Microtiming patterns and interactions with musical properties in samba music

March 2, 2011

Luiz Naveda (IPEM - Ghent University)
Fabien Gouyon (INESC - Porto)
Carlos Guedes (INESC - Porto)
Marc Leman (IPEM - Ghent University)

Abstract

In this study, we investigate the timing aspects of the Afro-Brazilian samba music. Instead of looking exclusively at the temporal aspects of musical events, we focus on the interaction between microtiming patterns and several musical properties: intensity, meter and spectral characteristics. Our study is based on computational analyses applied to a data-set of 106 musical audio excerpts. The data-set is processed through an auditory model and then segmented with respect to several spectral regions and metric levels. In the sequence, the resulting segments are described in terms of their musical properties, over which patterns of peak positions and their intensities are sought. A clustering algorithm is used to systematize the process of pattern detection.

Among our results, we confirm a phenomenon previously reported in the literature: the anticipations of the third and fourth semiquavers in a beat with respect to their quantized positions. We also argue that these patterns of microtiming deviations interact with different profiles of intensities that change according to the metrical structure and spectral characteristics. In particular, we suggest two new findings: (i) a small delay of microtiming positions in the lower end of the spectrum on the first semiquaver of each beat and (ii) systematic forms of *accelerando* and *ritardando* at a microtiming level covering 2-beats and 4-beats phrases. The results demonstrate that some characteristics of musical cultures such as the samba are beyond traditional representations of rhythm. It also demonstrates the importance of multidimensionality interactions with timing aspects of music, which might be only verifiable through computational approaches. More research is needed in order to find proper representations for rhythm and microtiming aspects in such contexts.

1   Introduction

The idea that some properties are invariant in a group of music examples is a primary assumption in the analysis of music styles (de Carvalho, 2000, p. 134). When it comes to

1

analysis of samba music, and in general music from the African diasporas, the majority of approaches concentrate on invariant properties of *rhythm*. However, rhythm involves a number of different aspects. It is claimed that rhythm conveys a combination of temporal structure, beat induction and timing (Honing, 2001), which interacts with a number of aspects encoded in sound, such as metric structures and dynamics (e.g., London, 2004; Palmer, 1997; Sethares, 2007), as well as aspects not encoded in sound such as motor-schemes (Todd, 1995; Palmer, 1997), kinematic models (e.g., Honing, 2003; Grachten and Widmer, 2009; Palmer, 1997; Todd, 1995) and other modalities such as dance (Naveda and Leman, 2009; Chernoff, 1991). How do the actual representations of samba style deal with such multi-dimensionality of rhythm? How could computer music help to detect meaningful invariants in this context?

The majority of analytical studies of samba is based on or modeled through symbolic representations (e.g., musical scores) that are designed to represent the perception of macro-structural characteristics of rhythms (i.e. happening at lower levels of the musical meter, London, 2004) such as the relative durations of musical events, bar, beat and phrases. For example, it has been claimed that samba music is characterized by a binary metric structure (binary bar) "muted" in the first beat and accentuated on the second beat (Chasteen, 1996; Galinsky, 1996; Moura, 2004). Other sources indicate that samba also exhibits a "polymetric rhythmic texture" or a musical texture in which different metric layers having different periodic lengths and metric phases coexist (Browning, 1995; Fryer, 2000). Great part of the literature refers to the general concept of syncopation or the figure of syncope (e.g., Sandroni, 1996; Sodré, 1979) as the main characteristic of samba. In particular, many authors have proposed rhythms figures that characterize rhythms or represent models for renditions of rhythms in samba. Examples of these propositions include the "tresillo" (Sandroni, 2001), the "characteristic syncope" (Andrade, 1991), the "tamborim cycle" (Araújo, 1992, quoted in Sandroni, 2001), the "samba rhythm necklace" (Toussaint, 2005), or the "Angola/Zaire Sixteen-pulse Standard Pattern" (Galinsky, 1996; Kubik, 1979).

In contrast with symbolic models, a number of studies highlight the fact that musical experience in samba is transmitted by means of *subjective texts* (texts, reports and interpretations from subjects) and informal contexts, based on oral traditions and social participation, rather than through more explicit knowledge or written documents (de Carvalho, 2000; Sandroni, 2001; Sodré, 1979). The strong link of samba with Afro-Brazilian religious rituals (de Carvalho, 2000) and social displays such as the *roda-de-samba* (Moura, 2004) indicates that samba cannot be easily detached from experience of dance, rituals, texts and from the experience in the social context as a whole. In this context, the action-perception loops experienced in the timing of activities such as dance (e.g., Sodré, 1979; Browning, 1995), manual labor (e.g., Fryer, 2000) or hand clapping have an active role in the ellaboration of music performances. From this socio-historical perspective, rhythm should be understood as a concept formed by a (1) number of modalities of dimensions of experience, hence hardly explainable as a composition of independent elements, and (2) as an experience that is strongly rooted on the perception of *timing* and action in time.

2

A small part of the literature deals with micro-time structures of rhythm. Some references are also made to the relevance of observing rhythmic phenomena occurring at the fastest level of the musical meter. This metric level is referred in the literature in a variety of forms: "tatum layer" (Bilmes, 1993), "valeurs opérationelles minimales" (Arom, 1989), "pulsation" (Polak, 1998) or "common fast beat" (Kauffman, 1980). A number of studies focus on small idiomatic deviations applied to the *tatum* level between instants where notes are actually played and their corresponding quantized positions. These deviations are referred to *microtiming*, here defined as a series of event shifts at a constant tempo (Desain and Honing, 1993; Bilmes, 1993).

Microtiming characteristics, and interactions with other musical features as pitch, phrasing or intensity, have also been observed in other music styles such as Jazz (Friberg and Sundstrom, 2002; Benadon, 2003, 2006, 2009), Norwegian traditional fiddle music (Johansson, 2005), Irish traditional fiddle music (Rosinach and Traube, 2006), or Viennese Waltz (where the second beat in a group of three is early and emphasized, Desain and Honing, 1989; Gabrielsson, 1985). Some links are also starting to be done between the presence of microtiming characteristics in music and groove perception or movement induction (Madison, 2006; McGuiness, 2006).

Few studies also exist on aspects of microtiming in samba music. Lindsay and Nordquist (2007) analyzed microtiming in recordings of samba instruments (*pandeiro*, *surdo* and *agogô*) using an improved spectrogram analysis based on Fulop and Fitz (2006) as basic signal representation combined with a manual annotation of events. They found systematic anticipations of the third and fourth semiquavers (within 1 beat) for the *pandeiro* recordings inside pairs of "short-long" onsets. They also found 4-beat patterns of onsets in progressive acceleration. Naveda et al. (2009) studied spontaneous vocalization of samba rhythms using a peak detection algorithm applied to auditory images (based on the auditory features proposed in Van Immerseel and Martens, 1992) and also found indications of systematic anticipations of the third and fourth semiquavers. Also using standard spectrogram analyses combined with manual annotation, Lucas (2002) found similar microtiming deviations in recordings collected in Minas Gerais state (Brazil) pertaining to the traditions of *Congado Mineiro*. Although the *Congado* traditions are accompanied by musical forms stylistically distinct from samba music, both *Congado* and samba share the same Afro-Brazilian roots. Gerischer (2006) collected several reports from musicians in the context of samba performed in Bahia (another Brazilian state). She realized a systematic analysis of microtiming based on field recordings and manual annotation. Gouyon (2007) analyzed commercial recordings os samba. He identified patterns of microtiming deviations by means of machine learning techniques applied to the "complex spectral difference", which was suggested in (Bello et al., 2004) as onset detection function. Results also indicated the existence of systematic anticipations of the third and fourth semiquavers. This overview accounts for evidences of systematic deviations that seem to occur on the third and fourth semiquavers in the beat level in samba music. However, most studies are based on small number of samples and most analyses rely on manual annotation of events or windowed FFT methods, whose temporal precision does not permit reliable analysis, specially in low-frequency components. Most
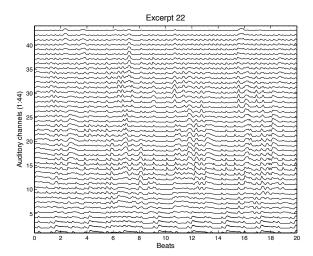
Figure 1: Example of loudness curves generated by auditory model. The 44 envelope curves represent a simulation of loudness on each auditory channel (for more details, see Van Immerseel and Martens, 1992).

importantly, most of these studies consider *only* (micro) temporal deviations and do not consider potential interactions with other musical features such as intensity, timbre, or meter.

We aim at studying, from a systematic point of view, and with a significant number of musical audio excerpts, microtiming characteristics of samba music and their interactions with different musical properties, namely intensity, meter and (estimations of) timbre.

The methodology is explained in Section 2, where we provide details on the data-set, on the extraction of low-level features from audio (accounting for an auditory model and segmentation of spectral regions and metric levels), on the method for computation of microtiming features, and finally on the method used for clustering the obtained information. Results are provided in Section 3, which examines the tendencies observed in the clustering groups. Finally, in Section 4 we discuss the results and possible impact in our hypotheses and Section 5 summarizes the contributions and implications of this study.

2 Method

2.1 Data-set

Our data-set consists in 106 excerpts of music collected from commercial CDs. The median of durations is 33 s. The range of genres includes music styles influenced by Rio de Janeiro's samba, such as *samba carioca*, *samba-enredo*, *partido-alto* and *samba-de-roda* (from Bahia). The excerpts were stored in mono audio files with a sample rate of 44100 Hz, 16 bits and normalized by amplitude.
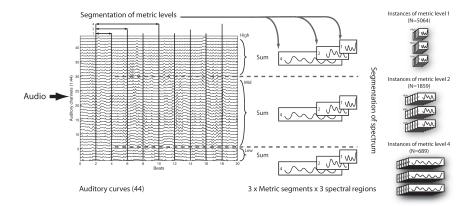
4

Figure 2: Two processes of segmentation of the auditory curves: metric levels and spectrum. The segmentation results in a collection of N instances for each metric level, divided by the spectrum region. The 3 spectral regions are also represented at this phase.

2.2 Extraction of low level features

2.2.1 Auditory model

We used an implementation of the auditory model described in Van Immerseel and Martens (1992) (.dlib library for Mac OSX). This auditory model simulates the outer and middle ear filtering and the auditory decomposition in the periphery of the auditory system. The results take a form of loudness curves representing the loudness on the auditory bands of the audible spectrum (for more details see Van Immerseel and Martens, 1992, p. 3514). The configuration used in this study provides 44 channels of loudness curves with sample frequency at 200 Hz, distributed over 22 critical bands (center frequencies from 70 Hz to 10.843 Hz). Figure 1 displays an auditory image (or loudness curves) generated from the auditory model of an example excerpt.

2.2.2 Segmentation

The segmentation of auditory curves involved two parts: (1) the process of segmentation of the spectrum range in the frequency domain, which averages auditory curves in 3 *spectrum regions* (low-, mid- and high-frequency spectrum) and (2) the process of segmentation of 3 metric levels in the temporal domain, which segments the features in segments with lengths 1, 2 and 4 beats. The processes of spectral and metric segmentation are illustrated in Figure 2.

Segmentation in spectral regions The data-set consists in excerpts of comercial polyphonic music which makes the separated instrumental sources unavailable. Current state-of-the-art source separation techniques are prone to bias and generation of artifacts that could disturb the detection of microtiming positions. However, current knowledge about indicates that percussion instruments of the samba ensamble have defined musical

functions and roughly defined spectrum signatures across the musical tessitura, as exemplified in Figure 3 . The musical function of each instrument is related to its timbre, which can be roughly represented by low-level descriptors in the frequency domain or, in our case, by loudness in time distributed in auditory channels.

In Figure 3, for example, the spectrum of the low-frequency samba drum, the *surdo*, is mostly concentrated in the lower part of the audible spectrum. *Tamborims*, *repiniques*, vocal parts and other instruments occupy the mid frequency region of the auditory spectrum. The spectrum signature of *Ganzás* and different kinds of shakers are concentrated in higher portions of the spectrum. Although the frequency components of these instruments overlap each other in the time and in the frequency domains (particularly during transients in the attacks points), the spectrum signature of each timbre is relatively discriminated from each other. Therefore, for each excerpt, we averaged the 44 loudness curves provided by the auditory model in 3 loudness curves that reflect estimated distributions of tessitura: low-frequency region (channels 1:6), mid-frequency region (channels 7:30) and high-frequency region (channels 31:44). For a similar procedure, see (Lindsay and Nordquist, 2007).
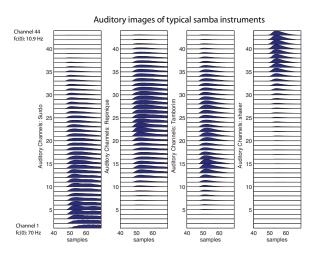

Figure 3: Response of the auditory model for attacks (80 samples = 0.4 s) of the instruments *surdo*, *repinique*, *tamborim* and shaker (from left to right). The graphs demonstrate how different instruments of a traditional samba ensemble have different responses or spectrum signatures.

Segmentation of metric levels   Current knowledge about samba also indicates that it has a well-defined, salient beat level (referred elsewhere as quarter-note), a binary bar structure (2 beats) and a fast metric level that divides the beat in four semiquavers (also referred as *tatum* level).

In order to identify the time points of the metrical accents, we performed the annotation of beat (1 beat) and bar (2 beats) levels for the entire dataset. Automatic beat
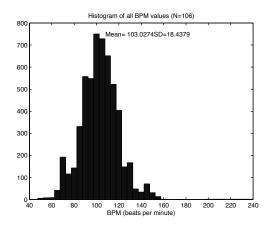
6

Figure 4: Histogram of BPM values of the beats for the whole data-set (106 excerpts, 5064 beats).

annotation using software such as BeatRoot (Dixon, 2007) and QMUL beat tracking plugins for Sonic Visualiser (Cannam et al., 2006) resulted in erractic beat tracking for this dataset. Therefore, we manually annotated the first and second quarter-note beats of each bar in the dataset (annotations were realized by three Brazilian musicians using the software Sonic Visualiser). This process results in a total of 5064 quarter-note beats in our data-set. Figure 4 shows the distributions of the BPM values for the data-set. Normality observed in the histogram shows a tendency of tempi towards 103 BPM (mean = 103.02, standard deviation = 18.42).

These annotations are used to define three different types of segments containing respectively 1, 2 and 4 beats (hence corresponding to three metric levels). Each will be subjected to a separate cluster analysis.

2.3   Computing microtiming features

For each musical excerpt, the previous segmentation steps yield a number of segments, corresponding to three different metric levels, and three spectral regions. Each of these segments is then subjected to an analysis of its microtiming deviations with respect to the mathematical semiquavers subdivision and is parameterized in order to compute microtiming features. This parameterization is described in Table 1, while Figure 5 provides an illustration of this process.

One should notice that the average between the position of the peaks of the three spectral regions in Phase 5 is of utmost importance. Manual beat annotation does not provide a precise beat segmentation (due to bias of the manual process) and differences of attacks between spectral regions do not offer common point for segmentation (due to discrepancies between attacks of instruments). Therefore, we opted to rely on the average position of the three spectral regions of the first semiquaver, which permits the

For *each excerpt*, for *each metric level* (i.e. 1-beat, 2-beat, or 4-beat) do:

Phase 1 Retrieve beat position and Inter-Beat Interval (IBI).

Phase 2 Retrieve strict semiquavers positions by generating a mathematical division of the beat (four steps of $\frac{1}{4}$ of the IBI).

Phase 3 Look for peaks in the proximity of first beat manual annotation, in *each spectral region*.

Phase 4 In each spectral region, select highest peak situated above threshold around first beat (if there are no peaks above threshold, retrieve NaN).

Phase 5 Compute average peak position of the 3 spectral regions.

Phase 6 Retrieve position and amplitude of the highest peak in close proximity of *each* semiquaver, in *each spectral region*.

Table 1: Pseudo-algorithm for the computation of microtiming features.

calculation of beat period and microtiming *relative* to this point. This does not affect the results because we rely on relative positions in relation to the IBI rather than absolute positions in seconds.

Finally, for a given metric level, we computed the following microtiming features for each semiquaver: (1) the position of the peaks with respect to the first beat in each spectral region (henceforth noted p) and (2) the intensity of the peaks (noted i). For n semiquaver in a given metric level, the instances will contain $p_{1,\cdots,n}$ peak positions and $i_{1,\cdots,n}$ peak intensities, in each spectral region $r_{1,\cdots,3}$. Table 2 specifies the structure of the instances further analysed in the next Section. Note that instances that feed the clustering algorithm combine information of spectrum, timing and intensities, and are clustered for three different metric levels or segments. The process of clustering applied to each metric level will lead to three different groups of results, displayed in Sections 3.2, 3.3 and 3.4 (metric levels 1- 2- and 4-beat, respectively). It is expected that different lengths of instances, or metric levels, will provide different configurations of clusters and reveal different patterns of interaction.

2.4   Clustering

In order to find common patterns between these instances, we carried out a k-means clustering based on an improved extension of the basic k-means algorithm, developed by Pelleg and Moore (2000) and implemented in Weka platform (Witten and Frank, Witten and Frank). This method searches for locations and numbers of clusters that efficiently improve the Bayesian Information Criterion (BIC) or the Akaike Information
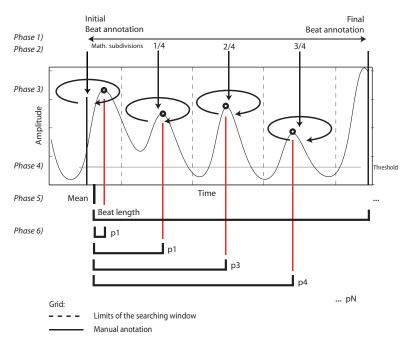
Figure 5: Description of the heuristic of calculation of microtiming deviations. Example of 1-beat metric level, and single spectral region only represented. See Table 1 for the explanation of each step of the algorithm.

Metric level 1-beat: $[p_{1,\cdots,4}, i_{1,\cdots,4}]_{r_{1,\cdots,3}}$

- 12 positions + 12 intensities =24 elements

Metric level 2-beat: $[p_{1,\cdots,8}, i_{1,\cdots,8}]_{r_{1,\cdots,3}}$

- 24 positions + 24 intensities =48 elements

Metric level 4-beat: $[p_{1,\cdots,16}, i_{1,\cdots,16}]_{r_{1,\cdots,3}}$

- 48 positions + 48 intensities =96 elements

Table 2: Description of the instances used in the the k-means process

Criterion (AIC) measure. The algorithm was configured to retrieve a minimum of 3 and a maximum of 5 clusters (arbitrary).
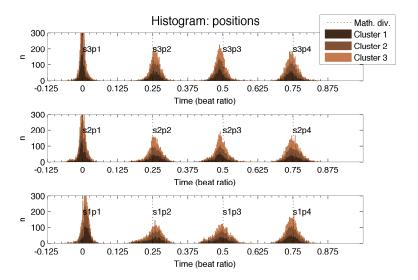
9

Figure 6: Distributions of the peak positions for metric level 1-beat, for all excerpts (N=5064, 106 excerpts). The shades of gray indicate the contribution of each cluster to the total distribution. The vertical grid indicates the mathematical subdivisions of the beat (0, $\frac{1}{4}$, $\frac{2}{4}$ and $\frac{3}{4}$ of the beat).

3 Results

We first provide results regarding average microtiming distributions in metric level 1-beat (Figure 6), while following sections examine the internal structure of these microtiming distributions by means of clustering analysis.

3.1 Microtiming distributions in metric level 1-beat

The results displayed in Figure 6 show an overview of the main microtiming tendencies for metric level 1-beat. We examined the deviations of all microtiming positions (4 positions × 3 spectrum regions) from the mathematical divisions of the beat using ANOVA.

The main observations derived from pair wise comparisons indicate that third and fourth semiquavers are significantly anticipated with respect to mathematical divisions of the beat ($F(10, 5064) = 422.39, p = 0$). This confirms results from previous studies (Naveda and Leman, 2009; Lindsay and Nordquist, 2007; Gouyon, 2007). Mean values for these anticipations are -0.026, -0.031 and -0.032 beats for the third semiquavers, in low-, mid- and high-spectrum regions respectively (i.e. 16, 18 and 19 ms in the case of excerpts with average tempo of 103 BPM), and -0.028, -0.018 and -0.027 beats for the fourth semiquavers, in low-, mid- and high-spectrum regions respectively (16, 11 and 16 ms for average BPM).

In addition, the first semiquaver in the low-spectrum region is delayed from its mathematical position. We have found a mean deviation of $+0.012$ beats, which represents
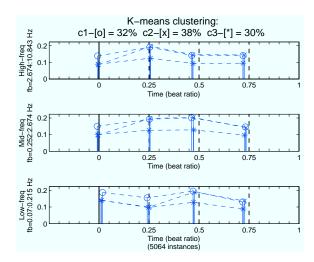
10

Figure 7: Cluster centroids c1, c2 and c3 for 5064 instances of metric level 1-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

7.3 ms when reported to the average BPM.

3.2 Clusters in metric level 1-beat

The cluster analysis of the 1-beat level resulted in three clusters for each spectrum region, displayed in Figure 7 (note that different clusters are represented by different stem markers, connected by traced lines, which facilitates the visualisation of the intensity profiles of the clusters). The representation of the cluster centroids confirms the observation made above: third and fourth semiquavers are anticipated in all three spectrum regions and in all three clusters, and the first semiquaver of the low-spectrum is slightly delayed.

In addition, analysis of intensities shows new information. Pair wise comparison after ANOVA (mean cluster intensities × 3 spectrum regions) shows that the second semiquaver is significantly accentuated in the mid- and high-spectrum in all clusters ($F(2, 3821) = 675.7201, p < 0$). In the high-spectrum, the clusters show flat intensities in the second half of the beat. Cluster c3 is generally less intense than the other clusters while cluster c1 is more intense. Cluster c2 seem to display a mixture of clusters c1 and c3: first peaks have the same properties of cluster c3 while the other peaks exhibit the same characteristics of cluster c1.

3.3 Clusters in metric level 2-beat

The cluster analysis of the metric level 2-beat resulted in five clusters. Figure 8a shows clusters c1, c2 and c5 for metric level 2-beat. The results exhibit the same systematic anticipations of third and fourth semiquavers, in every beat (or third, fourth, seventh and eight semiquavers, in a two-beat sequence). There is also a delay of the first (and
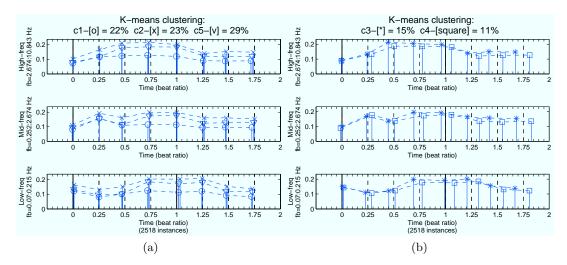
11

Figure 8: Cluster centroids c1, c2, c3, c4 and c5 for 2518 instances of metric level 2-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

fourth) semiquaver in the low-frequency region, as seen in the metric level 1-beat. This observation seems to affect both quarter-note beats at the bar level although differently: ANOVA shows that first semiquavers of the first and second quarter-note beats in the low-spectrum region are significantly delayed from their mathematical positions ($F(1, 802) = 15.2181, p < 0.0001$), by +0.0087 beats and +0.018 beats, respectively (recall that the deviation was +0.012 beat, close to the average of these two values, when focusing on the 1-beat level). The first semiquaver of the *second* beat is significantly more delayed than that of the first quarter-note beat. The former also shows an accentuation in intensity with respect to the latter, confirming the tendency to accentuation of the second beat, reported in the literature (Sandroni, 2001; Moura, 2004; Chasteen, 1996).

Peak intensities reveal more variability at this metric level. While the intensity peak of the second semiquaver (first beat) seems to be accentuated only in the mid-frequency region, the fourth semiquaver is accentuated in cluster centroids c2 and c5. In the second beat, peak intensities of the second to the fourth semiquavers are flattened. Cluster c1 has an overall low intensity and flat profile compared with the other clusters.

Figure 8b shows the results of the clusters c3 and c4. These results differ from those of clusters c1, c2 and c5 by the profile of increasing deviations accumulating in time. Cluster c3 shows an increasing anticipation in all regions and peaks. The anticipation increases until the last semiquaver of the second beat, which ends with almost 0.1 beat of anticipation from the mathematical position of the fourth semiquaver of the second beat. Cluster c4 shows the opposite pattern: an increasing delay from the first to the last semiquaver. The intensity patterns seem to be similar to the observed intensities in clusters c1, c2 and c5.
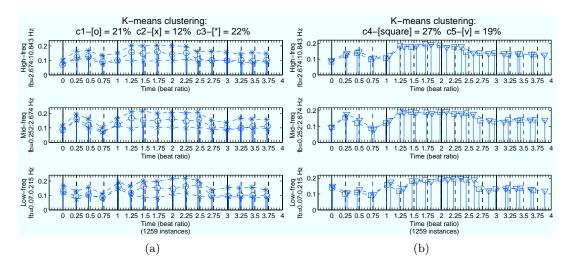
12

Figure 9: (a) Clusters c1, c2 and c3 (out of 5 clusters) calculated for 1259 instances of metric level 4-beat. (b) Clusters c4 and c5 (out of 5 clusters) calculated for 1259 instances of metric level 4-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

3.4   Clusters in metric level 4-beat

The clustering for the metric level 4-beat resulted in a solution of 5 clusters. Figure 9a shows the centroids of clusters c1, c2 and c3. Figure 9b shows the results for clusters c4 and c5. The metric level 4-beat includes all the main characteristics observed in metric levels 1- and 2-beat, with special attention to the deviations of peak positions. The profile of peak intensities seems to be quite similar for all clusters, including clusters c4 and c5. Clusters c1, c2 and c3 seem be discriminated by means of profiles of peak intensity. Cluster c2 seem to be more attenuated while cluster c1 and c3 display higher loudness curves.

Clusters c4 and c5 display the same pattern observed in the metric level 2-beat (Figure 8b). Results for cluster c4 indicates that 27% of the instances are grouped in a continuous acceleration profile that reaches up to -0.12 beats of anticipation in the last semiquaver in the high-spectrum region ($F(2, 597) = 2.8646, p < 0.057$). Although the deceleration pattern of cluster c5 represent only 19% of the instances the last peak position in this cluster reaches up to -0.18 beats in the last semiquaver (105 ms for the average tempo of 103 BPM).

Mean deviations (over 3 spectral regions) from mathematical positions for clusters c4 and c5 are displayed in Figure 10. The data shows a significant tendency of accelerations and decelerations but also an increasing level of variance. The microtiming positions that correspond to quarter-note beats (i.e. positions 1, 5, 9 and 13 on Figures 10a and 10b) show less tendencies of deviations, which may be attributed to a tendency to "keep track" of the quarter-note beat while applying accelerations and decelerations.
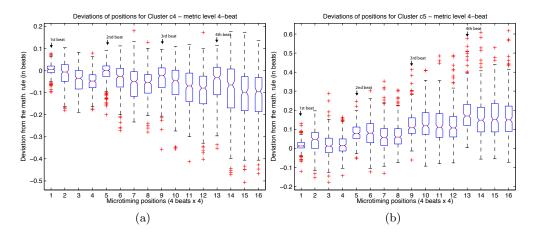
13

Figure 10: Mean deviations from the mathematical subdivisions for clusters c4 and c5.

4 Discussion

In this study we analyzed the interaction between microtiming, meter, intensity and spectral estimations of timbre.

The results confirmed the tendency of anticipations of the third and fourth semi-quavers at all metric levels (all quarter-notes) and spectral regions. This objectively confirms the existence of a systematic artifact described in previous studies about microtiming in samba music and other Afro-Brazilian musical traditions (Gerischer, 2006; Lindsay and Nordquist, 2007; Lucas, 2002; Gouyon, 2007).

We also provided indications of the existence of rhythmic devices that may characterize samba music which, to the best of our knowledge, have not been reported to date: (1) a small delay of instruments in the lower end of the spectrum on the first semiqua-ver of each beat, particularly on the second beat in a bar, and (2) systematic forms of *accelerando* and *ritardando* at a microtiming level.

These results raise several interesting hypotheses. The anticipation of the third and fourth semiquavers and the delay of the first semiquavers show a tendency of approximation of semiquaver rhythms towards triplet rhythmic figures. The coexistence of triplet rhythms with binary divisions are reported on several references on samba music (Daniel, 2006; Browning, 1995; Santos Neto, 2010; Kubik, 1990) and other musical cultures of the African diaspora (e.g., Schwartz and Fouts, 2003; Temperley, 2000). The effect of the coexistence of binary and ternary divisions could be a strategy to induce tension, ambiguity and flexibility in the rhythmic texture. Tension, for example, could be a way of bringing attention to specific performances and personal styles (see, for example, the concept of participatory discrepancies in Keil, 1987, 1995) or a mechanism of making the musical texture more interesting by creating a dialog between expected and unexpected rhythms. Ambiguity could reinforce the polymetric and polyrhythmic characteristics of samba music, which may act as an inductor of body movements (Browning, 1995; Sodré,

1979) or as an impulse to use dance gestures as a form of metrical disambiguation (see Naveda and Leman, 2009; Naveda, 2011). The temporal flexibility provoked by microtiming deviations could provide a temporal grid that is flexible enough to accommodate (and invite) participation of newcomers in the social displays of Afro-Brazilian practices but sufficiently challenging and idiosyncratic[1] to be recognized and performed in high-level performance renditions (see Vassberg, 1976; Chernoff, 1979, for a discussion about participation in African and Afro-Brazilian musical contexts) .

While it is well known that commetric beat patterns in samba are performed by percussion instruments such as *surdo* or *tantã* and accentuated in the second beat (Sandroni, 2001; Moura, 2004; Chasteen, 1996), (which is also reflected in our results), we were unable to find references to any systematic delay of such percussion instruments on the first semiquaver. Neither could we find references to the observation that, in a bar, the first semiquaver of the *second* quarter-note beat is significantly more delayed than that of the first quarter-note beat.

This hypothetical observation should be interpreted with caution however. The temporal range of delays in the low-frequency spectrum is very close to the sampling period of the auditory model (5 ms), which means that minimum significant delays found in the Figure 8a, for example, account for only 2-samples (10 ms) between the mathematical and actual peak positions. We focus on *relative* position, hence would argue that our observation does hold, however, more research would be welcome to support this observation.

With regards to *accelerando* and *ritardando*, we should consider that the computation of clusters may have merged two recurrent tendencies of outliers in the data-set. Nevertheless, the percentage of the instances represented by these clusters (c3-15% and c4-11%), similar cluster structures found in other metric levels above 2-beat (4-beat level), and the significance of these distributions (see Figure 9b), seem to indicate that they reflect real microtiming structures present in our data. If this hypothesis were to be confirmed, this could indicate that samba exhibits rhythmic devices similar to *accelerando* and *ritardando* forms, at microtiming level. Although these rhythmic artifacts are widely used to delimit phrases, endings and formal articulations in classical music (macro-time level), it is surprising that such devices would appear at the level of microtiming.

The variation of intensities demonstrate that microtiming in samba is subjected to interactions with accents and metrical structure. The flatness of semiquaver intensities observed in clusters in all metric levels, especially the 2-beat level, indicate the existence of artifacts that evidence the binary meter in the intensity profiles. While the first beat starts with a low-energy semiquaver in the low-frequency region and accents in the second (Figure 7) and fourth semiquavers (Figure 8a), the second beat starts with a characteristic strong bass accent, followed by flat and low intensity semiquavers. This

---

[1] For example, acculturated performers woud be recognized by their ability to perform systematic deviations (and interactions). This ability may be linked with subjective qualities attributed to skilled musicians or performances such as the "balanço" (balance), ginga (close to *groove* and related to body movements) or "suingue" (swing). See Gerischer (2006) for other examples.

oscillation of the interactions between beat positions may play an important role in the induction of metric properties.

The use of a psychoacoustically based feature as the main descriptor of the audio domain suggests that these observations may be available as proximal cues in the periphery of the auditory system. Moreover, the results show that microtiming can be understood as a temporal frame where a dynamic network of relationships among musical cues takes place in the performance of samba music. At the same time, microtiming creates tension by disrupting the flow of the *tatum* level it also keeps the metric structure organized via the interactions with patterns of intensity.

5   Conclusion

There are several indications that the perception and performance of timing are involves more information than what is encoded in the temporal structure of musical events. In this study, we used a systematic and explorative approach to reveal some aspects of this intricate node where time, accent, timbre and metrical properties converge. The application of computational approaches to our data-set of commercial samba music confirmed several characteristic microtiming deviations suggested in the literature and revealed other important interactions that enrich the knowledge about timing in samba and the knowledge about timing in the performance of music. The indications of the existence of other characteristics such as the delay of the first semiquaver (low spectrum) and the *accelerando* and *rallentando* microtiming patterns inaugurate new viewpoints on timing aspects that populate the tacit knowledge behind the performance of popular music styles. The multidimensional aspects of the knowledge that moves cultural forms such as the samba may reserve much more factors not easily depicted in traditional approaches to music (e.g., scores).

Note, however, that the present study does not claim an exhaustive overview of multidimensionality of microtiming structures in samba. Furher, the interactions in the context of samba should not be restricted only to musical dimensions encoded in sound. Samba is more than a musical style. It is a complex cultural environment, which inherits the relevance of experiencing timing from the "multiple experience flows" (Stone, 1985), present in the Afro-Brazilian religious rituals that form the background of samba culture (de Carvalho, 2000; Sandroni, 2001; Sodré, 1979). There, not only music and dance are involved, but also imagery, tradition, symbologies and other inter-textual components (Gerischer, 2006, p. 115). A typical description of an Afro-Brazilian Candomblé ceremony illustrates how this "original" experience of timing unfolds in the Afro-Brazilian music and ritual, which convey complexities that are far beyond timing and rhythm:

> "The dancers dance with great violence, energy, and concentration. Getting really involved in the rhythm and movement (. . .) The drummers (. . .) can play certain signals in the rhythmic pattern to cause the dancing to take a violent turn (. . .) One method is for one drum to syncopate the rhythm

slightly (another one maintaining it) such that a strong beat falls just before the main beat. (...) This gives a impression of increased speed when this is not really the case, and creates tension and feeling of imbalance in the listener or dancer" (Walker, 1973, quote in Fryer, 2000)

This example demonstrates how elaborated maps of timing and accents take part in an intricate system of metrical and rhythmic textures and forms of tension that tie together sound and movement. Samba music derives from this original combination of music and movement. How can computational musicology reach and reveal the elements behind these phenomena? How representations of meter and rhythm can adapted to new viewpoints provided by computational approaches to music, movement and image?

More research is needed to elucidate the interplay between descriptive characteristics of samba (e.g. microtiming characteristics in music, in dance) on the one hand, and the production of physical behavior (e.g. dancing, playing) on the other hand. More research is needed to provide reliable onset functions in polyphonic and better methods for the study of microtiming in non-Western music contexts. More studies should focus on the perceptual salience of microtiming structures and its relations to qualitative categories of music.

Acknowledgements

References

Andrade, M. d. (1991). *Aspectos da Música Brasileira*, Volume 11. Belo Horizonte: Villa Rica.

Araújo, S. (1992). Acoustic labor in the timing of everyday life; a social history of samba in Rio de Janeiro (1917-1980).

Arom, S. (1989). Time Structure in the Music of Central Africa: Periodicity, Meter, Rhythm and Polyrhythmics. *Leonardo 22*(1), 91–99.

Bello, J. P., C. Duxbury, M. Davies, and M. Sandler (2004). On the use of phase and energy for musical onset detection in the complex domain. *IEEE Signal Processing Letters 11*(6), 553–556.

Benadon, F. (2003). The expressive role of beat subdivision in jazz. In *Conference Proceedings of the Society of Music Perception and Cognition*, Las Vegas.

Benadon, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology 50*(1), 73–98.

Benadon, F. (2009). Time Warps in Early Jazz. *Music Theory Spectrum 31*(1), 1–25.

Bilmes, J. A. (1993). Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Expressive Timing in Percussive Rhythm. Master's thesis, MIT, Massachusetts.

Browning, B. (1995). *Samba: Resistance in Motion*. Indiana University Press.

Cannam, C., C. Landone, M. Sandler, and J. Bello (2006). The sonic visualiser: A visualisation platform for semantic descriptors from musical signals. In *Proceedings of the 7th International Conference on Music Information Retrieval*, pp. 324–327. Citeseer.

Chasteen, J. C. (1996). The prehistory of Samba: Carnival Dancing in Rio de Janeiro, 1840-1917. *Journal of Latin American Studies 28*(1), 29–47.

Chernoff, J. M. (1979). *African rhythm and African sensibility: aesthetics and social action in African musical idioms*. University of Chicago Press.

Chernoff, J. M. (1991). The Rhythmic Medium in African Music. *New Literary History 22*(4), 1093–1102.

Daniel, G. (2006). Educação musical a distância: tecnologia, velocidade e desaceleração. In *XVI Congresso da Associação Nacional de Pesquisa e Pós-graduação em Música*.

de Carvalho, J. J. (2000). *Afro-Brazilian Music and Ritual [s]*. Duke-University of North Carolina Program in Latin American Studies.

Desain, P. and H. Honing (1989). The quantization of musical time: A connectionist approach. *Computer Music Journal 13*(3), 56–66.

Desain, P. and H. Honing (1993). Tempo curves considered harmful. *Contemporary Music Review 7*(2), 123–138.

Dixon, S. (2007). Evaluation of the Audio Beat Tracking System BeatRoot. *Journal of New Music Research 36*(1), 39–50.

Friberg, A. and A. Sundstrom (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception 19*(3), 333–349.

Fryer, P. (2000). *Rhythms of Resistance: African Musical Heritage in Brazil*. London: Pluto.

Fulop, S. and K. Fitz (2006). Algorithms for computing the time-corrected instantaneous frequency (reassigned) spectrogram, with applications. *The Journal of the Acoustical Society of America 119*, 360.

Gabrielsson, A. (1985). Interplay between analysis and synthesis in studies of music performance and music experience. *Music Perception*, 59–86.

Galinsky, P. (1996). Co-option, cultural resistance, and Afro-Brazilian identity: A history of the pagode samba movement in Rio de Janeiro. *Revista de música latinoamericana 17*(2), 120–149.

Gerischer, C. (2006). O suingue baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion. *Ethnomusicology 50*(1), 99–119.

Gouyon, F. (2007). Microtiming in "Samba de Roda"—Preliminary experiments with polyphonic audio. In *Proceedings of the XII Simpósio da Sociedade Brasileira de Computação*, São Paulo, Brazil. Sociedade Brasileira de Computação Musical.

Grachten, M. and G. Widmer (2009). The kinematic rubato model as a means of studying final ritards across pieces and pianists. In *Proc. Sixth Sound and Music Computing Conference (SMC 2009)*, pp. 173–178.

Honing, H. (2001). From time to time: The representation of timing and tempo. *Computer Music Journal 25*(3), 50–61.

Honing, H. (2003). The Final Ritard: On Music, Motion, and Kinematic Models. *Computer Music Journal 27*(3), 66–72.

Johansson, M. (2005). Interpreting micro-rhythmic structures in norwegian traditional fiddle music. In *Rhythm and Micro-rhythm: Investigating musical and cultural aspects Rhythm and Micro-rhythm: Investigating musical and cultural aspects of groove-oriented music*, Oslo.

Kauffman, R. (1980). African Rhythm: A Reassessment. *Ethnomusicology 24*(3), 393–415.

Keil, C. (1987, August). Participatory Discrepancies and the Power of Music. *Cultural Anthropology 2*(3), 275–283.

Keil, C. (1995). The theory of participatory discrepancies: A progress report. *Ethnomusicology*, 1–19.

Kubik, G. (1979). Angolan Traits in Black Music, Games and Dances of Brazil. A Study of African Cultural Extensions Overseas. *Estudos de Antropologia Cultural. Junta de Investigaçoes Cientiacas de Ultramar Lisboa* (10), 1–55.

Kubik, G. (1990). Drum Patterns in the "Batuque" of Benedito Caxias. *Latin American Music Review/Revista de Música Latinoamericana 11*(2), 115–181.

Lindsay, K. and P. Nordquist (2007). Pulse and swing: Quantitative analysis of hierarchical structure in swing rhythm. *The Journal of the Acoustical Society of America 122*, 2945–2946.

London, J. (2004). *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford University Press, USA.

Lucas, G. (2002). Os sons do Rosário: o congado mineiro dos Arturos e Jatobá. *Belo Horizonte: Editora UFMG*.

Madison, G. (2006). Experiencing Groove Induced by Music: Consistency and Phenomenology. *Music Perception 24*(2), 201–208.

McGuiness, A. (2006). *Microtiming deviations in groove*. Ph. D. thesis, Australian National University, Canberra.

Moura, R. (2004). *No princípio, era a roda: um estudo sobre samba, partido-alto e outros pagodes*. Rio de Janeiro: Rocco.

Naveda, L. (2011, January). *Gesture in Samba: A cross-modal analysis of dance and music from the Afro-Brazilian culture*. Phd thesis, Ghent University.

Naveda, L. and M. Leman (2009). A Cross-modal Heuristic for Periodic Pattern Analysis of Samba Music and Dance. *Journal of New Music Research 38*(3), 255–283.

Naveda, L., M. Leman, F. Gouyon, A. Barbosa, and X. Serra (2009). Accessing structure of samba rhythms through cultural practices of vocal percussions. In *Proceedings of the 6th Sound and Music Computing Conference*, Portugal, pp. 259–264.

Palmer, C. (1997). Music performance. *Annual Review of Psychology 48*(1), 115–138.

Pelleg, D. and A. W. Moore (2000). X-means: Extending K-means with efficient estimation of the number of clusters. pp. 727–734. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.

Polak, R. (1998). Jenbe Music in Bamako: Microtiming as Formal Model and Performance Practice. *Iwalewa Forum 2*, 23–42.

Rosinach, V. and C. Traube (2006). Measuring swing in irish traditional fiddle music. In *Proc. International Conference on Music Perception and Cognition*, pp. 1168–1171.

Sandroni, C. (1996). Mudanças de padrao rtmico no samba carioca, 1917-1937. *Revista Transcultura de Musica 2*.

Sandroni, C. (2001). *Feitiço decente: transformações do samba no Rio de Janeiro, 1917-1933*. Jorge Zahar Editor: Editora UFRJ.

Santos Neto, J. (2010). Ginga: a Brazilian way to groove.

Schwartz, K. D. and G. T. Fouts (2003). Music Preferences, Personality Style, and Developmental Issues of Adolescents. *Journal of Youth and Adolescence 32*(3), 205–213.

Sethares, W. A. (2007). *Rhythm and transforms*. Berlin: Springer.

Sodré, M. (1979). *Samba, O Dono do Corpo*. Rio de Janeiro: Codecri.

Stone, R. M. (1985). In Search of Time in African Music. *Music Theory Spectrum*, 139–148.

Temperley, D. (2000). Meter and Grouping in African Music: A View from Music Theory. *Ethnomusicology 44*(1), 65–96.

Todd, N. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America 97*, 1940–1940.

Toussaint, G. T. (2005). The Euclidean algorithm generates traditional musical rhythms. *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, 47–56.

Van Immerseel, L. M. and J. P. Martens (1992). Pitch and voiced/unvoiced determination with an auditory model. *The Journal of the Acoustical Society of America 91*, 3511–3526.

Vassberg, D. E. (1976). African Influences on the Music of Brazil. *Luso-Brazilian Review 13*(1), 35–54.

Walker, S. S. (1973). *Ceremonial spirit possession in Africa and Afro-America: Forms, meanings, and functional significance for individuals and social groups*. Brill Archive.

Witten, I. H. and E. Frank. WEKA Software, v3. 5.2. *University of Waikato*.