MASTER

Visualizing consonance and dissonance in modal counterpoint

Rooij, J.P.

Award date:
2017

Department of Mathematics and Computer Science
Visualization Research Group

Visualizing Consonance and Dissonance in Modal Counterpoint

Master's Thesis for Computer Science and Engineering

Jan-Pieter Rooij

Supervisors:
dr.ir. H.M.M. van de Wetering

Final version

Eindhoven, July 2017

Abstract

The concepts of consonance and dissonance correspond roughly to the combinations of musical notes that are perceived by most to be pleasant and unpleasant, respectively. These concepts are some of the main cogs and gears available to composers for the purpose of introducing emotional tension into a piece of music. They are especially relevant in modal music, which roughly corresponds to the music of the renaissance and early baroque period and earlier. These types of music will then be the focus of the paper.

It is possible to extract information on consonance and dissonance from traditional sheet music, or staff notation. However, to do this one must be able to read sheet music. Even then, plucking out the relevant tidbits is not very efficient. The main goal is then to provide this information in a more accessible manner. Additionally, the information should be presented in such a way that it is mentally parsable in a much more efficient way than extracting the same information from traditional notation.

A piece of software called the "Counterpoint Analyzer" is created which accomplishes the aforementioned goals using several visualizations, each attuned to a different type of audience, including laymen, musicians and musicologists. Since each audience type has distinctly different demands, each visualization has some special features not found in the others.

Acknowledgements

I would like to thank my supervisor dr. ir. Huub van de Wetering for aiding me with completing this project. His expertise and his feedback during the development process were invaluable. Furthermore I would like to thank dr. ir. Jack van Wijk for providing additional insights into the problem at hand.

Contents

List of Figures

Chapter 1

Introduction & Background

1.1 Historical Background

The Western music tradition as we know it today was initiated at the start of the medieval period. From this point on, music history can be divided into several different periods, as seen in fig. 1.1.

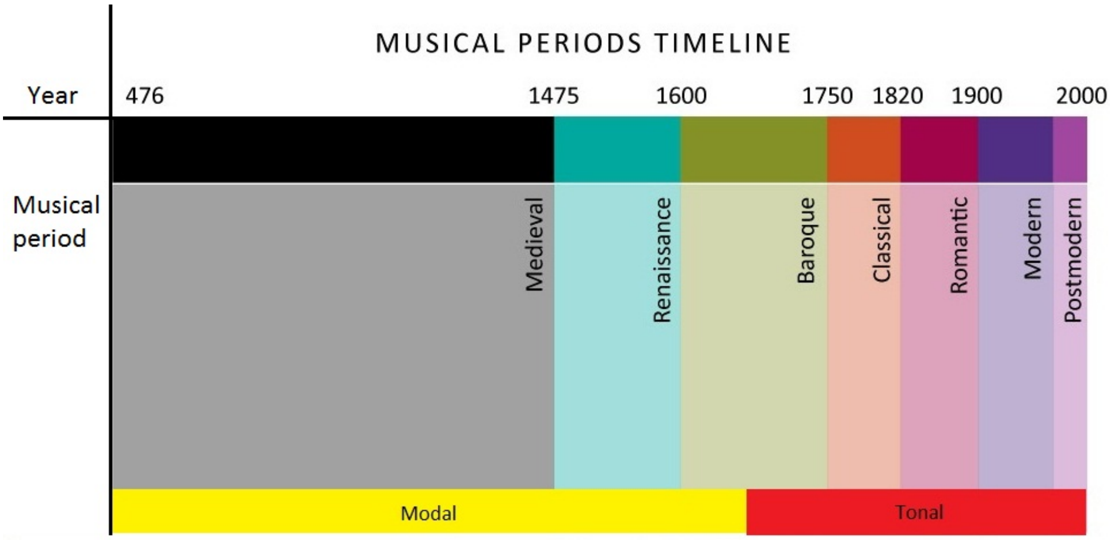

Figure 1.1: Timeline of musical periods [7]

One of the main divisions that can be made in known music history is that between **modal** music on the one hand and **tonal** music on the other. Modal music ranges from antiquity until roughly the first half of the Baroque period. After this point, most music can no longer be considered modal, but is instead called tonal, i.e. it can be in major or minor. The main focus in the paper will be on modal music. To explain these concepts accurately would require a fair amount of music theory. Therefore I will instead present the concepts in rudimentary form which should be accurate enough for its intended purposes.

1.2 Music Theory Background

Classical music often contains multiple melodies being played simultaneously. This is also known as **counterpoint**. Music that uses counterpoint is considered **contrapuntal**. The distance between two notes that are played at the same point in time is called an **interval**, as shown in fig. 1.2.

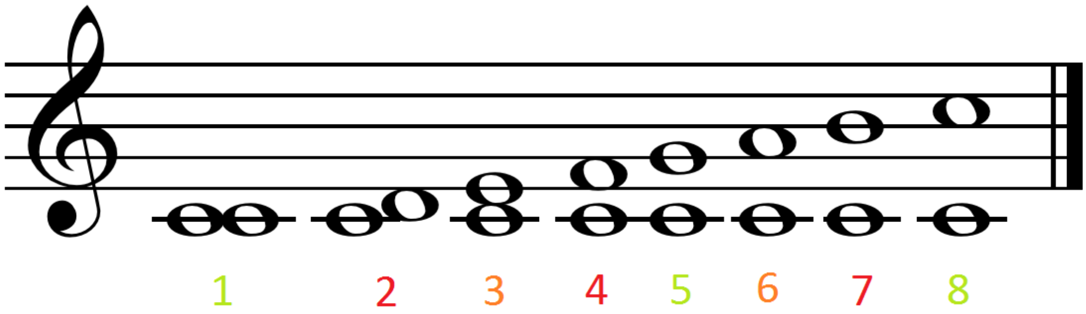

Figure 1.2: All modal intervals in an octave. The digits of perfect consonances are colored green, those of imperfect consonances are colored orange, and those of dissonances are colored red.

Intervals are expressed using integers, where the integer refers to the number of whole notes in between two notes playing together at some point in time. By matter of convention, and rather counter-intuitively, this number is then incremented by 1. To exemplify this: two identical notes sounding together have a distance of 1 between them, instead of zero. The same goes for all other intervals. Intervals can be either **consonant** or **dissonant**, depending on the size of the intervals. Intervals with sizes 1, 3, 5, 6 and 8 are considered consonant, or pleasant sounding. The complement of this set, i.e. intervals of sizes 2, 4 and 7 are considered dissonant, or unpleasant sounding. Note that conventions dictate that the distance between two notes with an interval of 1 is actually 0, even though one may expect the distance to be 1. Out of the consonant intervals, the intervals of sizes 3 and 6 are considered **imperfect** consonances and intervals of 1, 5 and 8 are considered **perfect** consonances.

In modal music, the only things of relevance are the intervals between multiple melodies. In tonal music, there is an overarching framework in which some notes fit and some notes don't. In other words, at any point in time, there is a subset of notes that can be considered to fit into the framework and a complement of this set, naturally containing notes that do not. The intervals between individual notes then become subservient as it were to the location and purpose of the notes in the overall framework.

Now that you are supplied with the necessary music theory, it is time to continue to the introduction.

1.3 Introduction & Problem Definition

Up until now, the lion's share of research in the academic field of music visualization has concerned tonal music. Comparatively little research has gone into modal music. This thesis attempts to remedy the aforementioned imbalance in the research field, in however slight a way, by providing a visualization method for displaying information about the presence (or lack thereof) of consonance and dissonance in any given modal piece.

Consonance and dissonance are one of the, if not the most important tools available to composers in order to introduce emotional content into music. So important in fact, that several theoretical systems were invented by music theorists during the early baroque period in order to codify a method for using consonance and dissonance to its greatest effect. The most well known of these works is Gradus ad Parnassus by Johann-Joseph Fux, which has been studied by almost all great composers of the common practice period of classical music. We can therefore conclude that understanding the role of the concepts of consonance and dissonance is key to comprehending the nuts and bolts behind a piece of modal music.

So, if this concept is so important for producing emotional response in an audience, are then consonances and dissonances not clearly audible when listening to a piece? And if they are, what use is there for a visualization?

Consonances and dissonances are in fact not so easily audible in general. There are several reasons for this. First of all, one has to be aware of the existence of the concept in order to direct one's attention to it while listening. This may seem like a moot point, but surprisingly many people do not know the concept, or do not realize its full importance. Then, even if aware of the concept, one needs to be able to hear the difference between intervals played in succession. Especially for laymen, this is not easy to do, and even for musicians this often requires some training (also known as solfege training).

Another reason is that most modal contrapuntal music tends to be fairly dense. What is meant by this is that in most contrapuntal pieces there are many things happening at the same time. Generally speaking, humans are notoriously bad at keeping track of multiple things simultaneously, and here it is no different. Even for the trained listener it is almost impossible to keep track of everything that's going on in a dense contrapuntal piece. For instance, in fig. 1.3 we see one bar from Johann Sebastian Bach's BWV 862, which is a fairly representative example of baroque counterpoint. Here, we see 4 independent voices moving simultaneously in different rhythms. Within a space of 2 seconds, over 30 intervals occur between the different voices. Hence, typically it is indeed useful to have a visualization at hand to help pick out the relevant intervals.

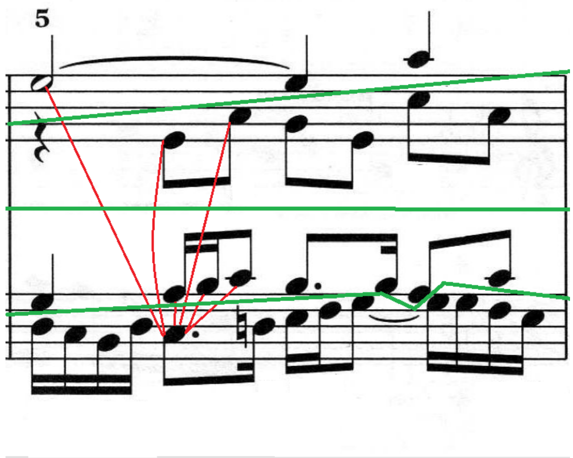

Figure 1.3: Excerpt from BWV 862. [12] This excerpt contains 4 voices. Green lines are used to indicate the separation between the voices. Red lines show the interval relations for a single bass note.

At this point, one may object that traditional music notation can already be considered a visualization from which relevant areas of consonances and dissonances can be extracted. Technically speaking this is true, however there are several reasons why another visualization is needed. Broadly speaking, the main problems with using traditional music notation for the purpose of extracting information on consonance and dissonance are that it is inefficient and that it is quite non-intuitive. One has to be able to read sheet music. Then, when this condition is met, a person has to determine the intervals between all combinations of notes sounding together across some passage of time. Even for a person who is fairly skilled at reading sheet music this is not an easy task. The goal here is then to find a way to show the locations of consonances and dissonances to users in a more intuitive manner. More concretely: what needs to be visualized is a set of differing types of relations between groups of notes that are played simultaneously at some point in time. The goal is to create a software package called the Counterpoint Analyzer that implements this functionality. Additionally, there will be extra functionality associated with the specific demands of different user groups. The exact specifications of these functions are discussed in later chapters.

1.4 Visualization Framework

In order to analyze the range of possible visualization methods in this thesis, I will use the oft-employed framework of marks and channels. Due to the widespread use of the framework, several slightly different versions have sprung up. The framework used in this thesis is the one described by Tamara Munzner [6].

Marks are geometric shapes that depict items from the data set and the connections between these items. Examples of marks include points, lines and areas of myriad shapes. Naturally, different marks are suitable for different types of situations.

Channels determine the look of marks. They can be used to add extra information that may not be readily incorporated through the use of marks alone. Examples of channels include position, color, shape, tilt, size, area and volume. Like with marks, there is a difference in usability and clarity between the channels. Although the relative usefulness of channels depends on the exact situation, Munzner does set up a general hierarchy of effectiveness for channels. Effectiveness is roughly defined as noticeability (or salience). More effective channels are therefore more noticeable. An example: curvature of marks is generally considered rather ineffective, while length and the position on a common scale tend to be rather effective. Additionally, shape tends to be less effective than color.

1.5 Structure

The structure of the thesis will proceed as follows. First, user groups and their respective requirements will be determined. Next, the general idea behind the main visualization will be discussed. Every user group gets its own "view": a modified version of the main visualization that takes into account the specific requirements of the user group in question. Each user group view is explained in a separate chapter in which the existing work is discussed and the marks and channels of the view evaluated. After the research and reasoning behind the views has been presented, we will go through some technical details in the implementation chapter and do a user evaluation. Finally, some concrete examples of possible applications of the final software product are discussed and conclusions are made.

Chapter 2

User Groups & Requirements

The Counterpoint Analyzer may be used by different user groups, each having different interests. This leads to differing requirements per user group. Below follows a rough description of the user groups that are likely to use the software.

Laymen: These are people with limited musical knowledge. They cannot read sheet music, nor do they know any music theory, nor do they have experience playing musical instruments, etc.

Musicians: Musicians are interested in performance aspects. They tend not to be especially interested in more advanced theoretical analysis.

Musicologists: These users are mostly interested in theoretical analysis, not in most aspects of performance analysis.

2.1 Functional requirements by user group

Below follows a general description of the functionality required for each different user group.

Laymen: The ability to see certain musical aspects without having to be familiar with sheet music or any music theory. These musical aspects include the degree of consonance and dissonance in certain sections of a piece, the general contours of melodies (relative pitch) and a clear distinction between different voices.

Musicians: The ability to view the musical aspects of the laymen user group with explicit music theoretical concepts added, such as absolute instead of relative pitch, the consonance/dissonance status of each individual note and bar divisions and numbers.

Musicologists: The ability to view the same musical aspects of the musicians user group with addition of more detailed musicological concepts, such as a more comprehensive view of the separation of voices, parallel intervals and some statistics such as the percentage of dissonances in a piece.

2.2 Precise requirements per user group

The precise requirements derived from the more general requirements described above are listed below per user group.

The traditional presentation of music in staff notation uses the vertical dimension to display pitch and the horizontal dimension to display time from left to right. In the next chapter, we will see that this is in fact a universal way to display time and pitch, which means it may be a good idea to keep these associations. Additionally, by using the same dimensional designation in the visualization it will become easier to process for those already familiar with sheet music.

2.2.1 General

1. The visualization should be usable at both a low and a high zoom level.
2. It should be possible to play the piece that is currently loaded through the speakers.
3. While playing, one should at all times be able to see the playback location relative to the visualization.
4. There should be a graphical interface to control main features.
5. Any given piece of modal music should be usable as input (provided in musicXML format).

2.2.2 Laymen

6. Users should be able to view the temporal locations of all consonances and dissonances in a piece of music.
7. The outline of melodies should be visible, with the relative pitch of notes intact.
8. The overall mood as influenced by consonances and dissonances should be clear at any time in the piece.

2.2.3 Musicians

9. On the horizontal axis, the time dimension should be preserved and the duration of notes should be visible.
10. Bar lines and bar numbers should be visualized.
11. The location of consonances and dissonances should be clear.
12. Consonances and dissonances should be visible without the user having to do any interval calculations.
13. The type of consonance/dissonance should be visible, i.e. perfect or imperfect.
14. The original musical score should be available.
15. The location in the score should be linked in some way to the location in the visualisation.

2.2.4 Musicologists

The requirements for the musicologists' group are identical to those of musicians, with addition of the following:

16. It should be clear to which voice each note belongs, i.e. the voices should be clearly delineated in the visualization.
17. It should be possible to see the exact size of intervals.
18. Several musicological statistics should be available, including total percentage of dissonant notes, the number of parallel intervals, et cetera.

Chapter 3

Piano roll visualization

In order to provide some initial starting points for the main visualization, we will first explore conceptual metaphors to find out about the intuitive ways in which people experience music. This will lead to an introduction of the most basic piano roll visualization, of which diverse manifestations are shown later.

3.1 Existing work

3.1.1 Conceptual metaphors

The theory of conceptual metaphors originally hails from the field of cognitive linguistics and says that human conceptual abilities have their roots in universal, sensory motor experiences of space, force, containment and orientation, although there is usually some socio-cultural influence. This theory has been applied to research in the field of music visualization by Wilkie, Holland and Mulholland (2010) [13] in order to to provide a well reasoned starting point and framework. Without this framework, it is easy to start flying blind since basically anything is possible when the link between traditional sheet music and the visualization drops away. Some common conceptual metaphors include ”virtue is up”, ”more is up”, ”bad is down”, etc.

In order to visualize music, two important concepts are pitch and time. The results for pitch are clear and apparently universal across culture: ”high pitch is up” and ”low pitch is down”. With regard to the length of notes and the development of a piece over time, we see the following metaphors: ”time is movement along a path” and ”music is movement along a path”. These metaphors may also be the reason why traditional sheet music evolved the way it did.

Now that we've established that the spatial dimensions of pitch and time are intuitive, there are good grounds to associate pitch with the vertical dimension and time with the horizontal dimension in the visualization even though there is no specific requirement to stick to traditional sheet music dimensions.

3.1.2 Piano roll

As described in the previous chapter and as confirmed by the research on conceptual metaphors, it would be preferable to maintain the intuitive horizontal and vertical dimensions of traditional sheet music in the musicians' visualization. A large portion of the research in the field has been focused on moving away from traditional sheet music instead of maintaining some of its features, for this reason there is a relative dearth of these types of visualizations in contemporary literature. This is perhaps why the work that is relevant to us is found at the very genesis of the field.

More accurately, this work comes in the form of the study of piano rolls (fig. 3.1). These rolls were used by player pianos in the late 19th century. Notes are ”visualized” if you will as holes in a two-dimensional chart, with time on the one axis and pitch on the other.

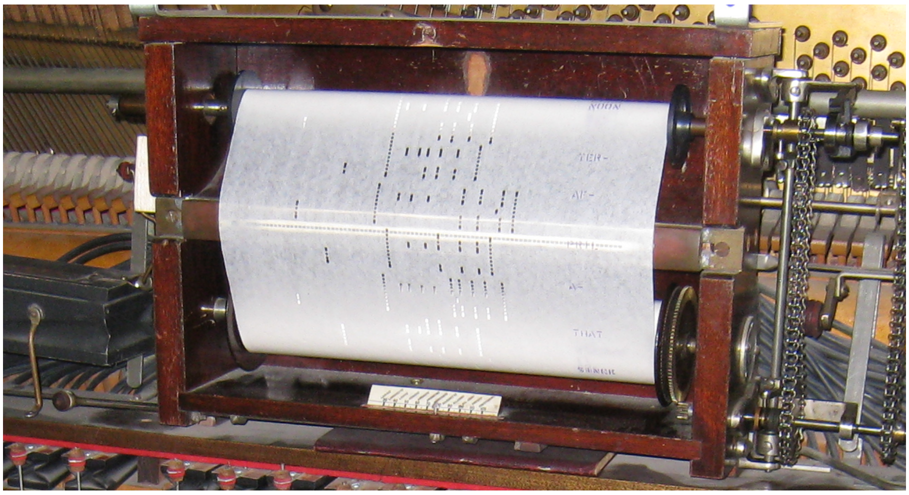

Figure 3.1: Physical piano roll [10] The roll is rotated by 90 degrees. This means the pitch is on the horizontal axis, the time on the vertical axis.

In modern music editing software, piano roll-type visualizations are still frequently used. Even so, the amount of information conveyed is often limited to merely showing the pitch and duration of notes. This can be seen in fig. 3.2. Putting it differently, most visual channels are not utilized in standard piano roll visualizations. In the following chapters, a number of different views will be explored, each corresponding to a particular user group and each having its own set of particular functions to fulfill the specific requirements of the group.

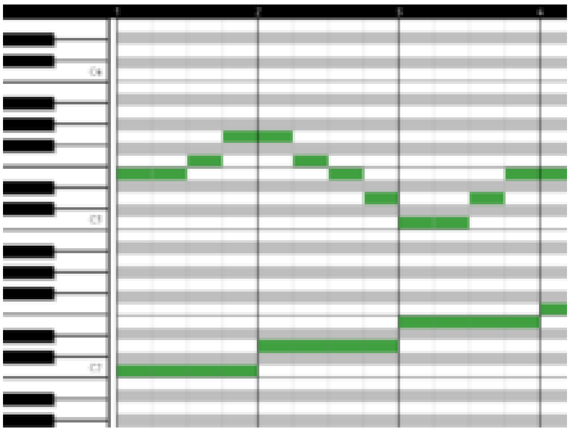

Figure 3.2: MIDI piano roll [11] The pitch is on the vertical axis, the time on the horizontal axis.

3.1.3 Research

Krawczyk [4] shows that the type of marks can be altered while still maintaining a piano roll-type visualization. In the first image of fig. 3.3 we see a standard piano roll, broken up vertically into 3 parts in the same way as individual voices are broken up in traditional sheet music. Immediately one of the limitations of Krawczyk's method becomes clear: the broken-up parts of the piano roll are not clearly delineated, and in fact it is not immediately clear that this is what we're seeing without forehand knowledge. In the second image, a grid structure is added, which is turned into a 3D in the final image. While visually attractive, one can argue how much the 3D view adds, since overlapping problems arise immediately, even in this rather simple example.

In fig. 3.4 Krawczyk applies different shapes to the standard box structure. In the right-most image, he succeeds at emphasizing the relative size of the intervals between the voices, arguably at the cost of clarity, since the overall view becomes rather messy. The delineation between different voices is also completely obfuscated.

In fig. 3.5 Krawczyk explores the uses of a circular piano roll. There are some limitations inherent in this method. The most notable of these is that the total length of the piece being visualized is quite limited. Even when displaying this short piece of only 24 bars, the surface of the marks is already so small that using a circular view with a larger piece would probably not be feasible while keeping the entire circle on screen. A circular visualization also gives the impression that the piece is repeatable, which is very often not the case. Nonetheless, this approach does have its advantages, such as being able to pack a relatively large amount of information in a small area.

Krawczyk's method has definite limitations, mostly in that a lot of musical information is lost, and the accuracy of that which is left is diminished. However, his research remains interesting, and it is useful to see all the ways in which one can mold the different facets of the standard piano roll.

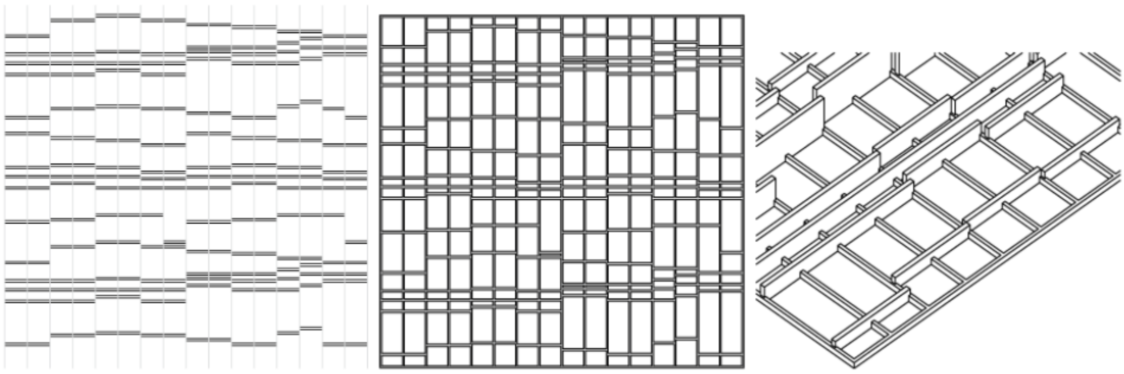

Figure 3.3: Left: Piano roll. Center: piano roll with added grid structure. Right: 3D piano roll. The piano roll contains 5 voices, split up vertically into 3 sections. [4]

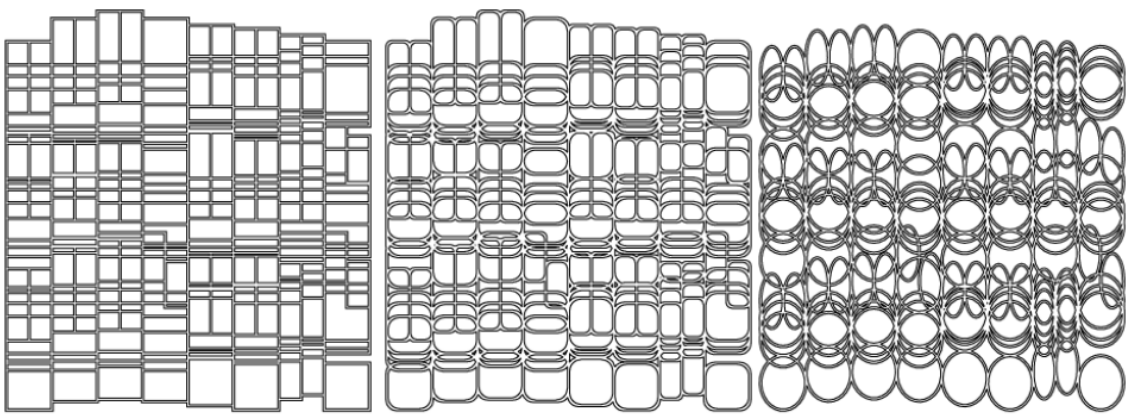

Figure 3.4: Piano roll with modified shapes. The relative difference between different interval sizes becomes progressively more exaggerated from left to right. [4]

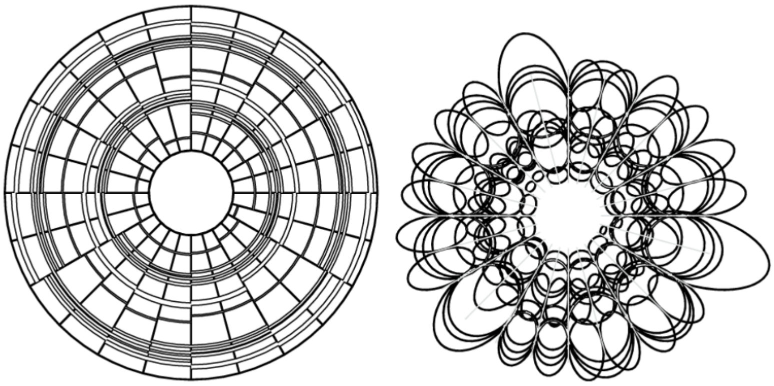

Figure 3.5: Circular piano roll, shown with both normal and modified shapes. [4]

3.2 Analysis of marks

One of the basic features of a piano roll visualization is that marks are used to represent notes. Channels can then be used to show additional information. For the purpose of seeing consonance and dissonance, it is important to see when notes are overlapping in time (i.e. sounding at the same time). The traditional system of music notation is thus not optimal since the overlap of notes is only indicated implicitly. It would be more useful to employ a mark that visualizes a note for the entire length of its duration, like in a piano roll (fig. 3.6). Aside from facilitating the visualization of overlap, indicating when the note sounds explicitly also removes an additional mental calculation needed to extract the note length, as is required when reading traditional sheet music. Out of the available marks (points, lines and areas), lines or areas would then seem to be the best options. Considering how there is much more horizontal space than vertical space on a computer screen with regular orientation, a horizontally elongated shape would be the most convenient when using areas. According to requirement 1, a line may not be the best option since it will not be easily visible at a low zoom level. The ideal mark for representing notes would then seem to be a long rectangle or oval area. In the following chapters, there will be some small changes in the exact shapes of the marks, however this general shape will consistently be used.

Note Length	Traditional Notation	Piano Roll Notation
Eight note (0.5 beat)	♪	
Quarter note (1 beat)	♩	
Half note (2 beats)	𝅗𝅥	
Whole note (4 beats)	𝅝	

Figure 3.6: In traditional sheet music, the length of notes is displayed implicitly, while in the piano roll visualization the length is displayed explicitly

In the next chapter, the visualization belonging to the first user group will be discussed. Channels will be used to add the required information specific to this user group onto the marks that were just discussed.

Chapter 4

Musicians' view

In this chapter, the design decisions behind the musicians' view will be discussed. Existing work will be discussed, followed by an analysis of visual channels.

4.1 Existing work

This section contains a description of a paper by Hayashi and a discussion of preattentives, a commonly used concept in visualization design.

Hayashi

Hayashi [1] shows that channels can be used to impart a greater amount of information than in a standard piano roll visualization (fig. 4.1), something which this thesis also attempts to do. To be more specific, Hayashi uses colored marks to separate thematic material in the individual voices and to show the recurrence of this material. This is mostly successful. Even at low zoom levels such as in fig. 4.1 most information is still visible. Furthermore, thematic overlap between different voices is also clearly visible, as seen through the green blocks on the right-hand side of the image. The main subject of Hayashi's visualization is thematic material. Although this is not directly compatible with the subject matter of this thesis, the paper still shows that color is potentially a very powerful channel when used in piano roll visualization.

Preattentives

The term preattentive refers to visual properties that humans can process without having to direct focused attention towards them. In other words, these properties can be distinguished almost immediately. Preattentives can be a very efficient way to bring the users' attention to one or multiple objects out of a field of objects. This is very useful for the research conducted here, since we are effectively trying to draw the users' attention to a few dissonant notes out of a field of mostly consonant notes. Not all channels have strong preattentive properties, and Munzner [6] has ranked them accordingly. With regard to preattentiveness, the strongest channels are color and shape, as seen below. Other channels like tilt and luminance are comparatively much less strongly preattentive.

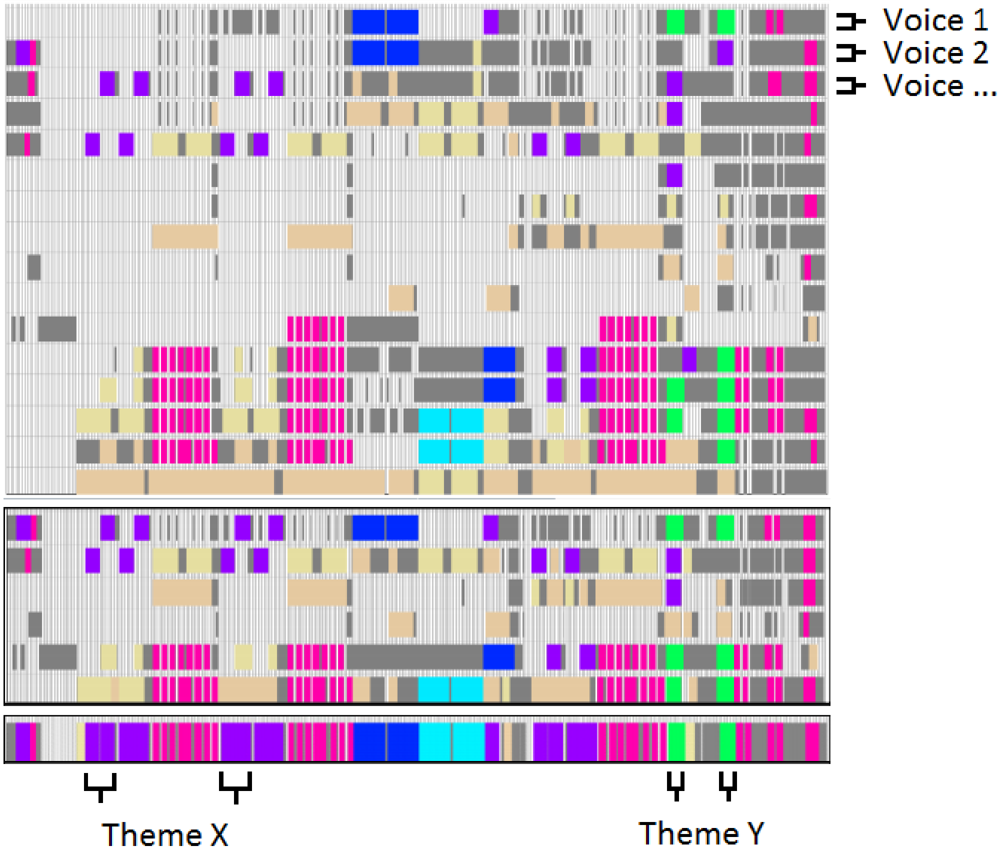

Figure 4.1: Hayashi's colorscore visualization of an orchestral score, with every instrument group occupying one vertical bar. Colors are used to indicate recurring thematic material both within and across voices. [1]

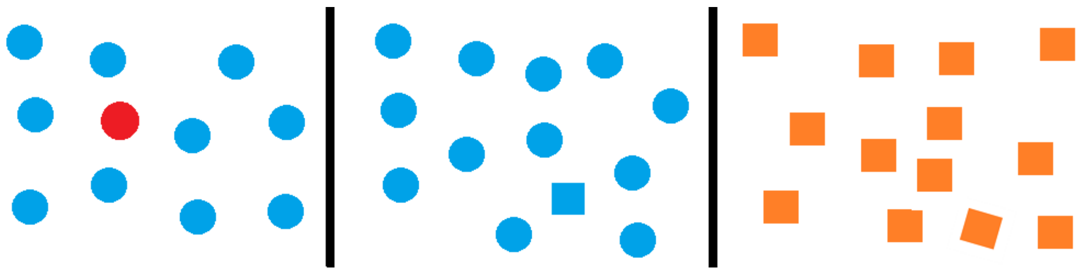

Figure 4.2: Color is the strongest preattentive (left). Shape is also a strong preattentive (middle). Tilt is less strong as a preattentive (right).

4.2 Analysis of channels

Next, the efficacy of different visual channels for realizing the requirements of the musicians' user group will be discussed. Available channels include position, color, shape, tilt and size.

Position

Requirements 6 and 7 put some significant constraints on the use of position, since they more or less fix the position of the marks used to represent notes. This channel is therefore unavailable.

Tilt

It may be possible to rotate a mark to indicate its consonance/dissonance status. We could for instance rotate a mark by 30 degrees when it is a dissonance and by 10 degrees when it is an imperfect consonance. This is shown in fig. 4.3.

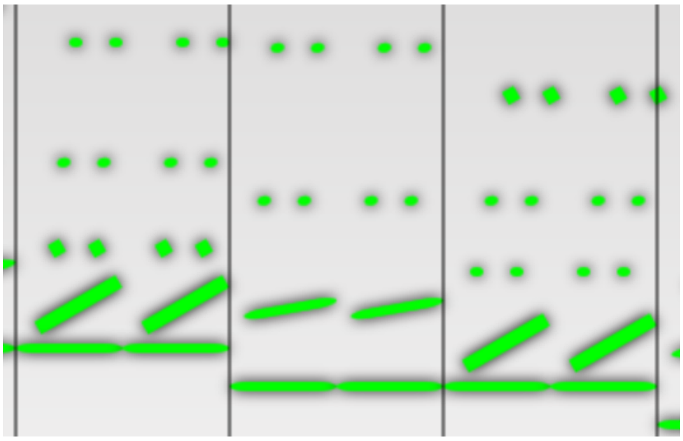

Figure 4.3: Marks representing dissonances are tilted

Since vertical space is rather limited, especially in very complex pieces, tilting the marks presents overlapping issues between marks when counterpoint gets dense. The issue is illustrated in fig. 4.4. This can be mediated to some extent by reducing the angle of rotation, but this does not solve the problem completely. Also, the tilted configuration of marks in this example seems to resemble something akin to a set of stairs, which is purely in the mind's eye and has no bearing on the actual data set. Using the tilting channel seems not to be a good option.

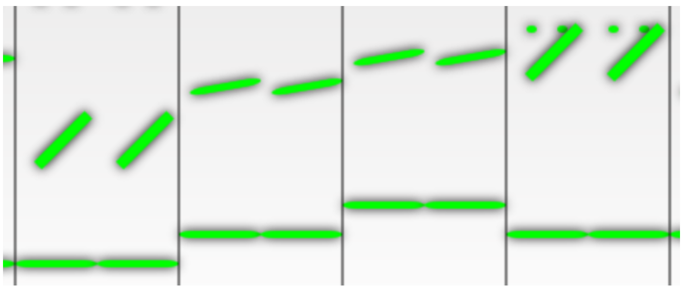

Figure 4.4: Overlap as a result of tilting dissonant marks

Color

Contrasting colors can be used to show the consonance/dissonance status of notes. In addition, opacity can be used. In fig. 4.5, the color red is used to indicate dissonance, the color orange to show imperfect consonance and the color green for perfect consonance. These specific colors are used because they have cultural connotations with negative and positive events, like for instance in traffic lights. Along with position, color is considered by Munzner to be the most powerful channel, and since it is still free we should use it. The observant reader will notice that only the notes belonging to the top voices are colored and the notes belonging to the bass voice are not. This is because notes are consonant or dissonant *with regard to* the bass note. In other words, the interval between a note and its corresponding bass note determines the status of a note belonging to one of the top voices. This means a bass note is itself neither consonant nor dissonant.

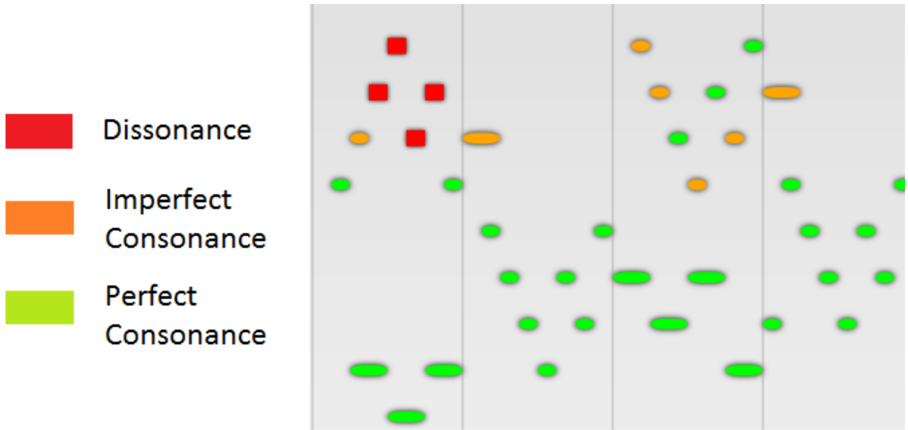

Figure 4.5: Colors are used to indicate perfect consonances (green), imperfect consonances (orange) and dissonances (red)

Shape

The shapes of the marks could be a good way to differentiate between consonant and dissonant notes. The changes that can be made to the marks in the standard piano roll view are however relatively minor, since the amount of space is rather constrained. Regardless, since shape is a fairly strong channel, even relatively minor changes in shape could already yield some results. As can be seen in the figures in this chapter, we have chosen to give the marks of consonant notes strongly rounded edges. This improves the delineation of notes that follow one another in quick succession, since if the normal rectangular shape were to be maintained, notes would be stuck together which would require an increase in padding. This is increasingly relevant as the zoom level decreases. Dissonant notes have been made 50 percent higher and have been given 90 degree corners in order to differentiate them from consonant notes. We've chosen to stick to fairly simple geometric shapes in order to keep the general overview of the piano roll view clear. Using for instance curved shapes, or shapes with irregular borders would make the overall view somewhat chaotic.

Horizontal Size

The horizontal size of marks is already used to indicate the duration of notes, as discussed earlier.

Vertical size

Using vertical size is an option, as long as the size is not increased to a point where overlap occurs between marks. The use of differential vertical sizing is illustrated in fig. 4.6, where dissonant notes are 50% higher than consonant notes.

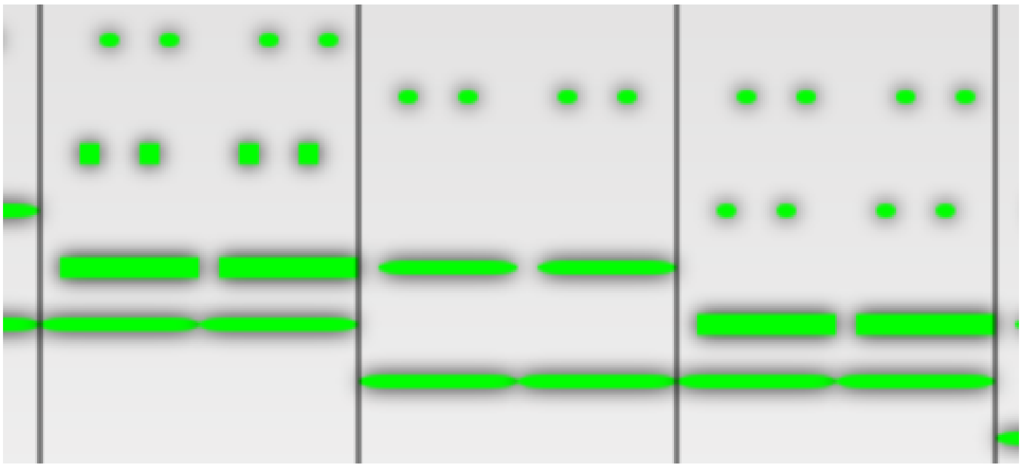

Figure 4.6: Vertical sizing of marks is used to differentiate between consonances and dissonances

4.3 Additional features

Krawczyk [4] uses boxes around each note to add some structure to the standard piano roll. This may be beneficial in itself, since it can be difficult for the human eye to judge the exact relative location of notes, especially when these notes are of different length and color. This can give the impression that notes are simply floating about. However, Krawczyk seems to simply add a box to every note, which is not so useful from a musical perspective because musicians may easily interpret these as being bar lines. I instead propose to add the actual bar lines, as can be seen in the images in this chapter.

4.4 Conclusion

To realize the requirements of the musicians' user group, color, shape and size were used. Tilt was found to be ineffective. The design decisions made in this chapter will be investigated later through user testing.

Chapter 5

Musicologists' view

Since the requirements of musicians and musicologists overlap to a great extent, the musicologists' view is heavily based on the musicians' view with some extra features tacked on to facilitate the greater amount of detail required for musicologists' aims. These features are specified in the requirements. Each requirement will be discussed separately.

5.1 Existing work

In traditional sheet music, each voice often has its own bar and is thus clearly delineated, as shown in figure 5.1. In a piano roll view this is not the case, which can lead to confusion since it is not always clear which note belongs to which melody.

Figure 5.1: Each voice has its own bar, which creates a clear visual separation of voices [8]

Malinowski [5] created a so-called Music Animation Machine. In contrast to the more generic visualization created in this thesis, Malinowski creates a separate visualization for each piece. Several of these visualizations feature separation of voices.

In the visualization shown in figure 5.2, every voice is assigned a certain color, and notes of that color belong to the respective voice. This is a rather effective approach, and it functions well at all zoom levels. It is clear which notes belong to which melody.

In the visualization in figure 5.3, a line is drawn between notes belonging to the same voice. This approach implies a greater sense of continuity in the melody, however when dealing with complicated and dense pieces this approach can in some cases lead to a messy overall view, with a multitude of intersections between lines. This effect increases as the zoom level decreases. Another

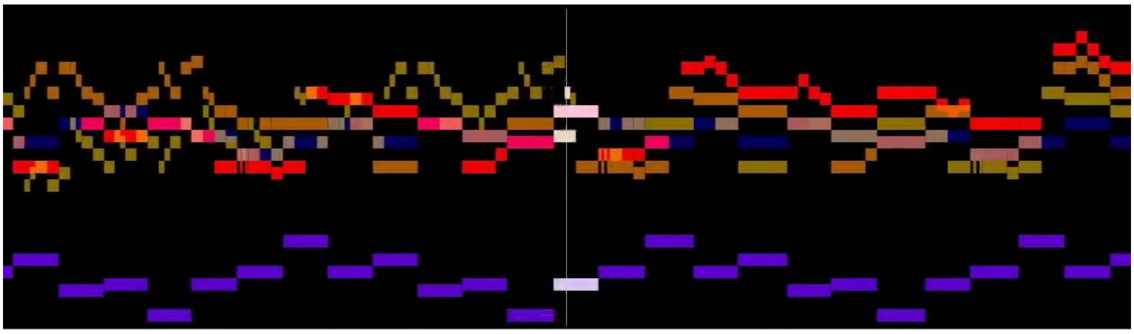

Figure 5.2: Notes are colored according to which voice they belong to [5]

downside is that whenever Malinowski uses lines between notes, he visualizes the notes using circles. This very quickly leads to overlapping issues which can be quite severe, also it is questionable how well vertical expansion works to express duration versus simple horizontal expansion. The latter would appear to be more intuitive according to conceptual metaphor theory, which states that time moves from left to right.

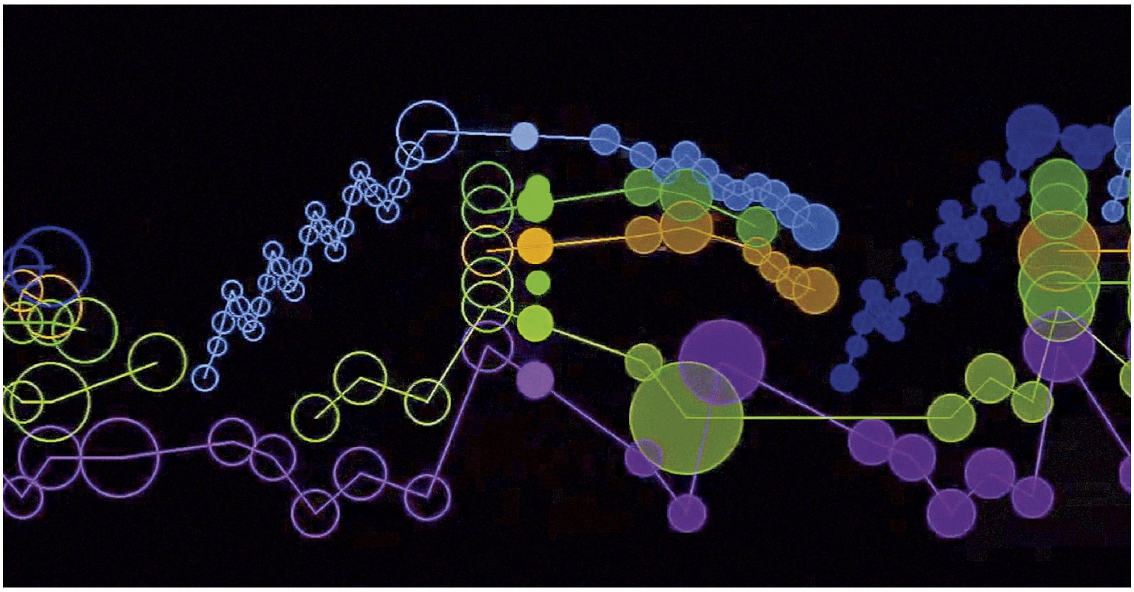

Figure 5.3: Lines are used to separate voices [5]

In figure 5.4, Malinowski uses the areas around the notes to show a separation of voices. This division is fairly clear, although there are some overlapping issues, and it is not always clear when voice crossing occur. In the bottom right quadrant of figure 5.4 we see for instance 3 colors in approximately the same location, but it is not quite clear which notes belong to which color. Due to the fact that the visualized piece was written for organ, the researcher chose to populate the areas around the notes with a sort of ribbed texture in order to provide a visual approximation of the timbre of an organ. The success of this endeavour is arguable, and in any case the texture of the areas does not add any musicological information. Additionally, there is some dead space in the middle of the areas. This highlights the exact location of the notes, but it does cost some valuable space. To summarize, the idea of using areas is interesting in itself, but Malinowski's implementation could be optimized when judging from a musicological perspective.

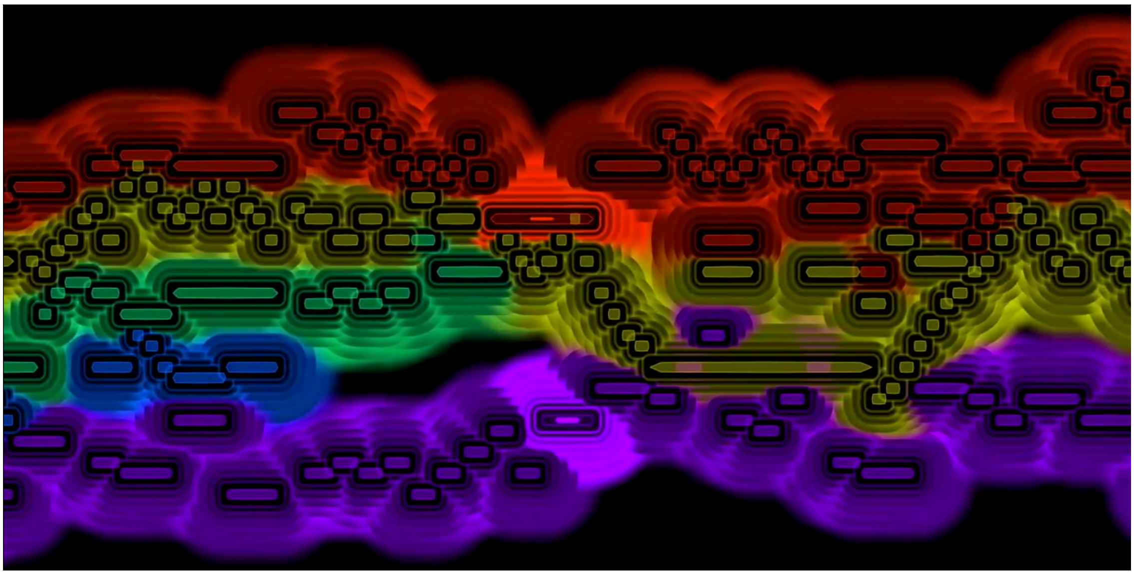

Figure 5.4: Notes are surrounded by colored areas. Each voice has a distinct color. [5]

5.2 Voice Separation Analysis

As seen in the existing work, out of the customary options of points, lines and areas, both lines and areas are viable options. For this reason, the channels of both types of marks will be analyzed below to see which is the superior option.

5.2.1 Channels for lines

Position

The main position of the lines is of course determined by the location of the notes between which they are drawn. This does leave the choice of where to attach the lines to the notes. The main options here would seem to be to attach the lines to either the ends of notes or to the middle of notes. In figure 5.5, the lines are attached to the middle of notes, and in the rest of the figures in this chapter lines are attached to the ends of notes.

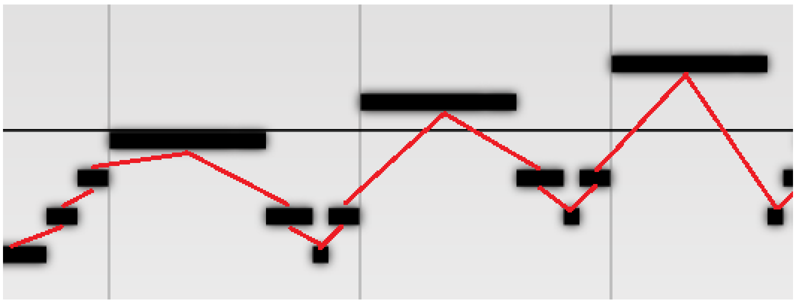

Figure 5.5: Lines are attached to the middle of notes.

Attaching the lines to the middle of notes draws a lot of attention away from the notes, due to the overall length of the lines being much greater. It also makes it seem like the pitch of notes is constantly gliding along between notes, instead of staying constant. For this reason we will choose to attach lines to the ends of the notes for now. The matter will however be evaluated by users later.

Tilt

The tilt of the lines is determined by the locations of the notes to which they are attached. Changing the tilt would thus cause the line not to be attached to at least one of the notes.

Size

Thicker lines are naturally more easily visible, especially at low zoom levels. Lines of greater thickness do however cause a larger amount of visual disturbance. Thick lines may also imply to the viewer that the lines themselves contain some type of important information, while they should of course be of secondary importance to the marks representing notes.

Color

Designating a different color to each line could help in differentiating the lines, especially when voices cross. It may also keep the color channel of the notes themselves unused. The colors that are used are bright and highly saturated since the area of the lines is quite small and the color coding should be clearly distinguishable. The color map is categorical. The exact colors are taken from www.colorbrewer2.org, a website that uses a multitude of perceptual guidelines to create proper color combinations. The colors are shown in figure 5.6.

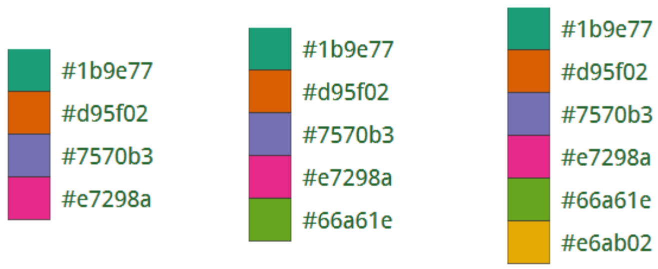

Figure 5.6: The color maps used for 4-, 5-, and 6-voiced pieces, respectively.

Shape

In order to minimize the amount of visual disturbance caused by the lines, it seems best to keep it simple and to use typical straight lines with equal width along the length.

Combination with consonance/dissonance

In order to show the separation of voices, several channels that were also used to show the consonance/dissonance status of notes were used. In order to show consonance, we have two main channels: color and shape. These options are explored in figure 5.8.

It is clear that the use of shapes is much clearer than the concurrent use of colors for showing dissonances.

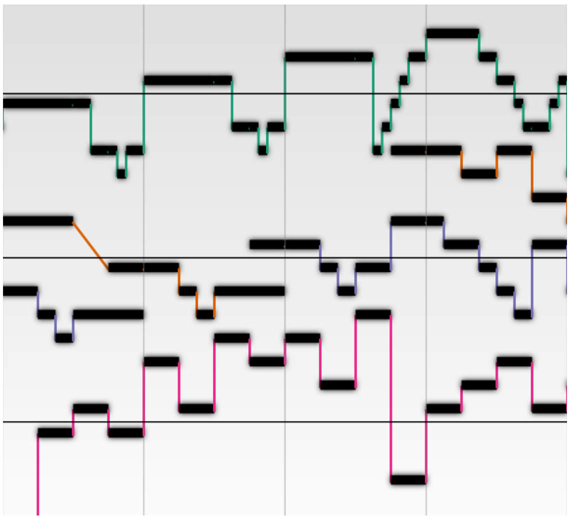

Figure 5.7: Lines are used to separate voices. No consonance/dissonance is shown. Lines belonging to different voices have different colors.

5.2.2 Channels for areas

Position

The only logical option here seems to be to have the area surround the notes.

Size

To determine the optimal size of the areas, the ratio between visibility on the one hand and overlap between voices on the other hand should be optimized. In the visualization by Malinowski in fig. 5.4, the areas are very big, which means there is a great deal of overlap. The areas could be made quite a bit smaller without sacrificing visibility, especially when the dead space around the note in the middle would be minimized.

Color

The areas surrounding each voice have differing colors per voice. The reasoning and color map are the same as earlier.

Opacity

In the example from the Music Animation Machine, the opacity of the areas is very high. While this causes the areas to be very clearly visible, it does cause severe overlap issues and draws almost more attention to the areas than to the actual notes themselves. Reducing the opacity can therefore be beneficial to remedy these issues, as shown in figure 5.9.

Combination with consonance/dissonance

In order to show the separation of voices, several channels that were originally used to show the consonance/dissonance status of notes were used. In order to show consonance, we have two main channels: color and shape. These options are explored in figure 5.11.

It is again clear that the use of shapes is much clearer than the concurrent use of colors for showing dissonances.

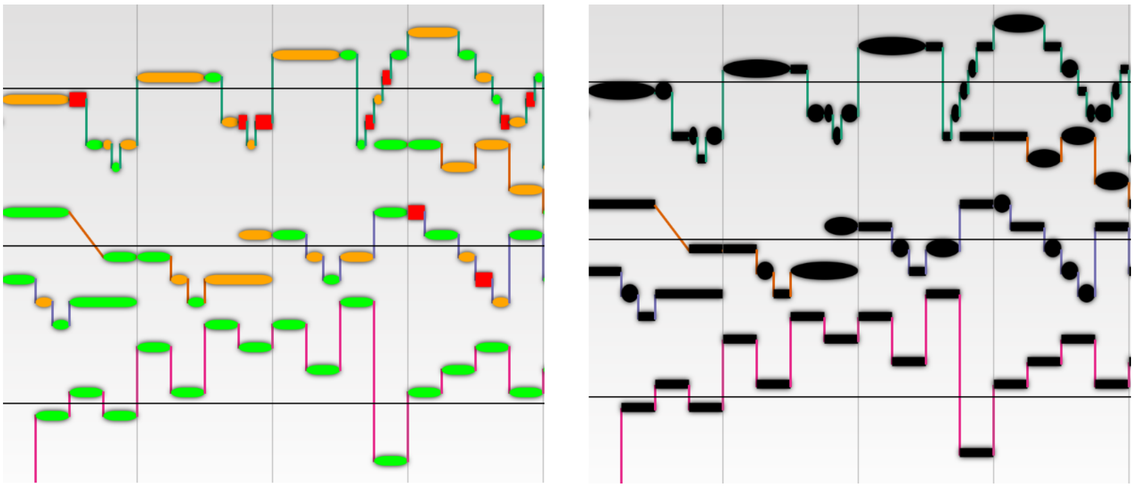

Figure 5.8: Lines are used to separate voices. On the left, consonance/dissonance is shown by filling the marks that represent notes with color. On the right, consonance/dissonance is shown using shapes.

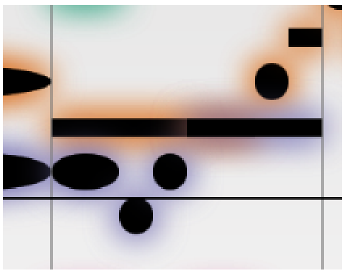

Figure 5.9: Reduced opacity solves overlap issues when notes are close together or even when notes are overlapping.

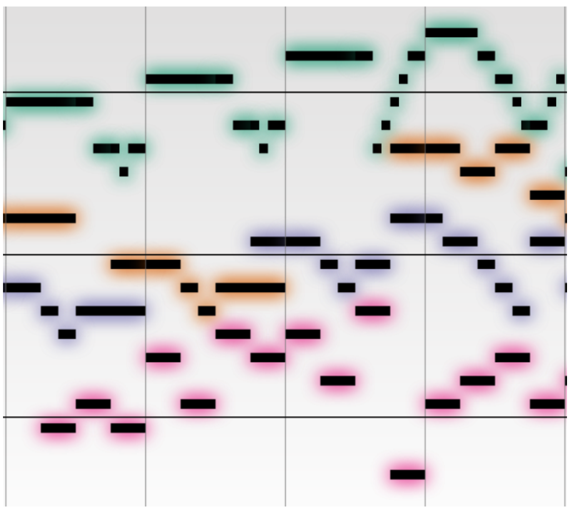

Figure 5.10: Areas are used to separate voices. No consonance/dissonance is shown.

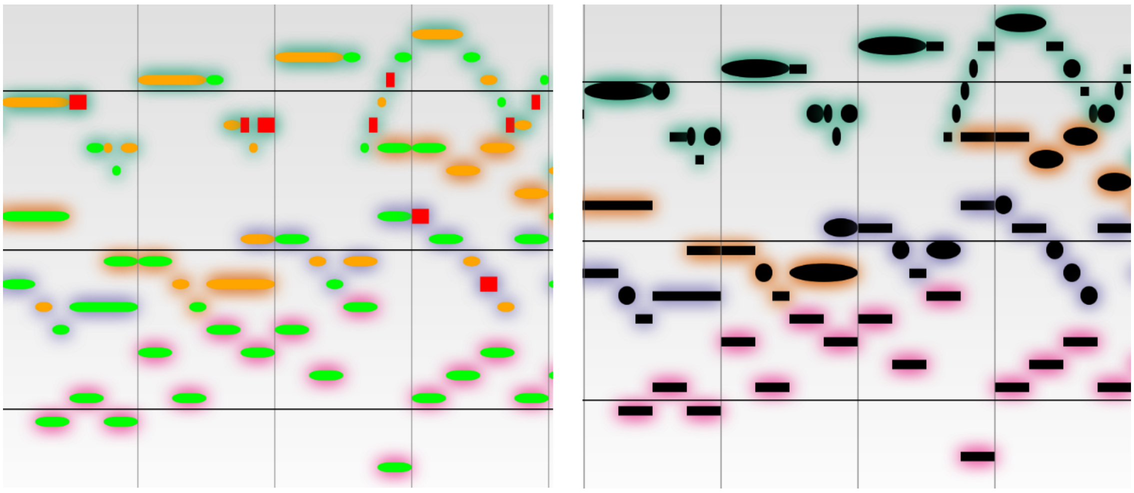

Figure 5.11: Areas are used to separate voices. On the left, consonance/dissonance is shown using color. On the right, consonance/dissonance is shown using shapes.

5.3 Additional features

In the musicians' view, bar lines were added as an additional feature. These function as guideposts to provide some perspective on time on a horizontal scale. For a musicologist, it is nice to have this perspective on a vertical scale as well. This is why the musicologists' view features octave lines that give some indication of the pitch range of notes.

Another requirement for the musicologist user group was that users should be able to see the exact size of intervals. This was done by implementing a tooltip which is activated by hovering over notes.

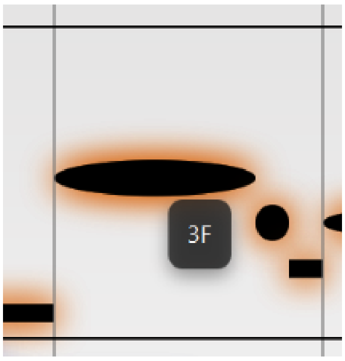

Figure 5.12: Tooltip containing letter of the note itself and the exact interval of a note relative to its bass note

5.4 Conclusion

To realize the requirements of the musicologists' user group, two different methods for voice separation were developed. Both methods will be investigated later through user testing, after which one will be chosen. The other option will remain available in advanced options of the Counterpoint Analyzer.

Chapter 6

Layman view

The purpose of the layman view is to let musical laymen get a feel for the concept of consonance and dissonance and how it influences the experience of music. In order to do this, the mood as influenced by consonance and dissonance should be visualized. At a low zoom level, this will lead to an "emotional signature" of the piece becoming visible. Additionally, the relative pitch of notes should be roughly visible such that users can see approximately where the melodies are and how many melodies are active at any given point.

Out of all views, the layman view is faced with the smallest amount of hard requirements. The average layman as specified in this thesis cannot read sheet music. Therefore, while in the other views the visualization needed to have some intuitive resemblance to traditional sheet music, this is not required here.

6.1 Existing Research

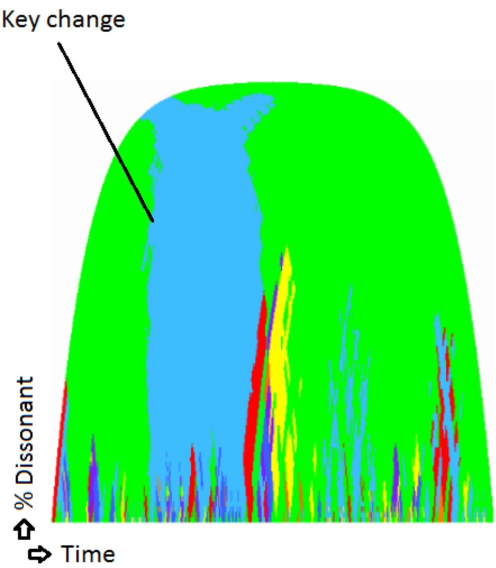

Figure 6.1: Harmonic visualization of Mozart Sonatina [9] The horizontal axis indicates time, the vertical axis indicates the amount of dissonance present. The coloured areas in the background show harmony.

As with most research in the music visualization field, visualization of mood has almost exclusively been focused on tonal music. There are many different opinions on what exactly constitutes the mood of a piece of music, and therefore each paper tends to use its own definition. There appears to be no existing research on visualizing mood in music purely based on the proportion of consonances and dissonances. The following examples are then mainly useful purely for the visualization of mood even though the definitions of mood may differ.

Sapp [9] created a system to visualize harmonic changes over time using colors. On the x-axis the time in the score is displayed while on the vertical axis the window of analysis is shown. The window of analysis refers to the percentage of notes that fit into the displayed key at a given point. To explain this concept further: sometimes composers will use notes that are outside of the dominant key at that point in time for effect. The vertical axis displays the extent to which this occurs. Every key has its own unique RGB value. The detection of harmony is performed automatically. In figure 6.1 the visualization of a Mozart Sonatina can be seen. It is clear that the piece starts and ends with C major sections (green) and has a middle section in G (blue). In between, Mozart uses notes outside of the key in many locations, as exemplified by the seaweed-like protuberances throughout the piece. Although the paper by Sapp is meant for a tonal instead of a modal context , the continuous visualization of harmony is similar to the continuously changing mood of a piece. Also, the protuberances that indicate notes outside of the harmonic context could be used to visualize dissonant notes.

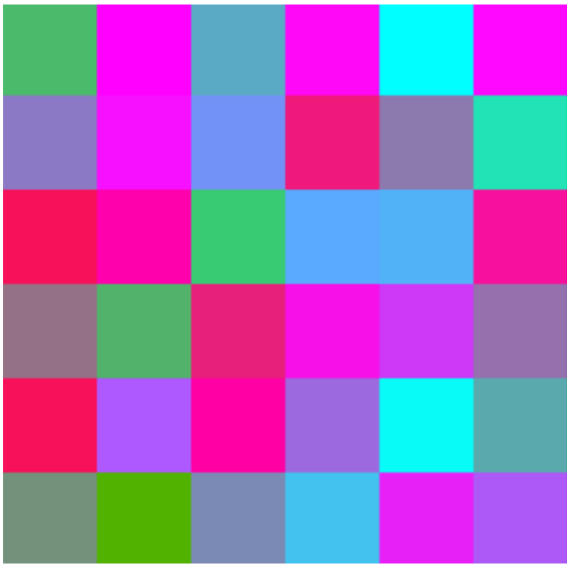

Figure 6.2: Visualization of the mood of the opening bars of Chopin's Op. 10 No. 3 [3]

Hiraga and Matsuda [3] used a set of colored areas arranged in a square to display the mood of a piece over a certain period of time. The exact mechanism is as follows. First, a section of a piece is selected. In figure 6.2 this concerns 2 bars of Chopin's Etude Op. 10 No. 3. Then, the notes are ranked in order of importance and an emotional value is calculated for each note, mainly based on performance information like articulation and dynamics. Each note has a small colored rectangle, with the most important note being shown on the top left, and the remaining notes shown in descending order of importance in a zigzag pattern. This visualization is less intuitive than I would like the layman view to be, since quite some background knowledge is needed to make sense of what's going on.

6.2 Analysis

6.2.1 Marks

According to the requirements, there are two main things that should be visualized: the mood of the piece and an indication of where notes are played, i.e. where the melodies are in the dimension of time.

Mood

The mood of a piece is present at any point in time. Mood, much like a melody, is a continuously and fluidly changing construct. This means points are not ideally suited for expressing it. Using lines would be possible, but using the same mark for both the melody and mood may not be the best thing to do because it may cause mix ups. One way to solve this problem would be to create delineated areas for the lines of the mood and melodies, respectively, however this would require the user to constantly switch their focus from one area to the other. Using some type of area would then seem to be the best option overall. This is echoed by the existing research, where areas are almost exclusively used for expressing mood and its change over time.

Melodies

Since laymen cannot read notes, the layman view is the only view where showing the exact pitches of notes is not a hard requirement. However, it can still be instructive to show the location and density of musical material, that is, where notes are being played and how many melodies are being played at a time. One way to accomplish this is to take the original piano roll view and to "flatten" it vertically, or in other words to keep the horizontal dimension and iron out the vertical dimension. An example of this is shown in the previously featured research of Hayashi in fig. 4.1, where for every voice it is shown when this voice is active without showing any specific tonal material. In figure 6.3 we can see the flattened view as implemented in the Counterpoint Analyzer, with the original view on the left and the flattened view on the right, both showing the same segment of music. In order to improve the clarity of the example, each voice has a different color.

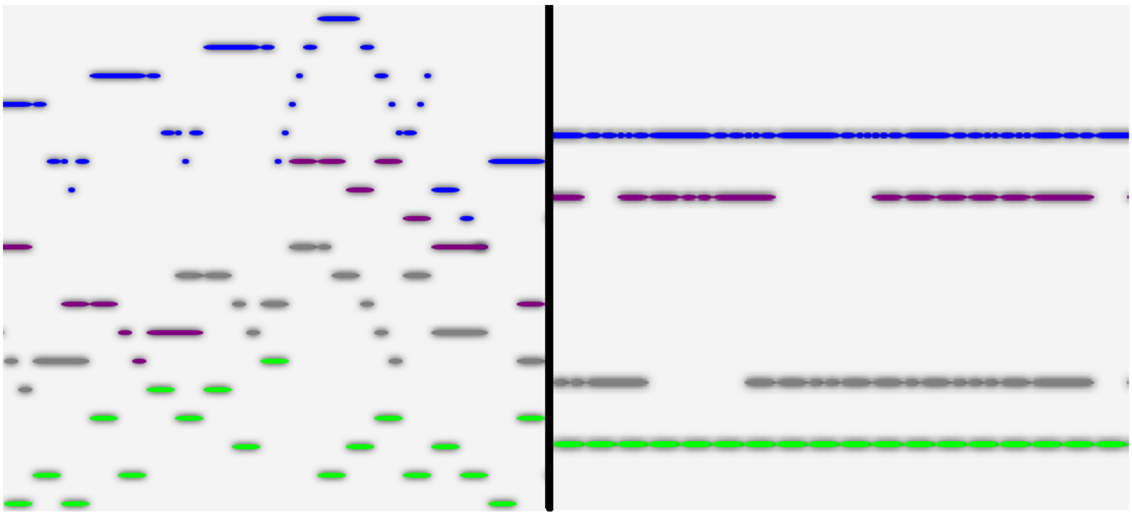

Figure 6.3: Comparison of regular piano roll visualization (left) vs. flattened piano roll (right)

The advantage in clarity of the flattened view over the piano roll view is quite big. An added advantage is that in the flattened view the voices are separated by design, much like in traditional sheet music notation (figure 5.1). Especially in cases where the color channel is not available this is an important benefit. The obvious disadvantage of the flattened view is that less information about the pitch of notes is available. Since one of the requirements of the layman view is to show at least some information related to the pitch of the notes, this is not optimal. Therefore, a semi-flattened view may be more appropriate (fig. 6.4). The semi-flattened view still allocates a specific range on the y axis to each different voice, however some pitch information is maintained in a flattened form. In the semi-flattened view, it is not feasible to accurately determine the pitch of notes like would be the case in the standard piano view, nonetheless it does provide a contour of the melodies which allows the layman user to more easily keep track.

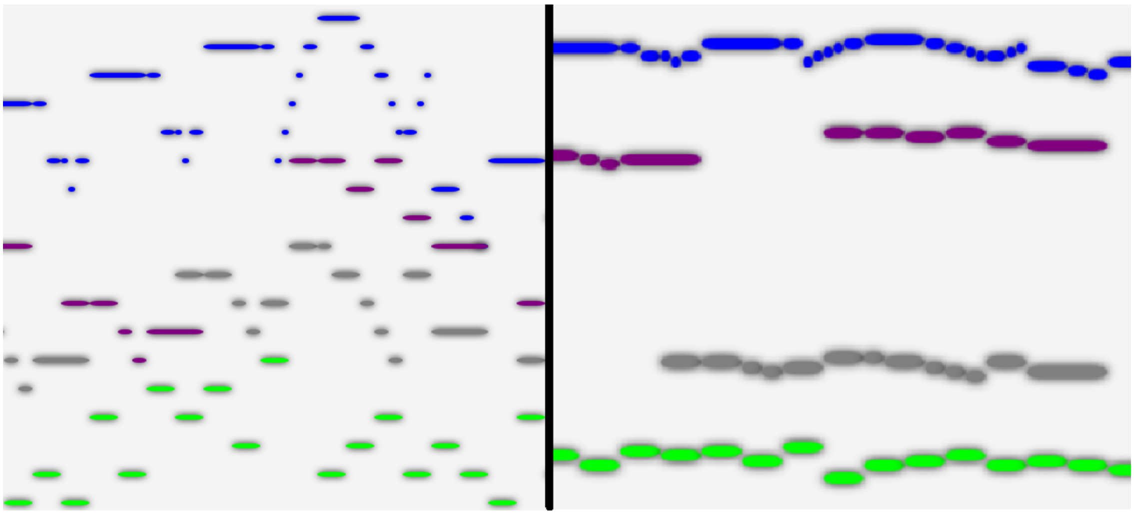

Figure 6.4: Comparison of regular piano roll visualization (left) vs. semi-flattened piano roll (right)

6.2.2 Channels for mood

Size

The mood of a piece is in essence made up of combinations of notes. At the same time it is an abstract construct that is present outside of said notes. Therefore, in order to visualize the mood, we can utilize all free space around the notes (or melody outlines, in this case). Since the mood is the main element of the visualization, it is allowed to take up most of the space.

Color and Shape

In the existing research, color is the channel that is most often used for visualizing mood. Here, it also seems like a viable option. One option is to calculate a value of overall consonance and dissonance at fixed intervals throughout the piece. This could for instance be every bar, or every beat. This value would incorporate the ratio of perfectly consonant vs. imperfectly consonant vs. dissonant notes. The exact calculation is specified in the implementation chapter. The color is then determined by locating the place of the calculated value on a predetermined color gradient scale. An example of a background containing several different gradient values is shown in figure 6.5. After this process, a color is associated with every temporal subdivision in the piece. A colored area could then be used to fill up the area related to every interval. When merging the transitions between the interval areas one would in effect create something resembling a heat map. At a low zoom level, this would create an emotional signature of a piece, i.e. a color code by which the piece is identifiable by its mood. This emotional signature could then be compared to that of other pieces.

With regard to the gradient, white is used to indicate consonance, since consonance is the default situation in the great majority of pre-20th century music and using a color would instead create the impression that something notable is occurring.

Position, Tilt and Shape

As discussed, mood areas are displayed across the whole length of the piece and form the backdrop of the entire visualization, position, tilt and shape are therefore not relevant.

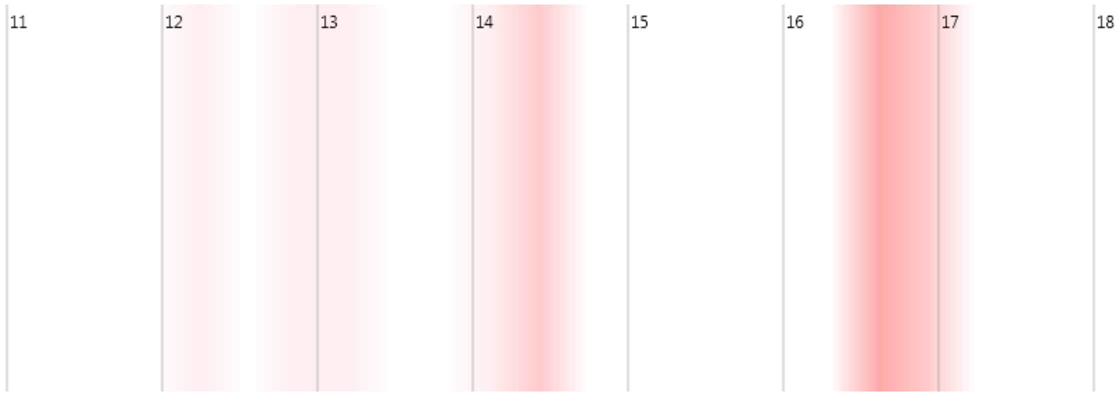

Figure 6.5: Background gradient showing density of dissonance

6.3 Conclusion

To realize the requirements of the laymen user group, the traditional piano roll view was modified to include information on the mood of the piece and to show relative pitch information instead of absolute pitch information. Several design decisions made in this chapter will be investigated later through user testing.

Chapter 7

Implementation

In this chapter, implementation details will be discussed. Although software development was one of the most time-consuming areas of the project, this is not a thesis on software engineering and the focus is therefore not on describing all details. Instead, only the most interesting aspects will be discussed.

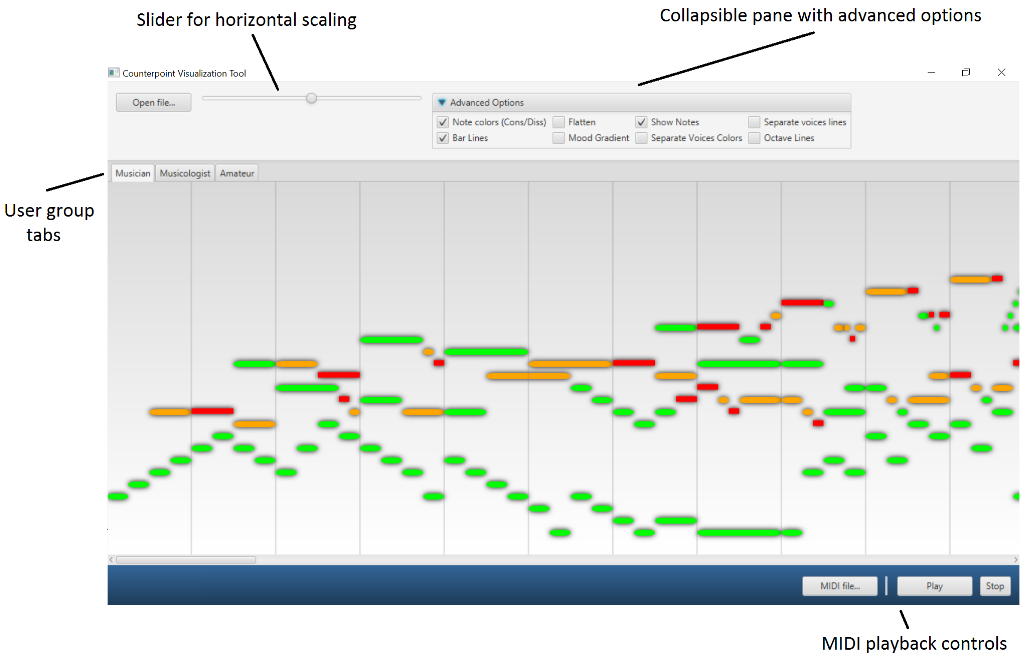

Figure 7.1: The interface

7.1 The interface

The interface features tabs that allow the user to switch between views. In addition, there is a collapsible pane that allows the user to set advanced options. The visualizations shown in the tabs consist of a combination of options from the advanced options pane that are automatically set upon switching tabs.

There is also a slider to set the horizontal scale. A scroll bar is used for horizontal navigation in cases where the full piece does not fit into the window, and vertical scaling happens automatically through a binding of the window size to the vertical padding between notes.

Moreover, there is an option to load a MIDI file and play or stop it. While playing the MIDI file, a vertical line moves across the visualization horizontally in order to show the location of playback.

7.2 Technology used

To provide the object oriented development framework and graphical interface, Java 8 and the JavaFX graphics package were used.

The final version of the Counterpoint Analyzer contains 3604 lines of code divided over 11 classes.

With regard to the musical data, in the field of visualization of music this data usually comes in two varieties, namely MIDI (Musical Instrument Digital Interface)-type data and typical waveform data. MIDI is a file format in which every note in a piece of music is encoded as a discrete event occurring at a specific time with a number of associated parameters such as pitch, octave, instrument type, length, etc. Most of the data contained in MIDI files cannot readily be extracted from waveform data. Using MIDI-type data would then seem to be the most appropriate for this thesis. To be more specific, the exact format used is MusicXML. Like MIDI, every note is encoded as an event with a set of associated parameters. MusicXML is a widely used format, and most of the commonly performed modal music repertoire is available in this format.

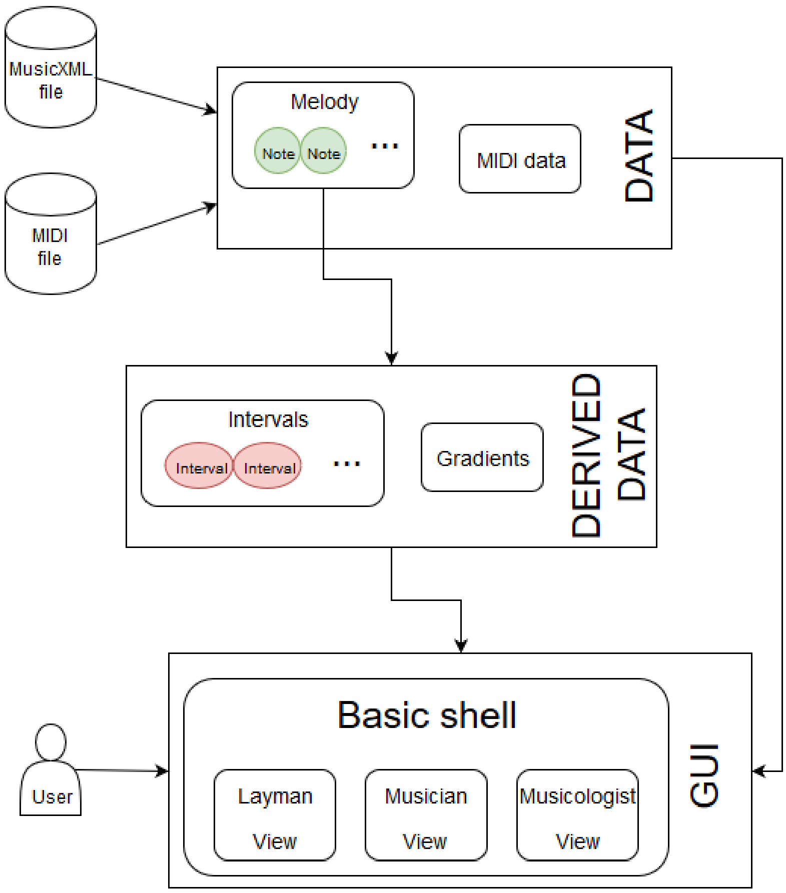

Figure 7.2: Diagram showing the main components of the software package and the ways in which they are connected

7.3 Obtaining data

In order to obtain the required data, it needs to be extracted from the MusicXML file. This is done using a library by Herremans [2], which provides simple text output containing all required information, one note per line. This library is the only external library that was used in the project. Apart from this, several of the default Java libraries are used. The rest of the code is thus original.

7.4 Storing data

The data delivered by Herremans' library is fed straight into an object oriented framework. In this framework, there are several objects that hold the data, as shown in figure 7.2. These objects will now be discussed.

The main data structure used for storing individual notes is the array. We want the size of this array to be as small as possible while still being able to store all notes. In order to do this, we start by finding the shortest note in the entire piece. Next, an array holding all notes in the piece is created, of which the size is equal to the total number of bars multiplied by the number of shortest notes that fit exactly into one bar. This way, every location in the array represents one note with the length of the shortest note, and we can be sure that it is possible to store every note while minimizing length. The arrays containing notes are stored in "Melody" objects. These objects are passed to an interval calculation algorithm, and the results are stored in "Interval" objects, combined with references to the notes to which the intervals belong and some assorted data functions. From these interval objects, gradients are derived. Both the interval and gradient calculations will now be discussed.

7.5 Calculating required information

One of the main pieces of required information are the intervals between the different voices. These are found by iterating through the combinations of the lowest voice and all other voices and calculating the distances between notes, excluding octave information. Accidentals (i.e. sharps and flats) are ignored, since the framework is designed to operate within a modal context. From these intervals, the consonance/dissonance status of notes is deduced and stored.

Below, the interval calculation algorithm is given in pseudo code. The algorithm is provided with an array list containing all voices. First, the top and bottom voices are assigned. The bottom voice stays constant, the top voices are iterated through. For each time slice, i.e. every array location, the interval is then calculated and stored as an integer.

Algorithm 1 Calculate intervals between voices

```
procedure IntervalCalculation(AllVoices)
    bottomVoice = AllVoices.get(0)
    //Iterate through all top voices
    for x=1 to AllVoices.length do
        topVoice = AllVoices.get(x)
        //Iterate through all time slices
        for y = topVoice.start to topVoice.length do
            if !topVoice.get(y).empty && !bottomVoice.get(y).empty then
                topNote = topVoice.get(y)
                //Calculate the interval
                if topNote.octave <bottomNote.octave then
                    tempInterval = bottomNote - topNote;
                else
                    tempInterval = topNote - bottomNote;
                end if
                //Make sure the interval stays in the right octave
                if tempInterval <0 then
                    tempInterval = tempInterval + 7;
                end if
            end if
        end for
    end for
    Store tempInterval
end procedure
```

7.6 Visualization

Musicians' view

The musicians' view was the first view to be developed, which means that in this stage the piano roll design was made. The piano roll visualization is the visualization on which all views are more or less based. The main challenge from an implementation perspective was then to keep it flexible and easily extensible, since many variations on the standard piano roll were likely in future development. Since at the start it was not completely clear which features were eventually going to be added, it was extra important to keep the basic framework as flexible as possible. One of the ways in which this was accomplished was by compartmentalizing the software to a maximum extent. This was accomplished by using several design patterns, including the MVC pattern and the observer pattern. Also, information that is passed between the view and the controller is encapsulated into several abstract data types that are easily extensible. For instance, an abstract data type containing all options and the status thereof is passed between the view and the controller.

Layman view

In the layman view, the mood is shown as a horizontally changing gradient. This gradient is calculated as follows. First, the piece is split up into a number of sections of equal length in time. Next, a color from the gradient needs to be assigned to each of these sections. This color is calculated by separately counting the number of notes considered perfectly consonant, imperfectly consonant and dissonant, and then dividing the sum of the amount of perfect and imperfect consonances by the total amount of notes in the section. Let g_i contain pc_i perfect consonances, ipc_i imperfect consonances and t_i notes in total, where i represents the section. The following formula then calculates the gradient variable:

$$g_i = \frac{pc_i + ipc_i}{t_i}$$

The color of the section is determined by applying a scale from 0 to 1 to the gradient and finding the color at the fraction. Finally, separate gradients are created between each vertical section to make the sections flow into each other. The full process is shown in figure 7.3. The size of the sections is equal to one bar in this example.

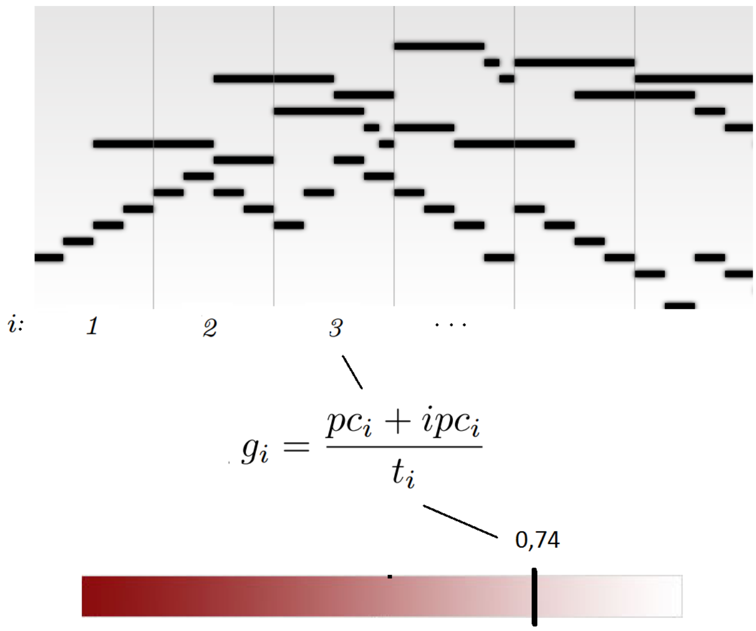

Figure 7.3: A piece is divided into a number of sections. For each section a value g_i is calculated, which is mapped onto a gradient.

Chapter 8

Qualitative Evaluation

8.1 Methodology

8.1.1 Assessment methods

Throughout this thesis, certain design decisions were reasoned to be the most logical. In this chapter, these decisions will be evaluated.

The visualizations in this thesis were designed with specific user groups in mind. Therefore, in this evaluation, test subjects from each of the user groups will be used. There will be a short training period before the start of the evaluation for all test subjects. During the training period, the main elements of the interface will be explained and the subjects will be able to try out the program as they wish. Every groups will be asked to perform several tasks to obtain objective data. Subjective data in the form of value assessment will also be obtained through a separate series of questions.

For the evaluation, the piece Nun Komm der Heiden Heiland by Johann Sebastian Bach will be used, unless indicated otherwise. This piece contains dense counterpoint which makes it suitable for the types of operations that are to be tested.

8.1.2 Test subjects & tasks

There are 6 test subjects in total, 2 for each user group. The test subjects in the musicians' group are conservatory students and graduates. The subjects in the musicologists' group include a professional conductor and a student in composition. The laymen category represents people who've never played an instrument and who cannot read sheet music.

Musicians

The main goal of the musician's visualization is to help musicians find consonances and dissonances in a piece.

Evaluation will be done by letting the test subjects perform several tasks. Since subjective experience also strongly influences the results, this will be inquired as well. Before each of the subjective questions, the test subjects will be asked to perform task T1.

- T1: Find the dissonances in the first 10 bars of the piece.
- S1: Do you think tilting the notes improves clarity?
- S2: Do you think voice separation adds useful information? Does it obfuscate the view?
- S3: Do you think increasing the vertical size of dissonances improves dissonance recognition?

Musicologists

For the musicologists' view, two methods for voice separation were presented, one using areas and one using lines.

The hypothesis is that the method that uses areas will be more effective.

Evaluation will be done by letting the test subjects perform several tasks. The tasks below are to be performed on two pieces of equal complexity. The first piece will be "Nun Komm der Heiden Heiland" by J.S. Bach and the second piece will be "Missarum Liber Nonus" by Palestrina. In order to make a comparison, the areas and lines methods will be assigned such that both pieces use different methods.

- T1: Find the number of voices.
- T2: Where is each of the voices introduced?
- T3: Find the total number of line crossings and their locations.
- T4: At several locations in the piece, one of the voices splits. Where does this occur?
- T5: What is the octave range used in this piece? First time the octave lines are hidden, the second time the octave lines are shown.
- S1: Do you like the method that uses areas or the method that uses lines more?
- S2: Do you prefer the notes to be attached to the ends of notes of to the middle of notes?

Laymen

The hypothesis is that the background gradient is more efficient at intuitively showing the location of dissonances than simply using channels on the marks of notes as is done in the other two visualization methods.

- T1: Find the place in the piece that has the highest degree of dissonance.
- T2: Find the voice with the largest pitch range.
- S1: Do you think the look of the semi-flattened or flattened voices is clearer?
- S2: Do you prefer the voices to be colored or black?

8.2 Results

Through trial and error, the participants were able to use the Counterpoint Analyzer fairly quickly without much external guidance. This shows that the interface is intuitively usable without having to consult documentation.

8.2.1 Musicians

- T1: The dissonances in the first 10 bars were easily found.
- S1: As expected, tilting the notes was not received positively by the participants. One participant said the that a series of tilted notes "resembled a set of stairs", and another claimed that it looked like the melody was moving upwards even though it was not. Moreover, participants thought the overall view became more "messy" when tilting notes. It should however be said that the overall clarity was not impacted and dissonant notes could still be found.

- S2: Both participants thought that although information about the separation of voices could be valuable, it did obfuscate the view to an extent where the downsides outweighed the upsides. This was especially pronounced when using the area method for voice separation due to the multitude of (occasionally clashing) colors, but it was also present in the method that uses lines, especially at low zoom levels.
- S3: Both participants agreed that increasing the vertical size of the marks made dissonances more noticeable, but the effect was very slight. Increasing vertical size by 25% on top of the initial increase improved the effect without increasing clutter.

8.2.2 Musicologists

- T1: The number of voices was found fairly quickly. Both participants zoomed out completely to get a full overview, after which they counted the number of colors.
- T2: To find where each of the voices was introduced, the participants again zoomed out completely. Using the area method, it was fairly easy to locate the introduction of the voices by simply looking for the first occurrence of the color corresponding to the voice. When using the line method this task took slightly more time because while the lines are also colored, the colored area of the lines is much smaller which makes it harder to recognize, especially at low zoom levels.
- T3: Voice crossings were easier to find in the visualization that employs lines. The explicit crossing of lines was easier for the participants to interpret than the altered vertical order of notes that occurs when voices cross in the visualization that uses areas.
- T4: Voice splitting is rare and only occurs a couple of times in the pieces. Therefore it took some time for the participants to locate them. Neither method was at a clear advantage here.
- T5: Without octave lines, participants were able to correctly deduce the octave range of the piece by looking for a dense area in the piece and counting the number of vertically stacked notes. This method did however take some time. With octave lines, the deduction of the octave range was almost instantaneous.
- S1: Although both methods proved to be effective, both participants had a small preference for the method that uses areas. They felt that both methods provided roughly the same information, while the method that uses areas had less visual clutter, especially at low zoom levels.
- S2: Both participants preferred the lines to be attached to the ends of notes, mainly because they thought attaching the lines to the middle of notes made the overall look feel cluttered, especially at low zoom levels.

8.2.3 Laymen

- T1: The location containing the highest degree of dissonance was easily found by zooming out completely and finding the darkest color.
- T2: The voice with the largest pitch range was also easily found by quickly iterating through the voices.
- S1: The semi-flattened look was preferred by both participants, mainly because they felt that too much information was being left out in the flattened view and adding more pitch information in the form of the semi-flattened view did not detract from the overall visualization. Moreover, one of the participants said that the lack of vertical movement in the flattened view made it look as if the same note was going to be played over and over.

- S2: Both participants preferred the color of the voices to be black. According to them, the vertical area dedicated to each voice was already enough to clearly segregate the voices, and colors thus added unnecessary visual information.

8.3 Conclusion

For the most part, the users' experiences validated the design decisions made in the preceding chapters. Differences between alternative solutions were usually small but significant. There were however tasks in which neither solution was at a clear advantage, such as using areas or lines to detect voice splitting. In some instances, the user evaluation prompted a changes to some aspects of the visualizations. For instance, the vertical height of dissonances in the musicians' view was increased.

Chapter 9

Results

9.1 Use case 1: Showing consonance and dissonance

The primary goal of the software developed in this paper is to show consonances and dissonances and their locations in any piece of music. Below, the final versions of the three visualizations are shown. The same section of the piece Nun komm, der Heiden Heiland by Johann Sebastian Bach is used for comparison.

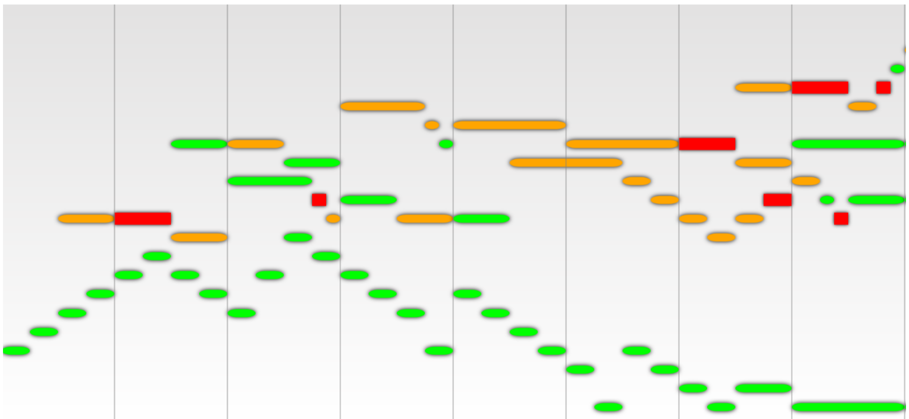

Figure 9.1: The final musicians' view

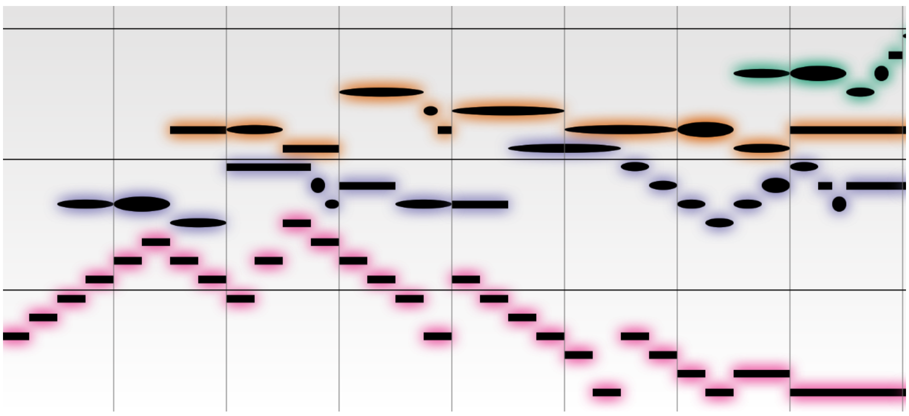

Figure 9.2: The final musicologists' view

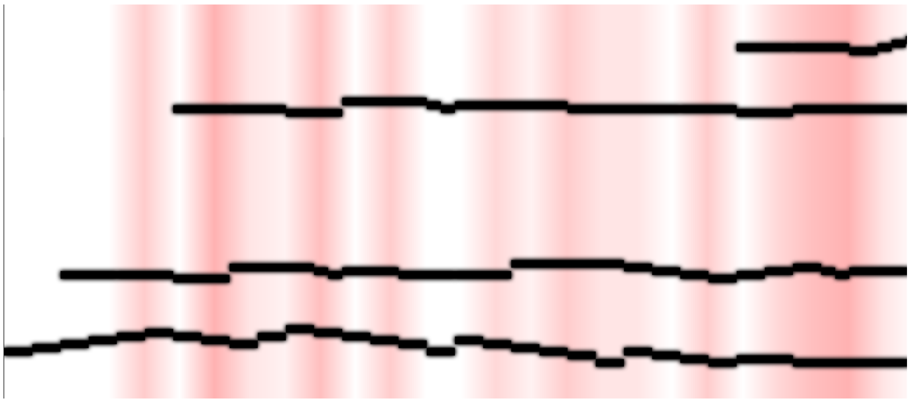

Figure 9.3: The final laymen view

9.2 Use case 2: The emancipation of dissonance

At the start of the 20th century, a group of prominent composers including Arnold Schoenberg and Anton Webern made the observation that from the start of the ancient Renaissance period on, music became increasingly dissonant through the centuries as people got more used to dissonant sounds. They dubbed this phenomenon the ”emancipation of dissonance”. It is possible to test this phenomenon by making a comparison between different historical periods and the amount of dissonance between them. The software developed in this paper can be used to do this. The layman view is particularly useful for this, since it allows us to get a rough view of the concentration of dissonance at a single glance. In the next few figures, sections of pieces from different time periods are visualized using the layman view. These pieces are representative for their time periods, since they contain the main features that are generally ascribed to the periods in which they were written.

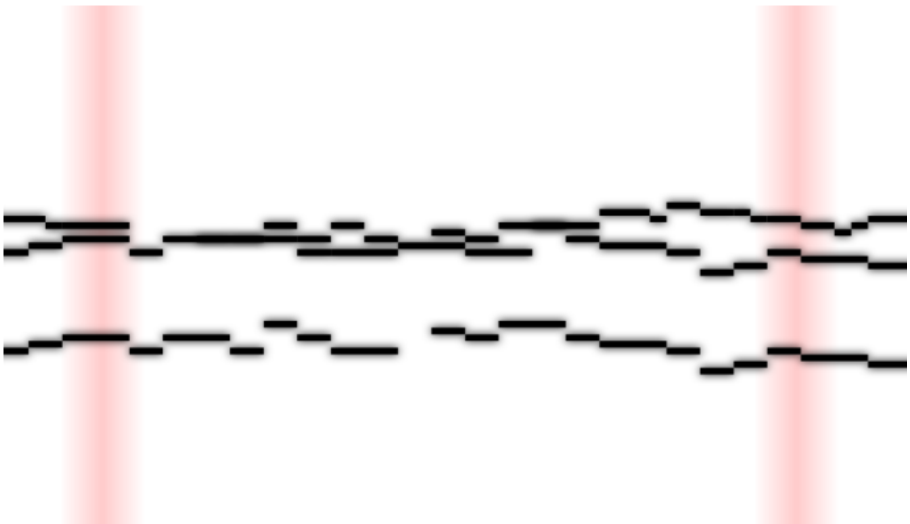

Figure 9.4: Early Renaissance music: Binchois (1441). Very limited amount of dissonance: the background is almost completely white.

As can be seen in the included figures, the degree of dissonance in the different pieces corresponds to what one would expect for the periods from which the pieces were taken. Also, the differences between the visualizations are large enough that given a random piece, one would in

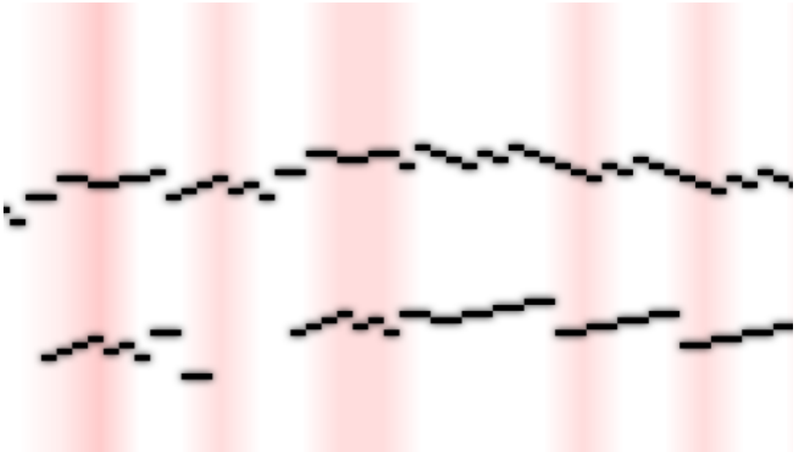

Figure 9.5: Late Renaissance music: Palestrina (1570). Moderate amount of dissonance, the background is colored in several places.

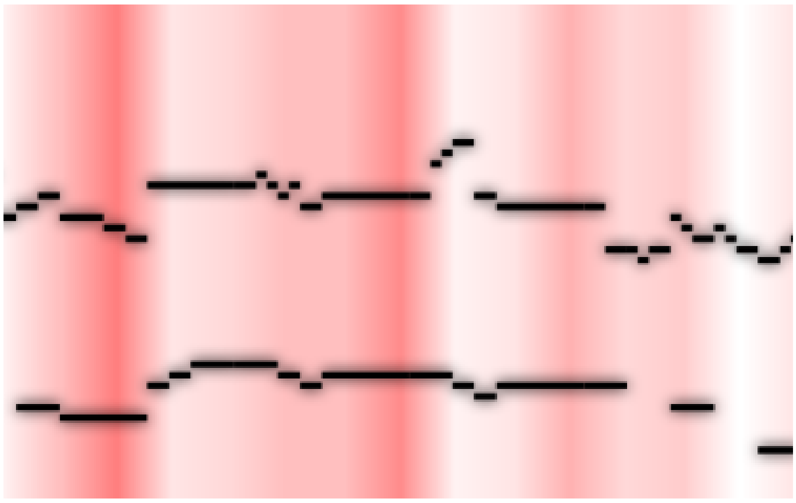

Figure 9.6: Late baroque music: J.S. Bach (1745). Heavy dissonance, most of the background is colored. Moreover, the colors are fairly dark, which indicates a high degree of dissonance.

most cases be able to make a fairly accurate estimate of the piece's time period without knowing any other information. Therefore, the emancipation of dissonance would appear to be a real phenomenon. This use case is merely meant as a demonstration of the software, since a thorough proof of Schoenberg's claim would of course require a wider sample from each time period to be visualized.

9.3 Use case 3: Separation of voices

In a traditional piano roll view, there is no information to indicate which note belongs to which voice, or even how many voices there are to begin with. This is displayed in figure 9.7. In figures 9.8 and 9.9, areas and vertical separation are used to display which note belongs to which voice. It is clear that these methods aide greatly in distinguishing the voices.

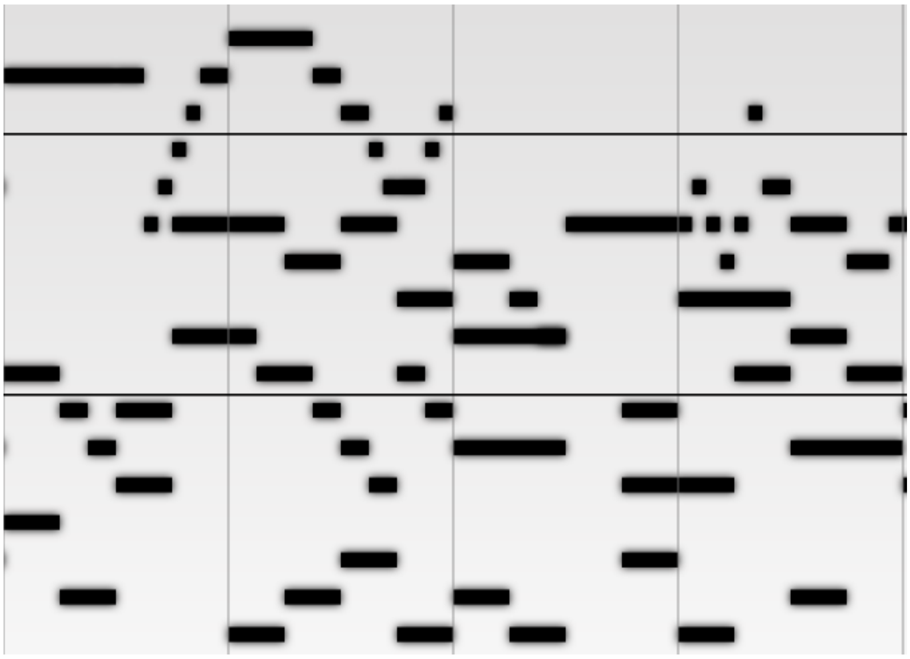

Figure 9.7: Bar 59-62 of Nun Komm der Heiden Heiland displayed using traditional piano roll visualization. Impossible to see which notes belongs to which voice.

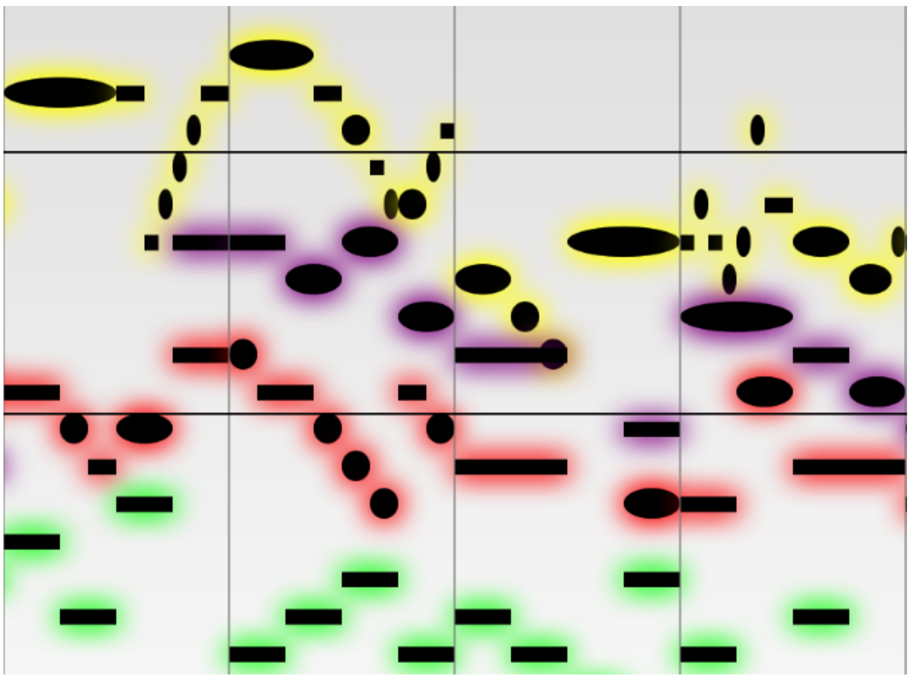

Figure 9.8: Bar 59-62 of Nun Komm der Heiden Heiland displayed using the musicologists' visualization. Colored areas are used to separate voices.

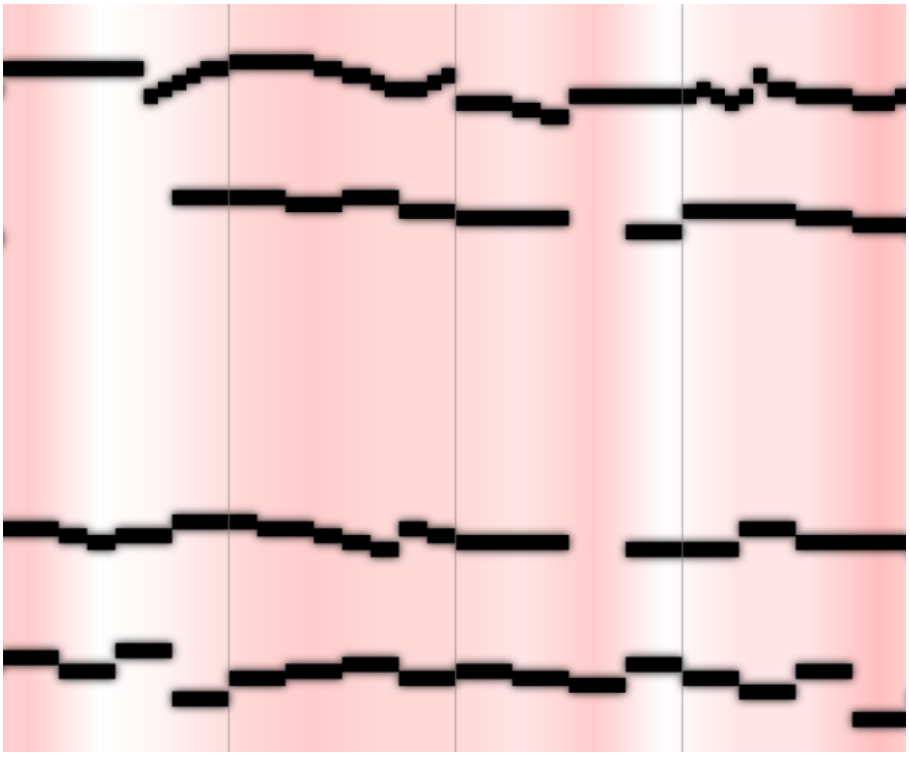

Figure 9.9: Bar 59-62 of Nun Komm der Heiden Heiland displayed using the laymen visualization. Delineated vertical areas are used to separate voices.

Chapter 10

Conclusions

In this thesis, the development of a visualization tool for the purpose of analyzing consonance and dissonance in modal music has been discussed. Several views corresponding to user roles were thoroughly investigated and implemented. For the analysis, the framework of marks and channels was used.

10.1 Requirements

At the start of the thesis, a number of requirements were postulated. The majority of these requirements have successfully been implemented over the course of the project, however some were not possible to realize. The requirements were realized using three different visualizations, each with their own special functions. In short, data from any musicXML file can be translated to a piano roll view with adjustable zoom level in a graphical interface. This piano roll view is modified to show the consonance and dissonance status of individual notes as well as the type thereof, and to which voice these notes belong. Furthermore, bar lines and octave lines are shown. Also, it is possible to load a MIDI file and play it, with the current playback location shown in the visualization.

Requirement 14 and 15 are concerned with adding a function which allows the user to see traditional sheet music alongside the visualization. However, it turned out that implementing a sheet music generator is not trivial, and in fact, making a good implementation could be a whole separate project. Libraries for generating sheet music from MIDI of MusicXML are few, and most of the libraries that are available at this moment do not function well enough to allow implementation into the project at this point, however this may be the case at some time in the future.

Requirement 19 is about the availability of statistics regarding for instance the percentage of dissonant notes and the number of parallel intervals. In a project like this, it is possible to add an almost infinite amount of functions. Therefore, at some point a decision has to be taken as to which of these proposed functions stray too far from the original design and thus provide information that is nice to know but doesn't necessarily serve the original purpose. The calculation of these types of statistics fits into this category, more so because it's not actually a visualization but simply a series of operations on data.

10.2 Contribution to field

The existing research in the field has been almost exclusively focused on tonal music, as opposed to the exploration of modal and atonal music. Additionally, although some research has gone into the analysis of consonance and dissonance, this has hitherto not been done in a visualization context. Moreover, visualization of contrapuntal music on a intervallic level has received relatively little attention. The main contribution to the field then lies in these areas.

10.3 Future Work

The Counterpoint Analyzer contains a framework that could rather easily be extended to include a far wider array of functions that are beyond the scope of this project. These functions could include purely visual extensions and improvements, but could also be purely related to data and inferences that can be made from this data. An example of this could be the calculation of certain statistics as described in requirement 21.

It may be nice to add extra functionality related to MIDI playback. Once could for instance imagine a special sound being played whenever a the MIDI player comes across a dissonance, or to assign distinct instruments to the different voices to increase the divide between their sonorities. The MIDI protocol works with binary code, and modifying it thus requires some not so straightforward binary manipulation. Furthermore, the quality of MIDI files varies considerably, since the encoding of classical pieces in MIDI is anything but standardized. One can for instance see multiple voices being lumped into a single voice, overlapping channels, etc. This creates additional challenges when trying to create a system that will universally accept these files.

Bibliography

[1] A. Hayashi. Colorscore – Visualization and Condensation of Structure of Classical Music, 2011. ixix, 11, 12

[2] D. Herremans. MusicXML parser in Java, 2015. 33

[3] N Hiraga, R; Matsuda. Graphical expression of the mood of music, 2004. xx, 26

[4] R.J. Krawczyk. Exploring the Visualization of Music, 2012. ixix, ixix, ixix, 8, 9, 10, 15

[5] Malinowski, S. Music animation machine, 2017. [Online; accessed February 28, 2017]. ixix, ixix, ixix, 17, 18, 19

[6] T. Munzner. Visualization Analysis Design, 2014. 4, 11

[7] Museotech. Timeline of musical periods, 2012. [Online; accessed January 12, 2017]. ixix, 1

[8] G. P. Palestrina. Missarum liber nonus, 1599. First edition, Venice: Erede di Girolamo Scotto, 1599. ixix, 17

[9] C Sapp. Harmonic visualisations of tonal music, 2001. xx, 25, 26

[10] Wikimedia Commons, the free media repository. Chase baker player piano, buffalo, ny, circa 1885, 2009. [Online; accessed January 12, 2017]. ixix, 8

[11] Wikimedia Commons, the free media repository. Computer music piano roll, 2009. [Online; accessed January 12, 2017]. ixix, 8

[12] Wikipedia, the free encyclopedia. Bachfuguebar, 2008. [Online; accessed January 11, 2017]. ixix, 3

[13] S.; Mulholland P. Wilkie, K.; Holland. What can the language of musicians tell us about music interaction design?, 2010. 7