# A Content Dependent Visualization System for Symbolic Representation of Piano Stream

# A Content Dependent Visualization System for Symbolic Representation of Piano Stream*

Alexander Adli[1] & Zensho Nakao[2]

[1] Graduate School of Engineering & Science, University of the Ryukyus, Okinawa 903-0213, Japan.
[2] Department of Electrical & Electronics Engineering, University of the Ryukyus, Okinawa 903-0213, Japan.
{alex, nakao}@augusta.eee.u-ryukyu.ac.jp

## Abstract

This paper provides an overview on the advances of music information retrieval in symbolic representation of music. Such musical aspects as key, tonality, bass, melody, dynamics, rhythm and patterns are considered as foundation for visualizing systems for the piano stream in classical music. The paper then describes the proposed visualizing system and Malinowski's music animation machine. It lays light on the challenges facing creating contemporary visualizing systems. It is supplied with a related references list for further study on the issue.

**Key Words**: Music content analysis, visualization, MIDI, piano stream, MIR.

## 1 Introduction

Most contemporary popular music player software (Windows Media Player, Winamp, iTune, etc.) contain a visualization module which provides a real time graphical complement for the performed digital music. These visualization modules rely mainly on amplitude and frequency information available in the audio files and use it as source information for the graphical module which usually performs complicated actions and creates complex patterns. In spite of this graphical complexity, it does not reflect the semantic of the music, and does not consider such musical concepts like tonality, bass and melody. A music information retrieval (MIR) system in audio files for those aspects is computationally expensive and still far from being able to satisfy the users (listeners/watchers).

The MIR community applies a lot of effort on analysis the musical contents both in digital audio and in symbolic representations. The symbolic representation like the MIDI files differs from the former in terms that it explicitly provides the information about the pedals, notes, and their timing. Different musical instruments are also divided into channels or tracks [9].

These characteristics of MIDI files makes them more suitable domain for content analysis than digital audio files and as a result visualizing systems for MIDI at the moment can be potentially more related to the musical content than their digital audio analogues [1]. Although MIDI files are popular, they are mostly used for composing aims or in musical databases but rarely among end-users for listening purposes. This is because MIDI requires high quality synthesizers which are too expensive or simply not good enough as their acoustic instruments equivalent.

As the quality of software synthesizers and the computing power of computers are rapidly improving, MIDI music listener auditory is expected to increase. A visualizing module for MIDI files based on musical content analysis might become am interesting contribution to the music industry.

In this paper the authors shed light on some of the music content analysis methods and techniques for MIDI; we will discuss such musical concepts such as key, tonality, melody and dynamics and their retrieval from MIDI data. In latter sections authors show how information about musical content can benefit visualizing systems in two examples: the authors' MIDI visualizer, and Malinowski Music Animation Machine (MAM). This is followed by conclusions and further research discussion, acknowledgment, and references to related works to both of visualizing music and content analysis.

* Part of the work was presented at the 11th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems KES 2007, Salerno, Italy, September 2007.

## 2 Music Content Analysis in MIDI Files

The MIDI representation of music contains symbolic information of performance events: pedals and notes (onset time, pitch, duration). In this sense it can be compared to musical score and hence, it can be analyzed by tools and rules of theory of music. There are however many differences between the musical score and MIDI; the key is not necessarily known in MIDI representation, also notes like C# and D♭ are indistinguishable. MIDI is a performance descriptor and not musical score for the musicians; this is why it is potential musical source for the end-user. We mentioned previously in the introduction the reasons why it is still not popular for this purpose.

A visualization system which can reflect graphically much of the musical contents in a way which makes sense for the listener/watcher must be able to extract musical aspects like key, tonality, bass, melody line, dynamics, rhythm, patterns, form, genre, emotions, etc. Behind every aspect there is at least few a algorithms and competitive methods for better extracting.

The authors will discuss the first seven aspects mentioned here and then give examples of implementing them into a visualizing system. They are discussed from the point of view of classical genre and single instrument – the piano.

### 2.1 Key and Tonality

Almost all classical tonal pieces are composed using one of the existed 24 tonalities of the diatonic music. This tonality is called the key. Tonality is one of the main musical attributes along with the meter, melody and harmony. The theory of music defines tonality as the hierarchy of importance that exists between tones in musical piece [12].

Tonality dynamically changes via progressions and modulations while the key rarely changes. The key mode (major or minor) reflects the overall "mood" of the piece. Parncutt [16] created a method for key detection in a sequence of chords. Later many techniques appeared for key detection by analyzing the composition beforehand and determining the key as the most frequent tonality with its tonic sound frequently appearing. The authors of this paper suggested a method for tonality recognition dynamically. This method is designed for monophonic piano stream in MIDI file. It divides the stream basing mainly on pitch salience values into segments where every segment represents a single tonality [8].

MIDI files contain a meta event command where the key information (meta command 59) can be stored. In case this information is available, visualizer can use it, and apply the tonality methods to detect tonality progressions and modulations. If not, then it has to apply some actions to determine the key. Fig.1 shows the tonalities in the famous Beethoven's "For Elise".

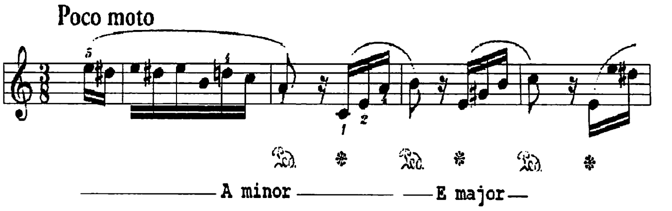


**Fig. 1.** Beethoven: *For Elise*. The piece is composed in A minor (the key). In this fragment the composition starts with its key tonality A minor, and then progresses to E major.

### 2.2 Bass and Melody Line

Basses are the foundation of the harmony, and usually are the lowest pitch in the harmonic line. Melody is the main part of the music, and it is the part people would sing if they were asked to express the music vocally. Most bass recognition algorithms are based on the assumption that the bass note is lowest than its neighbors. The authors in their work about tonality [8] also suggested an algorithm for bass recognition in piano stream based on the same principle.

Melody extraction is a difficult task. Uitdenbogerd and Zobel prpopsed the skyline technique [13]; it selects the highest pitch line of the note sequence into melody. However, even with later revisions of the technique such like in [14], it still fails to recognize complicated structures of melody or simply when the melody is in the lowest or middle pitches. Chopin's fragment in Fig. 2 shows the melody line where notes are circled. Skyline algorithm successfully classifies the higher notes as melodic. The lower notes "E" and "B" shown by arrows are basses and each one represents a harmony of E major and B major tonalities respectively.

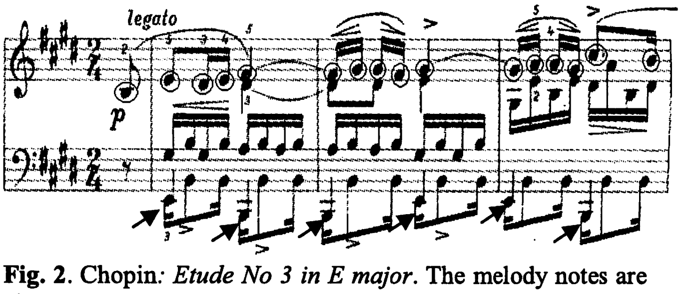


**Fig. 2.** Chopin*: Etude No 3 in E major*. The melody notes are circled. Arrowed notes are basses.

Another techniques are based on choosing the best MIDI channel which contains the melody [2][13]. Usually the melody channel has louder level than the average of all channels, less average distance between pitches, less rhythmic and so on. However, this technique is not much useful in classical piano music, as it is seldom in such music to divide into melody and harmony channels especially when both are played on the piano.

Recently a psychoacoustics models for melody extraction are in discussion.

### 2.3 Dynamics, Rhythm and Patterns

Although MIDI files are considered as a relatively convenient domain for the musical contents analysis, the loudness aspects of music or dynamics are relatively difficult to extract. In digital sound there are good deterministic methods for loudness extraction from sound wave amplitude [15]. In MIDI files, the loudness is reflected in the velocity parameter of the note which is related to the speed of the piano key pressing. The authors studied the factors affecting the loudness in piano [7][8]. The loudness can be estimated from notes velocities, pitch and piano amplitude envelope. The loudness then reflects the dynamics levels of the music (forte, piano, crescendo and diminuendo).

Rhythm is the variation of length of the notes. There are many two more relative aspects here: The meter, and the tempo. The rhythm is important as it reflects such features of music like accelerando and ritenuta. In 1983 Lerdahl and Jackendoff published their text, *A Generative Theory of Tonal Music* (GTTM). Since that time many methods based of music segmentation appeared for rhythm extraction.

The music segmentation however, depends on another feature of the music – the patterns. The pattern recognition in music leads to segmentations. In Fig. 3 the musical fragment can be classified into three parts with single pattern. The first part consists of 9 notes in C major starting in C which is the bass note here. The second and third repetitions repeat the note intervals of the first part but in different tonalities (G major and F major). Algorithms based on segmentation division by note interval repetition can extract the rhythm of the fragment in Fig. 3 successfully. However, in classical music segmentation and pattern recognition is more complicated process. Robert Rowe in his book *machine musicianship* [17] describes advanced methods formusic segmentation based on neural networks techniques.

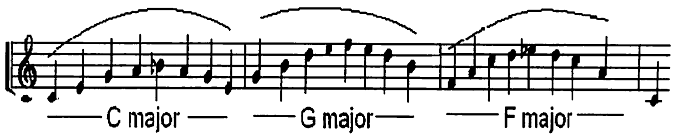


**Fig. 3.** Musical pattern been repeated in different tonalities

## 3 MIDI Visualizing systems

Visualizing music is a creative process. There might be infinite ideas of representing the contents of music via graphics. As we mentioned before, most of the popular contemporary players depends on the sound wave but not on the musical context. In this section we will describe two visualizers for MIDI files based on musical content (The authors visualizer and Malinowski Visualizer). We hope that describing the visualizing ideas here might inspire many graphic and music researchers and programmers to create their own original creative visualizing systems.

### 3.1 Author's MIDI Visualizer

This visualizing system was created cooperatively by Adli and Ermakov. The system can receive MIDI messages from MIDI files as well as directly from MIDI device like digital piano.

The visualizing idea is a space flight in real time with the music. Every note causes a small star to appear (fig. 4.) exactly in the onset time of the note. The small star starts move toward the user (in a 3D space) which makes the user feels as if he/she is flying forward. Bass notes cause bigger rotating stars which move exactly like the previous ones. The color of the small stars reflects the current tonality.

Minor tonalities are associated with dark and gloomy colors while the major ones are associated with bright and cheerful colors. The color and rotation speed of the bass stars depends on the pitch of the bass. The lower the bass the slower the rotation speed is. Chords when happen cause an explosion. The color of this explosion depends on tonality while the size depends on the overall loudness of all the notes of the chord. The dynamics is represented by the size of the bass stars; for quite music bass stars are small and vice versa. Rhythm is reflected by the speed of the flight. And when the system knows the key tonality it adds a trembling feature for all the objects whenever the dominant tonality appears to reflect the "anxious" feeling frequently provoked by this tonality.

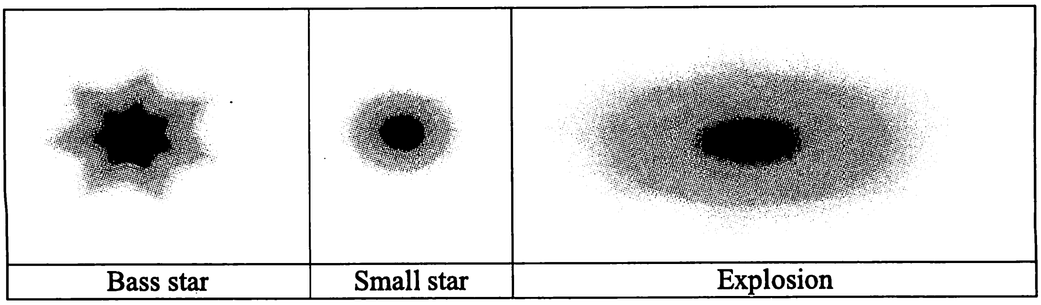


**Fig.4.** The shapes of MIDI visualizer. As the original background is black, all colors are inverted.

This visualizer is a real time one. And although it uses simplified algorithms for the content analysis, non-expert users noticed that the visualization performed by our visualizer were "closer" to the musical context than other popular visualization modules which were basing on frequency and amplitude extraction. The graphical and musical qualities however, were not necessarily better.

### 3.2 Malinowski Music Animation Machine (MAM)

Stephen Malinowski during many years created his own algorithms and own visulalizer which he called MAM [4]. MAM handles MIDI messages from file or directly via MIDI port. As for 2007 this MAM can visualize MIDI in 12 different ways. Below we will describe one of them.

Every note is represented by a circle (fig. 5). The "now line" divides the notes into two groups: the notes in the past are to the left of the line, and the future notes are to the right. The now line advances and the notes are played once the now line crosses the center of their circles. As X axis is the time axis, Y axis reflects the note number. The lower notes are located at the bottom of the screen at vice versa. The color of the circles reflects the pitch class of the note while the size reflects its duration. Once the now line "hits" the circle, its inner part starts following the line and decreases in size gradually until it disappears (both graphically and as a sound). The melody line and harmony line are also separated.

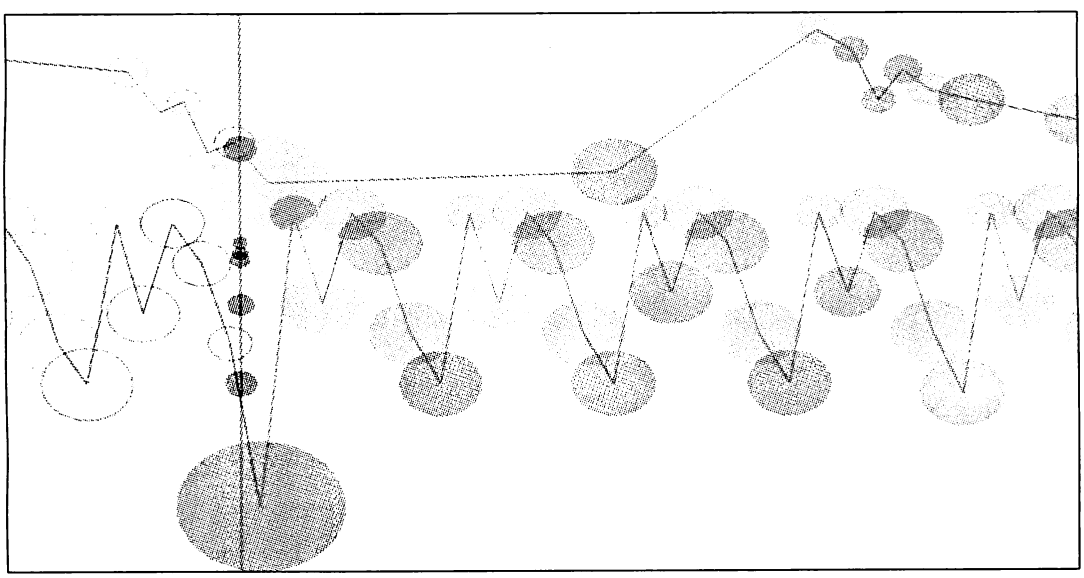

**Fig. 5.** The so called "part motion" visualization of Malinowski's MAM. It might be difficult to distinguish the different colors of the circles in grayscale.

### 3.3 Challenges in Visualizing systems

Real time music information retrieval system in MIDI is frequently late in detecting musical aspects exactly in time of their occurrence. However in most cases music is beforehand stored in a MIDI file, and the visualizing system can first analyze the music in advance.

As the MIDI files are not digital audio files, they rarely can produce high quality sounds on average platforms. There are however, many hybrid musical standards where music is stored in both audio and symbolic way. Such files can serve as good domain where the music is performed from the audio data, and the visualization uses the symbolic data.

Visualizing system can be used for other purposes like visual music representation, pattern recognition for performers or genres. For example the performance worm by Goebl et. al. [6] draws different shape "worms" in tempo-loudness space which helps to distinguish performers.

Visualizing systems are encouraged to create dynamically their own graphical patterns in real time relevant to the performed music.

## 4 Conclusions and Further work

MIDI files are expected to become more popular among the end users due to the improvements of both software and hardware synthesizers. Symbolic analysis of music in MIDI can provide more accurate information about the musical aspects like tonality, melody line and rhythm with less computational requirements if compared with the analysis in audio files.

Music visualizing systems are also becoming popular. The created visualization might be an aim by itself, or it can be used for different music classification applications. Visualizations in audio files are limited to amplitude and frequency information. There are however, many graphically advanced visualizators such as some Winamp plug-ins. Visualizations based on MIDI files can create visual patterns closely related to musical context as the musical information can be explicitly or implicitly obtained. The "MIDI visualizer" and "Music Animation Machine" are examples of such visualizing systems.

The authors wanted to invoke the MIR community creativity toward the potential visualizing possibility of music based on its contents.

### Acknowledgments.

This research was supported in part by Ministry of Internal Affairs and Communications (Japan) under Grant: SCOPE 072311002, for which the authors are grateful.

The authors of this paper would like to express their gratitude to A. Ermakov who made a lot of efforts to make our MIDI visualizer look more beautiful.

## References

[1] E. Isaacson.: What You See Is What You Get: On Visualizing Music, *Proceedings of International Symposium in Music Information Retrieval*, London, 2005.
[2] G. Ozcan, C. Isikhan, A. Alpkocak,: Melody Extraction on MIDI music Files, *Proceedings of the seventh IEEE International Symposium on Multimedia*, 2005.
[3] T. Li, M. Ogihara: Detecting Emotions in Music, *Proceedings of International Symposium in Music Information Retrieval*, 2003.
[4] S. Malinowsk: Music Animation Machine, 2005, URL http://www.well.com/user/smalin/ mam.html.
[5] Alexander Adli, Zensho Nakao, Yasunori Nagata: Calculating the expected sound intensity level for solo piano sound in MIDI files. *Proceedings of the SCIS & ISIS*, Tokyo Institute of Technology, Tokyo, September 2006.
[6] J. Langner, W. Goebl: Visualizing Expressive Performance in Tempo-Loudness Space, *Computer Music Journal*, Vol.27 No.4, Winter 2003, pp.69-83.
[7] Alexander Adli, Zensho Nakao, Toshiaki Yokoda, Yasunori Nagata: Piano Sound Characteristics: A Study On Some Factors Affecting Loudness In Digital And Acoustic Pianos, ICICIC, Japan, 2007, In Print.
[8] Alexander Adli, Zensho Nakao, Yasunori Nagata: A Score Based Algorithm for Tonality Recognition in Monophonic Piano Streams, Submitted.
[9] MIDI Manufacturers Association: www.midi.org
[10] K. Martin, E. Scheirer, B. Vercoe,: Music Content Analysis Through Models of Audition, *Proceedings of the ACM Multimedia Workshop on Content Processing of Music for Multimedia Applications*, Bristol, UK, 1998.
[11] D. Rizo, P. Ponce de Le´on, C. P´erez-Sancho, A. Pertusa, J. Inesta: A Pattern Recognition Approach for Melody Track Selection in MIDI Files, *Proceedings of International Symposium in Music Information Retrieval*, 2006
[12] D. Levitin: *This is your brain on Music*, Dutton, 2006.
[13] A. Uitdenbogerd, J. Zobel.: Manipulation of Music for Melody Matching", *Proceedings of ACM Internaional Multimedia Conference*", 1998.
[14] W. Chai: Melody Retrieval on the web, MS Thesis, Massachusetts Institute of Technology, Boston, 2000.
[15] H. Fastl, E. Zwicker: *Psychoacoustics : Facts and Models*, Springer, 3rd edition, 2006.
[16] Richard Parncutt: Revision of Terhardt 's psychoacoustical model for the roots of a musical chord, *Music perception 6*, 1988, pp. 65 – 94.
[17] Robert Rowe.: *Machine Musicianship*, The Mit Press, 2001.