Automatic Genre Classification of Latin Music Using Characteristic Rhythmic Patterns

First Author
Affiliation 1
email1@domain.com

Second Author
Affiliation 1
email2@domain.com

Third Author
Affiliation 2
email3@domain.com

Fourth Author
Affiliation 2
email4@domain.com

ABSTRACT

In the majority of musical genres, music is basically composed of repetitive rhythmic structures (patterns). Especially in Latin American music, particular styles can be distinguished through characteristics of these patterns. Therefore, the aim of the present work is the automatic classification of musical genres from Latin America using automatically extracted rhythmic patterns. The approach is based on setting up a knowledge base that consists of typical reference patterns for each genre. To obtain a tempo independent pattern representation, we apply both the scale transform and the log-lag autocorrelation function. Different distance measures were evaluated to measure the similarity between unknown patterns and reference patterns. Various tests with different preprocessing techniques were performed. For 9 distinct genres, a classification accuracy of $86,7$ % for tests with synthetic data and 47.9 % with real-world music was achieved. In addition, conclusions to rhythmic similarity of particular music styles were drawn. Dealing with non-western music, the work presents an operational method for genre classification in the research field of *Computational Ethnomusicology*.

1. INTRODUCTION

Rhythm can be characterized as being one of the most important dimensions of music. It describes various temporal aspects such as the timing of note-onsets, the duration of sounds, the rests within melodic constructions, or moments of silence within a musical piece. Almost everybody can perceive rhythm when tapping the foot or nodding the head synchronously with music. This perception of rhythm is mainly induced by percussive instruments that play typical repetitive structures called rhythmic patterns. Considering musical genres, one can observe that these patterns vary in different genres. Especially some of the Latin American music genres have unique patterns while others, depending on their historical derivation, are designated to have similar rhythmic structures. In Latin American music, most of the genres have prominent rhythmic patterns and representative percussion instruments which characterize the style. Understanding these constructions will help to improve computer-based systems to automatically classify unknown music. Typical examples for rhythmic patterns in Latin music are the clave patterns played with two sticks made of wood. The variant of the Son clave is depicted in Fig. 1.

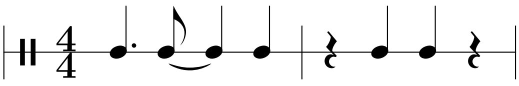

Figure 1: Son Clave pattern

The remainder of this paper is organized as follows. After outlining the goals and challenges of this publication in Sect. 2, we provide an overview of related work in Sect. 3. In Sect. 4, we illustrate our approach for genre classification and describe each processing step in detail. Different experiments are presented and discussed in Sect. 5. Finally, Sect. 6 concludes this work.

2. GOALS & CHALLENGES

The goal of this work is genre classification of Latin musical genres using solely information about rhythmic patterns of percussion instruments. We aim to detect significant patterns in musical pieces and estimate a genre on the basis of a priori known patterns of each genre. We also intend to discover rhythmic similarities as well as differences in rhythmic organization between the distinct styles. These latent information about musicological relations between the investigated genres can be useful for ethnomusicologists, which focus on music of Middle and South America and could contribute to the area of *Computational Ethnomusicology* [20].

Dealing with rhythmic patterns over time, one of the main challenges is that same patterns can occur in different tempi. It is crucial to obtain a tempo-independent representation in order to compare patterns and measure the similarity reliably. Another challenge is the variety of playing techniques of percussion instruments which is characteristic in all investigated musical genres. For example, common techniques of the conga drum as for instance open, closed, muffled, and slapped result in a large range of different sounds. For the construction of a knowledge base, it is important to incorporate all these techniques to allow their proper detection in unknown audio signals.

3. RELATED WORK

Genre classification is a popular task in Music Information Retrieval (MIR). Common approaches use pattern recognition techniques and obtain feature vectors to describe the content of audio signals. Various classification techniques are applied to classify unknown songs to different genres based on previously learned class models. Tzanetakis et al. [19] applied features for representing timbral texture, rhythmic content and pitch content. They propose the use of a *beat histogram* which is obtained by multiple peak picking of an enhanced autocorrelation. The authors achieved 61 % classification accuracy for 10 musical genres. Dixon et al. [7] extract prominent bar-length rhythmic patterns which represent the temporal position of events. They report a classification accuracy of 50 % using the rhythmic patterns and of 84 % using these patterns in conjunction with other features like meter, syncopation or swing factor.

To solve the problem of measuring rhythmic similarity independent of the tempo of musical pieces, different approaches are applied. Paulus et al. [15] make use of a Dynamic Time Warping (DTW). This algorithm is based on dynamic programming and sequential decision processes. The DTW technique is also used by Antonopoulos et al. [1] to compare rhythmic patterns of Greek and African music. Holzapfel et al. [12] pursue a tempo independent representation with the Scale Transform. It transforms the scaling factor between two equal patterns of different tempi into a phase contribution.

Most rhythm description systems focus on popular and western music. The reason for the small number of works focussing on traditional and folkloristic music styles is, according to [8], that the music data is not available and nonstandard. Recently, the ambition to build up ethnomusicological archives grew. [16] gives an overview of existing archives, [13] built a collection of Brazilian Lundu[1], and [18] created a database with 10 Latin American musical genres. An increasing amount of publications in the field of *Computational Ethnomusicology* have been published during the last years. Wright et al. [23] analyze Afro-Cuban music that is very complex in rhythm and organized around the prominent pattern of the clave. The clave is usually repeated over the whole musical piece. The authors use a matched filter approach to detect the clave rhythm and apply templates for estimating the tempo. Moreover, they highlight information about small timing deviations which are referred to as micro-timing. Also Gouyon [11] centers micro-timing and ascertains that the third and fourth 16th-note beats in the Brazilian genre *Samba de Roda* are slightly ahead of their corresponding quantized positions.

4. APPROACH

An overview over our approach is presented in Fig. 2. It is based on a knowledge base that consists of typical rhythmic reference patterns. The construction process is displayed on the left side of the diagram and comprises a collection of patterns with a subsequent audio synthesis using different sample libraries. Rhythm patterns are extracted and transformed to tempo-independent representations. The right side shows an unknown audio signal that should be classified automatically. After a segmentation, it is pre-processed to remove harmonic instruments of the signal. An extraction of rhythm patterns and transformation to tempo-independent representation is applied as well. Afterwards, the similarity values between the unknown rhythm patterns and the reference rhythm patterns are measured and exploited in a commonly-used Nearest Neighbor classifier. All presented processing steps will be explained in detail in the following subsections.

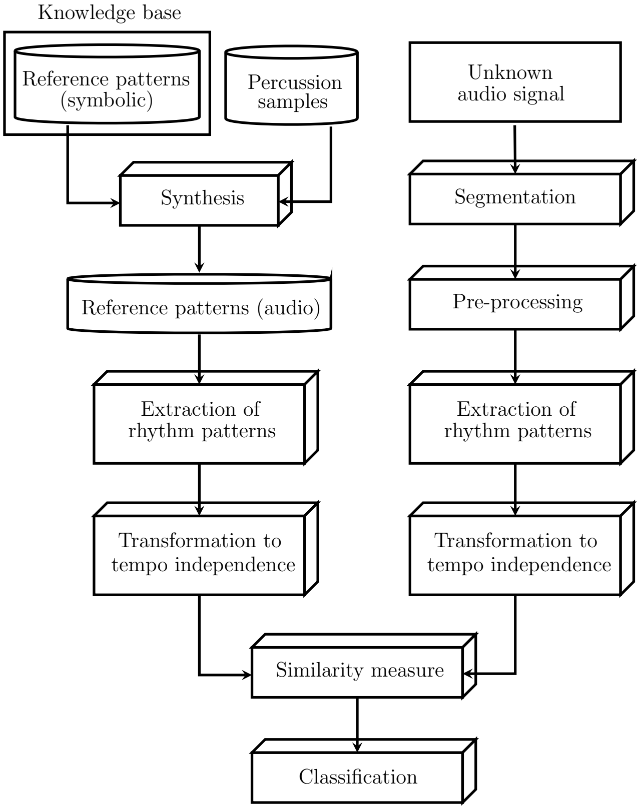

Figure 2: Overview

4.1 Data acquisition and Synthesis

In our work we used a taxonomy of 9 different genres from Latin America, presented in Tab. 1. We developed an audio database with the help of a renowned ethnomusicologist who verified our music to be typical for each of the considered genres. In a later process we want to classify every musical piece of this database.

A knowledge base for all Latin genres was constructed using MIDI for symbolic representation. Primary, samples of all typical percussion instruments for each genre were amassed. Thereafter, characteristic patterns for each instrument were collected from various instructional literature on Latin percussion, e. g. [3, 21, 22]. We analyzed the typical play-

---

[1] Brazilian popular musical form at the 19th century

	Genre	Abbr.	Origin	Songs	Ref. patterns
1.	Baião	BA	Brazil	47	12
2.	Bolero	BO	Cuba	37	10
3.	Bossa Nova	BN	Brazil	42	13
4.	Cha-cha-cha	CH	Cuba	34	11
5.	Samba-enredo / Batucada	EN	Brazil	54	68
6.	Mambo	MA	Cuba	45	14
7.	Merengue	ME	Dominican Republic	46	21
8.	Pagode	PA	Brazil	46	24
9.	Son	SO	Cuba	25	21
	Total			376	194

Table 1: Genre taxonomy with number of songs in audio database and distribution of reference patterns

ing techniques in the patterns and gathered a sample library from corresponding instrument sounds that consists of 26 isolated percussion instruments with 25 different playing techniques in total. Among others, we used instrument samples from the RWC database [10]. Consequently, the superposition of the single patterns leads to realistic mix patterns that are comparable to real-world recordings.

4.2 Segmentation and preprocessing

Generally, musical pieces consist of multiple segments which usually vary in instrumentation, harmony, rhythm and even genre. To analyze the typical rhythmic patterns of a song, we have to ensure that we select a segment with homogeneous rhythm that is representative for the observed genre. In a first step, we performed a segmentation related to the rhythm domain using a segmentation algorithm described in [5]. To obtain the ground truth data, we assign genre labels to characteristic audio segments of the rhythm domain in a second step. For this annotation, a software developed for the segment- and domain-based annotation of music signals[2] came into operation.

The majority of audio excerpts consist of both percussion instruments and harmonic sustained (melodic) instruments. For the reason that the reference patterns are built up with percussion instruments only, we applied two different preprocessing approaches to suppress any harmonic information in the audio excerpts. The first approach is described in [14] where horizontal lines in the spectrogram are interpreted as harmonic instruments and are removed successively from the spectrum. The second approach is derived from extracting fundamental frequencies in [2]. The detected frequencies are suppressed together with their harmonics. In our experiments, we tested with these two preprocessing steps as well as with the pure excerpt's audio signal.

The following processing steps explained in Sect. 4.3 and 4.4 are computed for both the audio reference patterns and unknown audio segments.

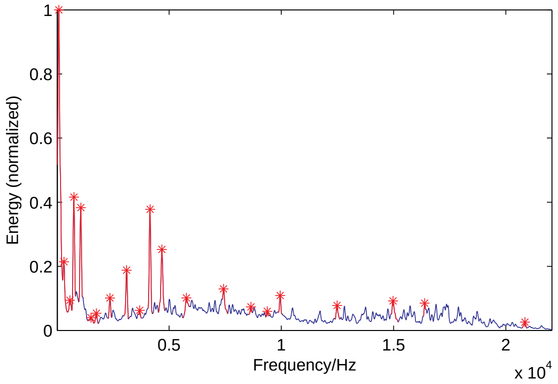

Figure 3: Energy spectrum with examined frequency bands

4.3 Extraction of rhythm patterns

For the extraction of rhythm patterns we compute the spectrogram of the audio segment using short-term Fourier transform with 25 ms block size and 1.5 ms hop size. An averaged spectrum of a synthesized mix pattern is depicted in Fig. 3. We pick up the most significant frequency peaks and define a frequency band to the adjacent local minima on the left and the right of each peak. Afterwards, we investigate the envelope curve that corresponds to each band and get typical structures as shown for a extracted Bossa Nova bass drum in Fig. 4. We pick up the onsets and set up a threshold computed using maximum and mean of the energy function. Consequently, we interpret the curve above the threshold as rhythm pattern.

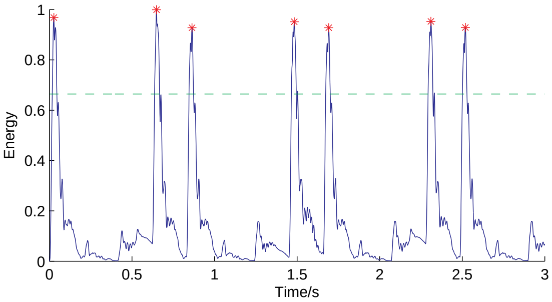

Figure 4: Extracted rhythm pattern from a bass drum

4.4 Tempo-independent representation

The problem of extracted rhythm patterns is their tempo dependency. Equal patterns played in different tempi exhibit a scaling factor and cannot be compared using conventional distance measures. We separately apply three different methods to transform the rhythm pattern to achieve a tempo-independent representation.

4.4.1 Log-lag autocorrelation function

The energy curve of the rhythm pattern is windowed using a window size of 8 seconds and a hop size of one second. For each window the autocorrelation function (ACF) is computed. We regard this function t as scaled version of t' with

---

[2] Tool to annotate music, developed in a project presented on http://www.project.domain

the scaling factor α. A logarithmic resampling of the lag axis is performed as presented in [4]. Due to

$$\log(t' * \alpha) = \log(t') + log(\alpha), \quad (1)$$

the scaling factor becomes an addend and can be removed for tempo independent representation.

4.4.2 *Scale transform*

After computing a windowed autocorrelation, we use the scale transform, that has been introduced by Cohen [6]. It is defined as

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(t)\, t^{-ic-1/2} dt \quad (2)$$

and its key property is the scale-invariance. This means that the scaling factor of scaled signals becomes a pure phase contribution and the magnitude distributions of both the scaled and the non-scaled version are equal.

4.4.3 *Scale transform using FMT*

In the third method the audio segment is filtered with the frequency band obtained from the extracted rhythm and only observed on the time areas of the onsets. The autocorrelated signal is used to build the Fast Mellin Transform (FMT) [17] as

$$M_f(c) = \int_{-\infty}^{\infty} f(e^t)\, e^{1/2t}\, e^{-ict} dt\ . \quad (3)$$

It computes the scale transform applying the Fourier transform of an exponentially warped signal weighted by an exponential window. The implementation works fine for scaled versions of the same patterns, but results in scaled versions when instruments with different pitch are used. This occurs because the same rhythmic pattern in another tempo is only exact scaled regarding the timing of onsets. If we observe the tone pitch of two scaled versions, we notice that the pitch changes. In contrast to that, variations of the same pattern with different tempi do not have a pitch change. To handle this problem, a logarithmic resampling of the coefficients $M_f(c)$ is done and we obtain a completely tempo-independent representation.

4.5 Measuring the similarity and classification accuracy

After obtaining tempo independence, the unknown rhythm patterns can be compared with rhythm patterns from the references. We implemented three different distance measures. The cross correlation

$$R_{xy}(\tau) = \frac{1}{N} \sum_{n=0}^{N-\tau} x(n) y(n+\tau) \quad (4)$$

is computed and represents the first measure. Because of the logarithmic resampling, two patterns are not aligned to compute other distance measures directly. So, the highest correlation coefficient is used to align two functions as shown in Fig. 5.

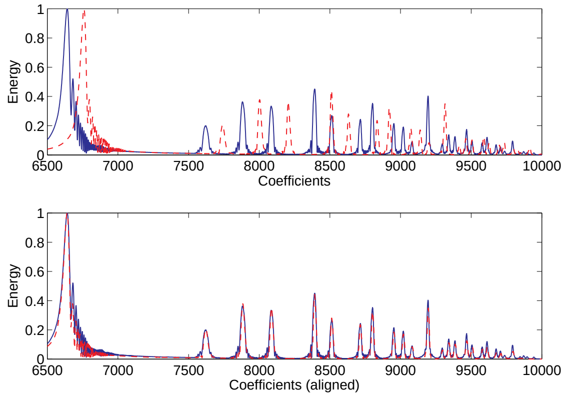

Figure 5: Alignment of two patterns using the cross correlation

After the alignment, the cosine distance

$$d_{cos}(x,y) = 1 - \frac{x \cdot y}{||x|| \cdot ||y||} \quad (5)$$

and the Euclidean distance

$$d_{euc}(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (6)$$

are computed. Moreover, the center frequencies f_x and f_y of the dominant frequency bands of two extracted rhythm patterns x and y are considered in the three distance measures using

$$d_{freq} = \frac{|\log f_x - \log f_y|}{\log f_{max} - \log f_{min}} \quad (7)$$

where $f_{min} = 20\,\text{Hz}$ and $f_{max} = 20\,\text{kHz}$ represent the human hearing range.

Generally, we used a simple Nearest Neighbor classifier approach to classify unknown patterns according to their distance to reference patterns. On the one hand, all extracted rhythm patterns were classified individually, denoted as accuracy K. Each musical piece is represented by all extracted rhythm patterns from that piece. So, the classification of complete songs is performed by calculating a mean similarity over all extracted rhythm patterns of each song towards each of the different genres, denoted as the accuracy $\bar{K}$.

5. EXPERIMENTS AND RESULTS

5.1 Automatic Pattern Classification

In our first experiment we tested the FMT in a task of pattern classification. We synthesized 87 distinct percussion patterns which represents 87 different classes. For each percussion pattern, various patterns are generated varying the tempo and the sound of the related instruments. We compute the FMT of 20 variants and use the mean to get one single representation of each pattern. We synthesized more 5 variants of the pattern with another tempo to be evaluated in the system.

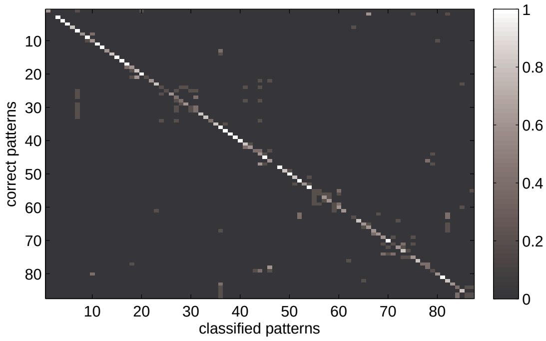

Figure 6: Confusion matrix from classification of 87 distinct patterns for R_{xy}

We receive best classification accuracy values of $R_{xy} = 63.9\%$, $d_{cos} = 63.0\%$ and $d_{euc} = 49.2\%$. For a 87 class problem these are very good results taking the random decision baseline of $1,15$ % into account. The Euclidean distance gives the inferior results, a result that was also found in [9]. The confusion matrix is visualized in Fig. 6 with a dominant diagonal line. Most of the confused patterns are different variants of a pattern like e. g. the Tumbao pattern of the conga played with one or two open tones on the end.

5.2 Automatic Genre Classification

In the genre classification we perform experiments in three different test scenarios: classification of synthetic mix patterns against synthetic references (5.2.1), classification of real-world music against synthetic references (5.2.2) and classification of real-world music against real-world references (5.2.3).

5.2.1 *Classification of synthetic mix patterns against synthetic references*

We generated 25 typical mix patterns of each genre using different sounds and combinations of participating percussion instruments. As references, 970 patterns were used containing 5 different sounds for each playing technique. The synthesis of references and mix patterns was done in different tempi. It was ensured that the tempo difference was larger than the half of the characteristic tempo range of the genre. We evaluated in a 5-fold cross validation and observed the best results for synthesized data using the FMT. For single rhythm patterns we achieve $K_{d_{cos}} = 51.1\%$ and for the complete mix pattern a classification accuracy of $\bar{K}_{d_{cos}} = 86.7\%$. These results were obtained for the cosine distance, but the cross correlation and the Euclidean distance showed good results near 85 % as well. Typical classification errors were made between the triangle in Baião with the Pandeiro in Samba.

5.2.2 *Classification of real-world music against synthetic references*

For classification of real-world music we applied the same set of reference patterns. From the audio database we pick up a randomly chosen set of five musical pieces for each genre. The results show accuracy near the baseline of 11.1% for 9 genres. Only some results are significantly better as $\bar{K}_{R_{xy}} = 32.0\%$ of the log-lag ACF. The problem is the difference between real-world patterns and synthesized patterns coming for instance from micro-timing deviations in real-world music. Another problem is the large number of percussion instruments and resulting diversity of existent patterns in the genre Samba-enredo. Thus, it contains many reference patterns with a high similarity to many of the extracted rhythm patterns of other genres as can be seen in Tab. 2.

		classified genre								
		BA	BO	BN	CH	EN	MA	ME	PA	SO
correct Genre	BA	7	3	13	9	46	6	9	7	2
	BO	2	12	15	12	36	8	5	5	7
	BN	3	6	16	6	40	6	3	5	15
	CH	8	7	14	14	36	7	3	-	11
	EN	3	2	18	9	40	12	6	8	3
	MA	7	1	8	13	42	4	6	10	10
	ME	11	3	24	13	32	4	1	4	7
	PA	9	7	13	13	49	1	4	-	3
	SO	9	4	10	5	46	9	4	5	7

Table 2: Typical confusion matrix (values in %) for the classification of real-world music using synthetic rhythm patterns (achieved for FMT with $K_{d_{cos}} =$ 11.9%)

5.2.3 *Classification of real-world music against real-world references*

To solve this problem we set up extracted rhythm patterns of real-world music as references. We observed the same data set as in the experiment before and evaluated with 5-fold cross validation. 4 of 5 parts of all songs related to a specific genre were used as references and the last part represented the test items. The results were significantly better and more robust relating to the method used. The log-lag ACF achieved 38 % of accuracy now, but especially the other methods show a remarkable enhancement. Our preprocessing techniques did not show a significant improvement. They filter to many important frequency components for percussion instruments with important pitch information like bells or membranophons. The methods for log-lag autocorrelation and direct computation of the scale transform achieve the best results. In a new experiment we evaluated these methods on the complete audio database with 5-fold cross validation obtaining the results shown in Tab. 3.

	Scale transform			log. autocorrelation		
	R_{xy}	d_{cos}	d_{euc}	R_{xy}	d_{cos}	d_{euc}
K	17.9	17.1	12.6	17.8	17.5	17.4
$\bar{K}$	31.4	$\boxed{40.2}$	14.4	42.3	39.9	$\boxed{47.9}$

Table 3: Accuracy values for real-world music (values in %)

We achieve the best result of $\bar{K}_{d_{euc}} = 47.9\%$ for the log-lag ACF. This result is useful for genre classification only based on the rhythmic dimension. Obviously, it is not sufficient for a reliable classification, but it has to be stressed again that we only used rhythmic patterns and omitted descriptors for harmony, melody and timbre.

By observing the confusion matrix of the best result in Tab.

		classified genre								
		BA	BO	BN	CH	EN	MA	ME	PA	SO
correct genre	BA	70	2	6	0	0	2	2	15	2
	BO	24	46	3	5	0	8	3	9	3
	BN	31	5	48	0	2	0	0	12	2
	CH	21	9	3	38	3	24	3	0	0
	EN	17	6	11	0	39	4	6	19	0
	MA	18	2	1	17	9	42	2	9	0
	ME	9	2	0	0	2	4	78	4	0
	PA	33	2	13	0	11	0	0	41	0
	SO	36	16	8	4	4	16	4	4	8

Table 4: Confusion matrix (values in %) for the best result of the log-lag ACF with $\bar{K}_{d_{euc}} = 47.9\%$

4, we can deduce interesting information on the rhythmic similarity between different genres. The diagonal line is present for correct classifications, but it is conspicuous that the genre Baião receives many classification of almost every genre. Many of Latin American music genres contain rhythmic patterns of this genre what could be an interesting information for ethnomusicologists. On the other side, Son does not contain many dominant rhythmic structures. There is mainly a clave rhythm and a maracas shaker. But clave also occur in other Cuban genres and because of its popularity, the shaker rhythm does not build a unique information for a specific genre.

The classification accuracy of Merengue is the highest. That could be traced to the significant Tambora bass drum whose onset occur on every beat as well as the unique pattern of the metal guiro. These patterns only occur in Merengue. A similarity can be determined between Cha-cha-cha and Mambo. This is understandable knowing their historical background. Both genres derived from the Cuban genre Danzón and contain similar rhythmic structures. Also, similarity was confirmed between the Samba styles Samba-enredo and Pagode. Bossa Nova is similar to these genres, too. It is an interesting confirmation knowing that Bossa Nova derived from the slower Samba style Samba-canção and that ethnomusicologists argue whether Bossa Nova is a Samba style or a completely independent musical genre.

6. CONCLUSION

In the present work we developed a system for the automatic classification of Latin American music. We created a knowledge base with characteristic patterns of typical percussion instruments for each genre and extracted rhythm patterns of unknown audio signals to measure the similarity with the references. We applied different techniques for tempo-independent representation on both synthetic data and real-world music. With synthetic data, we obtained a classification accuracy of 86, 7 % for the 9 investigated genres using the FMT in combination with the cosine distance. With real-world music, we achieved an accuracy of 47.9 % for the log-lag ACF with the Euclidean distance. Additionally, we obtained some interesting findings about Latin American music. We confirmed the derivation of certain styles and discovered rhythmic similarities of some specific genres.

7. ACKNOWLEDGEMENTS

This work has been partly supported by the research project ResearchProject[3] funded by the Ministry of Education and Research (ABCDEFGH: 01/23/45/67/89). Additionally, the regional Ministry of Economy, Employment and Technology supported this research by granting funds of the special Fund for Regional Development to the project ProjectName, enabling transnational cooperation between companies and their partners from other regions. In addition, the authors would like to thank the Ethnomusicologist Prof. Ethno Musicological for cooperation and substantial background information on the music genres we covered in this publication.

8. REFERENCES

[1] I. Antonopoulos, A. Pikrakis, S. Theodoridis, O. Cornelis, D. Moelants, and M. Leman. Music retrieval by rhythmic similarity applied on greek and african traditional music. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR)*, 2007.

[2] Author1. Thesis dealing with harmonic instruments. Master's thesis, A university in a country, Year.

[3] Author1 and Author2. *Book for typical Patterns in Latin American Music*. Book publisher, Year.

[4] Author1, Author2, Author3, and A. 4. Publication describing the log-lag autocorrelation function. In *Proceedings of the xth International Conference*, Year.

[5] Author1, Author2, Author3, and Author4. A publication about automatic segmentation of audio signals. In *Proceedings of a conference*, Year.

[6] L. Cohen. The scale representation. *IEEE Transactions on Signal Processing*, 41(12):3275–3292, 1992.

[7] S. Dixon, F. Gouyon, and G. Widmer. Towards characterisation of music via rhythmic patterns. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR)*, 2004.

[8] J. S. Downie. Music information retrieval. *Annual Review of Information Science and Technology (ARIST)*, 37:295–340, 2003.

[9] J. Foote, M. D. Cooper, and U. Nam. Audio retrieval by rhythmic similarity. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)*, 2002.

[10] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka. Rwc music database: Popular, classical and jazz music databases. In *Proceedings of the 3rd International Conference on Music Information Retrieval (ISMIR)*, 2002.

[11] F. Gouyon. Microtiming in *Samba de Roda* - preliminary experiments with polyphonic audio. In *Proceedings of the 11th Brazilian Symposium on Computer Music (SBCM)*, pages 197–203, 2007.

[12] A. Holzapfel and Y. Stylianou. A scale transform based method for rhythmic similarity of music. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 317–320. IEEE Computer Society, 2009.

[13] R. S. G. Lanzelotte, A. O. Ballesté, and M. Ulhoa. A digital collection of brazilian lundus. In *Proceedings of the 8th International Conference on Music*

---

[3] http://www.researchproject.domain

Information Retrieval (ISMIR), pages 163–164, 2007.

[14] N. Ono, K. Miyamoto, J. L. Roux, H. Kameoka, and S. Sagayama. Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram. In *Proceedings of the 16th European Signal Processing Conference (EUSIPCO)*, 2008.

[15] J. Paulus and A. Klapuri. Measuring the similarity of rhythmic patterns. In *Proceedings of the 3th International Conference on Music Information Retrieval (ISMIR)*, 2002.

[16] P. Proutskova. Musical memory of the world - data infrastructure in ethnomusicological archives. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR)*, 2007.

[17] A. D. Sena and D. Rocchesso. A Fast Mellin and Scale Transform. *EURASIP Journal on Advances in Signal Processing*, 2007(1):75–83, 2007.

[18] C. N. Silla Jr., A. L. Koerich, and C. A. A. Kaestner. The Latin Music Database. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)*, pages 451–456, 2008.

[19] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing*, 10(5):293–302, 2002.

[20] G. Tzanetakis, A. Kapur, A. W. Schloss, and M. Wright. Computational Ethnomusicology. *Journal of Interdisciplinary Music Studies (JIMS)*, 1(2):1–24, 2007.

[21] E. Uribe. *The Essence of Brazilian Percussion & Drum Set.* Warner Bros. Publications, distributed by Alfred Publishing Company, 1997.

[22] E. Uribe. *The Essence of Afro-Cuban Percussion & Drum Set.* Warner Bros. Publications, distributed by Alfred Publishing Company, 2000.

[23] M. Wright, A. Schloss, and G. Tzanetakis. Analyzing afro-cuban rhythm using rotation-aware clave template matching with dynamic programming. In *Proceedings of the 9th International Conference on Music Information Retrieval (ISMIR)*, pages 647–652, 2008.