MuSA: Music Information Processing

1. Research Team

Project Leader: Prof. Elaine Chew, *Industrial and Systems Engineering*

Graduate Students: Yun-Ching Chen

Undergraduate Students: Joann Emmanuel, Aaron Yang

2. Statement of Project Goals

The goal of MuSA is to create a robust and reliable system for recognizing and visualizing tonal patterns using the Spiral Array model for tonality [1]. This endeavor requires the developing of algorithms for abstracting and extracting pitch/time structures (see Figure 1) in music information. The resulting algorithms are fundamental tools for automated transcription, computer music analysis, composition, segmentation, synchronization and retrieval, and expression synthesis.

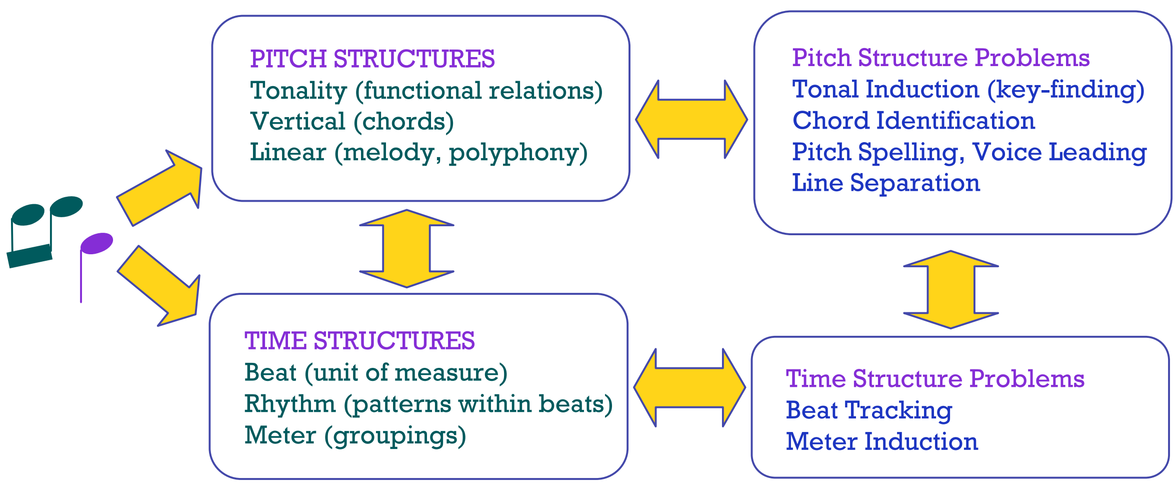

Figure 1: Abstracting pitch and time structures from music information.

3. Project Role in Support of IMSC Strategic Plan

The MuSA project impacts three of IMSC's vision areas: Entertainment, Communication and Education. One of the roles of musical expression, experience and creation is entertainment. It is an undeniable fact that, for many of us, music is a source of pleasure and amusement. Music has also been described as "the art of arranging sounds in time so as to produce a continuous, unified, and evocative composition, as through melody, harmony, rhythm, and timbre" (American Heritage Dictionary). This creative act can be better understood through systematic study using the computational resources available to us today. The underlying model in MuSA is the 3D representation for tonality called the Spiral Array [1]. The Spiral Array offers multiple ways to visualize music streams as trajectories or transformations in 3D space; and provides an effective tool for seeing what we hear when we listen to music. Apart from Information Management, the developing and results of MuSA are related to the User-Centered Sciences, Sensory Interfaces and Media Communication.

4. Discussion of Methodology Used

The MuSA application uses the Spiral Array model for tonality first proposed in the PI's dissertation [1]. The advantages of the Spiral Array model include: 1) entities from different hierarchical levels are represented in the same space; 2) musical knowledge embedded through distance – clustering of closely-related elements; 3) summarizing of musical information; and 4) provides a distance metric for comparing musical content.

The Spiral Array model has been shown to perform better at key finding than existing methods [5], and to provide a computationally viable method for detecting key boundaries [4]. In the CEG algorithm described in [5], each cumulative chunk of music is mapped to a point in the Spiral Array space, forming a trajectory that leads to the appropriate key region. The most likely key is found through a nearest-neighbor search.

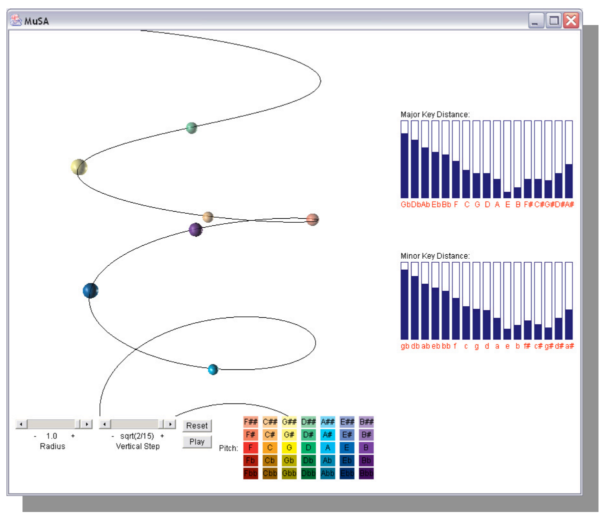

Figure 2: MuSA: music visualization using the Spiral Array

In MuSA, the CEG algorithm was ported to Java, and concurrent visualization of the computational analysis is provided using Java 3D (see Figure 2). The distance to each key is depicted on two bar charts, for all major and minor keys respectively. The notes (pitch events) are mapped to their positions on the pitch spiral, and their cumulative durations represented by the radii of the spheres. The orientation and aspect ratio of the spiral can be adjusted in real-time for best viewing.

5. Short Description of Achievements in Previous Years

MuSA is a young project that began in January 2002. All prior achievements are documented in the next section, 5a.

5a. Detail of Accomplishments During the Past Year

Implementation of the algorithms has led to new research questions and publications.

> Pitch spelling: A fundamental first step in analyzing digital music information such as MIDI is the assignment of the appropriate pitch name to a pitch index number. Each pitch is represented by a number in many digital formats. The same number could map to, say, A# or Bb, where the spelling is dependent on the key context. We proposed and implemented an algorithm for disambiguating pitch spellings. The first stage of testing for the algorithm is reported in [2]. Further tests are underway.

MuSA is has joined IMSC's list of demonstrations.

> MuSA: A Framework for Visualizing Music Content
> Created a 3-D visualization of tonal patterns in music using the Spiral Array model. Spatial proximity indicates degree of relatedness between objects. Each note maps to a sphere in the spiral, the pitch determines the position and the duration the radius of the sphere. The tonal center of a group of notes maps to a point representing the center of effect. The models allows distance-to-key to be charted accurately. (Prof. Elaine Chew, Principal Investigator) Yun-Ching Chen, Student (PHE 330).

A Music Engineering Laboratory (MEL) was established in PHE330. Apart from one PC and one Mac G4, the equipment in MEL include

> Music Engineering Lab Equipment
> Yamaha P80 88-key Digital Piano
> Roland Triton 61-key Music Workstation Sampler
> Genelec 1031A Near Field Producer Monitors (one pair)
> Mackie 24/8 24x8x2 Rec./PA Console
> Mackie MB-24 Meter Bridge for 24/8 Console
> Tascam DA-40 DAT Recorder
> Roland XV-5080 128 Voice Synthesizer
> Yamaha VL70M Virtual Acoustic Tone Generator
> Lexicon MPX 200 2-Ch Effects Processor
> Steinberg MIDex 8 x 8 USB Interface (2)
> PreSonus BlueMax Compressor/Limiter
> AKG K 240 K240M Headphones
> Emagic Logic Audio Platinum Digital Audio/MIDI Software

6. Other Relevant Work Being Conducted and How this Project is Different

The MuSA project uses the Spiral Array model. This successive embedding of spiral structures to represent musical entities is the Principal Investigator's invention [1]. The use of distance metrics generated by the Spiral Array model has been proven to be more efficient and to produce better tonal recognition results than existing models [4,5].

Our goal is to extract and visualize tonal content in music information. In this realm, [11] of Stanford's Center for Computer Assisted Research in the Humanities has proposed color maps for harmonic visualization. These two dimensional color images are in contrast to the three dimensional configuration of the Spiral Array.

In the cognitive science literature, researchers creating mental models of music have proposed spiral representations with alternative pitch arrangements [13]. Most recently, a spiral representation of key relationships like that in the innermost core of the Spiral Array model has been validated through neural probes that correlate brain signals with key modulations [14].

7. Plan for the Next Year

Further explorations are underway for developing new algorithms, and extending MuSA to applications for linking performance and visualization, and for generating musical expression.

7.1 MuSA: Further Explorations

Preliminary studies are underway to: 1) extend the tonal recognition capabilities of MuSA to chord recognition; and 2) to use MuSA to derive metrics for similarity assessment. Both applications are important components to content-based music information retrieval. See, for example, [10].

7.2 MuSA.RT: MuSA Real-Time

MuSA.RT is a research initiative integrating real-time music processing and content-based graphical rendering in interactive immersive environments. This is a collaborative project with Dr. Alexandre François and utilizes the Software Architecture for Immersipresence [8].

MuSA.RT Op.1 is a system for mapping a live MIDI stream (produced by a keyboard) to the Spiral Array (see Figure 3). Interactivity is provided through a gaming device for flying through the Spiral Array space. The algorithms in MuSA are being modified for data streams captured in real-time. The integration of live music, real-time computational analysis and interactive graphical representation is facilitated by the Modular Flow Scheduling Middleware [9].

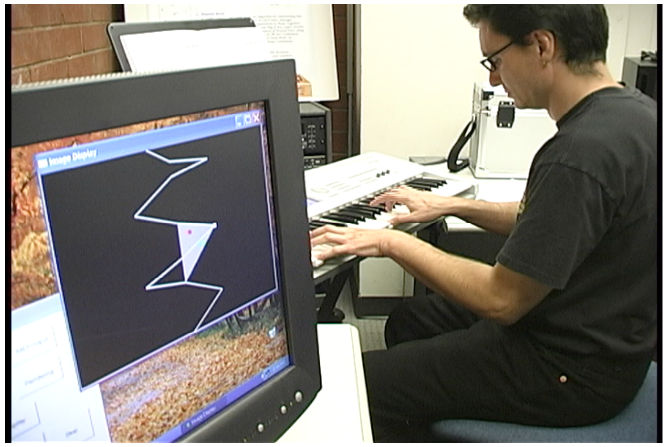

Figure 3: Preliminary rendering in the MuSA.RT Project

7.3 ESP: Expression Synthesis Project

Musical structure is closely tied to musical expression. The generating of convincing expression must relate structure to timing decisions. This new area provides an interface for exploring human emotion, intelligence, creativity and individuality [12]. One future research effort is to create user interfaces for controlling expression in music and to relate these decisions to musical structure.

8. Expected Milestones and Deliverables

MuSA offers a platform on which to develop and improve algorithms for music recognition, analysis and segmentation. The results will be published and disseminated at conferences such as the International Conference on Music and Artificial Intelligence (ICMAI), International Conference on Music Information Retrieval (ISMIR), International Computer Music Conference (ICMC) and Rencon (performance rendering contest). The MuSA software has become and will continue to be part of the IMSC technology demonstrations.

9. Member Company Benefits

N/A

10. References

[1] Chew, Elaine (2000). *Towards a Mathematical Model of Tonality*. Ph.D. dissertation. Operations Research Center, MIT. Cambridge, MA.

[2] Chew, Elaine & Chen, Yun-ching (2003). *Mapping MIDI to the Spiral Array: Disambiguating Pitch Spellings*. In H. K. Bhargava and Nong Ye, eds., Computational Modeling and Problem Solving in the Networked World, Kluwer, pp.259-275. Proceedings of the 8th INFORMS Computer Society Conference, ICS2003, Chandler, AZ, Jan 8-10, 2003.

[3] Chew, Elaine & Alexandre François (2003). *Real-Time Music Information Processing*. In Proceedings of the 31st International Conference for Computers and Industrial Engineering, ICCIE2003, San Francisco, CA, February 2-4, 2003.
[4] Chew, Elaine (2002). *The Spiral Array: An Algorithm for Determining Key Boundaries*. In C. Anagnostopoulou, M. Ferrand, A. Smaill (Eds.): *Music and Artificial Intelligence*. Springer LNCS/LNAI #2445, pp.18-31. Proceedings of the Second International Conference, ICMAI 2002, Edinburgh, Scotland, UK, September 12-14, 2002.
[5] Chew, Elaine (2001). *Modeling Tonality: Applications to Music Cognition*. In Proceedings of the 23rd Annual Meeting of the Cognitive Science Society, CogSci2001, Edinburgh, Scotland, UK, August 1-4, 2001.
[6] Cooper, Matthew and Foote, Jonathan (2002). *Automatic Music Summarization via Similarity Analysis*. In Proceedings of the International Conference on Music Information Retrieval, ISMIR 2002, Paris, France, October 13-17, 2002.
[7] Dixon, Simon, Goebl, Werner, and Widmer, Gerhard (2002) *Real Time Tracking and Visualization of Musical Expression*. In C. Anagnostopoulou, M. Ferrand, A. Smaill (Eds.): *Music and Artificial Intelligence*. Springer LNCS/LNAI #2445, pp.58-68. Proceedings of the Second International Conference, ICMAI 2002, Edinburgh, Scotland, UK, September 12-14, 2002.
[8] François, Alexandre R.J. (2002). *Components for Immersion*. In Proceedings of the IEEE International Conference on Multimedia and Expo.
[9] François, A. R.J. & Medioni, G. G. (2000). *A Modular Middleware Flow Scheduling Framework*. In Proceedings of ACM Multimedia 2000, p. 371-374.
[10] Pickens, Jeremy, Bello, Juan P., Crawford, Tim, Dovey, Matthew, Monti, Giuliano and Sandler, Mark (2002). *Polyphonic Score Retrieval Using Polyphonic Audio Queries: A Harmonic Modeling Approach.* In Proceedings of the International Conference on Music Information Retrieval, ISMIR 2002, Paris, France, October 13-17, 2002.
[11] Sapp, Craig (2001). *Harmonic Visualizations of Tonal Music*. Proceedings of the 2001 International Computer Music Conference, Havana, Cuba, pp. 423-430. Computer Music Association.
[12] Rencon (2002). *http://shouchan.ei.tuat.ac.jp/~rencon/top.html*
[13] Roger N. Shepard. (1982). *Structural representations of musical pitch*. In The Psychology of Music, Deutsch, Diana Eds. pp. 343-390. Academic Press.
[14] Zatorre, Robert J. and Krumhansl, Carol L. (2002) *Mental Models and Musical Minds*. Science Magazine 13 December 2002 issue, Vol 298: pp 2138-2139.