Capturing a Musician's Groove: Generation of Realistic Accompaniments from Single Song Recordings

Mathieu Ramona[1], **Giordano Cabral**[2], **François Pachet**[1]
[1]SONY CSL, 6 rue Amyot, 75005 Paris, France
[2]Centro Informática da UFPE, Av.Prof. Luiz Freire, s/n, CDU 50732-970, Recife - PE - Brazil
mathieu.ramona.pro@gmail.com, grec@cin.ufpe.br, pachetcsl@gmail.com

Abstract

This demonstration presents a concatenative synthesis engine for the generation of musical accompaniments, based on chord progressions. The system takes a player's song recording as input, and generates the accompaniment for any other song, based on the input content. We show that working on accompaniment requires a special care about temporal deviations at the border of the sliced chunks, because they make most of the rhythmic groove. We address it by discriminating accidental deviations against intentional ones, in order to correct the first while keeping the second. We will provide a full demonstration of the system, from the recording process to the generation, in various conditions, inviting the audience to participate.

1 The ReChord Player

1.1 Overview

The ReChord Player is a concatenative synthesis engine dedicated to the generation of accompaniment tracks for improvised music. It is carefully designed to reproduce the player's style, with a strong emphasis on rhythm and groove.

The system is built on top of the Lead Sheet Data Base [Pachet *et al.*, 2013] which contains over 10,000 standard songs of Jazz and Bossa Nova repertoires. A mobile application was written to allow musicians to record accompaniment, following the chord progression of any lead sheet in the database, and automatically upload the recordings to the remote server.

The musician can then browse in his personal recordings through a dedicated web page and generate in a few seconds a new accompaniment for any lead sheet in the database, built with the chunks extracted from a collection of his recordings, possibly only one song, as illustrated in Figure 1.

1.2 Previous Work in Concatenative Synthesis

Concatenative Synthesis has attracted a strong attention in the past [Schwarz, 2006]. Interestingly, most authors address the problem at the note scale, taken as a natural *musical atom*, or even at a lower scale in the case of granular synthesis. However, this approach is not relevant for accompaniment music, where most of a player's style lies in the rhythmic patterns

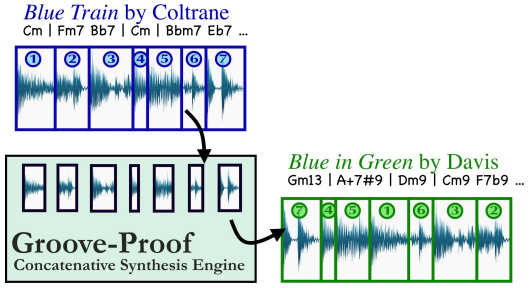

Figure 1: Overview of the ReChord Player engine.

used and in their progression. Because it relies on the prior knowledge of the lead sheet chord progression, the synthesis process in ReChord deals with the chords (described by their pitch, type and duration) to slice and concatenate the audio chunks. This chord-scale scheme raises two issues that are answered in ReChord :

- **Sparsity**: the corpus is naturally sparse. Only a small set of chord types may appear in the source song(s) and for only a few different root notes.
- **Temporal anticipations**: chunks must be carefully sliced and concatenated, because most of the *groove* consists precisely in slight temporal deviations, while some other deviations are purely accidental and should be corrected.

ReChord is an evolution of earlier attempts on Concatenative Synthesis, such as Virtual Band [Moreira *et al.*, 2013] which did not cope with anticipations because of the real-time constraint of live accompaniment.

Generation in ReChord is based on *Dynamic Programming* to deliver an optimal solution on the whole sequence, according to a combination of dedicated costs.

1.3 Sparsity

Similarly to VirtualBand [Moreira *et al.*, 2013], the sparsity of the training corpus is addressed by allowing 1) audio transpositions (pitch shifting) and 2) chord substitutions. Transpositions are penalized through a specific cost related to the transposition interval, while substitutions involve a cost based on the number of common notes between the original ad substituted chords.

1.4 Temporal anticipations

A natural approach to chunk segmentation consists in slicing audio chunks strictly by the meter (e.g. beats or bars). This works well only when 1) musicians play perfectly, i.e. with no unintentional delays or anticipations and 2) the music genre itself does not rely on anticipations or delays. In practice, this case is hardly ever met when dealing with most music genres. Examples of intentional anticipations can be found in numerous genres, such as Bossa Nova for instance. Discarding such metrical irregularities when generating accompaniments inevitably removes the specific groove of a musical genre, thereby impairing its overall expressiveness.

Such irregularities are precisely taken care of by the ReChord engine. Our approach includes the detection of metrical irregularities and their discrimination between errors and intentional (i.e. genre specific) anticipations. We propose a machine-learning approach for the detection issue, based on onset and harmonic features.

Erroneous deviations are corrected with localized time-stretching (as shown in Figure 2), while intentional anticipations are preserved in the sliced chunks.

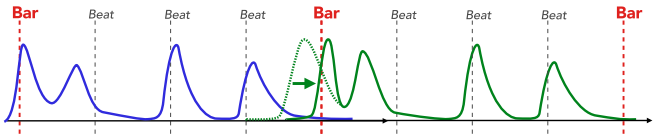

Figure 2: Accidental anticipation adjustment: the first note of the green pattern is synchronized with the beat through time-compression.

This approach is validated in [Ramona and Pachet, 2015] on Bossa Nova and Pop guitar strumming generation. A comparative evaluation shows that the proposed approach improves the expressiveness of the generation with more diverse and lively accompaniments.

2 Demo Scenario

The demonstration will be performed by the second author Giordano Cabral. The idea is simple: let the system learn one's style and play with it, in a variety of situations.

Recording Different recordings will be used as input:

- A collection of Brazilian songs played by confirmed Brazilian guitarists, and recorded with professional audio equipment, to show the audio fidelity of the system.
- The demonstrator will also play live a number of accompaniments, which will be used immediately for the generation of new songs. These will demonstrate that the system works on the fly, and is robust to ambient noise.
- People will then be invited to play accompaniments in their particular styles, no matter if experienced musicians or amateurs/novices.

Generation The generation process is performed remotely on our server, accessed through a web page[1]. This page allows the user to choose a song in the Lead Sheet DataBase, and generate an accompaniment in the style of any recording, possibly with an additional drum/percussion track. The possible combinations of such a system are endless and creativity emerges when combining very dissimilar genres (say, generating Penny Lane with Bossa Nova accompaniment).

A comparison of generations with and without temporal consistency will convince the audience of its importance.

The web page access will allow us to generate several songs in parallel, and let the user experiment with it on his mobile device.

Interaction In the end, *listening* to an accompaniment is not an obvious task. Accompaniment is played precisely to be at the background of a solo player/singer.

To show the generated accompaniments in real conditions, the demonstrator will then improvise on a few ones, mostly taken from the Brasilian repertoire. Interested people may be invited to play as well.

3 Requirements

A laptop for accessing the ReChord web page and a good WiFi connection. The rest is brought by the demonstrator.

4 Conclusion

This demonstration aims at convincing the audience about the importance of temporal deviations in preserving the groove of a gven accompaniment style and showing how style capture can be achieved with very sparse datasets. It also aims at showing an entertaining and creative application of AI.

References

[Moreira *et al.*, 2013] Julian Moreira, Pierre Roy, and François Pachet. Virtualband: Interacting with stylistically consistent agents. In *Proc. ISMIR '13*, 2013.

[Pachet *et al.*, 2013] François Pachet, Jeff Suzda, and Daniel Martin. A comprehensive online database of machine-readable leadsheets for jazz standards. In *Proc. ISMIR '13*, pages 275–280, Curitiba (Brazil), November 2013.

[Ramona and Pachet, 2015] Mathieu Ramona and François Pachet. Detecting and reproducing metrical irregularities in the generation of expressive audio accompaniments. *submitted to Computers in Entertainment (CIE) Special Issue on Musical Metacreation*, 2015.

[Schwarz, 2006] Diemo Schwarz. Concatenative sound synthesis: The early years. *Journal of New Music Research*, 35(1):3–22, March 2006.

[1] http://lsdb.flow-machines.com/rechord.php