A CHORD-SCALE APPROACH TO AUTOMATIC JAZZ IMPROVISATION

Junqi Deng, Yu-Kwong Kwok
Department of Electrical and Electronic Engineering
The University of Hong Kong
{jqdeng,ykwok}@eee.hku.hk

ABSTRACT

Jazz improvisation has been one of the biggest challenges in all kinds of musical goals. This late breaking article describes a machine to automate this process in a certain degree. This machine has three major components: a jazz chord estimator, a local scale tracker, and a improvised melody generator. Simple demo tracks have been generated in terms of chord-melody to showcase the effectiveness and potentials of this system.

1. INTRODUCTION

Jazz improvisation is considered one of the most difficult tasks among all kinds of instrumental performances. The difficulty is mainly due to the complex harmonic structure in jazz, the complicated licks and patterns, and the intricate musical relationship between these phrases and the harmonic context. In this article, we preliminarily approach these issues by three different modules, and to combine them as a unified machine to automate jazz improvisation at a beginning level.

As jazz is an improvisation based music style, extracting chord progressions and key modulations could help determine/imply *chord-scale* candidates to improvise over a given harmonic segment. A *chord-scale* is [5]:

> a linear rendering of a complex chord - an extended chord structure, with tensions and non-chord tones arrayed within an octave.

Each chord-scale is constructed by a root and a scale, where the root is the root of a chord.

Normally, a chord-scale is chosen based on the current musical key context, which is determined by the underlying *harmonic functional group* [5,6]. Table 1 shows some standard choice strategies suggested by a jazz guitar tutorial book [9]. Chord-scale system is a good way to analyze jazz harmony and melody. But in real improvisation, established jazz musicians seldom think of the chord-scale system. Instead, they memorize all these by heart and improvise using the variations of melody passages and patterns they learn through extensive practices [4]. To be precise, instead of generating notes from the scales, they create phrases that belong to the chord-scale. Infinity number of note sequences can be generated out of the chord-scale, but only some are acceptable to human musical aesthetics. In addition to create phrases within a single harmonic region, they also pay attention to the coherence along and across different phrases so as to make as much musical sense as possible.

Chord	Scale
7	Mixolydian, Phrygian-Dominant, Whole-tone
maj7	Lydian, Lydian-Dominant, Ionian, Ionian#5
min7	Dorian, Phrygian, Aeolian
min7b5	Locrian, Locrian#2
7b9	Phrygian-Dominant
maj7#11	Lydian
maj7#5	Ionian#5
dim7	Whole-half Diminished

Table 1. Chord-scale choices examples

Therefore a natural approach to automatic jazz improvisation is by estimating chord-scale sequence from the backings and generate notes within each chord-scale region.

2. JAZZ CHORD ESTIMATION

The chord is estimated by our jazz automatic chord estimation (ACE) system [3], whose work-flow is shown in Figure 1.

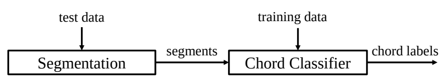

Figure 1. The jazz automatic chord estimation system.

3. JAZZ SCALE ESTIMATION

The above preliminary ACE work can be extended to scale estimation, thereby achieving a unified chord-scale estimation approach. Figure 2 shows the extended ACE system for local scale estimation. Instead of using a well-known key profile [10], binary "scale templates"(ST) are used to compute a local scale posterior probability surface from the "chord templates"(CT). Since each local scale cannot be determined by just one chord, a context window is introduced to collect a 3 chords neighborhood to facilitate

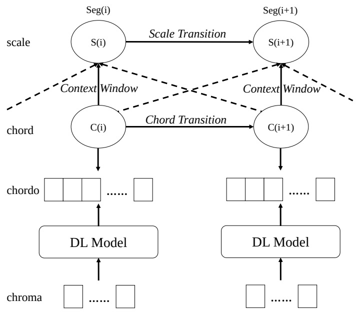

Figure 2. Template-based local scale tracking

the estimation:

$$\text{fit} = \sum_{i}^{12} ST_i \times \frac{\sum_{j=1}^{3} CT_{ij}}{3}, \quad (1)$$

For each segment, the system will decide a list of scale candidates from a look-up-table, then pick the one with the best fit scale:

$$scale = \max_{k}(\text{fit}_k). \quad (2)$$

The template-based scale tracking algorithm has been implemented during the Hackathon of ISMIR 2016. However since this is a relatively new direction in MIR, there is not any well-established labeled dataset available for testing. Therefore this module has not been formally evaluated under any objective measure yet.

4. NOTE SEQUENCE GENERATION

The two modules presented above, together with a proper note sequence generator, can be put together as a automatic jazz improvisation machine. There are two ways of generating notes:

Markov model note sequence generation — This model has two sets of parameters: the note prior probability matrix and the note transition matrix. These matrices can be trained from existing jazz solo MIDI datasets such as the Weimar Jazz Dataset [1] [1].

There is a preliminary implementation of this Markov model based note generator during the "Science of Music Hackathon" in August 2016 [2]. In this implementation (Figure 3), a simplified version of the model is built to generate sequence of "pitches" instead of "notes", where the former do not have duration information but the latter do. The code base of this project is available on-line [3]. It depends on the *pretty-midi* [8] library to do MIDI interfacing and perform simple sound synthesis. Several preliminary demos can also be found in the code base, showcasing simple chord-melody outputs from the system.

RNN-DBN automatic note sequence generation — To the other extreme, it could be a fully artistically automatic improvisation process if the "patterns of choices" from real jazz practice [4, 11] are modeled and learned. These are sequences of notes (each of them with both pitch and duration) within the context of chords or chord progressions. They can be well captured by sequential models such as an RNN using the chordal context as the starting cue, with the discriminative objective being whether the sequence of notes are artistically "acceptable" or not. This is the discriminative part of the model, and it is similar to the LSTM semantic analysis model by Maas et al. [7]. Note that this discriminative model cannot capture generative details. Hence there should be a generative model inserted into it, and thus it becomes a discriminative-generative model. Referring Boulanger's RNN-RBM based symbolic music generation system [2], if an instance of DBN (instead of RBM) is inserted between the output layer and the LSTM layer of the previous discriminative model, the resulting network will be able to both classify sequence to labels and generate sequence based on the prior labels.

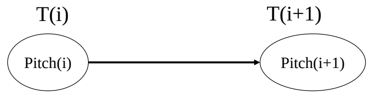

Figure 3. The Markov model for pitch sequence generation

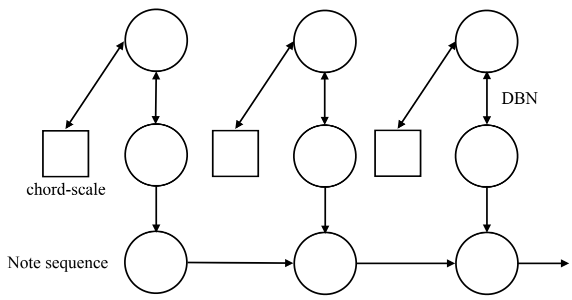

Figure 4. The RNN-DBN for note sequence generation

In this case, there needs to be a set of training cases for every chord-scale. By training the model with jazz lick patterns, a sequence of "acceptable" notes could hopefully be automatically generated by the chord-scale changes estimated from the backing track.

[1] http://jazzomat.hfm-weimar.de/dbformat/dboverview.html
[2] http://labrosa.ee.columbia.edu/hamr_ismir2016/
[3] https://github.com/tangkk/chordscale

5. REFERENCES

[1] Jakob Abeßer, Klaus Frieler, Martin Pfleiderer, and Wolf-Georg Zaddach. Introducing the jazzomat project-jazz solo analysis using music information retrieval methods. In *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR) Sound, Music and Motion, Marseille, France*, pages 187–192, 2013.

[2] Nicolas Boulanger-Lewandowski, Yoshua Bengio, and Pascal Vincent. Modeling temporal dependencies in high-dimensional sequences: Application to polyphonic music generation and transcription. *arXiv preprint arXiv:1206.6392*, 2012.

[3] Junqi Deng and Yu-Kwong Kwok. A hybrid gaussian-hmm-deep-learning approach for automatic chord estimation with very large vocabulary. In *ISMIR*, 2016.

[4] Garrison Fewell. *Jazz Improvisation For Guitar - A Harmonic Approach*. Berklee Press, 2010.

[5] Tom Hojnacki and Joe Mulholland. *The Berklee Book of Jazz Harmony*. Berklee Press, 2013.

[6] Mark Levine. *The jazz theory book*. " O'Reilly Media, Inc.", 2011.

[7] Andrew L Maas, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts. Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 142–150. Association for Computational Linguistics, 2011.

[8] Colin Raffel and Daniel PW Ellis. Intuitive analysis, creation and manipulation of midi data with pretty_midi. In *Proceedings of the 15th International Society for Music Information Retrieval Conference Late Breaking and Demo Papers*, 2014.

[9] Jeff Schroedl. *Hal Leonard Guitar Method - Jazz Guitar: Hal Leonard Guitar Method Stylistic Supplement Bk/online audio*. Hal Leonard, 2003.

[10] David Temperley. *The cognition of basic musical structures*. MIT press, 2004.

[11] Mark White. *The Practical Jazz Guitarist*. Berklee Press, 2012.