THE LATIN MUSIC DATABASE

Carlos N. Silla Jr.
University of Kent
Computing Laboratory
cns2@kent.ac.uk

Alessandro L. Koerich
Pontifical Catholic
University of Paraná
alekoe@ppgia.pucpr.br

Celso A. A. Kaestner
Federal University
of Technology of Paraná
celsokaestner@utfpr.edu.br

ABSTRACT

In this paper we present the Latin Music Database, a novel database of Latin musical recordings which has been developed for automatic music genre classification, but can also be used in other music information retrieval tasks. The method for assigning genres to the musical recordings is based on human expert perception and therefore capture their tacit knowledge in the genre labeling process. We also present the ethnomusicology of the genres available in the database as it might provide important information for the analysis of the results of any experiment that employs the database.

1 INTRODUCTION

The maturation of the music information retrieval (MIR) field has required that researchers move beyond work involving simplistic musical databases to more expansive studies that require much larger, more varied and carefully annotated collections [19]. McKay et al. [19] have suggested a list of desired features for the development of new music databases in order to have music databases that can be employed in different types of research instead of only a few experiments. The suggested list of desired features in new databases can be divided into three major groups: related to the musical recordings, related to the underlying framework to efficiently use the musical recordings and considering the accessibility and distribution of the database.

The desired features related to the musical recordings are: (1) The database should contain many different types of music; (2) The database should include many thousands of recordings; (3) The database should include a significant amount of commercial music; (4) Each recording should be annotated with as diverse a range of metadata fields as possible, in order to make the database usable as ground truth for as wide a range of research as possible; (5) Entire recordings should be accessible, at least indirectly, not just short excerpts; (6) Given that different compression methods can influence extracted feature values, the audio format(s) most commonly used by the public should be adopted in order to reflect realistic conditions.

The desired features related to the underlaying framework are: (7) It should ideally be possible to label segments of recordings as well as recordings as a whole; (8) Annotations of subjective fields such as genre or mood should include a wide range of candidate categories; allowing ten or so coarse categories is unrealistic; (9) It should be possible to assign multiple independent values to a single field so that, for example, a recording could be classified as both swing and blues; (10) The ability to construct ontological structures between fields could be useful; (11) Metadata should be made available to users in formats that are easy to both manually browse and automatically parse; (12) Automatic tools should be available to validate metadata and generate profiles of the database; (13) It should be easy to add new recordings and their metadata.

The last desired feature is concerned with the accessibility and distribution of the database: (14) Data should be freely and legally distributable to researchers.

Another question that often arises is how to assign a music genre in order to create a musical genre database. In the work of McKay and Fujinaga [18] this issue has been extensively covered. We consider the following definition of a musical genre: “a kind of music, as it is acknowledged by a community for any reason or purpose or criteria” [9].

In this paper we present the Latin Music Database (LMD), a 3.227 musical recordings database from ten different genres. This database was firstly developed to the task of musical genre classification, but it can be used by any MIR task since it provides not only the full recordings in the MP3 format but also information about the music title and performing artist. The main motivation behind the development of this database is due to the lack of ground-truth in the area. Most of databases normally contains few musical recordings, sometimes only excerpts of the full music recording, and also a small number of instances per class. One of the reasons that hinders the development of novel databases is the difficultly in distributing them due to copyright restrictions of musical recordings. However with the creation of the OmeN (On-demand Metadata Extraction Network) tool [17] researchers now have a tool to overcomes this limitation and make databases widely available to the MIR community. In this work an approach to assign musical genres that corroborates with the definition of [9] is used. We also present information concerning how the LMD attends this list of desired features along with a brief ethnomusicology

description of the musical genres used in the database.

This paper is organized as follows: Section 2 presents the related work; Section 3 presents the development process of the LMD; The concluding remarks and a brief discussion about future work are presented in the last section; Appendix A presents the ethnomusicology of the genres used in the LMD.

2 RELATED WORK

The work of Tzanetakis and Cook [22] have strongly motivated the research in automatic music genre classification. The contributions from this work were three fold: First, it showed the task of an automatic music genre classification as a pattern recognition problem; Second, it not only proposed a feature set for the task but also developed an open source software for feature extraction know as Marsyas; and Third, the database used in the experiments is available. This database contains 1.000 music pieces from 10 musical genres (*Blues*, *Classic*, *Country*, *Disco*, *Hiphop*, *Jazz*, *Metal*, *Pop*, *Reggae* and *Rock*). In the remainder of this work this database will be as referred the GTZAN database. From this work onwards, different methods were developed based on extracting features from the audio signal [12].

Current work in the field is trying to break through the possible glass ceiling that content-based audio classification using timbral features has reached [2]. In [3] the authors proposed a method for considering context attributes in order to improve classification accuracy. The experiments were performed with database containing 4.936 songs with 801 boolean attributes grouped in 18 categories, some of which being correlated with some acoustic aspect of the sound ("Main Instrument", "Dynamics"), while others seem to result from a more cultural aspect of the sound ("Genre", "Mood"). This method achieved an improvement of 5% accuracy in average when compared to using only features extracted from the music signal. In [16] a framework for augmenting genre classification accuracy using an ensemble that considers two different views of the same audio signal. The first view is based on the audio signal content, while the second is based on a symbolic form that is obtained using a transcription system. The experiments were performed on three databases: GTZAN, ISMIR Rhythm and ISMIR Genre. The achieved results shown improvements over the previous results on the same databases.

Another important aspect that concerns current research is how the evaluation of music genre classification have been performed. In [10] is verified that performing classification without artist filters (i.e. using music pieces from the same artist in both training and test sets) not only lower the classification accuracy but may also erode the differences in accuracies between different techniques. In [7] is presented a discussion on how to define ground-truth for the task of music genre classification. In this work they define ground-truth as an artefact of an individual response to music, not an artefact of the audio itself; to establish any ground-truth is therefore a cultural study. An approach that comply with this definition is presented by Gouyon et al. [12]. In this work they used the instantly ability that dancers have to recognize different ballroom musical genres (*Jive*, *Quickstep*, *Tango*, *Waltz*, *Viennese Waltz*, *Cha Cha Cha*, *Samba* and *Rumba*) to label their database. These approaches are similar to the one used to define ground-truth for the LMD, as we will see in section 3.1.

Despite the efforts on the development on new techniques for the task of automatic music genre classification, relatively less work have been done concerning the development of music databases that can be used as ground-truth. Besides the GTZAN database, other known databases are: the RWC database [11], the CODAICH database [19], the magnatune database which was used in the MIREX 2004 [8] genre contest, and the database developed by Homburg et al. [14].

3 THE DEVELOPMENT OF THE DATABASE

We start the development of the LMD [1] aiming to collect at least 3.000 musical recordings from ten Latin musical genres. However, there are many issues involved in the development of a new audio database, as we will see in the following.

3.1 The Process of Assigning Musical Genres

In the development of the LMD we used an approach motivated by the human perception of how to assign musical genres to the musical recordings. The genre is defined according to how human listeners dance the musical recordings. The musical recordings were labeled by professional teachers with over ten years of experience in teaching ballroom and Brazilian cultural dances. In the first stage of the process these professionals made a selection of the musical recordings that they judged representative of a specific genre, according to how that musical recording is danced. In the second stage of the process the first author of this work verified each one of the selected songs in order to avoid mistakes that were expected to happen due to the stress produced by manually listening and labeling each one of the songs. About 300 musical recordings were classified by month, and the total duration of the development of the LMD took a year.

As a result of this attempt, we developed the Latin Music Database which currently has 3.227 musical recordings from ten Latin music genres in MP3 format. Table 1 presents information concerning the informetrics of the LMD. The original source of the recordings are CD's (which were encoded using 256 kbps and 44 kHz) and from different websites.

[1] Information available on: http://www.ppgia.pucpr.br/∼silla/lmd/

			Track Duration		
Genre	Artists #	Tracks #	Min.	Max.	Length
Axé	37	313	01:22	08:37	19:09:07
Bachata	64	313	02:54	06:32	23:23:22
Bolero	99	315	01:50	05:45	19:19:44
Forró	27	313	01:17	07:50	18:38:16
Gaúcha	92	311	01:13	07:32	17:32:29
Merengue	96	315	01:39	07:30	20:21:39
Pagode	16	307	00:52	11:12	19:41:27
Salsa	54	311	01:55	09:12	28:53:48
Sertaneja	9	321	02:04	06:34	23:03:10
Tango	7	408	01:19	12:06	19:02:26

Table 1. Latin Music Database Informetrics

The employed protocol of human inspection based on the perception about how that musical recordings are danced contrasts with the suggestion given by Aucouturier and Pachet [1]. They suggest to use complete CDs from collections of a given genre or theme. We justify the approach because the latter proposal, in the case of Latin music, is highly inefficient for labeling. For example in the case of the four CD collection *Los 100 Mayores Exitos De La Musica Salsa* only half (50 out of 100) of the songs can be classified as *Salsa*, the remaining of the collection belongs to other music genres like *Merengue*, *Lambada* and even *Samba*.

Another approach that could be used for automatically labeling the music is the classification of albums based on the artist profile, a common practice in the area. However this approach does not seems to be adequate again. For example, if we want to add songs by Carlos Gardel (the famous Tango Composer), all his songs would be labeled as *Tango*. Although he has over 500 compositions, only about 400 of them are actually *Tangos*, so the straightforward labeling introduces unnecessary noise from a machine learning perspective. Even with other artists from a specific genre hardly ever all the tracks from an album are from the same genre. One interesting fact that was perceived during the database creation is that about one of three songs in each album are not from the main style of the Artist profile.

3.2 Storage and Retrieval of the Musical Recordings

Besides the acquisition and genre labeling of the music pieces, much effort was employed in order to develop a database with the following characteristics: (1) it must be easily used in other MIR tasks; (2) it must allow the reproducibility of the performed experiments; (3) it must avoids duplicity in the stored musical recordings; and (4) it must be easily extended, with the storage of new music pieces, new music genres or any other kind of desired metadata (i.e. mood). Taking this issues in consideration a prototype system for storing, labeling and handling the musical recordings was developed. The proposed system is in accordance with the list of desirable features proposed by McKay et al. [19].

The process of adding a new music recording to the framework is as follows. First, a musical genre is assigned to the recording following the procedure explained at Section 3.1. Next, the ID3v2 tag (which contains meta-information) of the musical recording is manually inspected to verify if the fields have been correctly fulfilled. Sometimes we need to correct / adapt them according to a standardization procedure, such as in the case of names that contain the special character &, indicating more than one artist in the same song. This procedure could be replaced by an automated process like the one suggested by McKay et al. [19] as it mainly involves converting the music titles and artists to a pre-defined standard. To add a new music to the database the mandatory fields in the ID3v2 tag are the artist and the song title. The reason is simple: even if only one person is working on adding new songs to the database, eventually albums from the same artist can contain musical recordings that also appear in other albums, e.g. in the case of albums with greatest hits of a particular artist. The contribution of this simple procedure is two-fold: first, it avoids having duplicate songs in the database; second, it enforces that all music added to the database have the valuable metadata of song artist and title.

Finally, the musical recording is stored in the database. In this step the framework obtains the metadata from the ID3v2 tag of the musical recording and verify if the song is not already present in the database, to avoid duplicity. If there is no duplicity the framework assigns the song an identifier code, associate this song to the defined genre and creates an internal copy of the song. The information concerning the music genre is stored into a relational database, because the Genre field of ID3v2 tag is not reliable. Another reason for this approach is flexibility, as it makes the assignment of multiple genres per musical recording as simple as adding another entry into the relational database.

The process of retrieval in the framework is straightforward: it is started by the user, who specifies how many recordings from each musical genre should be retrieved. The framework also implements other modules that execute specific tasks, such as the integration of the database with feature extraction frameworks – like Marsyas [22] – or to allow the user to retrieve the feature set in a specific data format – such as in the arff format for using the WEKA machine learning tool [24].

Another important aspect of the framework is related to the reproducibility of performed experiments. Since every musical recording in the database has information concerning the song artist and title, it is possible to create along the output of the feature sets in arff format a list of the songs in the same order in which they are used by classification modules. The file used to store this list is called SAL (Song Artist List). There are three reasons for representing this information in the SAL format: (1) Sometimes different artists interpret the same songs using the same or a different style;

Database	No. Songs	No. Genres	Commercial Music	Meta Data	Full Recordings
CODAICH	20.849	53	Yes	Yes	Yes
Homburg et al.	1.886	9	No	Yes	No
GTZAN	1.000	10	Yes	No	No
LMD	3.227	10	Yes	Yes	Yes
Magnatune	2.181	6	No	Yes	Yes
RWC	100	10	No	Yes	Yes

Table 2. Comparison of the existing databases for musical genre classification

(2) Using only the system assigned ID we cannot obtain a reliable information, because if for some reason (i.e. backup) the musical recordings need to be restored in the system, they will probably not be added in the same order they were originally; (3) It might be easier to understand and interpret the achieved results if one knows which songs / artists were used in the experiment.

3.3 Analysis of the Database

In this section we present an explanation on how the LMD deals with the list of desired features in new databases indicated in [19]. As discussed in Section 1, the list of desired features can be divided into three major groups: related to the musical recordings, related to the underlying framework to efficiently use the musical recordings and considering the accessibility and distribution of the database.

Concerning the features related to the musical recordings, the LMD satisfy all the desired criteria because: (1) it has 10 Latin music genres; (2) 3.227 music pieces; (3) every musical recording in the LMD is commercial; (4) the fields from artist, tile, and musical genre have all been carefully annotated; (5) only full recordings are available; (6) every recoding is in the MP3 format.

Concerning the features related to the underlying framework, some of them are not fully implemented yet. However, they are less critical than the features related to the musical recordings, because doing implementation improvements is easier than doing carefully and correct annotation. In the current framework the issues that can be handled are as follows: features (7) and (9) can be handled by creating a table in the relational database of the framework to consider the segments of the music recordings, or to allow more than one genre per musical recording; (8) Although only 10 Latin music genres are presented in the database, it is important to stress that the classification was done by human experts and not using website listings in order to achieve ground-truth. Therefore, we would like to encourage more approaches as the one presented in this paper, even if few genres are presented as the classification by human experts is a really time-consuming task; (10) is beyond the scope of the designed framework; (11) and (12) the framework outputs information in arff and SAL formats, and could be easily modified to deal with other formats as needed; (13) as seen in section 3.2 it is a straightforward process.

The last desired feature is concerned with the accessibility and distribution of the database. Due to copyright reasons, the LMD cannot be made freely available to the research community; however there are on-going efforts in two directions to disseminate the use of the database. The first direction is making the LMD available through the OmeN tool [17]. The second direction is contributing the LMD to the MIREX (Music Information Retrieval Evaluation eXchange) community [8].

A summary of the different databases according to the list of desired features related to the musical recordings as seen in Section 1 is presented in Table 2. An analysis of this table shows that besides CODAICH, only the LMD has all the desired features related to the musical recordings, due to its large number of entire commercial songs associated to metadata.

4 CONCLUDING REMARKS

In this work we present a novel and carefully developed database for the task of automatic music genre classification, namely the Latin Music Database (LMD). An analysis of the database shows that the LMD has all the desired features listed in the work of McKay et al. [19], with one exception: the legal questions related to music recording distribution. However two efforts are being made in order to make the LMD available to the research community: (a) contributing the database to the MIREX community; and (b) using it in the OMEN framework. The method for assigning musical genres although similar in concept to the one used in [12] differs from it for the following reasons: (1) the LMD has a greater number of musical recordings; (2) the genre labeling process is based on two human experts instead of assuming that the classification given by a website is the ground-truth; and (3) most of the genres used are different.

Another contribution of this work is that it presented briefly the ethnomusicology aspects of the genres used in LMD. This is an important aspect of the database, considering the current state of the art in the field of automatic music genre classification: understanding the ethnomusicology of the genres being worked is crucial to understand the social context to which the genres belong. This information along with the full metadata for the music pieces in the database allow that

context information may be explored in novel ways. One interesting aspect of future research is training a classifier based on the lyrics of the musical recordings, in order to provide further information about the context of specific genres or sub-genres. Another goal for future research is to expand the LMD and further classify the existing songs into their sub-genres categories.

5 REFERENCES

[1] Aucouturier, J. J.; Pachet, F. "Representing Musical Genre: A State of the Art", *Journal of New Music Research*, vol.32, no.1, pp.1-12, 2003.

[2] Aucouturier, J. J.; Pachet, F. "Improving timbre similarity: How high is the sky?", *Journal of Negative Results in Speech and Audio Sciences*, vol. 1, no. 1, 2004.

[3] Aucouturier, J. J.; Pachet, F.; Roy, P.; Beurivé, A. "Signal + Context = Better Classification", *Proc. 8th Intl Conf. on Music Information Retrieval*, pp.425–430, 2007.

[4] Austerlitz, P.; *Merengue: Dominican Music and Dominican Identity*. Temple University Press, 1997.

[5] Clark, W. A.; *From Tejano to Tango: Latin American Popular Music*. Routledge, 2002.

[6] Collier, S.; Cooper, A.; Azzi, M. S.; Martin, R.; *¡Tango!*. Thames and Hudson, 1997.

[7] Craft, A. J. D.; Geraint A.Wiggins, G. A.; Crawford, T. "How Many Beans Make Five? The Consensus Problem in Music-Genre Classification and a New Evaluation Method for Single-Genre Categorization Systems", *Proc 8th Intl Conf. on Music Information Retrieval*, 2007.

[8] Downie, J. Stephen. "The Music Information Retrieval Evaluation eXchange (MIREX)", *D-Lib Magazine*, vol.12, no.12, 2006.

[9] Fabbri, A. "Browsing Music Spaces: Categories and the Musical Mind", *Proc. of the IASPM Conference*, 1999.

[10] Flexer, A. "A Closer Look on Artist Filters for Musical Genre Classification", *Proc. 8th Intl Conf. on Music Information Retrieval*, 2007.

[11] Goto, M.; Hashiguchi, Hashiguchi, H.; Nishimura, T.; Oka, R. "RWC Music Database: Music Genre Database and Musical Instrument Sound Database".

[12] Gouyon, F.; Dixon, S.; Pampalk, E.; Widmer, G. "Evaluating rhythmic descriptors for musical genre classification", *Proc. AES 25th Int. Conf.*, pp. 196-204, 2004.

[13] Hernandez, D. P.; *Bachata: A Social History of a Dominican Popular Music*. Temple University Press, 1995.

[14] Homburg, H.; Mierswa, I.; Möller, B.; Katharina, M.; Wurst, M. "A Benchmark Dataset for Audion Classification and Clustering", *Proc. 6th Intl Conf. on Music Information Retrieval*, pp. 528–531, 2005.

[15] Koskoff, E.; *Music Cultures in the United States: An Introduction*. Routledge, 2005.

[16] Lidy, T.; Rauber, A.; Pertusa, A.; Iñesta; "Improving Genre Classification by Combination of Audio and Symbolic Descriptors Using a Transcription System", *Proc. 8th Intl Conf. on Music Information Retrieval*, 2007.

[17] McEnnis, D.; McKay, K.; Fujinaga, I. "Overview of On-demand Metadata Extraction Network (OMEN)", *Proc. 7th Intl Conf. on Music Information Retrieval*, 2006.

[18] McKay, K.; Fujinaga, I. "Musical Genre Classification: Is it worth pursuing and how can it be", *Proc. 7th Intl Conf. on Music Information Retrieval*, 2006.

[19] McKay, K.; McEnnis, D.; Fujinaga, I. "A Large Publicly Accessible Database of Annotated Audio for Music Research", *Proc. 7th Intl Conf. on Music Information Retrieval*, 2006.

[20] McGowan, C.; Pessanha, R. *The Brazilian Sound: Samba, Bossa Nova, and the Popular Music of Brazil*. Temple University Press, 1998.

[21] Steward, S.; *¡Musica!: Salsa, Rumba, Merengue, and More*. Chronicle Books, 1999.

[22] Tzanetakis, G.; Cook, P. "Musical genre classification of audio signals". *IEEE Trans. on Speech and Audio Processing*, vol.10, no.5, pp.293-302, 2002.

[23] Waxer, L. *Situating Salsa: Global Markets and Local Meanings in Latin Popular Music*. Routledge, 2002.

[24] Witten, I. H.; Frank, E. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2005.

A ETHNOMUSICOLOGY OF THE DATABASE

One important aspect of the research on music genre classification that is often forgot is how it can contribute to music-related fields like musicology. This might be due to the fact that most approaches still use only content based information extracted from the audio signal [12, 22]. However, as any other application in data mining, in order to better understand the results achieved by the state of art algorithms, prior knowledge about the data itself is needed. For the task of music genre classification, this can be achieved from the field of Ethnomusicology, which considers not only the origin of the music genres but also the social and cultural context in which it exists. In this section, we present a brief analysis of the music genres used in the database in the hope of: (a) a better understanding of any results using the database; (b) a briefly description of what some of the unknown genres are; (c) act as starting reference points to books that discuss the social and cultural aspects of the musical genres used in the LMD.

Axé: It is a Brazilian music genre, originated in Salvador, Bahia, where rhythm has been the key ingredient,

which typically feature simple melodies and harmonies. It is energetic, good-time party music, designed for dancing and Carnival. The lyrics usually focus on romance, sex, and Carnival [20].

Bachata: Closely associated with poor rural migrants residing in urban shantytowns in the Dominican Republic, was considered too crude, too vulgar, and too musically rustic to be allowed entrance into the mainstream musical landscape, since its emergence in the early 1960‘s. In 1991, Juan Luis Guerra and his group were responsible for transforming the genre into a success with the song *Bachata Rosa*. In his album only 4 of the 10 musics were considered as bachata because of the standard instrumentation (guitar, maracas and bongos) although the arrangements were enhanced by synthesizers and Guerra sang the romantic lyrics with the heightened emotion typical of bachata singers. Since 1990 the most significant change is bachata is the instrumentation, while still guitar led, bachata is increasingly being played with hollow-body electric guitars rather than acoustic guitars, which has dramatically changed its timbre[13].

Bolero: It has the same Afro-Cuban roots as the Rumba and is thought to have originated from Cuban or Spanish folk dances such as the Danzon and Beguine. However the Bolero is strongly associated with Mexico, because of its appropriation by the Mexican *Trios Romanticos*. Bolero poems often deal with themes of bittersweet, unrequited, betrayed or eternal love. The music is frequently arranged with Spanish vocals and a subtle percussion effect, usually implemented with guitar, conga or bongos [5].

Forró: Originated in 1946 when Luiz Gonzaga recorded the evolutionary song named *baião*, whose title become the name of a new genre. The *baião* derived from an old northeastern circle dance of African origin. It has a steady beat from beginning to end, making it easier to dance to. The basic instruments are the accordion, a triangle and a zabumba bass drum which are used to create a syncopated 2/4 rhythm. Around 1950, Luiz Gonzaga popularized others musical genres such as *Xaxado* and *Xote*. Nowadays much of his music is referred to as *Forró*, a word that originally meant a party or place to play music, however the etymology of the word is subject to much speculation [20].

Gaúcha: Is related to the southern states in Brazil and is related to its cultural identity. As with the Forró, this label is used to refer to different music genres such as the *Vaneira*, *Vaneirão*, *Buggio*, *Xote Gaúcho*, among others. The lyrics usually refer to a specific set of values like respect for the women, the love for the countryside, the culture and the animals.

Merengue: Is a century-old music has been used to tell news and gossip an to spread political propaganda since the birth and is also the national dance of the Dominican Republic. There is no mistaking its crisp.,zippy beat, hissed and scratched out on a metal grater quira in jaunty 2/4 time, or the bubbling triple beat roll of tambora rhythms and sharp twitter of interlocked saxophone vamps[21]. The mainstream Merengue is also called merengue cibaeño. There are other types of Merengue like the merengue palo eacho (or priprì) of the south and east. While few regional merengues are played today, a vital pri-prí culture persists. This music class for one singer, the single-headed basie drum, the guira, and the one-row accordion and its 12/8 rhythm is markedly different from the 4/4 of meregue cibaeño[4].

Pagode: It is a style of Brazilian samba, that surged in the mid-1970’s, when a group of musicians associated with the Carnival Block *Cacique de Ramos* started getting together for a *pagode*, a party where people played samba. In these informal get-togethers they introduced the tan-tan, a type of atabaque, which replaced the larger and heavier surdo. This was more practical for informal samba get-togethers, as the tan-tan could be more easily carried on buses, the mode of transportation for Rio de Janeiros‘ Working class. Almir Guineto added a banjo, which was louder than a cavaquinho and better for open-air gatherings. Ubirany started playing a hand-held repique, called *repique de mão* and dispensed with the usual use of drum sticks. And Bira played the pandeiro in unusual ways. The sound was intimate and earth with new percussive textures. Their lyrics were unpretentious, focusing on situations from their daily life [20].

Salsa: The term salsa did not come into common usage until the early 1970s, when it appeared in Latin New York magazine and was adopted as a category for the 1975 Latin New York Music Awards [15]. There are different theories around the origins of Salsa, one of them is that it surged in Cuba and later went to Puerto Rico. However, studies from the lyrics show that Puerto Rican musical identity and salsa are inseparable, because Puerto Ricans in New York made salsa a political movement in the 1970‘s. The most important aspect of the instrumentation in Salsa is the Clave as every musician in the band must understand it [23].

Sertaneja: A type of Brazilian country music, surged in national popularity and was the single biggest category in term of record sales. It is a pop-music version of *música caipira*, the rural folk music from Brazil‘s South, Southeast and central regions. Most *Sertaneja* artists are *duplas* (duos) who strum guitars and ten-strings violas, harmonizing plaintively as they croon about romance and rural life [20].

Tango: The etymology of the word tango cannot be traced completely, but it seems high probable that it reached the Western hemisphere with the slave-ships and on the lips of slaves. Regardless of the origin, in many parts of the Spanish-American empire, the word *Tango* acquired the standard meaning of a place where African Slaves assembled for the purpose of dancing. The dance as we know today come from a parodistic way of dancing the African-Argentine *tango* where the movements of *Quebradas* (an improvised dramatic contortion) and *Cortes* (a sudden break in the standard figures of the dance) were incorporated into dances in which the partners danced together [6].