceptibility” as I also demystified in Chapter 2.2.1, between outer parts and inner parts are again distinguishable within each of the composer, cf. Figure 54. Most compellingly, the correlation between piece-by-piece register width and susceptibility discriminates the three composers across all four parts of string quartet as sorted in Table 20 [17]. At the same time, the duplicate discrimination exists in the first part of their piano sonatas, too [17], as listed in Table 21. The distance between the correlation coefficients is, however, not so far.
Figure 54. Ranges and Means of Susceptibility.
B = Beethoven; H = Haydn; M = Mozart. 1–4 denote part numbers of string quartet.

Table 20. Correlations between Susceptibilities and Register Widths in Each Part of String Quartets [17].

<table>
<thead>
<tr>
<th>dataset</th>
<th>PCC</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart\Vn-2</td>
<td>-0.75</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mozart\Vc</td>
<td>-0.74</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mozart\Va</td>
<td>-0.73</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Mozart\Vn-1</td>
<td>-0.69</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Haydn\Va</td>
<td>-0.67</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Haydn\Vc</td>
<td>-0.63</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Haydn\Vn-2</td>
<td>-0.60</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Haydn\Vn-1</td>
<td>-0.48</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Beethoven\Vc</td>
<td>-0.45</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Beethoven\Vn-2</td>
<td>-0.42</td>
<td>&lt; 0.002</td>
</tr>
<tr>
<td>Beethoven\Vn-1</td>
<td>-0.36</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Beethoven\Va</td>
<td>-0.32</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table 21. Correlations between Susceptibilities and Register Width in the First Part of Piano Sonatas [17].

<table>
<thead>
<tr>
<th>dataset</th>
<th>PCC</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mozart</td>
<td>-0.90</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Haydn</td>
<td>-0.82</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Beethoven</td>
<td>-0.61</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>
3.2.2 Internal

Contrary to piece-by-piece observation, a correlation coefficient between local features can be measured within a piece. The correlation has the capability to represent the concurrence and non-concurrence as I reviewed in Chapter 2.2.4. Multiple values of a feature can be extracted from a piece through the aforementioned windowing technique as exemplified in Table 17 on page 72. This approach is unlike other existing “local feature” [75] and “event feature” [77]-[78] methods. They encode melody note by note in fashions which are not applicable to global feature. In contrast to those methods, windowing is just to extract individual “global feature” values from divided sections. Like White’s windowed key-profile analysis which was performed to extract the “local key” [57], I borrow the term “local” in order to stand for the features which are extracted via windowing.

The retrieval through cumulative-size windowing does not tell much, cf. Figure 55. Conversely, through fixed-size windowing, the internal correlations between local susceptibilities and local register widths are still distinguishable between outer parts and inner parts within each of the composers cf. Figure 56. Caution should be used that its implication is quiet dissimilar from the external perspective in Figure 53, Figure 54, and Table 20. The fact is that the two outer parts (first violin and violoncello) have less significant negative correlation between susceptibility and register width than the two inner parts (second violin and viola). To put it differently, the inner parts seems more vulnerable to “the effect of tessitura” [81]. An interest left to future composers might be how to reduce the vulnerability.
Figure 55. Ranges and Means of PCC between Local Susceptibilities and Register Widths by Cumulative-size Windowing. 
initial window size = 10 notes; hop size = 1 note. B = Beethoven; H = Haydn; M = Mozart. 1~4 denote part numbers of string quartet.

Figure 56. Ranges and Means of PCC between Local Susceptibilities and Register Widths by Fixed-size Windowing. 
fixed window size = 10 notes; hop size = 1 note. B = Beethoven; H = Haydn; M = Mozart. 1~4 denote part numbers of string quartet.
When a global feature is extracted from a piece, the contextual information is lost. By contrast, local features are capable of presentation of contextual variation. For instance, generally speaking, the local susceptibility values are relatively smaller in the development sections of the two respective piano sonatas by Mozart and Beethoven as delineated in Figure 57 and Figure 58. Certainly, there are more advanced models to recognize music structures. Instead of the MIR task, it again gives future composers a course to challenge.

**Figure 57.** Melodic Profile of the 1st Movement of Mozart’s Piano Sonata No. 16 in C major, K. 545. fixed window size = 20 notes; hop size = 10. Red vertical lines are manual annotations about sections.
Figure 58. Melodic Profile of the 1st Movement of Beethoven’s Piano Sonata No. 21 in C major, Op.53. fixed window size = 20 notes; hop size = 10. Red vertical lines are manual annotations about sections.

Apart from correlations and variations, distributions of local features are informative, too. With intent to examine such distributions, I implement a program to extract local music21 features. In my current program for this task, the first window always starts from measure 0 so that it can include pickup measure in the very beginning of piece. For those which have no pickup measure, there is not any note information in measure 0. Thereby, the first window size is practically one measure shorter than other windows in this kind of case.

The first violin part and the violoncello part in the 1st movement of Mozart’s string quartet N. 15 in D minor (K. 421) demonstrate opposite tendencies. In this case, he composed more descending than ascending intervals when he wrote less repeated
notes for first violin, cf. Figure 59; he employed more ascending than descending intervals when deploying fewer than 5% repetitions for violoncello, cf. Figure 60.

Figure 59. Distribution and Path of Local Features from 1st Violin Part in the 1st Movement of Mozart’s String Quartet No. 15 in D minor, K. 421. fixed window size = 10 measures; hop size = 1 measure.
Figure 60. Distribution and Path of Local Features from Violoncello Part in the 1st Movement of Mozart’s String Quartet No. 15 in D minor, K. 421.
fixed window size = 10 measures; hop size = 1 measure.

On the other hand, when repeated notes exceeded 40%, ascending intervals were always more than descending ones in the first violin part, cf. Figure 59. In contrast, the violoncello part always has more descending than ascending when consecutive notes consisted of over 40% repetitions, cf. Figure 60.
Just similar to the data sciences which were reviewed in Chapter 1.1, the internal perspective via local feature is able to depict a composer’s behavior and habit under certain circumstances. It specifies which features should be included or what parameters could be manipulated in order to generate more innovative algorithmic compositions.

3.3 Evaluation

On purpose to evaluate the effectiveness of the above-mentioned internal perspective, I conduct several experiments. I focus on concurrence and non-concurrence features from separate parts rather than all parts together. Tasks are limited to composer binary classifications. Balanced samples in an identical genre are picked from two composers in the same region and period. From a feature set, I select a reduced number of features which are appropriate to the dataset.

3.3.1 Balanced Dataset

I consider string quartets by two prolific Austrian composers, Franz Joseph Haydn (1732–1809) and Wolfgang Amadeus Mozart (1756–1791). Herlands et al. described that they both “present an ideal case of stylistically homogeneous composers who lived contemporaneously, wrote in the same Western classical style.” [71] Walthew also wrote that “Mozart always acknowledged that it was from Haydn that he learnt how to write String Quartets.” [90] Moreover, Sapp and Liu’s online experiments confirmed that distinguishing between their string quartet movements is a difficult task for human being as charted in Figure 61 [91]. Thereupon, the binary classification between those two composers is always interesting.
On the other hand, Walthew reminded us that “Although Haydn was thus swayed by Mozart, his own individuality not only persisted, but seemed to become even more pronounced [...] the beginning of his later Quartets which in its humour and sprightliness is extremely characteristic, and could hardly be mistaken for the perhaps more finely touched sentiment of Mozart.” [90] On that account, I select Haydn’s later string quartets and Mozart’s all ones so that the samples are not too challenging to classify merely upon no more than 32 features (cf. Chapter 3.3.2 and Chapter 3.3.3). Again, all files are in the Humdrum representation, cf. Chapter 1.4.3, and come from the KernScores [53] as I summarized in Chapter 1.6.2 and Chapter 2.1.2.

In the original dataset, there are 212 movements of Haydn’s string quartet. First, I exclude the opus 103 since it is an unfinished work. Next, I remove 2
movements which have more than 4 parts in the original file format. Finally, I keep the latest 80 available movements as listed in Appendix A3.

The original dataset consists of not only Mozart’s string quartets but also two flute quartets and one oboe quartet. Although those specific instruments replace the first violin (Vn-1) part, all other 3 parts are still the second violin (Vn-2), viola (Va), and violoncello (Vc) as common string quartet instrumentation. I include them to maximize the amount of Mozart’s samples as listed in Appendix A3. I remove another piece “Adagio and Fugue in C minor, K 546” to balance the portion of Haydn’s samples.

3.3.2 Feature Sets

I pick features from McKay's jSymbolic feature set, cf. Chapter 2.2.3. Features are extracted globally from a whole piece or locally from segments. Thus a correlation coefficient and a p-value can be estimated from every two series of local base feature. Below are the 5 different feature sets.

**G0032:** First of all, I particularly select 32 global features as listed in Appendix A1. They are related to melody and pitch because the melodic essence which was advised in Chapter 2.2.4 is especially meaningful in separate parts. Features about key, chord, tempo, duration in seconds, number of voices, etc. are excluded due to insufficient information within a single part in the Humdrum representation, cf. Chapter 1.4.3.

**C0496:** Second, I locally extracted features in the fixed sliding window size (5, 10, or 15 measures) and hop size (1 measure) across the whole piece. 496 correlation
coefficients between every two series of local base feature can be retrieved from the 
foregoing 32 global features.

**P0496:** Third, as a byproduct of previous feature set, 496 two-tailed p-values 
are also retrieved. This feature set represents the significance of above 496 
correlations. I regard it as another set of concurrence or non-concurrence feature.

**B0992:** Next, both the C0496 and P0496 feature sets are coupled. This one is 
solely added for comparison between window sizes in Chapter 3.3.3.

**J0528:** Last but not least, I combine the G0032 and C0496 feature sets to 
construct a comprehensive one. It has 528 features in total. 32 are the global features; 
496 are the internal correlation coefficients.

3.3.3 Classification

I extract the global features from each movement; every local feature is 
extracted from a specific length of segment. Besides, I compare 3 window sizes all 
with the same hop size. Classification tasks are repeated through a 10-fold cross 
validation. The accuracy baseline of this kind of balanced binary classification is 50%. 
After choosing the best window size, I compare the effect of supervised feature 
selection which was specified in Chapter 2.1.3. Feature sets fairly compete from 1 to 
32-dimensional feature space. Finally, I retrieve mean feature scores from the G0032, 
C0496, P0496, and J0528 feature sets respectively.

In my initial experiments, I set the window sizes 5, 10, and 15 measures (bars); 
the hop size is always 1 measure. I pick from the best feature to all ones and average 
all the classification accuracy. For instance, in the B0992 feature set (n = 992), I set
the number of feature from 1 to 992 for feature selection and evaluate the mean accuracy. Inter-feature-set mean accuracy and inter-part mean accuracy are both counted, too. Results disclose that the C0496 feature set has a slightly better overall performance (59.85%) than B0992 (58.05%) and P0496 (53.73%), cf. Table 22. Most importantly, the mean accuracy of 5-measure window size (59.81%) is better than 10-measure (57.09%) and 15-measure (54.73%). Therefore, I exclusively choose the 5-measure window size for the following experiments.

<table>
<thead>
<tr>
<th>part window size</th>
<th>B0992</th>
<th>C0496</th>
<th>P0496</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vn-1 5</td>
<td>57.72%</td>
<td>59.64%</td>
<td>54.86%</td>
<td>57.41%</td>
</tr>
<tr>
<td>Vn-2 5</td>
<td>64.02%</td>
<td>60.21%</td>
<td>60.97%</td>
<td>61.73%</td>
</tr>
<tr>
<td>Va 5</td>
<td>56.31%</td>
<td>56.18%</td>
<td>53.86%</td>
<td>55.45%</td>
</tr>
<tr>
<td>Vc 5</td>
<td>65.13%</td>
<td>71.75%</td>
<td>57.04%</td>
<td>64.64%</td>
</tr>
<tr>
<td>mean 5</td>
<td>60.79%</td>
<td>61.95%</td>
<td>56.68%</td>
<td>59.81%</td>
</tr>
<tr>
<td>Vn-1 10</td>
<td>58.33%</td>
<td>57.54%</td>
<td>57.29%</td>
<td>57.72%</td>
</tr>
<tr>
<td>Vn-2 10</td>
<td>56.03%</td>
<td>65.47%</td>
<td>47.39%</td>
<td>56.29%</td>
</tr>
<tr>
<td>Va 10</td>
<td>66.26%</td>
<td>61.90%</td>
<td>60.65%</td>
<td>62.94%</td>
</tr>
<tr>
<td>Vc 10</td>
<td>51.63%</td>
<td>54.21%</td>
<td>48.40%</td>
<td>51.42%</td>
</tr>
<tr>
<td>mean 10</td>
<td>58.06%</td>
<td>59.78%</td>
<td>53.44%</td>
<td>57.09%</td>
</tr>
<tr>
<td>Vn-1 15</td>
<td>57.26%</td>
<td>56.13%</td>
<td>58.64%</td>
<td>57.34%</td>
</tr>
<tr>
<td>Vn-2 15</td>
<td>56.60%</td>
<td>58.72%</td>
<td>53.17%</td>
<td>56.16%</td>
</tr>
<tr>
<td>Va 15</td>
<td>58.59%</td>
<td>61.36%</td>
<td>48.35%</td>
<td>56.10%</td>
</tr>
<tr>
<td>Vc 15</td>
<td>48.72%</td>
<td>55.08%</td>
<td>44.12%</td>
<td>49.31%</td>
</tr>
<tr>
<td>mean 15</td>
<td>55.29%</td>
<td>57.82%</td>
<td>51.07%</td>
<td>54.73%</td>
</tr>
<tr>
<td>overall N/A</td>
<td>58.05%</td>
<td>59.85%</td>
<td>53.73%</td>
<td>57.21%</td>
</tr>
</tbody>
</table>

In terms of the first violin part, the original global features in G0032 obviously surpass the concurrence feature sets C0496 and P0496. Meanwhile, the joint feature set J0528 broadly performs closely to G0032, cf. Figure 62. Speaking of the second violin part, in all feature sets except G0032, the classification accuracy roughly grows together with the number of feature, cf. Figure 63. When it comes to the viola part,
concurrence and non-concurrence features overwhelmingly promote the classification accuracy in the feature set J0528, cf. Figure 64. As far as the violoncello part is concerned, all other feature sets impressively exceed the global feature set G0032, cf. Figure 65. What is more, the concurrence feature set C0496 has the most excellent overall classification accuracy in this instrumental part.

![Figure 62](image-url)  
*Figure 62. Accuracy on 1st Violin.*  
x-axis: number of feature; y-axis: classification accuracy.
Figure 63. Accuracy on 2nd Violin.
x-axis: number of feature; y-axis: classification accuracy.

Figure 64. Accuracy on Viola.
x-axis: number of feature; y-axis: classification accuracy.
In many cases above, larger numbers of feature do not always result in better classification accuracy. Sometimes, they perform even worse as a result of overfitting, e.g. the G0032 feature set in the second violin part and the viola part. “Overfitting is the term used to mean that you used a dataset to estimate the parameters of your model, but your model isn’t that good at capturing reality beyond your sampled data.” [2] In my experiment, the balanced dataset only contains 160 movements, namely samples. In each round of the 10-fold cross validation, the linear kernel SVM (cf. 2.1.3) can merely learn from 144 samples with ground truth to classify the remaining 16 samples. The model may fit to the 144 samples too well to generalize its inference especially in higher-dimensional space. In consequence, it is not able to discriminate the other samples in a correct way.

Figure 65. Accuracy on Violoncello.

x-axis: number of feature; y-axis: classification accuracy.
Beyond the classification tasks above, I retrieve feature scores from a supervised feature selection via the scikit-learn library, cf. Chapter 2.1.3. All samples are fetched at once in pursuit of this inspection task. For every feature set, I average the scores of its all features. I additionally calculate the mean score of the best 32 features for the C0496, P0496, and J0528 feature sets. The outcome affirms that the concurrence and non-concurrence features tremendously contribute in the violoncello part as compared in Table 23. Moreover, in the same dimensionality, information from internal correlations in average outperforms the original features on most binary composer classification tasks except the first violin part.

<table>
<thead>
<tr>
<th>feature set</th>
<th>type</th>
<th>Vn-1</th>
<th>Vn-2</th>
<th>Va</th>
<th>Vc</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>G0032 all</td>
<td>global</td>
<td>8.43</td>
<td>1.63</td>
<td>1.59</td>
<td>2.30</td>
<td>3.49</td>
</tr>
<tr>
<td>C0496 all</td>
<td>concurrence</td>
<td>1.74</td>
<td>1.99</td>
<td>1.54</td>
<td>4.87</td>
<td>2.54</td>
</tr>
<tr>
<td>P0496 all</td>
<td>concurrence</td>
<td>1.39</td>
<td>1.31</td>
<td>1.08</td>
<td>3.00</td>
<td>1.70</td>
</tr>
<tr>
<td>J0528 all</td>
<td>joint</td>
<td>2.15</td>
<td>1.97</td>
<td>1.55</td>
<td>4.72</td>
<td>2.60</td>
</tr>
<tr>
<td>C0496 top 32</td>
<td>concurrence</td>
<td>4.24</td>
<td>5.57</td>
<td>3.43</td>
<td>19.94</td>
<td>8.30</td>
</tr>
<tr>
<td>P0496 top 32</td>
<td>concurrence</td>
<td>2.33</td>
<td>3.36</td>
<td>2.24</td>
<td>11.62</td>
<td>4.89</td>
</tr>
<tr>
<td>J0528 top 32</td>
<td>joint</td>
<td>8.43</td>
<td>5.19</td>
<td>3.77</td>
<td>18.82</td>
<td>9.05</td>
</tr>
</tbody>
</table>

I further sort out the 20 features with the best mean scores as sorted in Table 24. Among all of them, 5 are global features; 15 are concurrence or non-concurrence features, more precisely, the internal correlation coefficients. The best one from the feature set P0496 is ranked 27th, which is out of the scope here.
Table 24. Top 20 Features with the Best Mean Scores. cf. Appendix A1 for feature names.

<table>
<thead>
<tr>
<th>feature</th>
<th>set</th>
<th>type</th>
<th>Vn-1</th>
<th>Vn-2</th>
<th>Va</th>
<th>Vc</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>P8</td>
<td>G0032</td>
<td>global</td>
<td>35.01</td>
<td>3.68</td>
<td>4.65</td>
<td>8.79</td>
<td>13.03</td>
</tr>
<tr>
<td>M6+M18</td>
<td>C0496</td>
<td>concurrence</td>
<td>1.80</td>
<td>8.11</td>
<td>5.78</td>
<td>31.24</td>
<td>11.73</td>
</tr>
<tr>
<td>M19</td>
<td>G0032</td>
<td>global</td>
<td>38.64</td>
<td>1.16</td>
<td>1.22</td>
<td>4.73</td>
<td>11.44</td>
</tr>
<tr>
<td>P10</td>
<td>G0032</td>
<td>global</td>
<td>37.96</td>
<td>1.29</td>
<td>1.12</td>
<td>4.86</td>
<td>11.31</td>
</tr>
<tr>
<td>P2+P10</td>
<td>C0496</td>
<td>concurrence</td>
<td>2.00</td>
<td>7.84</td>
<td>6.80</td>
<td>27.94</td>
<td>11.14</td>
</tr>
<tr>
<td>M18+P1</td>
<td>C0496</td>
<td>concurrence</td>
<td>9.68</td>
<td>4.83</td>
<td>3.98</td>
<td>25.92</td>
<td>11.10</td>
</tr>
<tr>
<td>M6+M18</td>
<td>C0496</td>
<td>concurrence</td>
<td>7.15</td>
<td>9.50</td>
<td>2.39</td>
<td>24.58</td>
<td>10.90</td>
</tr>
<tr>
<td>P7</td>
<td>G0032</td>
<td>global</td>
<td>23.68</td>
<td>1.19</td>
<td>1.94</td>
<td>14.18</td>
<td>10.25</td>
</tr>
<tr>
<td>M17</td>
<td>G0032</td>
<td>global</td>
<td>14.50</td>
<td>5.16</td>
<td>16.69</td>
<td>4.33</td>
<td>10.17</td>
</tr>
<tr>
<td>P8+P9</td>
<td>C0496</td>
<td>concurrence</td>
<td>0.21</td>
<td>12.31</td>
<td>4.45</td>
<td>22.25</td>
<td>9.81</td>
</tr>
<tr>
<td>P4+P5</td>
<td>C0496</td>
<td>concurrence</td>
<td>1.43</td>
<td>2.77</td>
<td>3.77</td>
<td>28.37</td>
<td>9.08</td>
</tr>
<tr>
<td>M11+P5</td>
<td>C0496</td>
<td>concurrence</td>
<td>1.84</td>
<td>1.52</td>
<td>6.18</td>
<td>26.76</td>
<td>9.07</td>
</tr>
<tr>
<td>M19+P9</td>
<td>C0496</td>
<td>concurrence</td>
<td>1.82</td>
<td>6.10</td>
<td>5.85</td>
<td>21.42</td>
<td>8.80</td>
</tr>
<tr>
<td>M7+P1</td>
<td>C0496</td>
<td>concurrence</td>
<td>0.23</td>
<td>8.54</td>
<td>3.46</td>
<td>22.17</td>
<td>8.60</td>
</tr>
<tr>
<td>P9</td>
<td>G0032</td>
<td>global</td>
<td>9.39</td>
<td>9.72</td>
<td>9.23</td>
<td>5.71</td>
<td>8.51</td>
</tr>
<tr>
<td>M6+P9</td>
<td>C0496</td>
<td>concurrence</td>
<td>2.58</td>
<td>4.56</td>
<td>2.39</td>
<td>24.51</td>
<td>8.51</td>
</tr>
<tr>
<td>M2+M18</td>
<td>C0496</td>
<td>concurrence</td>
<td>11.14</td>
<td>4.70</td>
<td>2.70</td>
<td>15.26</td>
<td>8.45</td>
</tr>
<tr>
<td>P10+P10</td>
<td>C0496</td>
<td>concurrence</td>
<td>4.13</td>
<td>6.86</td>
<td>5.96</td>
<td>16.74</td>
<td>8.42</td>
</tr>
<tr>
<td>P9+P10</td>
<td>C0496</td>
<td>concurrence</td>
<td>2.24</td>
<td>5.26</td>
<td>6.38</td>
<td>19.51</td>
<td>8.35</td>
</tr>
<tr>
<td>M18+P5</td>
<td>C0496</td>
<td>concurrence</td>
<td>5.50</td>
<td>0.30</td>
<td>2.52</td>
<td>24.43</td>
<td>8.19</td>
</tr>
</tbody>
</table>

Again, most features which have exceptional scores in the first violin part come from the global feature set G0032; those have better scores in the violoncello part come from the concurrence feature set C0496. Several features have high scores in more than one part.

3.3.4 Discussion

The first violin part is always the easiest one to classify in previous research using numerous global features sets [73]-[74]. By contrast, it is the most difficult part to be classified through concurrence and non-concurrence features as displayed in Figure 62 on page 87. Since the first violin part usually has the most dynamic and predominant melodic line in classical string quartets, I assume that traditional global
features are quite capable of discrimination between Haydn’s and Mozart’s outstanding melodies. Conversely, the other 3 parts have to sacrifice their freedom more often than the first violin. The selected global feature set G0032 is not able to facilitate the classification as much as my concurrence or joint feature sets, i.e. C0496, P0496, or J0524 even in the same n-dimensional feature space. Thereby, the inner states of concurrence and non-concurrence are informative features.

Although I have validated the discrimination capabilities which are boosted by concurrence and non-concurrence features in Chapter 3.3.3, it is worth to point out the state of the art of composer classification of separate string quartet parts. Hillewaere, Manderick, and Conklin selected 112 of 168 global features from 3 sets to classify 112 Haydn's and 95 Mozart's string quartet movements in the period 1779–1790. The accuracy baseline was 54.1%. After the leave-one-out cross validation (LOOCV), they at most achieved 74.4% (Vn-1), 66.2% (Vn-2), 62.8% (Va), and 75.4% (Vc) classification accuracy in either RBF-kernel SVM or 3-gram models [73].

With the same features, dataset, and LOOCV, Taminau et al. reached 73.0% (Vn-1), 62.8% (Vn-2), 68.6% (Va), and 69.6% (Vc) classification accuracy through subgroup discovery methods [74]. Given my aforementioned experiment results, I have the reason to believe some classification accuracy could be improved after corresponding concurrence and non-concurrence features are considered.

Apart from the discrimination capability, I inspect some internal correlations between two local base features in case of any abnormal circumstance. The most significant correlations across the two composers are reported in Table 25. The most
positive correlation is always between the features “size of melodic arcs” and “range”.

The most negative correlations are between the features “importance of middle
register” and “importance of high register”, or between “importance of bass register”
and “importance of middle register”. All above correlations are fairly reasonable.

Table 25. The Most Significant Internal Correlations.

<table>
<thead>
<tr>
<th>feature</th>
<th>part</th>
<th>PCC</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M19 and P10</td>
<td>Vn-1</td>
<td>≈ 0.968</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>M19 and P10</td>
<td>Vn-2</td>
<td>≈ 0.953</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>M19 and P10</td>
<td>Va</td>
<td>≈ 0.982</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>M19 and P10</td>
<td>Vc</td>
<td>≈ 0.995</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>P14 and P15</td>
<td>Vn-1</td>
<td>≈ -0.961</td>
<td>≈ 0.0008</td>
</tr>
<tr>
<td>P14 and P15</td>
<td>Vn-2</td>
<td>≈ -0.959</td>
<td>≈ 0.0063</td>
</tr>
<tr>
<td>P13 and P14</td>
<td>Va</td>
<td>≈ -0.833</td>
<td>≈ 0.0116</td>
</tr>
<tr>
<td>P13 and P14</td>
<td>Vc</td>
<td>≈ -0.879</td>
<td>≈ 0.0011</td>
</tr>
</tbody>
</table>

Speaking back of exploring unexploited compositional space, the two
examples below have small portions of empty area. Their distributions are not
condensed, cf. Figure 66 and Figure 67. At first glance, there is limited unexploited
space. On the contrary, chances sometimes exist beyond the range of coordinates. In
Figure 66, the “direction of motion” feature’s neutral value is 0.5, namely a status of
an equal number of ascending and descending intervals. The distribution reveals a
favor to smaller values. The boundaries on x-axis reach about 0.35 in the left-hand but
less than 0.60 in the right-hand. If the borders should be balanced, the range from 0.60
to 0.65 is a whole new territory. Similarly, every neutral correlation coefficient is zero.
In Figure 67, the y-axis ranging from -0.6 to 0.4 is unbalanced. Besides, a majority of
the samples has positive values in x-axis.
Figure 66. 1st Violin Distribution of M19 and M17 in Table 24.

Figure 67. Violoncello Distribution of P2+P10 and M6+M18 in Table 24.

This kind of exploration is to penetrate the genre frontier. It is unlike the intercultural or universal frontier as I encountered in Chapter 3.1. All the samples here
belong to one genre because my primary goal in Chapter 3.3 is to evaluate the improvement in classification tasks with the internal perspective. Nonetheless, concurrence and non-concurrence features themselves are cross-level, intercultural, and even universal as I reminded in Chapter 2.3. They are applicable to any time series data.

### 3.4 Composition

On purpose to create something in the unexploited space, the first step is to understand the features. Second, inspecting some outliers may point a direction to try. Third, new compositions need to be analyzed by the same retrieval program so that their locations in the dimensional feature space can be validated. In this sub-chapter, I will exclusively explore the space of “register width” and “susceptibility” features.

#### 3.4.1 Fundamentals

As I elaborated in Chapter 2.2.1, the susceptibility denotes the sensitivity of energy transfer. The energy ratio is based on the audio frequency ratio. The frequency difference between pitches is a nonlinear relationship; the interval between consecutive pitches is, however, a linear representation. Hence, a melody’s susceptibility is sometimes contrary to intuition or to common-sense expectation.

For instance, an ascending diatonic scale in any octave (cf. Figure 68 and Figure 69) has a larger susceptibility than a descending diatonic scale (cf. Figure 70 and Figure 71) no matter in which octave. This fact is also true on chromatic scales as notated in Figure 72 and Figure 73.
Figure 68. Ascending Diatonic Scale.
register width = 1; susceptibility = 7.67.

![Ascending Diatonic Scale Diagram]

Figure 69. Melodic Pitch Contour (middle line) of Figure 68.
The first point of middle line denotes the original melody’s central frequency, cf. Chapter 2.2.1.

![Melodic Pitch Contour Diagram]

Figure 70. Descending Diatonic Scale.
register width = 1; susceptibility = 6.21.

![Descending Diatonic Scale Diagram]
Figure 71. Melodic Pitch Contour (middle line) of Figure 70.
The first point of middle line denotes the original melody's central frequency, cf. Chapter 2.2.1.

Figure 72. Ascending Chromatic Scale.
register width = 1; susceptibility = 7.67

Figure 73. Descending Chromatic Scale.
register width = 1; susceptibility = 6.21.

By contrast, an ascending melody contains alternating minor third intervals and
minor second returns has a smaller susceptibility than a descending melody as
exemplified in Figure 74 and Figure 75.
Therefore, a melodic inversion always changes the original susceptibility, but it may be either increasing or decreasing.

3.4.2 Outliers

After looking into the outliers in Table 16 on page 67, I found an interesting characteristic across most cases which have both intermediate or wide register widths and large susceptibilities. They do not reach higher registers until the very end. For example, the four Mazurkas are all relatively much longer pieces which do not touch their highest registers until the end. Besides, the two China melodies consist of a short sample and a longer sample which both end in their top pitches as illustrated in Figure 76 and Figure 77. Finally, the only opera aria sample is the tenor part (Alfredo) of the piece. Although he mostly sings around one octave, the melody suddenly culminates in the very end of duo with the soprano part (Violetta).
3.4.3 Experiments

I manually write and select some pieces. In my first example below, I try to mimic the characteristic which was inferred in previous paragraph (Chapter 3.4.2).

The melodic pitch contour is not dynamic. Nevertheless, it abruptly climbs in the end as shown in Figure 78. Its register width and susceptibility is quite away from the
main cluster and is close to an instrumental outlier (Mazurka) and several vocal ones, cf. Figure 79. This unusual sample locates in the unexploited space but not far beyond the frontier.

![Figure 78. Experimental Composition 1. register width = 2.08; susceptibility = 9.22.](image)

**Figure 78.** Experimental Composition 1. register width = 2.08; susceptibility = 9.22.

![Figure 79. Distribution of Register Width and Susceptibility from All Samples in Figure 47 and Four Unprecedented Samples. coordinates of four new samples: Figure 78 (9.22, 2.08); Figure 80 (17.1, 2); Figure 81 (12.74, 5.58); Figure 82 (3.34, 4.58).](image)

**Figure 79.** Distribution of Register Width and Susceptibility from All Samples in Figure 47 and Four Unprecedented Samples. coordinates of four new samples: Figure 78 (9.22, 2.08); Figure 80 (17.1, 2); Figure 81 (12.74, 5.58); Figure 82 (3.34, 4.58).

In my second example, I arrange a short phrase which originally stays within a perfect fifth range. I make the melody rise steeply in the last measure, as displayed in Figure 80. This sample locates in an emptier area, cf. Figure 79 above. Its register
width is a little narrower than the first example; its susceptibility is, however, much larger.

\[ \text{Figure 80. Experimental Arrangement.} \\
\text{register width} = 2.00; \text{susceptibility} = 17.10. \]

My third example pushes the border of register width while maintaining a quite large susceptibility. It begins in bass register and concludes in soprano register as notated in Figure 81. This example penetrates the frontier and reaches an unclaimed territory, cf. Figure 79 on page 100. Of course, its melodic contour does not match most of common cases. In other words, the cluster in Figure 79 represents our general impression on melody. If an outlier locates further from the cluster, its style is more novel according to these two features.

\[ \text{Figure 81. Experimental Composition 2.} \\
\text{register width} = 5.58; \text{susceptibility} = 12.74. \]

My final example is an unsuccessful attempt. I compose a melody from the high register to the low register as exhibited in Figure 82. Its register width is wide, but its susceptibility is fairly small. This relationship corresponds with the trend line.
This, this sample is still in the cluster even though the nearby population is sparse, cf. Figure 79 on page 100.

![Figure 82. Experimental Composition 3.](image)

My experiments demonstrate the manipulation of the register width and the susceptibility. The challenge is how to control the two features together by a well-designed algorithm. In my old compositional model which was mentioned in Chapter 2.2.1, the “range” (convertible into register width) and “susceptibility” parameters are the maximal possible values to be allowed by the program, cf. Figure 83. Its generative results do not necessarily reach the maxima. Thereby, I need to rework the algorithm in the future.
Figure 83. Screenshot of My Old Program [11].
4. Conclusion

One benefit of generative systems for artists is the ability to explore the variety of output from a designed process; the artistic act can be found in the design of the system, as well as the selection from its output.

Philippe Pasquier, Arne Eigenfeldt, Oliver Bown, and Shlomo Dubnov [92]
In this final chapter, I will remark my contributions and the applications. There are still unanswered questions due to the limitation in my research. On that account, I will propose some future works, too.

4.1 Contribution

First, I initiate innovative praxis. I not only conceptualize it but also implement my own mathematical model and retrieval programs which were elaborated in Chapter 2.1.3 and Chapter 2.2.1. The praxis may bring benefits to AC, MIR, and CMA as I stated in Chapter 1.2.3. While most of MIR scientists seek the best features which are perfect for discrimination between samples in dissimilar classes, I prefer features which have the capability to unveil the incompletely exploited space. They like to see distributions where samples from distinctive classes are clearly divided. By contrast, I am more interesting in distributions where almost all samples cluster somewhere with a tiny ratio of diverse outliers spreading out. CMA and MIR often adopt culture-aware and domain-specific approaches, which are able to precisely retrieve information from samples. Conversely, I favor intercultural and cross-level ways so that the constraints from the original styles do not limit my imagination and freedom in AC as I described in Chapter 2.3.

Second, I extract the susceptibility features (cf. Chapter 2.2.1) from numerous datasets. The 18009 samples in 91 sub-datasets come from various genres in Renaissance, Baroque, Classical, Romantic, and modern eras across three continents as I informed in Chapter 1.6 and Chapter 2.1.2. The intercultural and cross-level retrieval confirms my hypothesis in previous research [11] that vocal styles broadly
have larger susceptibilities than instrumental styles in average. It also reveals the significant correlation between susceptibility and register width as I depicted in Chapter 3.1.2.

Third, I realize the inner states of concurrence and non-concurrence in Chapter 2.2.4. They have not been quantitatively examined. I propose to extract the internal PCCs and p-values from individual pieces. In Chapter 3.3, I implement other programs in order to retrieve such information and to evaluate their performance and improvement on the binary composer classification of separated string quartet parts. Results disclose that, in the same dimensionality, information from internal correlations outperforms the original features on most binary composer classification tasks except the first violin part, which has been shown to have the best accuracy in previous research. This method extends common bag-of-feature approaches and is directly applicable to any feature set specifically for time series data.

Fourth, I manipulate the register width and susceptibility on purpose to explore the unexploited area in the bidimensional feature space. My composition examples demonstrate different successes. The unpresented samples reach the unclaimed territories. They have unusual melodic contours which are distinct from most of common cases. Those experiments bring me a challenge to rework my old compositional algorithms so that the new model is able to reach both the two maximal values of feature in the same piece.
4.2 Application

The exploration of unexploited compositional space is a novel way to learn and observe music. It can be applied to any kind of music. One may consider including culture-aware and domain-specific features if the goal only is to penetrate the genre frontier, in contrast with intercultural and universal frontiers, as I distinguished in Chapter 3.3.4. When exploring higher-dimensional space, caution should be exercised to watch out for the “curse of dimensionality” [87], cf. Chapter 3.1.3.

The method of concurrence and non-concurrence extends common bag-of-feature approaches and is directly applicable to any feature set specifically for time series data. Concurrence or non-concurrence features themselves are cross-level, intercultural, and even universal. They rely on the base features in whatever levels. Hence, the internal PCC and p-value can be extracted from any two series of local features even if they belong to different levels.

4.3 Limitation

In this dissertation, all my explorations are approached by visually examining the distributions on two-dimensional areas. How to discover the unexploited high-dimensional space remains a challenge for me. I believe that there must be a mathematical way to describe the non-empty sparsity (with diverse outliers) which I have been looking for.

Besides, more intercultural and cross-level base features need to be added into my exploration. I have started converting some jSymbolic features (cf. Chapter 2.2.3) into more versatile ones so that the shrunk new edition of feature set could be applied
to more types of dataset. There are also several high-level feature sets [75] which I should consider including. Thus, my current study is merely a very small step of the long-term, large-scale, intercultural and cross-level exploration.

Finally, Wiggins has argued that “Music (as opposed to sound) can be studied only in a context which explicitly allows for, and is built on, (albeit de facto) models of human perception; to do otherwise is not to study Music at all.” [25] The unexploited space cannot always guarantee a psychologically or psychoacoustically valid feature, let alone a controllable parameter in compositional algorithms as I reminded in Chapter 1.2.3. It just illuminates some possibilities which composers have to find out what and how to pursue next.

4.4 Expectation

There are several other symbolic datasets as I recommended in Chapter 1.6.2. I am supposed to import some of them into my future works to enrich the data diversity. Additionally, I hope to seek monophonic audio recordings due to the insufficient non-Western symbolic music datasets (cf. Chapter 1.6.2 and Chapter 1.6.3) and thanks to my successful experiments on cross-level features as I reported in this dissertation.

Aside from the most significant correlations as I discussed in Chapter 3.3.4, there are many other significant ones which are worth to analyze. One can even respectively look into Haydn’s and Mozart’s internal correlations to compare their distinction, especially for the concurrence and non-concurrence features which have top scores as I listed in Table 24. This information may hopefully “expand composer’s point of view to design creative compositional algorithms.” [16]
Concurrence and non-concurrence features should exist in not only symbolic but also sub-symbolic data provided that investigators find the appropriate unit and length of window size for audio samples. Even if separate parts are not accessible, from the mixed audio I can at least extract series of local feature in any kind and afterwards retrieve the internal correlation coefficients and p-values. Thereupon, I propose to include the internal concurrence and non-concurrence as a valuable musical feature for further investigation.

As I informed in Chapter 1.2.3, researchers may deploy state-of-the-art ANN to automatically learn features and detect unexploited compositional space. Contemporary composers have to, however, select controllable features and devise AC algorithms as well as parameters unless they do not want to compose by their own. My ultimate goal is to evolve toward the “mutualism” [17]. Composers learn from the dimensional feature space and distributions through MIR and CMA with intent to devise better algorithms and parameters to manipulate in AC. Then AC has the capability to promote better techniques and features for MIR and CMA, which again stimulate composer’s imagination and creativity.
Appendix

A1. Selected jSymbolic Features from [15]

Sixteen Pitch Features

- P1: Most Common Pitch Prevalence
- P2: Most Common Pitch Class Prevalence
- P3: Relative Strength of Top Pitches
- P4: Relative Strength of Top Pitch Classes
- P5: Interval Between Strongest Pitches
- P6: Interval Between Strongest Pitch Classes
- P7: Number of Common Pitches
- P8: Pitch Variety
- P9: Pitch Class Variety
- P10: Range
- P11: Most Common Pitch
- P12: Primary Register
- P13: Importance of Bass Register
- P14: Importance of Middle Register
- P15: Importance of High Register
- P16: Most Common Pitch Class

Sixteen Melody Features

- M2: Average Melodic Interval
- M3: Most Common Melodic Interval
- M4: Distance Between Most Common Melodic Intervals
- M5: Most Common Melodic Interval Prevalence
- M6: Relative Strength of Most Common Intervals
- M7: Number of Common Melodic Intervals
- M9: Repeated Notes
- M10: Chromatic Motion
- M11: Stepwise Motion
- M12: Melodic Thirds
- M13: Melodic Fifths
- M14: Melodic Tritones
- M15: Melodic Octaves
- M17: Direction of Motion
- M18: Duration of Melodic Arcs
- M19: Size of Melodic Arcs
A2. Susceptibilities of All Datasets

Part One of Three (mean ≥ 6)

**Table 26. Susceptibilities of the Datasets (mean ≥ 6).**

<table>
<thead>
<tr>
<th>dataset</th>
<th>samples</th>
<th>max</th>
<th>mean</th>
<th>min</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono\EFC\Europa\jugoslav</td>
<td>115</td>
<td>20.15</td>
<td>8.13</td>
<td>4.97</td>
<td>2.50</td>
<td>0.31</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\kinder</td>
<td>213</td>
<td>14.54</td>
<td>7.62</td>
<td>4.97</td>
<td>2.08</td>
<td>0.27</td>
</tr>
<tr>
<td>Mono\Friuli (Italy)</td>
<td>80</td>
<td>12.78</td>
<td>7.54</td>
<td>4.79</td>
<td>1.88</td>
<td>0.25</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\variant</td>
<td>26</td>
<td>9.64</td>
<td>7.17</td>
<td>4.79</td>
<td>1.11</td>
<td>0.15</td>
</tr>
<tr>
<td>Mono\EFC\Europa\czech</td>
<td>43</td>
<td>12.78</td>
<td>6.88</td>
<td>5.4</td>
<td>1.51</td>
<td>0.22</td>
</tr>
<tr>
<td>Mono\EFC\Europa\schweiz</td>
<td>93</td>
<td>29.38</td>
<td>6.86</td>
<td>4.79</td>
<td>2.82</td>
<td>0.41</td>
</tr>
<tr>
<td>Mono\EFC\Europa\rossiya</td>
<td>37</td>
<td>12.78</td>
<td>6.85</td>
<td>4.97</td>
<td>1.66</td>
<td>0.24</td>
</tr>
<tr>
<td>Mono\EFC\Europa\nederland</td>
<td>85</td>
<td>11.95</td>
<td>6.81</td>
<td>4.79</td>
<td>1.24</td>
<td>0.18</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\altdeu1</td>
<td>309</td>
<td>14.54</td>
<td>6.72</td>
<td>4.79</td>
<td>1.28</td>
<td>0.19</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\185\Alto</td>
<td>185</td>
<td>10.24</td>
<td>6.67</td>
<td>4.79</td>
<td>1.17</td>
<td>0.18</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\371\Alto</td>
<td>370</td>
<td>10.24</td>
<td>6.63</td>
<td>4.79</td>
<td>1.05</td>
<td>0.16</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\altdeu2</td>
<td>316</td>
<td>18.85</td>
<td>6.62</td>
<td>4.64</td>
<td>1.20</td>
<td>0.18</td>
</tr>
<tr>
<td>Mono\EFC\Europa\polska</td>
<td>25</td>
<td>9.02</td>
<td>6.61</td>
<td>5.31</td>
<td>1.10</td>
<td>0.17</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\371\Soprano</td>
<td>370</td>
<td>10.61</td>
<td>6.59</td>
<td>4.97</td>
<td>1.07</td>
<td>0.17</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\185\Soprano</td>
<td>185</td>
<td>10.24</td>
<td>6.59</td>
<td>4.97</td>
<td>0.99</td>
<td>0.15</td>
</tr>
<tr>
<td>Mono\EFC\Europa\magyar</td>
<td>45</td>
<td>11.95</td>
<td>6.52</td>
<td>4.97</td>
<td>1.52</td>
<td>0.23</td>
</tr>
<tr>
<td>Mono\EFC\Europa\italia</td>
<td>8</td>
<td>7.9</td>
<td>6.52</td>
<td>5.78</td>
<td>0.71</td>
<td>0.11</td>
</tr>
<tr>
<td>Mono\EFC\Europa\sass</td>
<td>91</td>
<td>12.78</td>
<td>6.47</td>
<td>4.79</td>
<td>1.51</td>
<td>0.23</td>
</tr>
<tr>
<td>Mono\EFC\Europa\danmark</td>
<td>9</td>
<td>8.1</td>
<td>6.44</td>
<td>5.17</td>
<td>0.84</td>
<td>0.13</td>
</tr>
<tr>
<td>Mono\EFC\Europa\france</td>
<td>14</td>
<td>8.4</td>
<td>6.42</td>
<td>5.68</td>
<td>0.86</td>
<td>0.13</td>
</tr>
<tr>
<td>Mono\EFC\Europa\lothring</td>
<td>71</td>
<td>9.02</td>
<td>6.42</td>
<td>5.31</td>
<td>0.78</td>
<td>0.12</td>
</tr>
<tr>
<td>Mono\EFC\Europa\romania</td>
<td>28</td>
<td>9.64</td>
<td>6.41</td>
<td>4.97</td>
<td>1.07</td>
<td>0.17</td>
</tr>
<tr>
<td>Mono\EFC\Europa\oesterrh</td>
<td>104</td>
<td>12.39</td>
<td>6.40</td>
<td>4.17</td>
<td>1.55</td>
<td>0.24</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\ballad</td>
<td>687</td>
<td>13.43</td>
<td>6.40</td>
<td>4.79</td>
<td>1.09</td>
<td>0.17</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\erk</td>
<td>1700</td>
<td>18.85</td>
<td>6.39</td>
<td>3.85</td>
<td>1.25</td>
<td>0.20</td>
</tr>
<tr>
<td>Mono\EFC\Europa\svrige</td>
<td>11</td>
<td>8.1</td>
<td>6.34</td>
<td>5.17</td>
<td>0.80</td>
<td>0.13</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\185\Tenor</td>
<td>185</td>
<td>12.78</td>
<td>6.33</td>
<td>4.79</td>
<td>0.92</td>
<td>0.14</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\zuca</td>
<td>616</td>
<td>11.95</td>
<td>6.32</td>
<td>4.26</td>
<td>1.04</td>
<td>0.16</td>
</tr>
<tr>
<td>Harm\Bach\Chorale\371\Tenor</td>
<td>370</td>
<td>12.78</td>
<td>6.30</td>
<td>4.79</td>
<td>0.86</td>
<td>0.14</td>
</tr>
<tr>
<td>Mono\EFC\Europa\luxemburg</td>
<td>8</td>
<td>7.08</td>
<td>6.26</td>
<td>4.97</td>
<td>0.81</td>
<td>0.13</td>
</tr>
<tr>
<td>Mono\EFC\Europa\ukraina</td>
<td>13</td>
<td>7.9</td>
<td>6.25</td>
<td>5.26</td>
<td>0.88</td>
<td>0.14</td>
</tr>
<tr>
<td>Mono\pentatonic</td>
<td>67</td>
<td>14.54</td>
<td>6.21</td>
<td>4.79</td>
<td>1.30</td>
<td>0.21</td>
</tr>
<tr>
<td>Mono\EFC\Europa\misc</td>
<td>30</td>
<td>11.95</td>
<td>6.18</td>
<td>4.12</td>
<td>1.77</td>
<td>0.29</td>
</tr>
<tr>
<td>Mono\American\Pawnee</td>
<td>86</td>
<td>14.54</td>
<td>6.18</td>
<td>4.19</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Mono\EFC\Europa\england</td>
<td>4</td>
<td>8.1</td>
<td>6.14</td>
<td>5.38</td>
<td>1.14</td>
<td>0.19</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\dva</td>
<td>106</td>
<td>11.95</td>
<td>6.12</td>
<td>4.32</td>
<td>1.24</td>
<td>0.20</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\faller\kbd</td>
<td>110</td>
<td>11.95</td>
<td>6.05</td>
<td>4.79</td>
<td>0.96</td>
<td>0.16</td>
</tr>
<tr>
<td>Nottingham\Christmas</td>
<td>13</td>
<td>7.3</td>
<td>6.04</td>
<td>4.79</td>
<td>0.63</td>
<td>0.10</td>
</tr>
</tbody>
</table>
### Table 27. Susceptibilities of the Datasets (5 ≤ mean < 6).

<table>
<thead>
<tr>
<th>dataset</th>
<th>samples</th>
<th>max</th>
<th>mean</th>
<th>min</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mono\EFC\Europa\deutsch\boehme</td>
<td>704</td>
<td>15.54</td>
<td>5.99</td>
<td>4.14</td>
<td>0.94</td>
<td>0.16</td>
</tr>
<tr>
<td>Mono\British children's songs</td>
<td>38</td>
<td>8.89</td>
<td>5.97</td>
<td>4.5</td>
<td>1.05</td>
<td>0.18</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\fink</td>
<td>566</td>
<td>11.95</td>
<td>5.96</td>
<td>4.26</td>
<td>0.91</td>
<td>0.15</td>
</tr>
<tr>
<td>Mono\Nova Scotia (Canada)</td>
<td>152</td>
<td>15.54</td>
<td>5.96</td>
<td>4.5</td>
<td>1.24</td>
<td>0.21</td>
</tr>
<tr>
<td>Mono\EFC\China\han</td>
<td>1222</td>
<td>14.67</td>
<td>5.95</td>
<td>4.1</td>
<td>1.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Harmo\Deutscher Liederschatz</td>
<td>231</td>
<td>9.9</td>
<td>5.93</td>
<td>3.91</td>
<td>0.88</td>
<td>0.15</td>
</tr>
<tr>
<td>JRP\Secular Song</td>
<td>157</td>
<td>12.39</td>
<td>5.62</td>
<td>4.37</td>
<td>1.05</td>
<td>0.17</td>
</tr>
<tr>
<td>Mono\EFC\China\natmin</td>
<td>206</td>
<td>9.02</td>
<td>5.58</td>
<td>4.16</td>
<td>0.86</td>
<td>0.15</td>
</tr>
<tr>
<td>Mono\EFC\China\xinhua</td>
<td>10</td>
<td>7.9</td>
<td>5.58</td>
<td>3.84</td>
<td>1.57</td>
<td>0.28</td>
</tr>
<tr>
<td>Mono\American\Sioux</td>
<td>245</td>
<td>11.95</td>
<td>5.54</td>
<td>3.79</td>
<td>1.06</td>
<td>0.19</td>
</tr>
<tr>
<td>Nottingham\Playford</td>
<td>15</td>
<td>7.4</td>
<td>5.51</td>
<td>4.26</td>
<td>0.72</td>
<td>0.13</td>
</tr>
<tr>
<td>Nottingham\Slip Jigs</td>
<td>11</td>
<td>6.87</td>
<td>5.50</td>
<td>4.27</td>
<td>0.76</td>
<td>0.14</td>
</tr>
<tr>
<td>Nottingham\Morris</td>
<td>31</td>
<td>7.08</td>
<td>5.49</td>
<td>4.79</td>
<td>0.68</td>
<td>0.12</td>
</tr>
<tr>
<td>Mono\EFC\Europa\deutsch\test</td>
<td>12</td>
<td>6</td>
<td>5.46</td>
<td>4.97</td>
<td>0.41</td>
<td>0.08</td>
</tr>
<tr>
<td>Mono\EFC\Europa\tirol</td>
<td>14</td>
<td>7.1</td>
<td>5.45</td>
<td>4.56</td>
<td>0.76</td>
<td>0.14</td>
</tr>
<tr>
<td>Mono\EFC\China\shanxi</td>
<td>802</td>
<td>9.9</td>
<td>5.42</td>
<td>4.17</td>
<td>0.60</td>
<td>0.11</td>
</tr>
<tr>
<td>Nottingham\Hornpipes</td>
<td>65</td>
<td>9.02</td>
<td>5.40</td>
<td>4.5</td>
<td>0.67</td>
<td>0.12</td>
</tr>
<tr>
<td>Nottingham\Ashover</td>
<td>46</td>
<td>7.37</td>
<td>5.38</td>
<td>4.32</td>
<td>0.75</td>
<td>0.14</td>
</tr>
<tr>
<td>Symb\Tri\Vocal</td>
<td>1833</td>
<td>15.07</td>
<td>5.31</td>
<td>3.93</td>
<td>0.90</td>
<td>0.17</td>
</tr>
<tr>
<td>Mono\American\Ojibway</td>
<td>42</td>
<td>7.9</td>
<td>5.30</td>
<td>4.88</td>
<td>0.88</td>
<td>0.17</td>
</tr>
<tr>
<td>Nottingham\Waltzes</td>
<td>52</td>
<td>6.82</td>
<td>5.29</td>
<td>4.27</td>
<td>0.59</td>
<td>0.11</td>
</tr>
<tr>
<td>Nottingham\Reels</td>
<td>461</td>
<td>9.02</td>
<td>5.27</td>
<td>3.79</td>
<td>0.73</td>
<td>0.14</td>
</tr>
<tr>
<td>JRP\Motet</td>
<td>178</td>
<td>7.5</td>
<td>5.22</td>
<td>4.17</td>
<td>0.59</td>
<td>0.11</td>
</tr>
<tr>
<td>Keyboard\Mazurka</td>
<td>55</td>
<td>16.3</td>
<td>5.14</td>
<td>3.42</td>
<td>2.44</td>
<td>0.47</td>
</tr>
<tr>
<td>Nottingham\Jigs</td>
<td>340</td>
<td>7.9</td>
<td>5.13</td>
<td>3.92</td>
<td>0.61</td>
<td>0.12</td>
</tr>
<tr>
<td>Harmo\Bach Chorale\371\Bass</td>
<td>370</td>
<td>6.87</td>
<td>5.11</td>
<td>4.14</td>
<td>0.54</td>
<td>0.11</td>
</tr>
<tr>
<td>JRP\Mass</td>
<td>411</td>
<td>7.08</td>
<td>5.09</td>
<td>4.17</td>
<td>0.50</td>
<td>0.10</td>
</tr>
<tr>
<td>Harmo\Bach Chorale\185\Bass</td>
<td>185</td>
<td>6.87</td>
<td>5.08</td>
<td>4.14</td>
<td>0.52</td>
<td>0.10</td>
</tr>
<tr>
<td>Aria</td>
<td>177</td>
<td>9.02</td>
<td>5.05</td>
<td>3.82</td>
<td>0.65</td>
<td>0.13</td>
</tr>
</tbody>
</table>
Part Three of Three (mean < 5)

Table 28. Susceptibilities of the Datasets (mean < 5).

<table>
<thead>
<tr>
<th>dataset</th>
<th>samples</th>
<th>max</th>
<th>mean</th>
<th>min</th>
<th>SD</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SymbTr\Instrumental</td>
<td>187</td>
<td>7.3</td>
<td>4.67</td>
<td>3.82</td>
<td>0.52</td>
<td>0.11</td>
</tr>
<tr>
<td>MedleyDB\Vocal</td>
<td>61</td>
<td>7.65</td>
<td>4.43</td>
<td>3.24</td>
<td>0.72</td>
<td>0.16</td>
</tr>
<tr>
<td>Keyboard\Chopin Prelude</td>
<td>24</td>
<td>6.87</td>
<td>4.42</td>
<td>3.07</td>
<td>0.85</td>
<td>0.19</td>
</tr>
<tr>
<td>Keyboard\Bach WTC Fugue</td>
<td>44</td>
<td>5.2</td>
<td>4.34</td>
<td>3.52</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td>String Quartet\Haydn\Va</td>
<td>201</td>
<td>7.08</td>
<td>4.29</td>
<td>3.69</td>
<td>0.44</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Mozart\Va</td>
<td>82</td>
<td>5.7</td>
<td>4.24</td>
<td>3.6</td>
<td>0.37</td>
<td>0.09</td>
</tr>
<tr>
<td>String Quartet\Mozart\Vn-2</td>
<td>82</td>
<td>5.68</td>
<td>4.23</td>
<td>3.52</td>
<td>0.44</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Beethoven\Vn-2</td>
<td>71</td>
<td>14.96</td>
<td>4.22</td>
<td>3.37</td>
<td>1.34</td>
<td>0.32</td>
</tr>
<tr>
<td>String Quartet\Haydn\Vn-2</td>
<td>201</td>
<td>6.07</td>
<td>4.17</td>
<td>3.52</td>
<td>0.38</td>
<td>0.09</td>
</tr>
<tr>
<td>String Quartet\Mozart\Vc</td>
<td>82</td>
<td>5.62</td>
<td>4.15</td>
<td>3.28</td>
<td>0.44</td>
<td>0.11</td>
</tr>
<tr>
<td>String Quartet\Beethoven\Va</td>
<td>71</td>
<td>5.17</td>
<td>4.14</td>
<td>3.69</td>
<td>0.33</td>
<td>0.08</td>
</tr>
<tr>
<td>MedleyDB\Instrumental</td>
<td>47</td>
<td>5.8</td>
<td>4.12</td>
<td>3.17</td>
<td>0.63</td>
<td>0.15</td>
</tr>
<tr>
<td>WJazzD</td>
<td>453</td>
<td>6.23</td>
<td>4.09</td>
<td>3.24</td>
<td>0.42</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Haydn\Vc</td>
<td>201</td>
<td>6.13</td>
<td>4.04</td>
<td>3.42</td>
<td>0.39</td>
<td>0.10</td>
</tr>
<tr>
<td>Keyboard\Mozart Sonata</td>
<td>69</td>
<td>6.21</td>
<td>4.03</td>
<td>3.09</td>
<td>0.58</td>
<td>0.14</td>
</tr>
<tr>
<td>Keyboard\Clementi Sonatina</td>
<td>17</td>
<td>4.68</td>
<td>4.00</td>
<td>3.55</td>
<td>0.35</td>
<td>0.09</td>
</tr>
<tr>
<td>Keyboard\Scott Joplin</td>
<td>47</td>
<td>5.13</td>
<td>3.97</td>
<td>3.3</td>
<td>0.41</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Mozart\Vn-1</td>
<td>82</td>
<td>5.31</td>
<td>3.94</td>
<td>3.42</td>
<td>0.42</td>
<td>0.11</td>
</tr>
<tr>
<td>Keyboard\Scarlatti Sonata</td>
<td>59</td>
<td>5.2</td>
<td>3.91</td>
<td>3.3</td>
<td>0.38</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Haydn\Vn-1</td>
<td>201</td>
<td>6.07</td>
<td>3.89</td>
<td>3.27</td>
<td>0.38</td>
<td>0.10</td>
</tr>
<tr>
<td>String Quartet\Beethoven\Vc</td>
<td>71</td>
<td>4.91</td>
<td>3.81</td>
<td>3.42</td>
<td>0.34</td>
<td>0.09</td>
</tr>
<tr>
<td>Keyboard\Haydn Sonata</td>
<td>23</td>
<td>4.33</td>
<td>3.76</td>
<td>3.09</td>
<td>0.36</td>
<td>0.09</td>
</tr>
<tr>
<td>String Quartet\Beethoven\Vn-1</td>
<td>71</td>
<td>4.85</td>
<td>3.74</td>
<td>3.3</td>
<td>0.39</td>
<td>0.10</td>
</tr>
<tr>
<td>Keyboard\Beethoven Sonata</td>
<td>103</td>
<td>4.61</td>
<td>3.57</td>
<td>3.07</td>
<td>0.36</td>
<td>0.10</td>
</tr>
</tbody>
</table>
A3. Selected Quartets from [53] for Chapter 3.3

Haydn’s String Quartets (80 movements)

- Op. 54, No. 2: Mov. 1, 3, 4.
- Op. 54, No. 3: Mov. 1, 2, 3, 4.
- Op. 55, No. 2: Mov. 1, 2, 3, 4.
- Op. 55, No. 3: Mov. 1, 2, 4.
- Op. 64, No. 1: Mov. 2, 4.
- Op. 64, No. 2: Mov. 3, 4.
- Op. 64, No. 3: Mov. 1, 3, 4.
- Op. 64, No. 4: Mov. 1, 2, 3, 4.
- Op. 64, No. 5: Mov. 2, 3, 4.
- Op. 64, No. 6: Mov. 1, 2, 3, 4.
- Op. 71, No. 1: Mov. 1, 2, 3, 4.
- Op. 71, No. 2: Mov. 1, 2, 3, 4.
- Op. 71, No. 3: Mov. 1, 2, 3, 4.
- Op. 74, No. 1: Mov. 1, 2, 3, 4.
- Op. 74, No. 2: Mov. 1, 2, 3.
- Op. 74, No. 3: Mov. 2, 3, 4.
- Op. 76, No. 1: Mov. 1, 2, 3.
- Op. 76, No. 2: Mov. 2, 3.
- Op. 76, No. 3: Mov. 2, 3.
- Op. 76, No. 4: Mov. 1, 2, 3, 4.
- Op. 76, No. 5: Mov. 1, 2, 3, 4.
- Op. 76, No. 6: Mov. 2, 3, 4.

Mozart’s String and Other Quartets (80 movements)

- No. 1 in G major, K 80: Mov. 1, 2, 3, 4.
- No. 2 in D major, K 155 /134a: Mov. 1, 2, 3.
- No. 3 in G major, K 156/134: Mov. 1, 2, 3.
- No. 4 in C major, K 157: Mov. 1, 2, 3.
- No. 5 in F major, K 158: Mov. 1, 3.
- No. 6 in Bb major, K 159: Mov. 1, 2.
- No. 7 in E♭ major, K 160/159a: Mov. 2, 3.
- No. 8 in F major, K 168: Mov. 1, 2, 3, 4.
- No. 9 in A major, K 169: Mov. 1, 3, 4.
- No. 10 in C major, K 170: Mov. 2, 3, 4.
- No. 11 in Eb major, K 171: Mov. 1, 2, 3, 4.
- No. 12 in Bb major, K 172: Mov. 1, 2, 4.
• No. 13 in D major, K 173: Mov. 1, 3, 4.
• Flute Quartet in D major, K 285: Mov. 1, 2, 3.
• Flute Quartet in A major, K 298: Mov. 1, 2, 3.
• Oboe Quartet in F major, K 370/368b: Mov. 1, 2.
• No. 14 in G major, K 387: Mov. 1, 2, 3, 4.
• No. 15 in D minor, K 421/417b: Mov. 1, 2, 3, 4.
• No. 16 in Eb major, K 428/421b: Mov. 1, 2, 3.
• No. 17 in Bb major, K 458: Mov. 1, 2, 3.
• No. 18 in A major, K 461: Mov. 2, 3, 4.
• No. 19 in C major, K 465: Mov. 1, 2, 3, 4.
• No. 20 in D major, K 499: Mov. 1, 3, 4.
• No. 21 in D major, K 575: Mov. 2, 3.
• No. 22 in Bb major, K 589: Mov. 1, 3, 4.
• No. 23 in F major, K 590: Mov. 1, 2, 3, 4.
Bibliography


[82] P. Ammirante and F. A. Russo, “Low-Skip Bias: The Distribution of Skips Across the Pitch Ranges of Vocal and Instrumental Melodies is Vocally


