PST: Pitch Spelling Technology

1. Research Team

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

Graduate Students: Yun-Ching Chen

2. Statement of Project Goals

Pitch spelling is the process of assigning contextually correct letter names to numeric representations for musical pitch such as MIDI numbers. Each pitch is represented by a number in MIDI and other digital formats, and the same number, say, 58 could map to A# or Bb depending on the key context. Accurate pitch spelling is a critical component of systems for automated transcription, computer music analysis, segmentation, synchronization and retrieval, and expression synthesis. Our goal is to develop robust pitch spelling algorithms for the purpose of recognizing and visualizing tonal patterns using the Spiral Array model for tonality [6]. The algorithms are implemented in Java as part of the MuSA project initiated in 2002. MuSA stands of Music on the Spiral Array and consists of a set of software tools for music analysis based on the Spiral Array.

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. 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 underlying model in MuSA is the 3D representation for tonality called the Spiral Array [6]. 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. 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].

The Spiral Array represents pitch classes as points on a spiral and summarizes music information using an interior point called the CE (Center of Effect). We use this CE as a proxy for the key context. Contextually consistent pitch names are determined through a nearest neighbor search for the most appropriate pitch class representation in the Spiral Array space (see Figure 1).

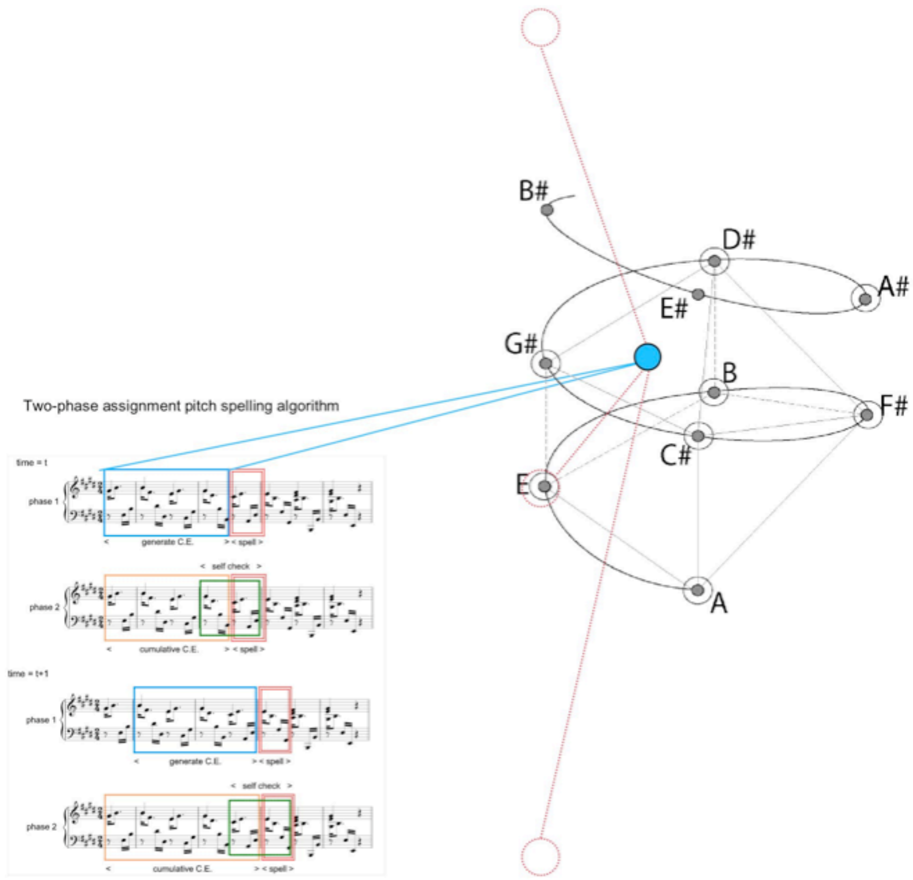
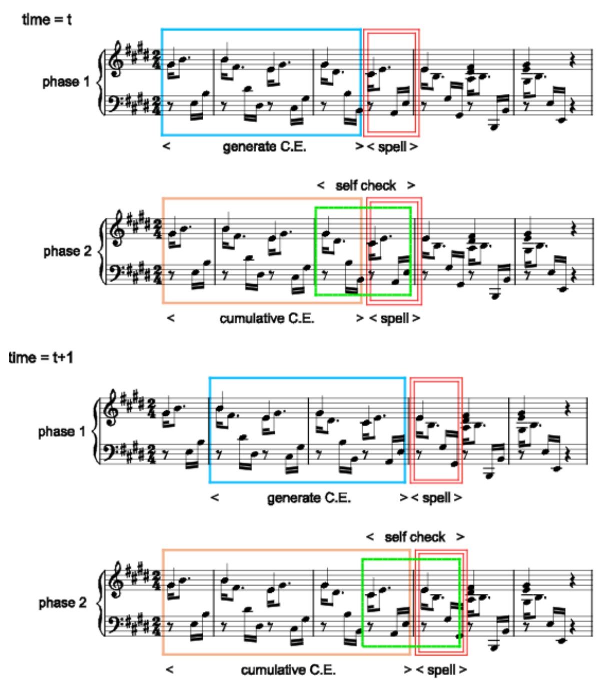


Figure 1: Assigning pitch names using a nearest-neighbor search.

Using various strategies for determining local and global contexts, we proposed and compared algorithms for pitch spelling. Figure 2(a) shows the cumulative window approach, Figure 2(b) the sliding window approach, and Figure 2(c) a two-phase bootstrapping approach. The algorithm depicted in Figure 2(a) and its results are reported in [3]. The computational results of the latter two algorithms are presented in [2]. The bootstrapping approach was shown to be most sensitive to varying tonal contexts and to be able to capture sudden contextual changes.

(a) a cumulative window approach

(b) a sliding window approach



(c) a two-phase bootstrapping approach

Figure 2: Strategies for determining tonal contexts.

5. Short Description of Achievements in Previous Years

MuSA is a young project that began in January 2002. It has grown from a system for visualizing music content to a platform for developing and testing music analysis algorithms.

MuSA: A Framework for Visualizing Music Content
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).

MuSA has joined IMSC's list of demonstrations.

5a. Detail of Accomplishments During the Past Year

Implementation of the algorithms has led to new research questions and publications. The following paper describes the algorithm depicted in Figure 2(a):

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.

The following paper presents computational results for the algorithms in Figure 2(b) and (c):

Chew, Elaine & Chen, Yun-Ching (2003). *Determining Context-Defining Windows: Pitch Spelling using the Spiral Array*. In Proceedings of the 4th International Conference for Music Information Retrieval, ISMIR, Baltimore, MD, Oct 26-30, 2003.

The following paper has been submitted to the *Computer Music Journal* for review:

Chew, Elaine & Chen, Yun-Ching. *Pitch Spelling using the Spiral Array*.

We have also filed an invention disclosure for the pitch spelling algorithms. The USC office of technology and licensing has filed for a provisional patent for our pitch spelling technology:

Invention Disclosure and Provisional Patent
Pitch Spelling Using the Spiral Array
USC Ref. 3532 / MWE Ref. 28080-117
Inventor: Elaine Chew; Co-Inventor: Yun-Ching Chen (MS ISE)

Figure 1 appears as an illustration in a poster for a call for participation for the "Context-Driven Topologies" project led by Deborah MacPherson, an independent museum designer, and Johns Hopkins University's Digital Knowledge Center.

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 [6]. 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]. The model's key representations have also been demonstrated to be neurologically linked to the mental representation of key relationships (see [10]).

Temperley [9] proposed an algorithm for pitch spelling that finds the assigns pitch names through a nearest neighbor search on the line of fifths. Cambouropoulos [1] has conducted numerous experiments in pitch spelling using an algorithm that prioritizes more likely pitch interval structures. Meredith [8] has proposed another algorithm, called ps13, for pitch spelling.

Until now, the challenge in pitch spelling research has been the lack of an appropriate test set with spelling solutions. Meredith has amassed a large-scale database of over 2,000,000 notes and is in the process of testing and comparing all algorithms. We sent him an executable file of our pitch spelling algorithm for testing using his database in January 2004.

7. Plan for the Next Year

We continue to work with David Meredith of City University, London, to assess the efficacy of our pitch spelling algorithms in comparison to the state-of-the-art.

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

The algorithms for pitch spelling will be ported to the sister project, MuSA.RT (MuSA. Real-Time) – described in another report. Distinct from MuSA, which is a platform for testing analysis algorithms implemented in Java, MuSA.RT focuses on developing real-time implementations of the algorithms for live capture of MIDI input (for example, in a performance) and concurrent visualization of musical structures, and is implemented using SAI [7].

The algorithms for pitch spelling will also be used in an ongoing NSF-ITR project on music information retrieval (No. 0219912: *A User-centric Content-based Approach to Indexing, Query and Retrieval of Music through Signal Processing and Knowledge-based Methods*) – described in another report. Pitch spelling is necessary for determining the key of each piece in the database in order to normalize the data for similarity assessment.

Pitch spelling is also necessary in the analysis algorithms underlying the Expression Synthesis Project (ESP) – described in another report.

8. Expected Milestones and Deliverables

MuSA offers a platform on which to develop and test 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), Computer Music Modeling and Retrieval (CMMR), International Conference on Music Information Retrieval (ISMIR). The MuSA software has become and will continue to be part of the IMSC technology demonstrations.

9. Member Company Benefits

Our pitch spelling technology can be implemented as plug-ins to music notation or music analysis software.

The pitch spelling algorithms will help automated transcription systems achieve greater accuracy in the musical scores. For example, Figure 3 shows a leading music notation software system's spelling errors (circled) and our algorithm's spelling errors (arrowed) in a particularly challenging test set with numerous key changes and shifting contexts. Our algorithm showed only two errors, compared to the notation software's 28 in the same passage.

Figure 3: Pitch spelling errors in a challenging dataset.

The pitch spelling algorithms will also assist in producing more accurate analyses in systems for music analysis.

10. References

[1] Cambouropoulos, Emilios. *Pitch Spelling: A Computational Model*. *Music Perception*, Summer 2003, Vol.20, No.4, 411-429.

[2] Chew, Elaine & Chen, Yun-Ching, *Determining Context-Defining Windows: Pitch Spelling using the Spiral Array*. In Proceedings of the 4th International Conference for Music Information Retrieval, ISMIR, Baltimore, MD, Oct 26-30, 2003.

[3] Chew, Elaine & Chen, Yun-Ching, *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.

[4] Chew, Elaine, "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, "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] Chew, Elaine, "Towards a Mathematical Model of Tonality." Ph.D. thesis, MIT, Cambridge, MA, 2000.

[7] François A., "A Hybrid Architectural Style for Distributed Parallel Processing of Generic Data Streams," Proceedings of the International Conference on Software Engineering, Edinburgh, Scotland, UK, May 2004.

[8] Meredith, David, Pitch Spelling Algorithms. In Proceedings of the Fifth Triennial ESCOM Conference, pp.204-207, Hanover University of Music and Drama, Germany, 2003.

[9] Temperley, David (2001). *The Cognition of Basic Musical Structures*. MIT Press: Cambridge, MA.

[10] Zatorre, Robert J. and Krumhansl, Carol L., Mental Models and Musical Minds. *Science Magazine* 13 December 2002 issue, Vol 298: pp 2138-2139.