This article was downloaded by: [b-on: Biblioteca do conhecimento online INESC-Porto], [Fabien Gouyon]
On: 02 November 2011, At: 05:46
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Journal of New Music Research
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/nnmr20

Microtiming Patterns and Interactions with Musical Properties in Samba Music

Luiz Naveda [a] , Fabien Gouyon [b] , Carlos Guedes [b] & Marc Leman [a]
[a] Ghent University, Belgium
[b] INESC Porto, Portugal

Available online: 19 Oct 2011

To cite this article: Luiz Naveda, Fabien Gouyon, Carlos Guedes & Marc Leman (2011): Microtiming Patterns and Interactions with Musical Properties in Samba Music, Journal of New Music Research, 40:3, 225-238

To link to this article: http://dx.doi.org/10.1080/09298215.2011.603833

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Journal of New Music Research
2011, Vol. 40, No. 3, pp. 225–238

Routledge
Taylor & Francis Group

Microtiming Patterns and Interactions with Musical Properties in Samba Music

Luiz Naveda[1], Fabien Gouyon[2], Carlos Guedes[2] and Marc Leman[1]

[1]Ghent University, Belgium; [2]INESC Porto, Portugal

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

Abstract

In this study, we focus on the interaction between microtiming patterns and several musical properties: intensity, meter and spectral characteristics. The data-set of 106 musical audio excerpts is processed by means of an auditory model and then divided into several spectral regions and metric levels. 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. The results confirm previously reported anticipations of the third and fourth semiquavers in a beat. 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 at 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 two-beat and four-beat phrases. The results demonstrate the importance of multidimensional interactions with timing aspects of music. However, 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 (Carvalho, 2000, p. 134). When it comes to analysis of samba music, and in general music from the African diasporas, the majority of approaches concentrate on studying invariant properties of *rhythm*. However, rhythm involves a number of different aspects. It is said 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. Palmer, 1997; London, 2004; Sethares, 2007), as well as aspects not encoded in sound such as motor-schemes (Todd, 1995; Palmer, 1997), kinematic models (e.g. Todd, 1995; Palmer, 1997; Honing, 2003; Grachten & Widmer, 2009) and other modalities such as dance (Chernoff, 1991; Naveda & Leman, 2009). How do the actual representations of samba style deal with such a multi-dimensionality of rhythm? How could computer music help to detect meaningful invariants in this context?

The majority of analytical studies of samba are based on or modelled on 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 by some researchers that samba music is characterized by a binary metric structure (binary bar) 'muted' in the first beat but with a stress 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 with different periodic lengths and metric phases coexist (Browning, 1995; Fryer, 2000). A great part of the literature refers to the general concept of syncopation or the figure of syncope (e.g. Sodré, 1979; Sandroni, 1996)

Correspondence: Luiz Naveda, IPEM, Ghent University, Blandijnberg 2, Ghent, 9000 Belgium. E-mail: luiznaveda@gmail.com

http://dx.doi.org/10.1080/09298215.2011.603833 © 2011 Taylor & Francis

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

as the main characteristic of samba. In particular, many authors have proposed rhythm 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' (Kubik, 1979; Galinsky, 1996).

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 by means of more explicit knowledge or written documents (Sodré, 1979; Carvalho, 2000; Sandroni, 2001). The strong link between samba and Afro-Brazilian religious rituals (Carvalho, 2000) and social displays such as the *roda-de-samba* (Moura, 2004) indicates that samba cannot be easily detached from the experience of dance, rituals, texts as well as from its 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 labour (e.g. Fryer, 2000) or hand clapping play an active role in the ellaboration of music performances. From this socio-historical perspective, rhythm should be understood as a concept formed by (1) a 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 in the perception of *timing* and action in time.

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 to in the literature as 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 as *microtiming*, and are defined as a series of event shifts at a constant tempo (Bilmes, 1993; Desain & Honing, 1993).

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

Several studies examine aspects of microtiming in samba music. Lindsay and Nordquist (2007) analysed microtiming in recordings of samba instruments such as *pandeiro*, *surdo* and *agogô*. Lindsay and Nordquist used an improved spectrogram analysis inspired in Fulop and Fitz (2006) as basic signal representation, which was combined with a manual annotation of events. They found systematic anticipations of the third and fourth semiquavers (within one beat) for the *pandeiro* recordings inside pairs of 'short-long' onsets. They also found four-beat patterns of onsets in progressive acceleration. Naveda, Leman, and Gouyon (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)). They 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*. Even though the *Congado* traditions are accompanied by musical forms stylistically different 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) analysed commercial recordings of samba. He identified patterns of microtiming deviations by means of machine learning techniques applied to the 'complex spectral difference', which was suggested in Bello, Duxbury, Davies, and Sandler (2004) as an onset detection function. Results also indicated the existence of systematic anticipations of the third and fourth semiquavers. This overview accounts for evidence of systematic deviations that seem to occur on the third and fourth semiquavers at the beat level in samba music. However, most studies are based on a small number of samples and most analyses rely on manual annotation of events or windowed FFT methods, whose temporal precision does not permit a reliable analysis to be made, specially in low-frequency components. Most importantly, most of these studies *only* consider (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.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 09:40 02 November 2011

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 in Section 5 the contributions and implications of this study are summarized.

2. Method

2.1 Data-set

Our data-set consists of 106 excerpts of music collected from commercial CDs. The median of duration 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 44,100 Hz, 16 bits and normalized by amplitude.

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 within 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 a sample frequency at 200 Hz, distributed over 22 critical bands (centre 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 consists of two parts: (1) the process of segmentation of the spectrum range in the frequency domain, which averages out auditory curves to three *spectrum regions* (low-, mid- and high-frequency spectrum) and (2) the process of segmen-

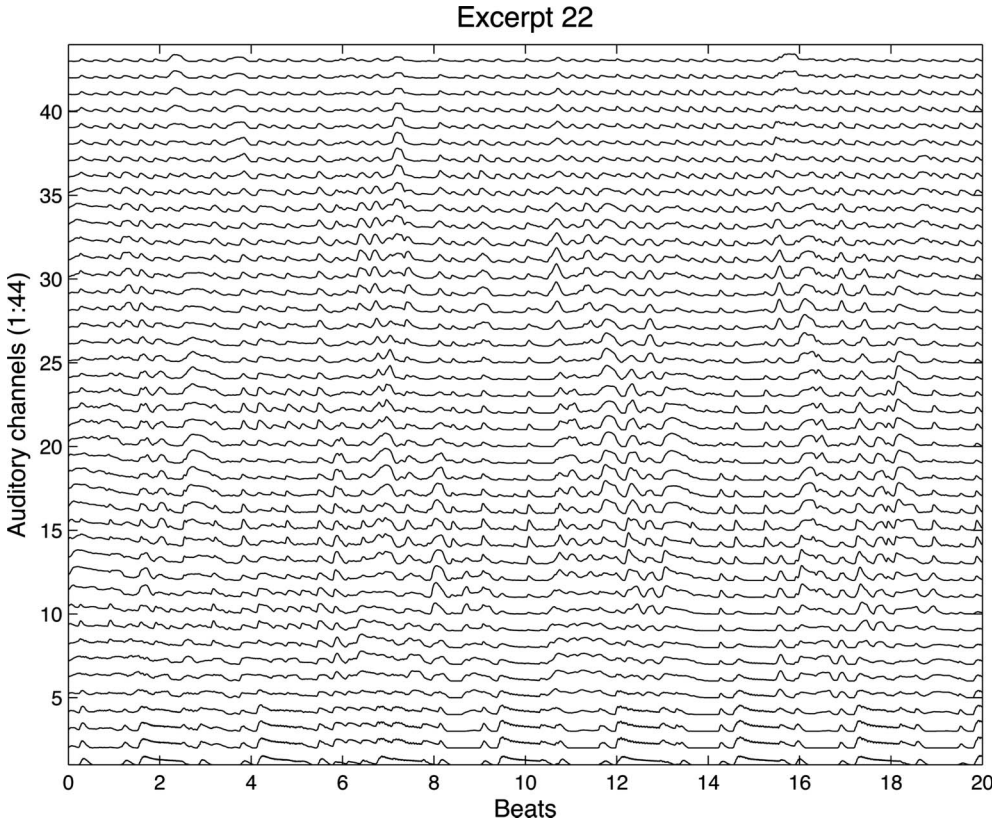

Fig. 1. Example of loudness curves generated by auditory model. The 44 envelope curves represent a simulation of loudness on each auditory channel (for more details, see Van Immerseel and Martens (1992)).

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

tation of three metric levels in the temporal domain, which segments the features into segments of lengths 1, 2 and 4 beats. The processes of spectral and metric segmentation are illustrated in Figure 2.

2.2.2.1 Segmentation in spectral regions. The data-set consists of excerpts of commercial polyphonic music which makes the separate instrumental sources unavailable. Current state-of-the-art source separation techniques are prone to bias and to generation of artifacts that could disturb the detection of microtiming positions. However, current knowledge indicates that percussion instruments used in the samba ensemble 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 within the frequency domain or, in our case, by loudness in time distributed among auditory channels.

In Figure 3, for example, the spectrum of the low-frequency samba drum (the *surdo*) is 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 signatures of *Ganzás* and different kinds of shakers cover higher portions of the spectrum. Although the frequency components of these instruments overlap with each other in the time and frequency domains (particularly during transients in the attack points), the spectrum signature of each timbre is relatively discriminated from each other. Therefore, for each excerpt, we averaged out the 44 loudness curves provided by the auditory model to three loudness curves that mirror estimated distributions of instruments in the musical tessitura: low-frequency region (channels 1:6), mid-frequency region (channels 7:30) and high-frequency region (channels 31:44). To read more about a similar procedure, see Lindsay and Nordquist (2007).

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

In order to identify the time points of the metrical accents, we performed the annotation of beat (one beat) and bar (two beats) levels of all audio excerpts in the dataset. Automatic beat annotation using software such as BeatRoot (Dixon, 2007) and QMUL beat tracking plugins for Sonic Visualiser (Cannam, Landone, Sandler, & Bello, 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 present in our data-set. Figure 4 shows the distributions of the BPM values within the data-set. Normality observed in the histogram shows a tendency of tempi towards 103 BPM (mean = 103.02, standard deviation = 18.42).

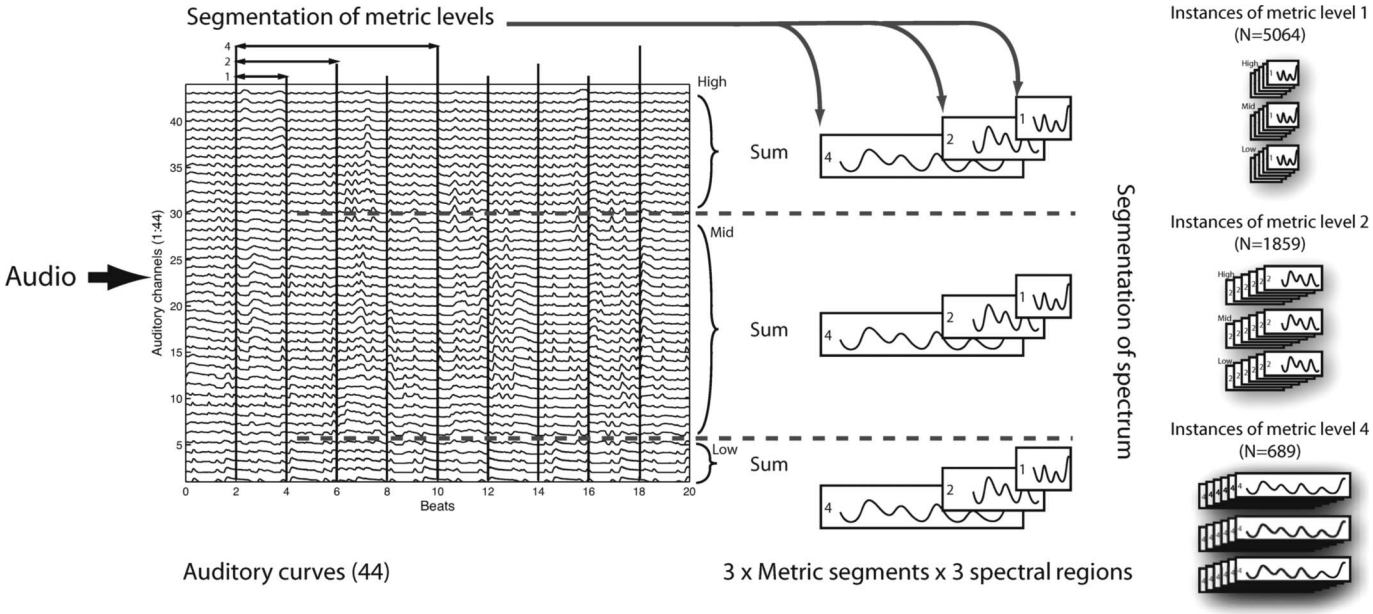

Fig. 2. Two processes of segmentation of the auditory curves: metric levels and spectrum. The segmentation results in a collection of *N* instances for each metric level, divided by the spectrum region. The three spectral regions are also represented in this phase.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 09:40 02 November 2011

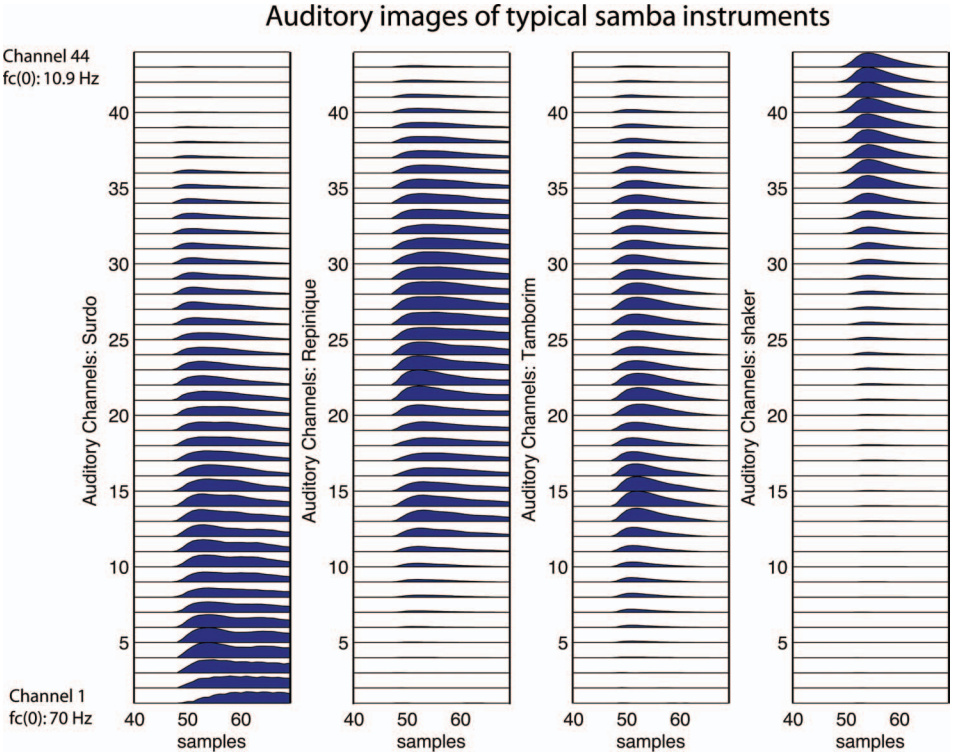

Fig. 3. Response of the auditory model for attacks (80 samples = 0.4 s) of the instruments *surdo*, *repinique*, *tamborim* and shaker (from left to right). The graphs demonstrate how the auditory model responds to different instruments of a traditional samba ensemble.

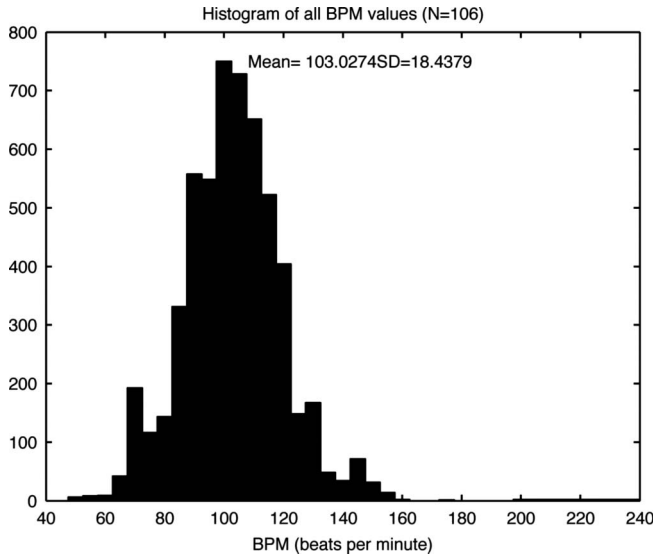

Fig. 4. Histogram of BPM values of the beats for the whole data-set (106 excerpts, 5064 beats).

These annotations are used to define three different types of segments containing respectively one, two and four 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 semiquaver 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. On the one side, manual beat annotation does not provide a precise segmentation of the beat (due to bias of the manual process). On the the other side, the automatic detection of onset attacks of the first semiquaver results in three different onset points for each spectrum region (which respond to discrepancies between attacks of instruments). Therefore, we opted to set the beat position in the average position of the three spectral regions of the first semiquaver, which permits the calculation of the beat period and microtiming segments *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.

Downloaded by [b-on: Biblioteca do conhecimento online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

Finally, for every 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 denoted p) and (2) the intensity of the peaks (denoted i). For n semiquavers in a given metric level, the instances will contain $p_{1,\ldots,n}$ peak positions and $i_{1,\ldots,n}$ peak intensities, in each spectral region $r_{1,\ldots,3}$. Table 2 specifies the structure of the instances further analysed in the next section. Note that instances which feed the clustering algorithm combine information on spectrum, timing and intensities, and are processed in three different metric levels or segments. The process of clustering applied to each metric level leads to three different groups of results, displayed in Sections 3.2, 3.3 and 3.4 (metric levels one- two- and four-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.

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

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 within the proximity of the first beat manual annotation, in *each spectral region*.
Phase 4	In each spectral region, select highest peak situated above threshold around the first beat (if there are no peaks above threshold, retrieve NaN).
Phase 5	Compute average peak position of the three spectral regions.
Phase 6	Retrieve position and amplitude of the highest peak in close proximity of *each* semiquaver, in *each spectral region*.

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 the Weka platform (Hall et al., 2009). Using this method, one is able to search for locations and numbers of clusters that efficiently improve the Bayesian

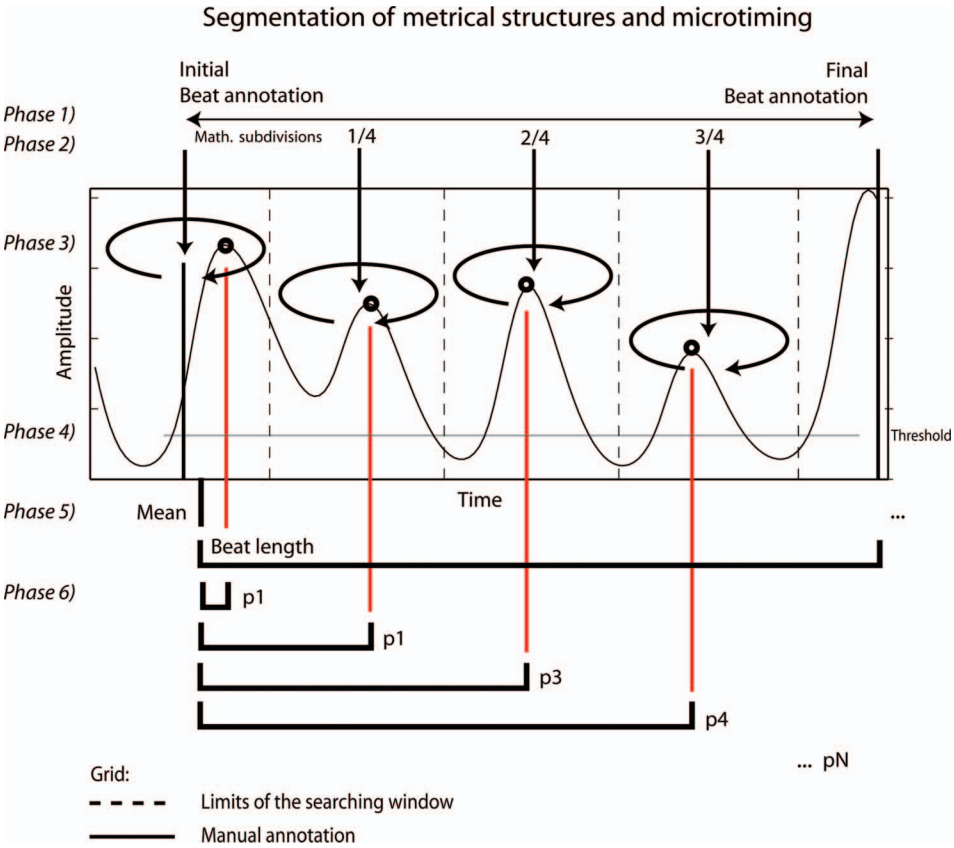

Fig. 5. Description of the heuristic of calculation of microtiming deviations. Example of one-beat metric level, and single spectral region only represented. See Table 1 for the explanation for each step of the algorithm.

Downloaded by [b-on: Biblioteca do conhecimento online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

Information Criterion (BIC) or the Akaike Information Criterion (AIC) measure. The algorithm was configured to retrieve a minimum of three and a maximum of five clusters (arbitrary).

3. Results

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

3.1 Microtiming distributions in metric level one-beat

The results displayed in Figure 6 show an overview of the main microtiming tendencies for metric level one-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 (Gouyon, 2007; Lindsay & Nordquist, 2007; Naveda & Leman, 2009). 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 (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 7.3 ms when reported to the average BPM.

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

Metric level 1-beat: $[p_1, \ldots, {}_4, i_1, \ldots, {}_4]_{r_{1,\ldots,3}}$
• 12 positions + 12 intensities = 24 elements
Metric level 2-beat: $[p_1, \ldots, {}_8, i_1, \ldots, {}_8]_{r_{1,\ldots,3}}$
• 24 positions + 24 intensities = 48 elements
Metric level 4-beat: $[p_1, \ldots, {}_{16}, i_1, \ldots, {}_{16}]_{r_{1,\ldots,3}}$
• 48 positions + 48 intensities = 96 elements

3.2 Clusters in metric level one-beat

The cluster analysis of the one-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 visualization 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

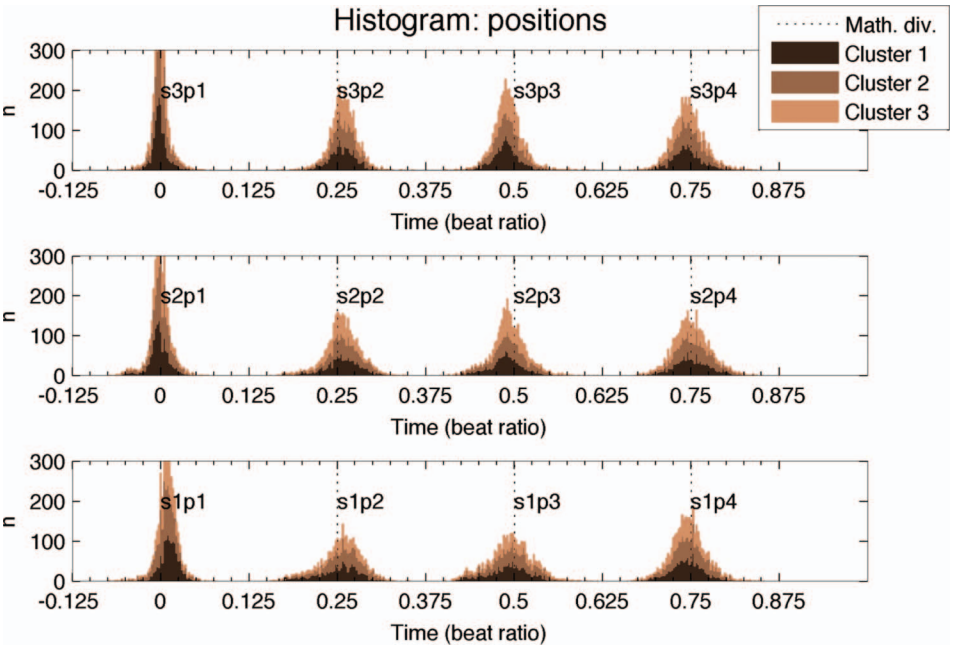

Fig. 6. Distributions of the peak positions for metric level one-beat, for all excerpts ($N = 5064$, 106 excerpts). The shades of grey indicate the contribution of each cluster to the total distribution. The vertical grid indicates the mathematical subdivisions of the beat ($0, \frac{1}{4}, \frac{2}{4}$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

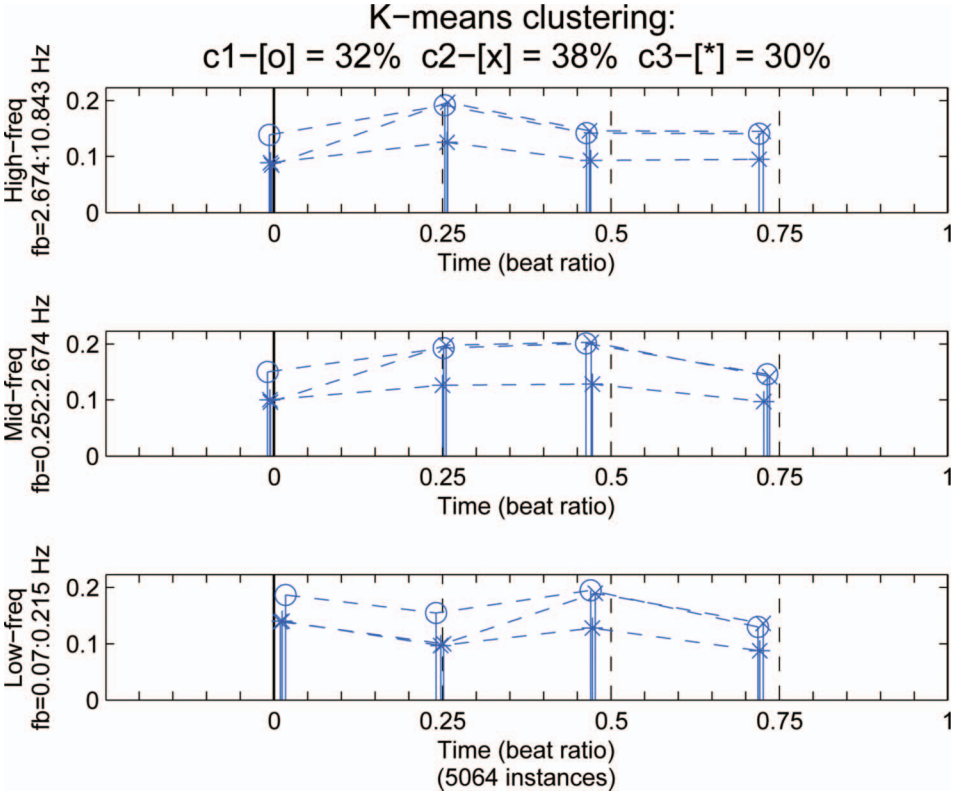

Fig. 7. Cluster centroids c1, c2 and c3 for 5064 instances of metric level one-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

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 seems 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 two-beat

The cluster analysis of the metric level two-beat resulted in five clusters. Figure 8(a) shows clusters c1, c2 and c5 for metric level two-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 fourth) semiquaver in the low-frequency region, as seen in the metric level one-beat. This observation seems to affect both quarter-note beats at the bar level but in different ways: 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 one-beat level). The first semiquaver of the *second* beat is significantly more delayed than that of the first quarter-note beat. The former also has more intensity with respect to the latter, confirming the tendency to accentuate the second beat, as mentioned in the literature (Chasteen, 1996; Sandroni, 2001; Moura, 2004).

Peak intensities reveal more variability at this metric level. While the intensity profile of the second semiquaver (first beat) seems to be accentuated only in the mid-frequency region, the profile of 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 to the other clusters.

Figure 8(b) shows the results of the clusters c3 and c4. These results differ from those of clusters c1, c2 and c5 because they show increasing deviations accumulating over 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.

3.4 Clusters in metric level four-beat

The clustering process applied to the instances of the metric level four-beat resulted in a solution of five clusters. Figure 9(a) shows the centroids of clusters c1, c2 and c3. Figure 9(b) shows the results for clusters c4 and c5. The metric level four-beat includes all the main

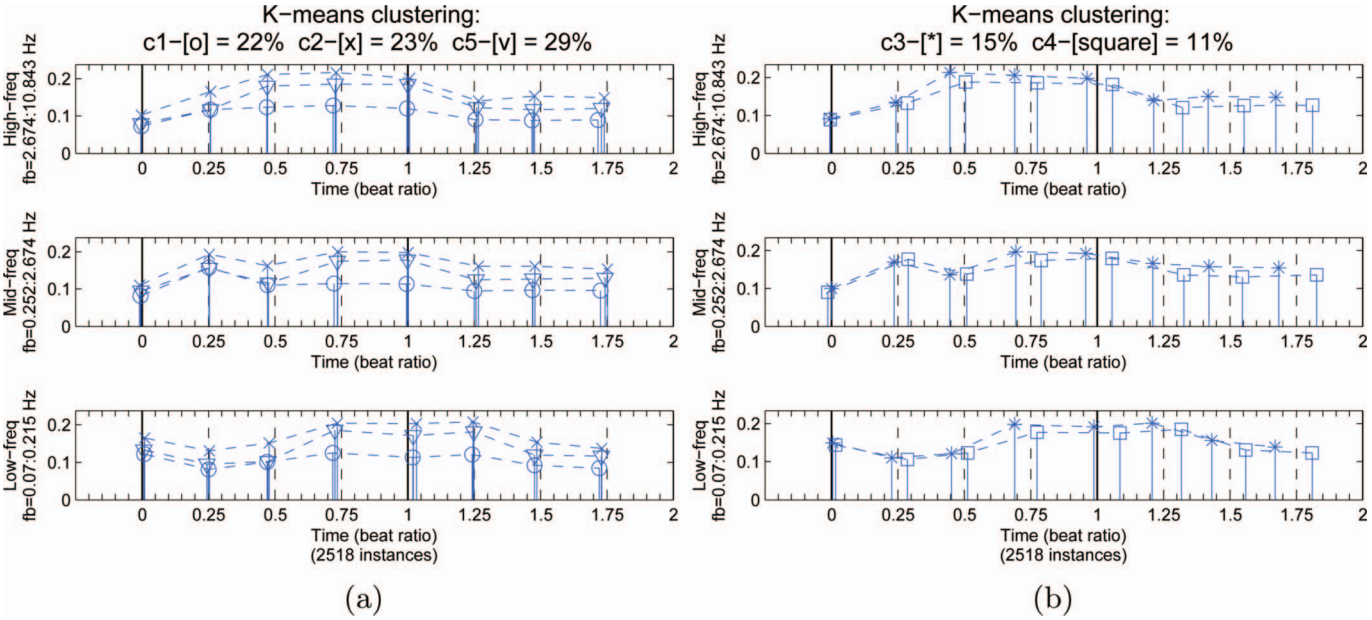

Fig. 8. (a) Cluster centroids c1, c2 and c5 (out of five clusters) calculated for 2518 instances of metric level two-beat. (b) Cluster centroids c3 and c4 (out of five clusters) calculated for 2518 instances of metric level two-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

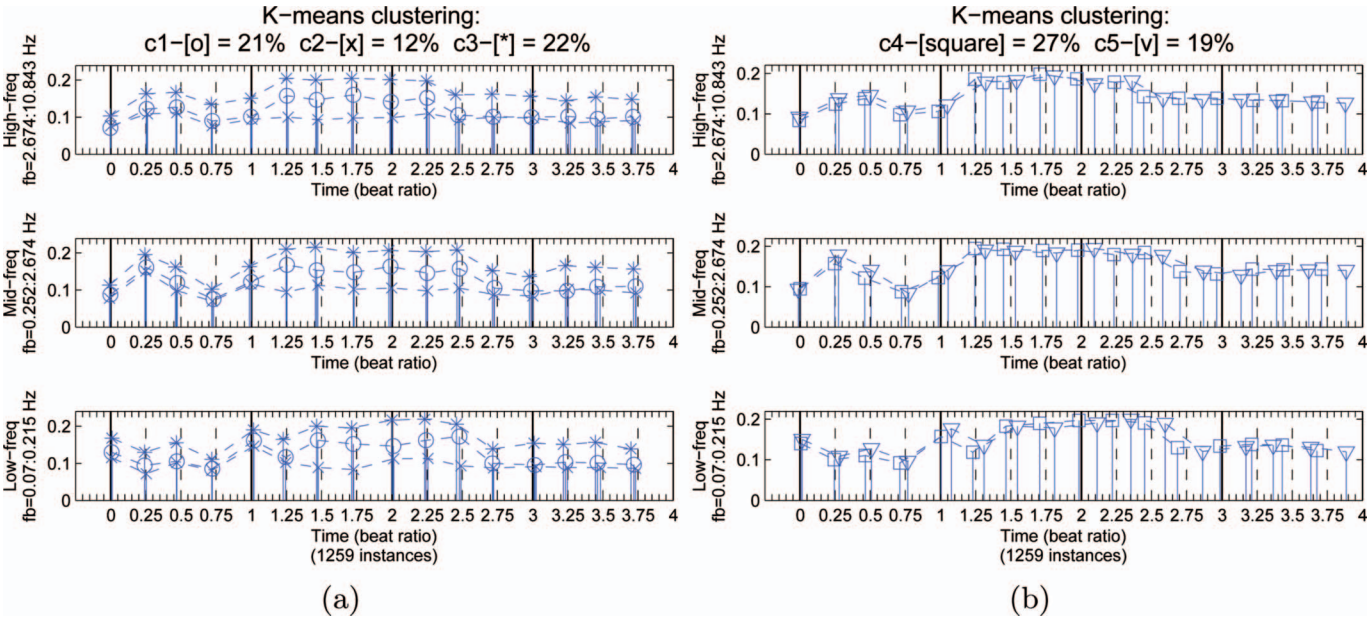

Fig. 9. (a) Cluster centroids c1, c2 and c3 (out of five clusters) calculated for 1259 instances of metric level four-beat. (b) Cluster centroids c4 and c5 (out of five clusters) calculated for 1259 instances of metric level four-beat. Ticks represent 0.05 beats. Vertical traced lines indicate mathematical divisions of the beat.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

Downloaded by [b-on: Biblioteca do conhecimento online INESC-Porto], [Fabien Gouyon] at 03:40 02 November 2011

characteristics observed in metric levels one- and two-beat, especially 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 differentiated by their profiles of peak intensity. Cluster c2 seems to be more attenuated while clusters c1 and c3 display higher loudness curves.

Clusters c4 and c5 display the same pattern observed in the metric level two-beat (Figure 8(b)). Results for cluster c4 indicate 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 represents only 19% of the instances, the last peak position in this cluster reaches up to namely -0.18 beats in the last semiquaver (105 ms for the average tempo of 103 BPM).

Mean deviations (over three spectral regions) from mathematical positions for clusters c4 and c5 are displayed in Figure 10. The data shows a significant tendency towards acceleration and deceleration but also an increase in the level of variance. The microtiming positions that correspond to quarter-note beats (i.e. positions 1, 5, 9 and 13 on Figures 10(a) and (b)) show less tendencies towards deviations, microtiming. which may be attributed to a tendency to signal quarter-note beats during the processes of acceleration and deceleration.

4. Discussion

In this study, we analysed the interaction between microtiming, meter, intensity and spectral estimations of timbre. The results confirmed the tendency towards anticipations of the third and fourth semiquavers 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 (Lucas, 2002; Gerischer, 2006; Gouyon, 2007; Lindsay & Nordquist, 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 at the lower end of the spectrum on the first semiquaver 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 put forward several interesting hypotheses. The anticipation of the third and fourth semiquavers and the delay of the first semiquavers may indicate a tendency of approximation of semiquaver rhythms towards triplet rhythmic figures. The coexistence of triplet rhythms with binary divisions are mentioned in several references on samba music (Kubik, 1990; Browning, 1995; Daniel, 2006; Santos Neto, 2010) and other musical cultures of the African diaspora (e.g. Temperley, 2000; Schwartz & Fouts, 2003). 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 enhancing attention to specific performances and personal styles (see, for example, the concept of participatory discrepancies in Keil (1987, 1995)). It could also provide a mechanism for making the musical texture more interesting by creating a dialogue between expected

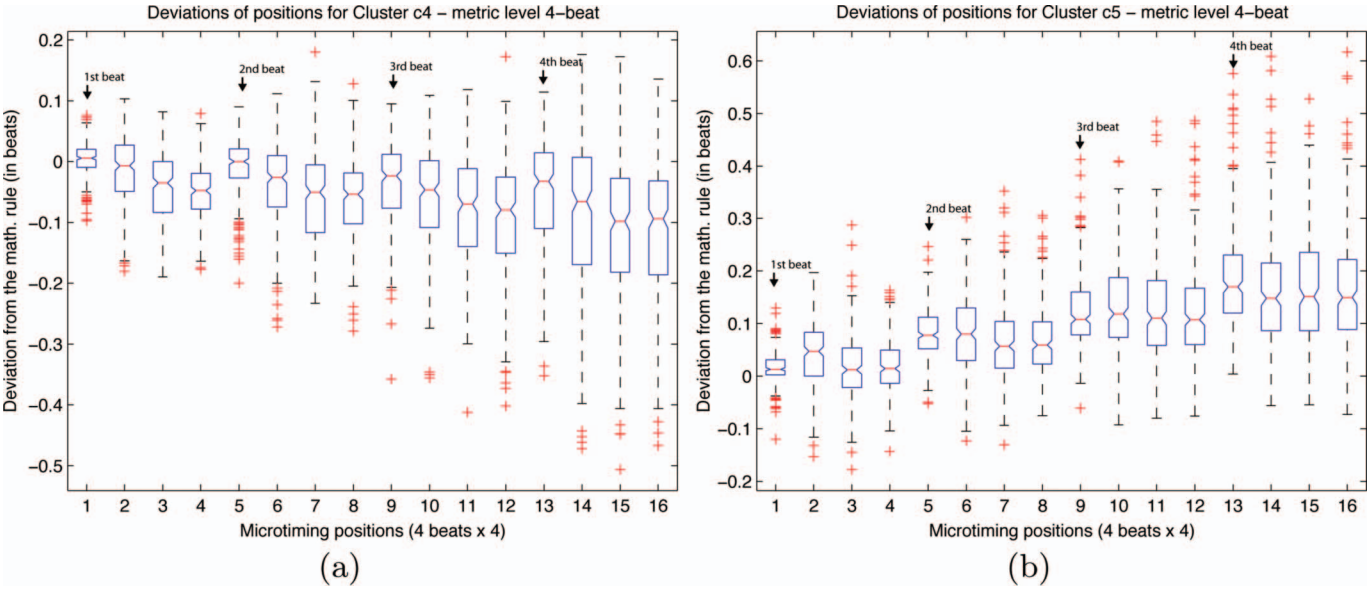

Fig. 10. Mean deviations from the mathematical subdivisions for clusters c4 (a) and (b) c5.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

and unexpected rhythms. Ambiguity could enhance the polymetric and polyrhythmic characteristics of samba music, which may act as an inductor of body movements (Sodré, 1979; Browning, 1995) or as an impulse to use dance gestures as a form of metrical disambiguation (see Naveda & Leman, 2009; Naveda, 2011). The temporal flexibility caused 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) and 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 (Chasteen, 1996; Sandroni, 2001; Moura, 2004), (which is also reflected in our results), we were unable to find previous 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 however be interpreted with caution. 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 Figure 8(a), for example, account for only two samples (10 ms) between the mathematical and actual peak positions. Because we focus on *relative* position, we would argue that our observation does stand on its own. However, more research would further support this observation.

With regards to *accelerando* and *ritardando*, we should consider that the computation of clusters may have merged two recurring tendencies resulting from 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 two-beat (four-beat level), and the significance of these distributions (see Figure 9(b)), 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 the 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 at all metric levels, especially the two-beat level, indicate the existence intensity profiles that induce metrical properties related to the binary meter. 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 8(a)), the second beat starts with a characteristic strong bass accent, followed by flat and low intensity semiquavers. This 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 and 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 contain 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 code where time, accent, timbre and metrical properties converge. The application of computational approaches to our data-set of commercial samba music confirmed the existence of several characteristic microtiming deviations suggested in the literature and revealed other important interactions that enrich our knowledge about timing in samba and the knowledge about timing in the performance of music. The discovery of other characteristics such as the delay of the first semiquaver (low spectrum) and the *accelerando* and *ritardando* microtiming patterns stimulates new viewpoints on timing aspects that populate the tacit knowledge behind the performance of popular music. Multidimensional aspects of the knowledge that move cultural forms such as samba may include much more elements not easily depicted in traditional approaches to music (e.g. scores).

[1]For example, acculturated performers would 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.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

Note, however, that the present study does not claim an exhaustive overview of multidimensionality of microtiming structures in samba. 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 (Sodré, 1979; Carvalho, 2000; Sandroni, 2001). There, not only music and dance are involved, but also imagery, tradition, symbols and other intertextual 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 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, p. 19, quoted in Fryer, 2000, p. 19)

This example demonstrates how elaborate maps of timing and accents are part of an intricate system of metrical and rhythmic textures and forms of tension that unite sound and movement. Samba music is derived from this original combination of music and movement. How can computational musicology reach and reveal the elements behind these phenomena? How can representations of meter and rhythm be adapted to new standpoints on 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 behaviour (e.g. dancing, playing) on the other hand. More research is needed to provide reliable onset functions in polyphonic audio 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

This study was supported by a Short Term Scientific Mission financed by the COST IC0601 Action on Sonic Interaction Design, by a grant from Ghent University (Belgium) and partly by CAPES (Brazil). The authors wish to thank Inaê Benchaya Duarte for the annotation files.

References

Andrade, M.d. (1991). *Aspects of Brazilian Music* (original title: Aspectos da Música Brasileira) (Vol. 11, in Portuguese). 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)*. Urbana-Champaign: Universidade de Illinois.

Arom, S. (1989). Time structure in the music of central Africa: Periodicity, meter, rhythm and polyrhythmics. *Leonardo*, *22*(1), 91–99.

Bello, J.P., Duxbury, C., Davies, M., & Sandler, M. (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. (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, MA, USA.

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

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

Carvalho, J.J. de. (2000). *Afro-Brazilian music and ritual[s]*. Lecture presented in the Graduate Seminar in Ethnomusicology at the University of Wisconsin-Madison.

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*. Chicago: 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*, Brasilia, pp. 27–31.

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

Desain, P., & Honing, H. (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.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 05:40 02 November 2011

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

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 11th Brazilian Symposium on Computer Music (XI SBCM)* (pp. 197–206). São Paulo, Brazil: Sociedade Brasileira de Computação (SBC).

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

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. (2009). The WEKA data mining software: An update. *ACM SIGKDD Explorations Newsletter*, *11*(1), 10–18.

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 *Conference on Rhythm and Micro-rhythm: Investigating Musical and Cultural Aspects of Groove-oriented Music*. Oslo, Norway. Retrieved from http://www.hf.uio.no/imv/english/research/projects/rhythm/dok/

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

Keil, C. (1987). 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*, *39*(1), 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 10). Lisboa: Junta de Investigações Científicas do Ultramar.

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., & Nordquist, P. (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*. New York: Oxford University Press.

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* (PhD 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). *Gesture in samba: A cross-modal analysis of dance and music from the Afro-Brazilian culture* (PhD thesis). Ghent University, Belgium.

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

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

Pelleg, D., & Moore, A.W. (2000). X-means: Extending K-means with efficient estimation of the number of clusters. In *Proceedings of the 17th International Conference on Machine Learning (ICML '00)* (pp. 727–734). San Francisco, CA.

Polak, R. (1998). Jenbe music in Bamako: Microtiming as formal model and performance practice. *Iwalewa Forum*, *2*, 23–42.

Rosinach, V., & Traube, C. (2006). Measuring swing in Irish traditional fiddle music. In *Proceedings of the International Conference on Music Perception and Cognition*, Bologna, Italy, pp. 1168–1171.

Sandroni, C. (1996). Mudanças de padrão rítmico no samba carioca, 1917–1937. *TRANS-Transcultural Music Review*, *2* (article 12).

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

Santos Neto, J. (2010). Ginga: A Brazilian way to groove. Retrieved from http://www.jovisan.net/Educational PDFs/Ginga.pdf

Schwartz, K.D., & Fouts, G.T. (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*, *7*, 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. In *Proceedings of BRIDGES: Mathematical Connections in Art, Music and Science*, Banff, Alberta, Canada, pp. 47–56.

Van Immerseel, L.M., & Martens, J.P. (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*. Leiden: EJ Brill.

Downloaded by [b-on: Biblioteca do Conhecimento Online INESC-Porto], [Fabien Gouyon] at 09:40 02 November 2011