VISUALIZING MUSIC: TONAL PROGRESSIONS AND DISTRIBUTIONS

Arpi Mardirossian and Elaine Chew
University of Southern California Viterbi School of Engineering
Epstein Department of Industrial and Systems Engineering
Integrated Media Systems Center
Los Angeles, CA 90089, USA

ABSTRACT

This paper presents a music visualization tool that shows the tonal progression in, and tonal distribution of, a piece of music on Lerdahl's two-dimensional tonal pitch space. The method segments a piece into uniform time slices, and determines the most likely key in each slice. It then generates the visualization by dynamically showing the sequence of keys as translucent, growing discs on the two-dimensional plane. The frequency of a key is indicated by the size of its colored disc. Each color and position corresponds to a key, and related keys are shown in proximity with related colors. The visual result effectively presents the changing distribution of the keys employed. The proposed visualization is an improvement over more basic charting methods, such as histograms, and it maintains standards of information design in the form of added dimensionality, color, and animation. We show that the visualization is invariant under music transformations that preserve the piece's identity. We conclude by illustrating how this method may be used to visually distinguish between tonal progression and distribution patterns in western classical versus Armenian folk music.

1 INTRODUCTION

Music visualization literature can be broadly grouped into two categories: visualization of individual pieces of music (our focus), and of collections of pieces. It can be said that the first form of music visualization created for individual pieces was music notation itself. An experienced musician can often look at the score of a piece and "see" what the music sounds like. Music notation cannot be used readily as a mainstream form of visualization because it can take years of training to learn to decipher the subtleties of the encoded information.

Our goal is to create a more intuitive visualization that reveals important features of the music that may not be readily audible to the inexperienced ear. The challenge with developing such a visualization is that music is complex, consisting of multiple inter-related features. A successful visualization must strike a balance between simplicity and comprehensiveness. We aim to create imagery that is both intuitive and informative.

In this paper, we propose a visual interface based on Lerdahl's Tonal Pitch Space [12], which portrays all major and minor keys on a two-dimensional (2D) plane. The distributions of keys are indicated as growing colored discs, where the colors correspond to the keys detected, and the size of the discs to the key frequency. Figure 1 shows the visual interface. The user selects the piece and the granularity of analysis through the graphical user interface.

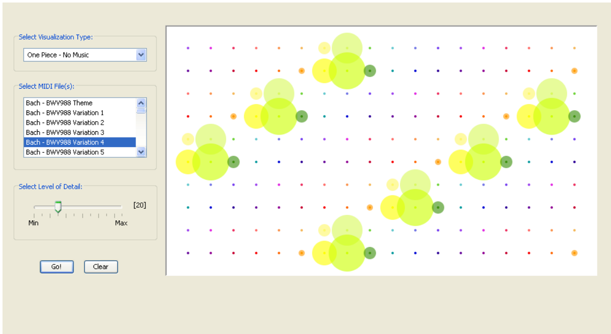

Figure 1. Snapshot of Visualization Interface

In our previous work ([14, 15]), we investigated how key distributions could be successfully used to assess similarity between pieces, demonstrating that key distributions, although a summarization of the musical content, can serve as good representations of pieces. The current visualization method is an extension and improvement of the key distribution approach, expanding and adding richness to the simple histogram representation through an increase in dimensionality, addition of color, and animation.

According to Tufte [21], an acknowledged expert in information design and visual literacy, increasing the number of dimensions of a visualization sharpens the information resolution. In the histogram, the keys were shown on a one-dimensional line, while in the new visual interface, the keys (all major and minor keys) are shown on a 2D plane, thus capturing the network of inter-relations amongst keys. The frequency of the keys (the third dimension) is shown in the size of the discs. Furthermore, the progression of disc growth shows the range of movement of keys within the piece over time. Hence, we have essentially four dimensions of information captured in a dynamic 2D interface.

Tufte states that representations that progress, such as the proposed animated visualization, can be referred to as 'small multiple designs' and "answer directly [the question of 'compared to what?'] by visually enforcing comparisons of changes, of the differences among objects, of the scope of alternatives." The proposed visual interface incorporates these ideas of small multiple design by taking a sequence of keys and showing the evolution frame-by-frame. This dynamic visualization allows one to see the sequential progression of keys, an important component in communicating with music.

Third, Tufte states the fundamental uses of color in information design as being: to *label* (color as noun), to *measure* (color as quantity), to *imitate reality* (color as representation), and to *enliven or decorate* (color as beauty). In our visualization, color labels by distinguishing between keys, measures by displaying the amount of time spent in each key, imitates reality by showing the relationship between keys, and decorates since the same visualization in black and white would not be nearly as visually pleasing.

The remainder of the paper is organized as follows: Section 2 presents related work in music visualization, Section 3 describes our visualization system, followed by validation of the visualization through translation analysis in Section 4, and demonstration of the visualization system in Section 5. Section 6 presents our conclusions.

2 RELATED WORK

This section reviews a selection of the many music visualization systems developed so as to put the work presented here in perspective. As noted above, visualizations can be broadly categorized into visualizations of collections and individual pieces. Since our work does not consider collections, this review will be limited to visualizations of individual pieces. These systems may be further subcategorized as follows: representations of direct versus interpreted data, and static versus dynamic presentations. Direct data refers to data that is extracted directly from the music (such as pitch and onset time), while interpreted data refers to information that must be determined from extracted data (for example, tempo and key).

Let us consider visualizations of direct data. The most basic visualizations in this category are 2D waveforms and spectrograms which usually show time on the x-axis, and have primary values of interest on the y-axis. Additional mappings of these primary values are often shown using color or greyscale ranges. Misra, Wang, and Cook [17] present such visualizations (real time) with some added features and dimensionality. Another example of direct data visualization is Malinowski's "Music Animation Machine" [13], which dynamically shows notes in a simplified piano roll representation.

We now turn our attention to systems that visualize interpreted data. We first review static systems. One approach to music visualization is to create self-similarity maps. In the work developed by Cooper & Foote [8], the acoustic similarity between all instants of an audio recording is calculated and displayed on a 2D grid. Similar or repeating elements are visually distinct, allowing identification of structural and rhythmic characteristics. Another self-similarity visualization by Wattenberg et. al. [19] displays musical form as a sequence of translucent arches. Each arch connects two repeated, identical passages of a composition. By using repeated passages as landmarks, the maps reveal deep structures in musical compositions.

An early work by Cohn [7] established mappings of music onto the harmonic network (also known as the *tonnetz*). We now transition to visualizations of interpreted data that are also dynamic. Related to the harmonic network visualization is Toiviainen & Krumhansl's [20] visualization of listeners' continuous ratings of tonal contexts on a toroid representation of keys (shown in 2D). Their work measures and models real-time responses using self-organizing maps. Chouvel [5] showed tonal analyses of a number of pieces on a hexagonal version of the *tonnetz*. Gomez & Bonada [10] developed a tool to visualize the tonal content of polyphonic audio signals. This tool includes different views that may be used for the analysis of tonal content of a music piece through visualization of chord and key estimation, and tonal similarity assessment. Sapp [18] developed a multi-timescale visualization technique for displaying the output from key-finding algorithms. In his visualization, the horizontal axis represents time in the score, while the vertical axis represents the duration of an analysis window used to select music for the key-finding algorithm. Each analysis window result is colored according to the determined key.

The following works also maintain history information. Langer & Goebl [11] introduced a method for displaying tempo and loudness variations of expressive music performance. In this visualization, a dot moves through a 2D space representing tempo (x-axis) and loudness (y-axis), leaving behind a trace of the recent trajectory that may be interpreted as the performance path. Chew & François [4] developed a visual environment in which tonal information from musical performances are mapped, in real time, to a three-dimensional representation of tonal space. Their MuSA.RT analysis and visualization system also portrays musical memory as a trajectory that touches on the recently visited tonal regions. Our current approach can be considered a 2D counterpart of this work, with the difference that it shows not only the keys as they unfold, it also portrays the cumulative key information as dynamically varying spatial distributions of colored discs.

3 SYSTEM DESCRIPTION

This section describes the components of our music visualization method, which displays the progression of the tonal content of a music piece. We begin by slicing a piece of music into segments of uniform time length, and determining the key for each segment using a key-finding algorithm. We then map the sequence of keys onto a 2D space that contains points representing all possible keys.

3.1 Segmentation

We begin by segmenting each piece into a given number of segments, m, of uniform length. The value m is selected by the user by moving the slider on the left-hand-side of the display panel. It controls the level of detail, and degree of stability, of the visualizations. As m increases, so does the level of granularity of the information displayed. Note that there are alternate methods of segmentation, including natural and sliding window methods, that may be considered in future work.

3.2 Key Determination

Once a piece is segmented, the key of each segment must be determined. While any key-finding algorithm may be invoked to identify the keys (see [16] for references to key-finding algorithms), we utilize the Spiral Array Center of Effect Generator (CEG) algorithm [2, 1]. The Spiral Array is a geometric model for tonality that represents tonal elements, such as pitch classes and keys, using a set of nested helixes. The collection of pitches of a piece are mapped to their corresponding positions in the pitch class spiral using a pitch spelling algorithm [3]. An aggregate position of these positions is obtained by weighting each pitch class representation by its proportional duration in the segment. The key is then determined through a nearest neighbor search for the nearest key presentation on the major and minor key helixes. This key finding algorithm can be used for both MIDI and audio input [16, 1, 6]. Even though we illustrate the visualization using MIDI data, the method extends to audio music visualization as well.

3.3 Tonal Pitch Space

In music theory, *pitch spaces* model relationships between pitches based on the degree of relatedness among them, with closely related pitches placed near one another, and less closely related pitches placed farther apart. Models of pitch space may be in the form of graphs, groups, lattices, or geometrical figures such as helixes. For our visualization method, we use Lerdahl's 2D representation of major and minor keys in his Tonal Pitch Space [12].

Refer to Table 1 for a depiction of Lerdahl's key space; major keys are notated in capital letters while minor keys are not. In this arrangement of keys, the circle of fifths is placed on the horizontal axis while relative and parallel major/minor relationships alternate along the vertical axis. Recall that the *circle of fifths* depicts relationships among the 12 pitch classes comprising the scale. The *relative minor* of a particular major key (or the *relative major* of a minor key) is the key which has the same key signature but a different tonic. The *parallel minor* of a particular major key (or the *parallel major* of a minor key) is the minor key with the same tonic. The *tonic* is the first note of a musical scale. Note that the Tonal Pitch Space may be extended infinitely as we cycle through all keys.

d♯	g♯	c♯	f♯	b	e	a
F♯	B	E	A	D	G	C
f♯	b	e	a	d	g	c
A	D	G	C	F	B♭	E♭
a	d	g	c	f	b♭	e♭
C	F	B♭	E♭	A♭	D♭	G♭
c	f	b♭	e♭	a♭	d♭	g♭

Table 1. Tonal Pitch Space

3.4 Color Selection

Every possible key is assigned a different color for visualization. The circle of fifths and the color wheel are merged to determine the color assignments. Figure 2 depicts the circle of fifths with each key assigned to a color from the color wheel. Keys on the outer ring represent major keys while keys on the inner ring represent minor keys. The main idea of this color assignment is to have keys that are considered to be close one to another be assigned colors that are also related. For example, C Major and A Minor are assigned a dark and light green respectively.

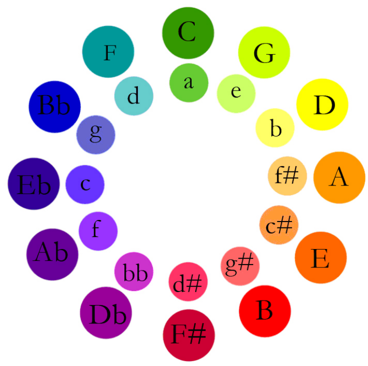

Figure 2. Color Assignments for Major and Minor Keys

3.5 Animation

This section outlines the way the animated visualization looks and progresses. The background of the visualization contains points that represent the keys in the Tonal Pitch Space. Each point is a different color according to the coloring scheme outlined above. The visualization is synchronized with the music. As a music piece progresses, the disc over the key of the present segment grows by one unit, indicating the key of that segment, and the cumulative information of the key distribution. Each time a key is re-visited, the disc over that point grows. At the end of the piece, the visualization displays a 2D version of the distribution of keys for the piece, with the size of discs representing the frequency of the keys.

3.6 User Interface

The visualization method outlined above has been implemented in an intuitive user interface to promote ease-of-use and to encourage the process of exploration and discovery. Refer to Figure 1 for a snapshot of the interface. The user can select to view the visualization synchronized with the music, or without music replay, and a set delay between each frame. The user may also select the piece to visualize by clicking on the desired piece in the menu. The last parameter controlled by the user is the segmentation size, selected by moving the slider, the value of which ranges from 5 to 60. The user may obtain any key name by placing the mouse over a point on the grid of keys.

3.7 Example

Consider the first variation of Beethoven's *32 Variations in C Minor* (WoO80). Refer to Figure 3 for a frame-by-frame illustration of the visualization of this piece. The segmentation parameter, m, was chosen to be 8, the number of bars in the piece. The sequence of identified keys for the slices is as follows: C Minor, F Major, C Minor, C Major, C Minor, C Minor, F Minor, C Minor. Each frame shows the up-to-date analysis of each slice. In each frame, the disc corresponding to the key of the current segment grows in size. For example, we know from the visualization that the piece begins and ends in the key of the piece (C Minor) because, in both the first and last frame, the disc corresponding to the C Minor point grows in size. Additionally, recall that the Tonal Pitch Space has each key repeated such that the window on the grid dictates which keys will be shown multiple times. In this particular example, there are no repeats because of the relatively small size of each frame. In contrast, there are many repeated keys (and key distribution patterns) in Figure 1.

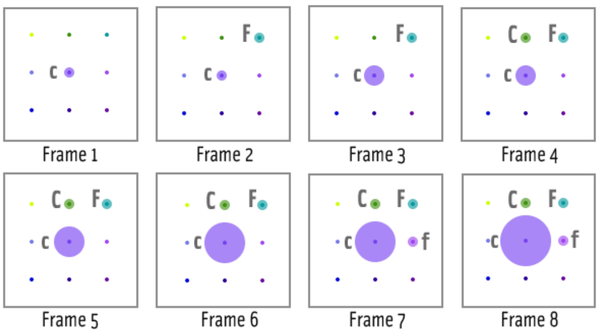

Figure 3. Frame-by-Frame Visualization of Beethoven's WoO80 First Variation

4 VALIDATION

This section presents a formal validation of our visualization method. If a music visualization method aims to go beyond being simply aesthetically pleasing, and strives to transform music into a visual medium, then it must share certain important characteristics with the music. We test whether our proposed visualization method is in fact a good mapping of music onto a visual space by considering its invariance under the transformations outlined by Dorrell in [9], namely, pitch and octave translation, time and amplitude scaling, and time translation. These are the types of changes in music that do not influence human perception in the recognition of a piece. For this analysis we consider the theme of Mozart's *Ah, Vous Dirai-je, Maman* (K265). The piece is segmented into 9 slices for the visualizations; Figure 4 shows the last visualization frame.

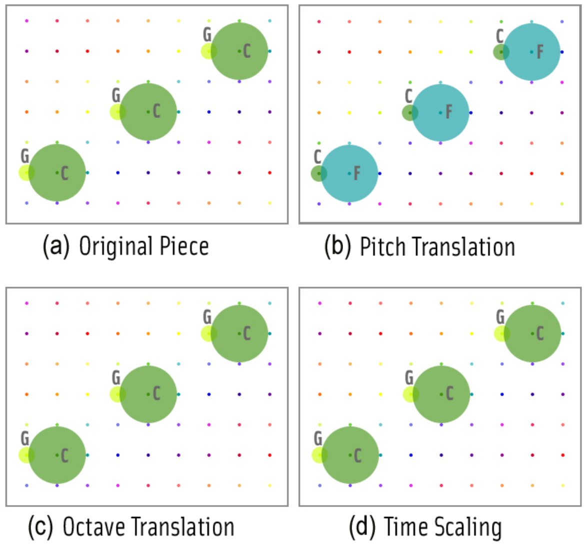

Figure 4. Last Frame of Visualization of Mozart's K265 Theme - Original Piece and Alterations

4.1 Pitch Translation Invariance

Pitch Translation transposes a piece into a different key. Transposition does not alter the musical quality of a piece in any significant way. In fact, we do not normally consider a piece transposed into a different key as being a different piece. The patterns revealed by our visualization method remain intact, and are simply shifted over to the area of the new key. Consider again the example of Mozart's K265 theme which is originally in the key of C Major. We transposed it to the key of F Major. Refer to Figures 4(a) and 4(b) for the last frame of the visualization of the original and transposed piece respectively. In contrast, a visualization method that uses only color and not spatial position to label a key would result in less similarity between the original and transposed pieces.

4.2 Octave Translation Invariance

Octave Translation refers to the transposition of a piece into a different octave. It does not alter the quality of the music either, and could be considered a special type of pitch transposition. Refer to Figure 4(c) for the last frame of the visualization of the example piece transposed down one octave. Notice that since the points representing the keys on the Tonal Pitch Space do not distinguish between octaves, the visualization is identical to the original. Octave translation bears different similarities to the original than other transpositions. This is reflected in the visualization, where octave translation has no effect while other transpositions are indicated by a spatial translation.

4.3 Time Scaling Invariance

Time Scaling refers to the changing of the tempo. If a piece is played faster or slower, we recognize it as being the same piece. This is translated into the visualization in Figure 4(d), which shows a time-scaled version of Mozart's K265. We sped up the original piece by doubling its tempo. Since each piece is segmented into an equal number of segments, time-scaling has no effect on the visualization. For both the original and fast version, each segment has the exact same content.

4.4 Amplitude Scaling Invariance

Amplitude Scaling refers to changing the volume of a piece. This simply states that turning the volume up or down does not change the music. This could however have an effect on certain computation methods. Because our visualization method is based on tonal features, the amplitude has no effect.

4.5 Time Translation Invariance

Time Translation refers to the time at which a piece is played. This is perhaps the most obvious of all the symmetries. A piece is exactly the same if it is played now, in five minutes, or in a year. Our visualization will also look the same for the same piece no matter when it is invoked.

5 DEMONSTRATIONS

We now demonstrate the functionality of our visualization method with several examples. The ability to see the high level tonal progression of a piece over time, and its usage of different tonalities, could provide insight into the deep structures and nature of individual pieces, as well as different genres of music.

5.1 Classical Music

Classical and popular western music have a common structure that we have come to expect. In general, classical pieces begin in the key of the piece, then travel through the terrain of various other keys, and ultimately return to the original key at the end of the piece. These pieces can be thought of as having a center 'star' around which the piece revolves even though there is variation in how far a piece will stray from this center, and how often it will return to visit it through the course of the piece.

As an example, reconsider the visualization of the first variation of Beethoven's *Variations in C Minor* (WoO80) shown in Figure 3. Notice, in the first frame, that the piece begins in C Minor (the key of the piece). The key then travels to F Major, revisits C Minor, travels to C Major, revisits C Minor again, travels to F Minor, and finally returns to C Minor in the last frame.

5.2 Armenian Music

In contrast to the general visual sequence and patterns laid out by classical music, Armenian traditional music generates a different pattern. Instead of having a center of interest, the visualization tool reveals a sequential pattern of key progression that does not return to the original key. Typically, a piece begins in and stays in one key for a period of time, and then moves to a neighboring key. The piece typically does not end in the key in which it began. We provide two examples of this tonal behavior.

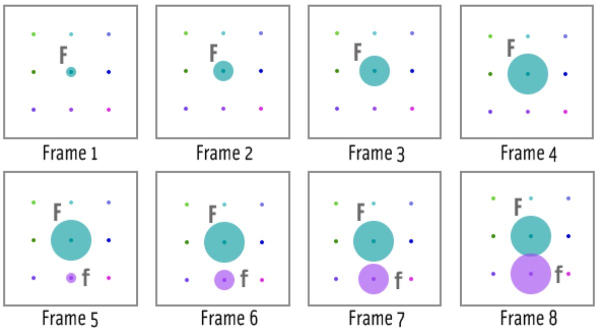

Figure 5. Frame-by-Frame Visualization of Armenian folk song 'Amber Goran'

Consider the Armenian folk song entitled 'Amber Goran' ('Lost Clouds'). Refer to Figure 5 for a frame-by-frame view of the visualization of this piece with $m = 8$. Notice how the piece begins in F Major and remains there from frames 1 through 4, and then travels to F Minor, and stays there for the remainder frames.

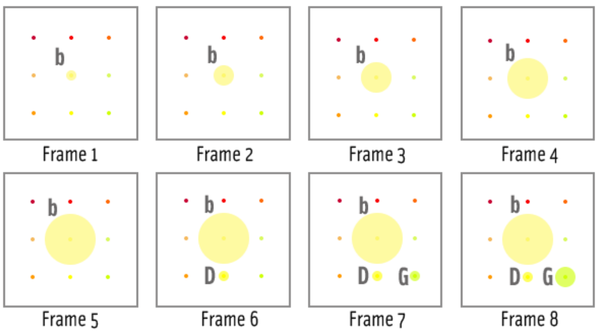

Figure 6. Frame-by-Frame Visualization of Armenian dance song 'Mbarerk'

Now consider a traditional Armenian dance song entitled 'Mbarerk' ('Dances'). The visualization of this piece ($m = 8$) is shown in Figure 6. The piece stays in B Minor for frames 1 to 5, then travels to D Major for frame 6, and ends by traveling to G Major for frames 7 and 8.

6 CONCLUSION

Music has movement that is intertwined with the component of time. The visualization method we have developed strives to exhibit this same characteristic. Our method segments a piece of music into a given number of slices, determines the keys for each slice, and dynamically displays the keys on a 2D space. It also uses color to identify keys, thus adding further richness to the visualization.

The advantage of the proposed method is that it not only indicates the tonality of each segment as it occurs, it also shows the cumulative distribution of tonalities used thus far. The spatial arrangement and the animated progression capture contextual change in a piece as well as the relative contextual change. We have shown how this method maintains standards of information design, and is a successful translation of music onto a visual space. We have also given examples of how the visualizations may be used for the comparison of pieces, for example, of differing genres.

7 ACKNOWLEDGEMENTS

This material is based upon work supported by a University of Southern California (USC) Digital Dissertation Fellowship, and by the National Science Foundation (NSF) under grant No. 0347988. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors, and do not necessarily reflect the views of the USC Graduate School or NSF.

8 REFERENCES

[1] Chew, E. "Modeling Tonality: Applications to Music Cognition", *Proceedings of the 23rd Annual Meeting of the Cognitive Science Society*, 2001.

[2] Chew, E. *Towards a Mathematical Model of Tonality*. Ph.D. thesis, Massachusetts Institute of Technology, 2000.

[3] Chew, E. and Chen, Y.-C. "Real-Time Pitch Spelling Using the Spiral Array", *Computer Music Journal*, 2005.

[4] Chew, E. and François, A.R.J. "Interactive Multi-Scale Visualizations of Tonal Evolution in MuSA.RT Opus 2", Newton Lee (ed.): Special Issue on Music Visualization and Education, *ACM Computers in Entertainment*, 2005.

[5] Chouvel, J.-M. "Représentation harmonique hexagonal toroïde", *Musimédiane – revue audiovisuelle et multimédia d'analyse musicale*, [Web site] December 2005, URL: `www.musimediane.com`

[6] Chuan, C.-H. and Chew, E. "Fuzzy Reasoning in Pitch Class Determination for Polyphonic Audio Key Finding", *Proceedings of the 6th International Conference for Music Information Retrieval*, 2005.

[7] Cohn, R. "Neo-Riemannian Operations, Parsimonious Trichords, and Their 'Tonnetz' Representations", *Journal of Music Theory*, 1997.

[8] Cooper, M. and Foote, J. "Visualizing Musical Structure and Rhythm via Self-Similarity", *Proceedings of the International Conference on Computer Music*, 2001.

[9] Dorrell, P. *What is Music? Solving a Scientific Mystery*. Phillip Dorrell, 2005.

[10] Gomez, E. and Bonada, J. "Tonality Visualization of Polyphonic Audio", *Proceedings of International Computer Music Conference*, 2005.

[11] Langer, J. and Goebl, W. "Visualizing Expressive Performance in Tempo-Loudness Space", *Computer Music Journal*, 2003.

[12] Lerdahl, F. *Tonal Pitch Space*. Oxford University Press, 2001.

[13] "Music Animation Machine", [Web site] 2007, URL: `www.musanim.com`

[14] Mardirossian, A. and Chew, E. "Key Distributions as Musical Fingerprints for Similarity Assessment", *Proceedings of the 1st IEEE International Workshop on Multimedia Information Processing and Retrieval*, 2005.

[15] Mardirossian, A. and Chew, E. "Music Summarization Via Key Distributions: Analyses of Similarity Assessment Across Variations", *Proceedings of the 7th International Conference on Music Information Retrieval*, 2006.

[16] "1st Annual Music Information Retrieval Evaluation eXchange", [Web site] 2006, URL: `www.music-ir.org/mirex2005`

[17] Misra, A., Wang, G. and Cook, P. R. "sndtools: Real-Time Audio DSP and 3D Visualization", *Proceedings of the International Computer Music Conference*, 2005.

[18] Sapp, C. "Harmonic Visualizations of Tonal Music", *Proceedings of the International Computer Music Conference*, 2001.

[19] "The Shape of Song", [Web site] 2007, URL: `www.turbulence.org/Works/song`

[20] Toiviainen, P. and Krumhansl, C.L. "Measuring and Modeling Real-Time Responses to Music: The Dynamics of Tonality Induction", *Perception*, 2003.

[21] Tufte, E. *Envisioning Information*. Graphics Press, Cheshire, CT, 1990.