

ON RHYTHMIC PATTERN EXTRACTION IN BOSSA NOVA MUSIC

Ernesto Trajano de Lima Geber Ramalho

Centro de Informática (CIn)—Univ. Federal de Pernambuco

Caixa Postal 7851—50732-970—Recife—PE—Brazil

{etl, glr}@cin.ufpe.br

ABSTRACT

The analysis of expressive performance, an important research topic in Computer Music, is almost exclusively devoted to the study of Western Classical piano music. Instruments like the acoustic guitar and styles like Bossa Nova and Samba have been little studied, despite their harmonic and rhythmic richness. This paper describes some experimental results obtained with the extraction of rhythmic patterns from the guitar accompaniment of Bossa Nova songs. The songs, played by two different performers and recorded with the help of a MIDI guitar, were represented as strings and processed by FIEXPAT, a string matching algorithm. The results obtained were then compared to a previously acquired catalogue of “good” patterns.

1 INTRODUCTION

It is common sense that playing music in the exact way it is written in the score results in a mechanical and uninteresting succession of sounds. To make written music interesting, the musician is required to make variations on low level musical parameters, such as: Local tempo (*accelerandi*, *ritardandi*, *rubato*); dynamics; notes articulation (*staccati*, ligatures, etc.); micro-silences between the notes, etc. [12].

Several researchers stress the importance, as well as the difficulties, of studying this phenomenon, also known as *expressive performance* [14]. These researches, in general, focus on building relationships between different musical elements (such as harmony and melody or even the low level parameters previously mentioned), and expressive performance itself. These relationships can be described in many ways, from different points of view or *levels of abstraction*, and including various musical parameters. Examples of such relationships are rules like “*lengthen a note if it is followed by a longer note and if it is in a metrically weak position*” or “*stress a note by playing it louder if it is preceded by an upward melodic leap larger than a perfect fourth*” [13].

With some exceptions [3, 6], the role of rhythm in expressive performance has not been thoroughly studied so far, despite its intuitive importance. Moreover, the research is almost exclusively devoted to the Western Classical Mu-

sic composed for the piano. We are interested in studying the *Música Popular Brasileira* (Brazilian Popular Music)—MPB, represented by artists like João Gilberto, Tom Jobim, Caetano Veloso, Gilberto Gil, etc. We are particularly interested in the guitar accompaniment, that is, in how the guitar player accompanies the singer or solo instrument.

This paper presents an experiment that focus on the discovery of rhythmic patterns in Bossa Nova music. For this, two different performers played several songs on a MIDI guitar, which were processed in the form of strings. These strings were then processed using FIEXPAT, a pattern matching algorithm [8], and the results were then compared to a catalogue of patterns that reflects the rhythmic patterns used by João Gilberto, the “inventor” of the Bossa Nova style.

The remainder of the paper is organized as follows: In Section 2, we discuss what motivated us to try such an experiment. In Section 3, we describe how the data was acquired and the representation we used. In Section 4, we present the experiment itself and discuss the results we obtained. Finally, in Section 5, we present some conclusions and future directions for this work.

2 MOTIVATION

Apart from obvious differences (instrument, genre/style and player’s role), there is a much more fundamental difference between research focusing on Western Classical Music and research that deals with MPB: While in the Western Classical Music case there is some sort of “official” notated version of musical pieces (the score, namely), in MPB there is no such thing. What may be available is the chord grid (chord sequence that should be played), and, in some rare cases, the score of the melody. Even when the chord grid is available, the musician is usually allowed to change the harmony and play something different from what was initially notated. Considering that, it becomes clear that the guitar player has a major role in the accompaniment’s construction.

This importance becomes even clearer when the rhythm is taken into consideration, because, although the musician may have the chord sequence notated (*i.e.*, information about the harmony may be somehow specified), there is no indication whatsoever of the rhythm the musician should

play. It is entirely up to him/her to decide about the rhythm.

Some studies pointed out, however, that the guitar accompaniment in styles like *Bossa Nova* and *Samba* is built by the concatenation of certain rhythmical groups or *patterns* [4, 9]. There are, however, several aspects of the accompaniment construction that are only known by practitioners of these styles. Moreover, the knowledge about the accompaniment construction is mainly subjective. Due to this lack of formalized knowledge, there are many open questions such as:

- Are there rhythmic patterns that are preferred by a certain performer or required for a certain musical style? In which situations and in which frequency do they show up?
- Are there variations of these patterns? Is it possible to group these variations in meaningful way? Which variations (timing, dynamics, etc.) are acceptable within a pattern?
- Is it really the case that everything is a pattern, i.e., are there parts that are not recurrent?
- Is it possible to justify the choice of a pattern in terms of other musical features (melody, harmony, tempo, musical structure, style, etc.)?
- Is it possible to build a dictionary of patterns for a given player? Does this dictionary changes when the style changes (*Bossa Nova* and *Samba*, for instance)? Do different players have different dictionaries?
- Is it possible to build a grammar or a set of rules that is able to describe formally how the patterns are chained and/or the rhythmical transformations done by a given performer? If so, what are the relations between grammars from performers p_1 and p_2 ?

More general questions could also be posed: To which extent patterns used in *Bossa Nova* music are different from patterns used in *Samba*? How different is *Samba* today (in terms of patterns and pattern usage) from *Samba* in the 1920's or 1930's? Is *Bossa Nova* today still played as it was in the 1950's, when it was created? How can we describe those differences, if any?

3 DATA ACQUISITION AND REPRESENTATION

For the experiment, two different players, referred to from now on as $G1$ and $G2$, were invited to record the accompaniment of some *Bossa Nova* songs on a MIDI guitar¹. $G1$

¹ The equipments we used in the recordings were the following ones: An acoustical guitar with an RMC Poly-Drive II pick-up installed (<http://www.rmcpickup.com/polydriveii.html>) that was connected to a Roland GR-33 guitar synthesizer responsible for the pitch to MIDI conversion (<http://www.roland.com/products/en/GR-33/index.html>).

performed the following songs: *Bim Bom*, *O Barquinho*, *Insensatez* (How Insensitive), *Garota de Ipanema* (Girl from Ipanema), *Só Danço Samba*, and *Wave*. From $G2$, we recorded *A Felicidade*, *Chega de Saudade*, *Corcovado*, *Desafinado*, *Eu Sei Que Vou Te Amar*, *Samba de uma Nota Só*, *Garota de Ipanema*, *Só Danço Samba*, *Insensatez*, *Tarde em Itapoa*, and *Wave*. In the total, we collected 16 recordings (ca. 30 minutes of music). It was requested for the performers to play the songs according to a provided notation (the chord grid as notated by Chediak [1]).

The acquired data was, however, not ready for usage. Probably due to technological restrictions, the resulting MIDI files were noisy and it was necessary to clean the collected songs before using them. Noisy files contained notes that were not played by the guitarist and these notes could be grouped into two basic types: Very short notes (usually high pitched) and notes with very low volume (loudness). There was yet a second type of problem, namely events that were somehow misplaced by the recording equipment (usually a semitone up or down the actually played note). The first type of noise was removed automatically, but the second one required manual correction, which was done with the help of the recording's audio files². After this step, the data was beat tracked at the eighth note level using *Beat-Root* [3], an interactive system that outputs the MIDI file beat tracked.

As we are interested in the discovery of rhythmic patterns, an essential information is the moment when a note is played or its *onset*. It would be very hard, however, to use this information only in a meaningful way. We should, then, associate some other information with onsets in order to better represent the songs. Pitches may come to the reader's mind as the most relevant information that could be used with onsets, but they are not intrinsically linked to the rhythm. If we, however, observe how sound is produced by the guitar player, we may link each onset to the finger that generated it. So, the finger used to pluck the string and produce sound may prove a much more interesting *abstraction* than, for instance, the pitches.

But, this abstraction was not readily available in the files we collected³. So, we had to develop an algorithm to automatically determine the right hand *fingerings* [11]. Roughly speaking, we first introduced the concept of *hand position*, that is, the fingers' position regarding the string or strings they are about to pluck. Then, we created a *hand position set* that contains all relevant hand positions. Considering that the fingering is formed by transitions between consecutive hand positions, we assigned a cost to each transition⁴.

² We recorded, at the same time, MIDI and audio information.

³ Note that the MIDI files we had at hand contained no information at all about the fingering. But, we collected them in a way that each string was recorded separately on its own MIDI channel.

⁴ This cost represents, in fact, the amount of effort required to change from hand position HP_i to hand position HP_{i+1} .

But, what are the patterns João Gilberto plays? In the literature we were able to find several transcriptions of patterns played by João Gilberto [4, 9]. So, we have built a catalogue containing 21 different patterns (labeled P1, P2, etc.), all manually transcribed by musicologists from João Gilberto’s recordings. Examples of patterns in this catalogue are showed in Figure 3.

Figure 3. Examples of patterns in the catalogue

The experiment was, then, to explore the data set we had acquired, verifying if patterns from this catalogue could be found in the data set. For the experiment the algorithm FIEXPath [8] was chosen. It is an inexact string matching algorithm that was inspired by algorithms from the Computational Biology field, but that also incorporates results from previous research on musical similarity [7]. Given an input (here the simplified string previously described) and using the edit distance as its similarity measure⁹, the algorithm outputs a collection of patterns, organized in classes. Each class has a prototype, that is the most representative pattern of the class, and several occurrences, possibly inexact, of this prototype. It is also possible for the user to provide the algorithm with some constraints, such as the maximal and minimal length of patterns, the similarity threshold, the maximum difference between two candidates to be compared, etc.

In our first experiment FIEXPath was used “as is”, that is, we used the algorithm without any sort of modifications and/or adaptations. Although the algorithm found patterns from the catalogue, including patterns with some small modifications, the results, in this case, were not so good as we previously expected. Table 1 summarizes the results we obtained with FIEXPath.

The algorithm’s configuration we used was the following: $m_{min} = 17$ (minimal pattern length), $m_{max} = 34$ (maximal pattern length), and similarity threshold of 0.75 (normalized values). The lengths m_{min} and m_{max} correspond to patterns varying from one to two measures¹⁰. On

⁹ Dynamic programming is used to compute it efficiently.

¹⁰ It does not mean, however, that the patterns necessarily start at the

Pat	G1	G2	Pat	G1	G2
P1	×	×	P12		
P2	×		P13		
P3			P14		
P4	×		P15	×	
P5			P16		
P6			P17		
P7			P18		
P8			P19		
P9	×	×	P20		
P10	×	×	P21		
P11			/	/	/

Table 1. Patterns from the catalogue found in data set

average, each song has 48 classes of patterns (smallest number of classes was 40 and greatest number was 78).

In song *Garota de Ipanema*, played by G2, the algorithm was able to identify the following pattern: |A---P---B-p-+---|P---P---Bp---+|. The corresponding pattern in the catalogue is |A---P---B-p-+---|A---P---B-p-+-+|. Looking closer at the modifications done by the performer, it is possible to notice that he usually substitutes A by P in the second measure of the pattern, and that he anticipates both final chords in this same measure. Figure 4 depicts these variations.

Figure 4. Pattern found by FIEXPath

In song *Wave*, as played by G1, FIEXPath identified as the prototype of class 18 the following pattern: |A---P---B-p-+|-|B---P---Bp---+p-. The corresponding pattern in the catalogue is |A---P---B-p-+-|-|B---P---B-p-+-|-|. As in the previous case, the modifications done by the performer are mainly anticipations.

The main problem we found was that the extracted patterns, many times, start from the middle of a measure, such as in ---B---B---+---|P---B---B---+---|P (FIEXPath also found the expected pattern, |P---B---B---+---|P---B---B---+---). This brings some problems to the evaluation of the results: Due to the great number of patterns with this kind of structural malformation, it becomes difficult to validate the patterns.

beginning of the measure. There is no way to specify such a constraint in FIEXPath.

This problem is even bigger when most of a class is formed by such patterns (it happened many times, unfortunately).

From these results we may rush and conclude that FIEXPAT performed poorly, since it found just a few patterns from the catalogue in the data set and because these patterns were grouped into too many classes. However, as pointed out in the previous paragraph, there was a sort of structural malformation in many patterns found by FIEXPAT. So, what would happen if we could “help” the algorithm, somehow providing it with some structural information about the patterns?

Unfortunately, it is not possible to describe for the algorithm the desired aspect of the pattern: We can not say that it should only consider patterns that begin on the measure bar, for instance. We can only describe the minimum and maximum pattern lengths, as well as the minimum similarity threshold. If we, however, look more carefully at the patterns in the catalogue, it is possible to identify some common structure: They are all two measures long and they start either at the measure bar or at a “puxada” right before the measure bar¹¹.

We then introduced a step just before the actual pattern matching where all songs were segmented according to this structure. It is important to note that with this modification, we changed the abstraction level at which the pattern extraction happens: Instead of comparing event by event (basses or chords), now we compare groups of events (i.e., the segments).

After we adjusted our implementation of the algorithm to accept the segments described before as input, we ran the experiment again. In this case, differently from the previous one, the algorithm found systematically in every song patterns from the catalogue. This time, more than half of the catalogue’s patterns was found, which significantly improves the results we obtained before. On average, each song had 14 classes of patterns (smallest number was 6 and greatest number was 34). Table 2 shows the patterns found in this second part of the experiment.

After running the experiments and having in mind that we have used a small data set, it is now possible to discuss a little bit the results we obtained. First of all, let us examine the patterns used by the performers. As described by Dahia et al. [2], patterns in the catalogue belong to one of four groups: cyclical (*P2* to *P8*), beginning (*P9* to *P12*), special (*P13* and *P14*) and fill-in patterns (*P15* to *P21*). Pattern *P1* is the main Bossa Nova pattern and forms its own group. From the results presented before (Tables 1 and 2), it is possible to see that the players used patterns belonging to all groups, but the special one. This group, however, contains patterns that are rarely used and it would not be a problem if they do not show up. Given that the catalogue contains Bossa Nova patterns, we can say, then, that the per-

Pat	G1	G2	Pat	G1	G2
P1	×	×	P12		
P2	×	×	P13		
P3	×	×	P14		
P4	×		P15		×
P5	×		P16	×	×
P6	×		P17		
P7			P18		×
P8			P19	×	
P9	×	×	P20		
P10	×	×	P21		
P11			/	/	/

Table 2. Patterns found by the modified version of FIEXPAT

formers do follow the Bossa Nova style in the rhythmic accompaniment construction.

There are occasions, however, that the pattern used does not belong to the catalogue. *G1*, for instance, uses the pattern |P---B---B---+---|P---B---B---+--- in song *Eu Sei Que Vou Te Amar*. We can interpret this fact in many ways, such as: the player deliberately used a non Bossa Nova pattern, the catalogue may not be complete, or even that Bossa Nova style has changed over the time and players started to use or create their own patterns.

5 FINAL REMARKS AND FUTURE WORK

This paper described an experiment that dealt with the extraction of rhythmic patterns from Bossa Nova songs. Sixteen beat tracked MIDI files, representing the recording of several songs by two different players, were represented as a string and thereafter processed by a string matching algorithm called FIEXPAT. The objective was to identify in the data set patterns from a previously acquired catalogue of patterns.

First results showed that, although FIEXPAT could find some patterns, many patterns, including some that could be clearly heard, were not found. After examining the structure of the catalogue’s patterns, we were able to pre-segment the songs and use the pre-segmented version of the song as input for the algorithm. This time the results were much better, and more than half of the patterns was found by the algorithm.

It is important to stress the relevance of the abstractions we have built. First of all, we could reduce the complexity of the pattern matching process by rewriting a polyphonic song as a monophonic line, with minimum information loss. This first abstraction was, however, not sufficient, as can be seen with the first experiment we made, and we had to build a second abstraction, going up one structural level: From single events to groups of events (*segments*). Only due to these

¹¹ It means that all segments look like |<pattern> or like 1*|<pattern>, where the * symbol represents a sequence of zero or more minuses only.

abstractions we could explore the data set we collected.

Of course, there are several points for improvement. One that is particularly important is the representation. We used a very simple representation of the events. Attributes like tempo, structural or harmonic information are not represented. The more attentive reader may have even noticed that the actual duration of the events is not represented at all. We just used the onset information, which turned out to work for this experiment. To further investigate the particularities of the patterns, however, we surely need to represent appropriately the duration of each event. Another point related to the representation is the following: what kind of results would we have if we had represented the downbeat explicitly? Would they be better than the ones we presented here?

The evaluation of the algorithm's results was done in an *ad hoc* manner: Results were compared, one by one, to the patterns in the catalogue. This procedure takes too much time, is error prone, and, therefore, must be improved. We plan to implement a tool to help us with this task.

The data acquisition is an important and non-trivial problem. It is important because if we want to draw relevant and significant conclusions about Bossa Nova (and any musical style, in fact), we must have a much more representative data set. And it is non-trivial because there are many factors involved varying from the availability and willingness of certain performer to record for us, to copyright issues of the collected material. One possible way to remedy this problem is to use audio recordings as start point, but, unfortunately, there are other problems (separating guitar and singing voice signals most notably).

FIEXPath's problem (structural malformation) should be examined more carefully, since, instead of a problem, it can mean another thing. Several patterns of the catalogue have a common substructure (i.e., sub-parts of these patterns are equal). It may be the case that the second measure of patterns whose second measure are equal, are frequently concatenated with patterns whose first measure are similar. So, it may be the case that these concatenations are so typical that they are "promoted" to patterns by the algorithm.

Finally, we hope that the questions previously formulated were provocative enough. We do believe that answers to them will bring much more understanding of how the musicians bring their genres to life, as well as register how they, musicians and styles, evolved through the years.

6 REFERENCES

- [1] Chediak, A. editor. *Songbook: Bossa Nova*, volume 1–5. Lumiar Editora, Rio de Janeiro, 1990.
- [2] Dahia, M., Santana, H., Trajano de Lima, E., Sandroni, C., Ramalho, G., and Cabral, G. "Using patterns to generate rhythmic accompaniment for guitar", In *Proc. of Sound and Music Computing (SMC'04)*, pages 111–115, 2004.
- [3] Dixon, S. "Automatic extraction of tempo and beat from expressive performances", *Journal of New Music Research*, 30(1):39–58, 2001.
- [4] Garcia, W. *Bim Bom: A Contradição sem Conflitos de João Gilberto*. Editora Guerra e Paz, 1999.
- [5] Meredith, D., Lemström, K., and Wiggins, G. "Algorithms for discovering repeated patterns in multidimensional representations of polyphonic music", *Journal of New Music Research*, 31(4):321–345, 2002.
- [6] MMM. Music, mind, machine group, 2003. <http://www.nici.kun.nl/mmm/>. Last access date Mar 12 2004.
- [7] Mongeau, M., and Sankoff, D. "Comparision of musical sequences", *Computer and the Humanities*, 24:161–175, 1990.
- [8] Rolland, P.-Y. "FIEXPath: Flexible extraction of sequential patterns", In Nick Cercone, Tsau Young Lin, and Xindong Wu, editors, *Proceedings of the 2001 IEEE International Conference on Data Mining (ICDM'01)*, pages 481–488. IEEE Computer Society, 2001.
- [9] Sandroni, C. *Feitiço Decente: Transformações do Samba no Rio de Janeiro (1917–1933)*. Jorge Zahar Editor, Rio de Janeiro, 2001.
- [10] Sayegh, S. "Fingering for string instruments with the optimum path paradigm", In Peter M. Todd and D. Gareth Loy, editors, *Music and Connectionism*, pages 243–255, Cambridge (MA), 1991. The MIT Press.
- [11] Trajano de Lima, E., Dahia, M., Santana, H., and Ramalho, G. "Automatic discovery of right hand fingering in guitar accompaniment", In *Intern. Comp. Mus. Conf. (ICMC)*, pages 722–725, 2004.
- [12] Widmer, G. "Applications of machine learning to music research: Empirical investigations into the phenomenon of musical expression", In R. S. Michalski, I. Bratko, and M. Kubat, editors, *Machine Learning, Data Mining and Knowledge Discovery: Methods and Applications*. Wiley & Sons, Chichester (UK), 1998.
- [13] Widmer, G. "Machine discoveries: A few simple, robust local expression principles", *Journal of New Music Research*, 31(1):27–50, 2002.
- [14] Widmer, G., Dixon, S., Goebel, W., Pampalk, E., and Toubudic, A. "In search of the Horowitz factor", *AI Magazine*, 24(3):111–130, 2003.