AN AUDIO-VISUAL DATABASE OF CANDOMBE PERFORMANCES FOR COMPUTATIONAL MUSICOLOGICAL STUDIES

Martín Rocamora, Luis Jure, Bernardo Marenco, Magdalena Fuentes, Florencia Lanzaro, Alvaro Gómez

School of Music and Faculty of Engineering, Universidad de la República, Uruguay

{rocamora,lj}@.eumus.edu.uy

ABSTRACT

The development of computational tools for musicological studies and musical analysis has been a very active area of research in recent years. Some of these tools operate on symbolic representations of music, while others process the audio from a recorded performance. In both cases, an annotated dataset representative of the corpus under study is essential for the research. This paper reports the generation of a database of audio and video recordings of Candombe performances, including detailed descriptions of the process of curation and production, technical aspects, musical content, and annotations. It is part of an ongoing interdisciplinary research effort for the development of tools for computer-aided music analysis. This encompasses several research problems concerning the development of technologies capable of extracting musically meaningful information from a recorded performance. A significant part of the recordings and annotations is being released for the research community. Of course, apart from its research purposes, the database has also a documentary value.

1. INTRODUCTION

The development of computational tools for musicological studies and musical analysis has been a very active and promising area of research in recent years [23]. Some of these tools operate on symbolic/semantic representations of music, specially when dealing with Western notated musics. Different systems exist for symbolic music representation, like MusicXML, Humdrum, ABC, Musedata and MIDI, among the most frequently used.

When dealing with musics that do not rely primarily on music notation, on the other hand, computational musicology can take advantage of the techniques provided by MIR, a relatively new field of research that pursues the development of technologies capable of extracting musically meaningful information from digital audio signals [23]. In either case, a representative annotated dataset of the corpus under study is essential for the research, both for the development and testing of the techniques and tools, and for the proper musicological analysis.

In this context, several datasets for development and evaluation of MIR applications were developed, from the pioneering and widely used RWC database [12, 13, 11] to more recent endeavours such as the Million Songs Database [3] or the MendelyDB [4]. While most of them are devoted to audio recordings, the ENST-Drums database [10] is a collection of audio and video recordings for drum signal processing. The availability of suitable annotations–such as beat structure, chords or melody line–is a critical issue about the usefulness of such datasets [14]. Even though some software tools have been developed to alleviate the process [5], producing those labels is usually a very time consuming task, manually done by music experts.

This paper introduces an audio-visual database of Candombe performances intended for computational musicological studies, including annotations of metrical information (beat and downbeat), and temporal location and type of strokes. Besides, a discussion on the technical requirements and tackled challenges is provided. The goal was to generate data to analyze a percussion performance in an efficient, affordable, and non-invasive manner. Although different approaches were considered such as motion-capture systems and various types of sensors [20], audio and video recordings were favoured as non-intrusive means of capturing the performance. Thus, five renowned Candombe drummers were recorded on a multi-track audio system and simultaneously filmed by three video cameras. Besides, in order to properly register the fast movements of the percussionists, high frame rate videos were also produced, by means of two additional affordable cameras adapted for this purpose. The paper also outlines possible applications of the dataset and gives some examples of research works that make use of it.

The remaining of this document is organized as follows. Section 2 is devoted to introducing Candombe drumming and the group of performers that took part in the session. In Section 3 a description of the recording session, including some technical considerations, is provided. Section 4 explains how each type of content was rendered, as well as their intended purposes. Section 5 describes some of the efforts that are being conducted regarding the annotation of the dataset. The paper ends with some discussion on the present work and ideas for future research.

2. MUSICAL SETTING

2.1. Candombe drumming

Candombe, one of the most characteristic and defining components of Uruguayan popular culture, has its actual roots in the traditions brought from Africa in the 18th century by the victims of the slave trade. In 2009 it was inscribed on the Representative List of the Intangible Cultural Heritage of Humanity by UNESCO, taking into account, among other things, that it is "a source of pride and a symbol of the identity of communities of African descent in Montevideo" [21]. In constant evolution for over 200 years, it was, however, gradually embraced by the whole society [9] [1].

Drumming is the essential component of this tradition; sharing many traits with other musics of the Afro-Atlantic cultures, the Candombe rhythm can be difficult to interpret correctly for listeners not familiar with it. It is also challenging for MIR techniques of beat detection, and it has been shown that standard algorithms perform poorly when processing recordings of Candombe drums [17].

The instrument of Candombe is simply called *tambor*, which is the Spanish generic word for "drum". There are three sizes: *chico*, *repique* and *piano*, each with a distinctive sound and pitch range, and characteristic rhythmic patterns. The Candombe rhythm results from the interaction of the patterns of the three drums, and a minimal group (*cuerda*) must have at least one of each. An additional important pattern, shared by the three types of drum, is the *madera* or *clave*. Played with the stick on the shell of the instrument, it has a similar function to the *clave* or bell in Afro-Atlantic musics. The most traditional manifestation of Candombe drumming is the *llamada de tambores*, a drum-call parade taking place specially on weekends and public holidays, when groups of players —typically between approximately 20 and 60—meet at specific points to play the drums marching on the street.

In the past, Candombe was practiced almost exclusively in some neighbourhoods of the city by members of the community. Each neighbourhood or *barrio* had a distinctive and recognizable style of performing the rhythm, the three most important being *barrio Sur* (*Cuareim*), *barrio Palermo* (*Ansina*) and *barrio Cordón* (*Gaboto*). In the last decade of the 20th century, Candombe drumming grew in popularity, and today groups of performers can be found all over Montevideo, as well as in other cities of the country. During Carnival, groups representing different *barrios* are organized in *comparsas*, comprising not only drums but also dancers and traditional characters in costume, to take part in the very popular *Desfile de Llamadas*.

Candombe has also influenced several genres of popular music in Uruguay, and while its rhythm has been adapted to the drum set and other percussion instruments like congas, it is common to find bands that include a small *cuerda de tambores* as rhythmic accompaniment.

2.2. Group of performers

Several factors had to be taken into account when selecting the group of players to participate in the session. The first criterion was to pick players of the highest level, and representative of the most authentic tradition of Candombe drumming. But in addition to the individual quality of each player, the balance of the ensemble as such also had to be carefully planned. The first basic condition was that all the drummers had to belong to the same *barrio*, in order to guarantee stylistic compatibility; but musical and personal understanding and affinity among all the performers was also essential. One final consideration was including some performers proficient in more than one type of *tambor*, in order to have a wider range of combinations with the limited number of players that could be convened.

Eventually, a group was assembled consisting of five performers, all of them belonging to families of long-standing tradition in the community of *barrio Palermo* (Ansina style): Gustavo Adolfo Oviedo Gradín (b. 1953), Fernando Silva Pintos (b. 1955), Sergio Ariel Ortuño Priario (b. 1966), Héctor Manuel Suárez Silva (b. 1968), and José Luis Giménez García (b. 1969). Besides being widely acknowledged as outstanding players in one or more types of *tambor*, and having led the *cuerda de tambores* of several *comparsas*, they all have vast experience as professional musicians and performing artists in different settings of popular and orchestral music. [1]

Gustavo Oviedo, one of the most important and influential players of *tambor piano* of the last decades, has been for many years the leader of the drums of *barrio Palermo*. Together with his brother Edinson "Palo" Oviedo, he has directed two important *comparsas* of the neighbourhood: Concierto Lubolo and Sinfonía de Ansina.

Although a strong player of *tambor piano* in a *llamada de tambores*, Fernando "Hurón" Silva is mainly known as a virtuoso player of *repique* in small groups. Together with the Oviedo brothers, he co-led the *comparsa* Concierto Lubolo in the 1980s and early 1990s. At that time, the three of them formed a legendary trio that participated in many important recording sessions of Uruguayan popular music.[2] During the session, Oviedo and Silva only played *piano* and *repique* respectively.

The three remaining performers belong to a younger generation, and the three are known as accomplished players of the three types of drum. Luis Giménez played *chico* and *repique*, while both Sergio Ortuño and Héctor Manuel Suárez (primay *repique* and *piano* respectively), played the three drums in different takes (See Table 3).

[1] Most notably in the "Suite de Ballet según Figari" (1952) by Uruguayan composer Jaurés Lamarque Pons (1917–1982), an orchestral piece that includes Candombe drums in the last movement.

[2] Among many others, with noted Uruguayan musician Jaime Roos.

3. RECORDING SESSION

In order to encompass both the documentary and research purposes, the recording session had to be carefully prepared. It was devised taking into account a wide range of possible MIR applications, such as drum event detection, audio source separation, automatic transcription, beat-tracking and audio-visual music analysis.

Several pilot studies were carried out to evaluate technical requirements and to test different solutions. A preliminary recording session was conducted in studio, which allowed to foresee some difficulties and to identify shortcomings beforehand.

The selected venue was a modern medium-size concert hall, called Zavala Muniz, which is part of the Solís Theatre building, located in Montevideo. Among other facilities, it provided appropriate lighting and acoustic conditions. The recording session took place on September 6th 2014, and lasted for about 6 hours, involving a crew of a dozen people.

3.1. Audio recording

The audio setup was selected to produce two main different outcomes: a stereophonic recording of the ensemble and separate audio channels of each drum. The aim of the former is to provide a realistic spatial sound image of the scene, with good localization of sources plus the effects of the room acoustics [2]. Conversely, the purpose of the separate channels is to record only the direct sounds coming from a given drum, with no interference from the others. This is intended to facilitate the analysis and transcription –either automatic or manual– of the individual performances. Additionally, in this way, the whole set of channels constitutes an appropriate research framework for sound source separation.

The selected microphone setup allows for different options to create a stereophonic audio mixture. Two coincident microphone techniques [8] (in which spatial image is only based on intensity differences) were adopted: a pair in mid-side (MS) configuration and a pair in XY pattern. The XY microphone pair was mainly intended as a backup and was connected to a separate portable digital audio recorder. Unlike the fixed spatial image of the XY pattern, the MS technique gives some flexibility to adjust the width of the stereo spread after the recording is finished. Besides, the mid MS channel grants monoaural compatibility, whereas collapsing the XY tracks into mono can result in some phase cancellation in high frequencies. Finally, since both pairs were placed close to each other, this audio redundancy can be used to study the influence of the microphones on problems such as source separation or sound recognition.

In addition, a spaced microphone technique [8] (which also captures time-of-arrival differences) was applied, using a pair of omnidirectional microphones in AB configuration. The AB pair was placed in a T-shape array with respect to the coincident stereo pairs, so that they form a three-point pick-up pattern called *Decca tree* [8] (after the record company). By combining the coincident and spaced microphone techniques into a stereo downmix, a good compromise between spatial sense, stereophonic image and center definition can be obtained. Furthermore, the low frequency response is improved by the contribution of the omni microphones, compared to that of directional microphones such as the ones used in the coincident pairs.

Separate audio channels were obtained by close-miking each drum and recording them to independent tracks. Yet, achieving good source separation given the high sound pressure level produced by the drums turned out to be challenging. Previous tests were conducted with acoustic panels standing in-between musicians (called gobos) to reduce spill from one instrument into the other spot microphones. Though effective, they interfere with a natural performance and were discarded. Eventually, a set of dynamic microphones–tailored for percussion instruments–yielded the best results among the tested options. They are able to handle high pressure levels and their moderate sensitivity prevents them from picking too much sound from other drums.

Table 1, summarizes the microphone setup and gives additional details, including manufacturer and model of microphones and recorders. All the microphones involved in recording the spatial sound image are condenser transducer type, because of their higher sensitivity to capture distant sounds, reverberation and nuances. Audio was recorded at a sampling rate of 48 kHz and 24-bit precision. An outlook of the stage is shown in Figure 1, during the performance of a five-drum ensemble.

Sound Devices 788T-SSD				
channel 1	Schoeps	MK 4	cardioid	mid
channel 2	Schoeps	MK 8	figure 8	side
channel 3	Sennheiser	E604	cardioid	spot
channel 4*	Sennheiser	E604	cardioid	spot
channel 5	Sennheiser	E604	cardioid	spot
channel 6	Sennheiser	E604	cardioid	spot
channel 7	AKG	C414	omni	left
channel 8	AKG	C414	omni	right
Tascam HD-P2				
channel 1	Neumann	KM 184	cardioid	left
channel 2	Neumann	KM 184	cardioid	right

* Neumann TLM 103 in isolated strokes and solo performances.

Table 1. Input list and microphone setup for audio recording.

Figure 1. Stage outlook during a five-drum performance.

3.2. Video recording

For documentary purposes a conventional video filming set was used. It comprises three video cameras equipped with professional grade lenses, two Canon 7D and a Sony alpha-99. Video was recorded at 24 frames per second (fps) and 1920x1080 pixel resolution (HD, H.264). One of the cameras always captured a wide shot of the ensemble, while the other two focused on closed-up views of the *repique* and *piano* players, as depicted in Figure 2. A traditional clapperboard was used for embedding a visual and aural reference to synchronize different audio and video sources.

Due to the very fast movements of an accomplished percussionist playing Candombe, the computational analysis of a performance based on video records at conventional frame rates (e.g. 24 fps) is quite limited. For this reason, the use of high-speed cameras was also considered. From a survey of commercially available cameras of this type, the GoPro HERO3+ Black Edition was chosen as an affordable option. Coincidently, it has been recently reported to be used for similar purposes [22]. It is a rugged and compact action camera, primary associated with outdoor sports, often attached to helmets or surfboards. Thus, various technical issues had to be tackled to adapt it for the current application.

Based on previous tests, a frame rate of 240 fps and 848x480 pixel resolution was selected. This produced reasonably smooth data and low blur of moving objects, and proved to be suitable for automatic processing. At such high frame rates, fluctuations in the lighting conditions are critical and the power-line flicker effect can arise. Experiments conducted in advance in the concert hall suggested that the change in brightness of the existing incandescent lamps was not troublesome. However, moderate lighting fluctuations (almost not perceivable to the naked eye) are present in the recorded videos, and can hinder the performance of video processing algorithms if not dealt with properly.

Two cameras of the same type were used simultaneously, which –apart from hardware backup– provided a stereo recording that can be used for 3D analysis and scene construction (see Figure 3). This entailed a calibration process with a checkerboard pattern during the recording session.

By rotating the wide angle field of view by 90°, a more adequate framing was obtained. The camera has no view finder, so it had to be connected to a display screen in order to check the correct framing of each video take. A remote control allows for basic operation of the device, but the wireless communication has to be enabled, which is a power demanding feature. For this reason, battery endurance was not enough and connection to a power supply became mandatory. Consequently, the need of wiring the camera, for monitoring and powering, prevented the use of the standard housing and mounting accessories. Besides, only relying on the remote control turned out to be not sufficiently robust and access to the buttons had to be granted. Similarly, access to the SD card slot was also necessary (which is not possible with the housing), to readily transfer the large files generated at high frame rate. Given all these requirements, a custom camera mounting had to be made, capable of holding two cameras in upwards position, with an adjustable distance between them and providing access to all the necessary connections and buttons. This mounting can be seen in Figure 4, attached to a standard tripod. Because of the fixed focal length of the camera, the tripod had to be placed not very far from the subject to get a detailed view. As a result, only one musician was effectively recorded at high-speed in ensemble performances.

The scene was slightly prepared for research purposes, to aid the application of automatic video processing techniques and to simplify the evaluation of algorithms. This was done in a non-invasive manner, taking care not to alter the sound of the drums or disturb the performer. As shown in Figure 3, the stick and the contour of the drumhead were painted to ease their automatic segmentation. In addition, some fiducial paper markers were attached to the floor, the drum and the performer's body for evaluation purposes. All the drums used were 3D scanned using a Kinect system to accurately register shape data in case it could be useful in future research.

4. DATASET

During the recording session, three different types of content were registered, namely, isolated strokes, solo performances and drum ensembles. The dataset comprises 51 takes, totaling a duration of nearly an hour and a half. This section provides a short description of how each type of content was rendered, as well as their intended purposes.

Some musicians were asked to change role during the session and played different drums. Besides, for the sake of data variability, different instruments of the same drum type were used. To do that, two instruments of each type (i.e *chico*, *repique* and *piano*) were utilized. This set of drums was previously prepared, by painting the drumhead contour

Figure 2. Video framing examples, a wide shot of the ensemble (left) and close-up view of the *repique* performer (right).

Figure 3. Matching frames of the stereo recording pair.

Figure 4. High-speed cameras in stereo pair with custom-made mounting, wired to two screens and power supply.

and attaching some fiducial paper markers to the shell and drumhead (as shown in Figure 3). But in addition, the performers were asked to bring their own instruments–which where obviously not prepared–and part of them were involved in recording isolated strokes and drum ensembles. The number of takes of each content type is shown in Table 2, and this information is expanded in Table 3 by giving the number of takes in which each performer is involved.

Strokes		**Solo**		**Ensembles**	
chico	4	chico	8	three drums	9
repique	4	repique	10	four drums	3
piano	3	piano	8	five drums	2

Table 2. Number of takes of each content type.

	Strokes			**Solo**			**Ensembles**		
	chico	repique	piano	chico	repique	piano	chico	repique	piano
Silva		1			3			8	
Suárez	1	1	1	3	2	3	2	1	6
Ortuño	2	1	1	2	2	2	2	8	1
Oviedo			1			3			9
Giménez	1	1		3	3		10	2	

Table 3. Number of takes by each performer.

4.1. Isolated drum strokes

The musicians were asked to produce the sound of individual strokes separately, and were recorded in turns playing different drums. These strokes are supposed to be the same they would use in a real performance. Therefore, they were requested to render some particular stroke types, but also to include those that belong to their personal repertoire.

The set of isolated drum sounds is intended for two main different uses. Firstly, to provide a database of strokes in order to train and evaluate sound recognition algorithms, which are typically part of automatic transcription systems. Secondly, to build a set of audio samples of each stroke type to be used in sample-based audio synthesis. In previous research, software tools were implemented to produce rhythmic patterns and synthetic performances. To that end, several samples of each type of stroke are randomly selected and located according to predefined rules. To precisely synthesize audio files from scores in this way, proved to be a valuable research tool, useful for performing tests in a controlled framework. For instance, it avoids some aspects concerning a real performance, such as micro-timing variations, and simplifies the creation of ground-truth labels.

4.2. Solo drum performances

Despite that Candombe is –above all– a collective form of music, part of the session was devoted to recording each musician alone executing the rhythmic patterns of a certain drum. They were asked to play at different tempos and in different styles, and freely developed their improvised parts.

The aim of these recordings is to avoid any interference from other drums, something which is not completely fulfilled by the separate channels of an ensemble recording. In this way, automatic tasks such as onset detection can be carried out without the need to deal with spurious events from other drums. In addition, this allows, for instance, to study spectral timbre features for sound recognition in a more realistic situation compared to that of the isolated strokes. Ultimately, these performances can be contrasted with that of ensembles, in order to investigate to what extent musical behaviours are alike.

4.3. Drum ensembles

Ensembles were recorded in groups of three, four and five performers. The first case corresponds to a Candombe ensemble in its minimal form, that is, one of each of the three drums. In the case of four performers, an extra *repique* was added to the ensemble. Finally, groups of five drums consist of two *piano*, two *repique* and one *chico*.

There are 14 complete ensemble recordings, which last from about 2 to 4 minutes each, for a total of 40 minutes. The different groups, which involved several combinations of the same musicians (as well as some of them in different roles), yielded various sorts of interactions and performance types. For instance, different characteristics typologies were executed for beginning the performance, including *piano* anacrusis and several variations of the *clave* pattern. In the same way, different dynamics and tempo–as well as variations of them–were also performed.

Care was taken to alternate the type of drum in front of the high-speed cameras, in order to have a balanced record of *repique* and *piano* performances in ensemble. The resulting separate audio channels are adequate for several automatic tasks such as onset detection, transcription and source separation. The amount of interference from other drums depends on the distance between performers. For this reason, the best results were obtained for the recordings of three drums, in which the performers were farther apart.

Evidently, the most important aspect provided by ensemble performances compared to the previous recordings is the interplay between musicians. This encompasses call-and-response interactions, alternations of musical roles, variations in dynamics, temporal synchronization and collective modulations of tempo, just to name a few. Among the various aspects that can be studied from these recordings, the different forms of interaction and the entrainment processes involved are some of the most appealing. Besides, the different sorts of musical embodiment are also of great interest (for example, foot tapping can be observed in the wide shot of the ensemble, Figure 2-left).

5. ANNOTATIONS

Manual labeling of music is notably a laborious process, for which software tools have been developed [5] and methodologies and file formats have been proposed [14]. The annotations are very useful for both, musicological studies and the development of music information technologies. This section describes some of the efforts that are being conducted regarding the annotation of this dataset.

5.1. Beat and downbeat annotation

The location of beats and downbeats was annotated by one of the authors for the 14 recordings of ensembles. The annotations are available for the research community,[3] together with the corresponding stereo audio files. These recordings and labels are part of the dataset used in a previous research work on beat and downbeat tracking [17]. Besides, the annotation of solo performances is currently being carried out and will be also released. It has already been completed for the piano recordings, to be used in an ongoing research.

The annotations are provided as comma-separated values files (.csv) in which data is stored as plain text, as shown in Figure 5. The files contain two columns, and each line corresponds to a beat. The values in the first column are the time instants of the beats in seconds. The numbers on the second column are two values separated by a dot, which indicate the bar number and the beat number within the bar, respectively. For instance, 1.1, 1.2, 1.3 and 1.4 denote the four consecutive beats of the first bar. Hence, each label ending with .1 corresponds to a downbeat.

[3] Available from `http://www.eumus.edu.uy/candombe/datasets/ISMIR2015/`

```
1.78875,   1.1
2.31252,   1.2
2.83536,   1.3
3.35711,   1.4
3.87728,   2.1
4.39545,   2.2
4.91217,   2.3
5.42815,   2.4
...    ,   ...
```

Figure 5. Example of the format of the beat annotation files.

5.2. Annotation of onset and stroke type

There is another effort underway to annotate the location of each onset and its stroke type. To this effect, solo performances and separate channels of the ensemble recordings are used. The annotation process is facilitated by automatically locating events through a standard onset detection algorithm based on the Spectral Flux [7]. The resulting events are manually validated and/or corrected, by inspecting the audio and video files. Finally, a certain class label representing the stroke type is manually assigned to each onset.

These annotations can be used in research on onset detection, sound recognition and automatic transcription. The semi-automatic annotation process described above has been applied to *repique* solo performances from this dataset (as well as other recordings from the preliminary studio session). The annotations, audio files and high-speed videos, were used for the development of an audio-visual transcription system in previous work [15]. Eventually, all this material will be available to the research community.

5.3. Miscellaneous annotations

Several other sorts of annotations are being generated, which depend on the type of research problem addressed and the particular characteristics of the dataset. For example, a recent experiment was conducted attempting to identify those temporal segments of a *repique* performance when the *clave* pattern is played [18]. This involved manual annotation of separate tracks from ensemble recordings and yielded very informative data about the interaction of two *repique* performers playing together.

With regards to the high-speed videos, the precise annotation of the most important objects that appear in the scene–namely, the drumhead, the stick and the hand of the performer–is necessary for the development and assessment of automatic detection algorithms [15]. This is a very time consuming task, and efficient labeling mechanisms are being developed based on the fiducial markers.

6. DISCUSSION AND FUTURE WORK

Apart from the documentary value and its potential in educational activities, there are different research approaches that can benefit from this dataset. To gain insight into the possibilities that it offers, some examples of previous research works that make use of it were given. A detailed description of the data generation process was also provided, in the hope that it can be useful for other researchers involved in producing datasets for computational analysis of music. A continuous effort to extend the existing annotations of the dataset has to be undertaken, in order to expand its possible applications. The whole data including the annotations, will be ultimately available to the research community to foster reproducible research.

In future work, audio recordings and metrical annotations will be used for the analysis of rhythmic patterns, as in [19], which disclose relevant information about the style of different traditions and individual players. Besides, microtiming studies will be conducted [16], by combining the metrical information provided by beat and downbeat labels with the temporal location of events given by the onset annotations. Among other research strands, the analysis of entrainment using this dataset is part of the future work perspectives [6].

7. ACKNOWLEDGMENTS

The authors wish to acknowledge the performers and all the people involved in the creation of the dataset. Thanks to Rafael Barbé, Pablo Cancela, Andrés Costa, Juan Ignacio Fernández Hoppe, Juan Martín López, Guillermo Rocamora, Federico Sallés and Mateo Soler from the recording crew. Our gratitude to the percussionists that took part in various pilot tests and in the preliminary recording session: Victoria Bonanata, Juanita Fernández, Néstor Ferreira and Ricardo Gómez. And thanks to Juan Cardelino for fruitful discussions. The custom camera mounting was crafted by Roberto Rodríguez and Sergio Beheregaray. Editing and synchronization tasks are being conducted by Guillermo Carter.

This work was partially funded by Comisión Sectorial de Investigación Científica (CSIC) and Agencia Nacional de Investigación e Innovación (ANII), Uruguay.

8. REFERENCES

[1] G. Andrews, *Blackness in the White Nation: A History of Afro-Uruguay*. Chapel Hill: The University of North Carolina Press, 2010.

[2] B. Bartlett and J. Bartlett, *Recording Music on Location*, 2nd ed. Oxford: Focal Press, 2007.

[3] T. Bertin-Mahieux, D. P. W. Ellis, B. Whitman, and P. Lamere, "The million song dataset," in *Proc. of*

the 12th International Society for Music Information Retrieval Conference (ISMIR 2011), Miami, Florida, USA, Oct. 2011, pp. 591–596.

[4] R. M. Bittner, J. Salamon, M. Tierney, M. Mauch, C. Cannam, and J. P. Bello, "MedleyDB: A multi-track dataset for annotation-intensive MIR research," in *Proc. of the 15th International Society for Music Information (ISMIR 2014)*, Taipei, Taiwan, Oct. 2014, pp. 155–160.

[5] C. Cannam, C. Landone, M. B. Sandler, and J. P. Bello, "The sonic visualiser: A visualisation platform for semantic descriptors from musical signals," in *Proc. of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, Oct. 2006, pp. 324–327.

[6] M. Clayton, R. Sager, and U. Will, "In time with the music: the concept of entrainment and its significance for ethnomusicology," in *European meetings in ethnomusicology., 11 (ESEM Counterpoint 1)*, 2005.

[7] S. Dixon, "Onset detection revisited," in *Proc. of the 9th International Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada, Sep. 2006, pp. 133–137.

[8] J. Eargle, *The Microphone Book*, 2nd ed. Oxford: Focal Press, 2004.

[9] L. Ferreira, "An Afrocentric Approach to Musical Performance in the Black South Atlantic: The Candombe Drumming in Uruguay," *TRANS-Transcultural Music Review*, vol. 11, p. 12, 2007.

[10] O. Gillet and G. Richard, "Enst-drums: an extensive audio-visual database for drum signals processing," in *Proc. of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, Oct. 2006, pp. 156–159.

[11] M. Goto, "AIST annotation for the RWC music database," in *Proc. of the 7th International Conference on Music Information Retrieval (ISMIR 2006)*, Victoria, Canada, Oct. 2006, pp. 359–360.

[12] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC music database: Popular, classical and jazz music databases," in *Proc. of the 3rd International Conference on Music Information Retrieval (ISMIR 2002)*, Paris, France, Oct. 2002.

[13] ——, "RWC music database: Music genre database and musical instrument sound database," in *Proc. of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*, Baltimore, Maryland, USA, Oct. 2003.

[14] M. Lesaffre, M. Leman, B. D. Baets, and J. Martens, "Methodological considerations concerning manual annotation of musical audio in function of algorithm development," in *Proc. of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*, Barcelona, Spain, Oct. 2004.

[15] B. Marenco, M. Fuentes, F. Lanzaro, M. Rocamora, and A. Gómez, "A multimodal approach for percussion music transcription from audio and video," in *Proc. of the XX Iberoamerican Congress on Pattern Recognition (CIARP)*, Montevideo, Uruguay, Nov. 2015, p. (submitted).

[16] L. Naveda, F. Gouyon, C. Guedes, and M. Leman, "Microtiming patterns and interactions with musical properties in samba music," *Journal of New Music Research*, vol. 40, no. 3, pp. 225–238, 2011.

[17] L. Nunes, M. Rocamora, L. Jure, and L. W. P. Biscainho, "Beat and downbeat tracking based on rhythmic patterns applied to the Uruguayan Candombe drumming," in *Proc. of the 16th International Conference on Music Information Retrieval (ISMIR 2015)*, Málaga, Spain, Oct. 2015, p. (to appear).

[18] M. Rocamora and L. W. P. Biscainho, "Modeling onset spectral features for discrimination of drum sounds," in *Proc. of the XX Iberoamerican Congress on Pattern Recognition (CIARP)*, Montevideo, Uruguay, Nov. 2015, p. (submitted).

[19] M. Rocamora, L. Jure, and L. W. P. Biscainho, "Tools for detection and classification of piano drum patterns from candombe recordings," in *Proc. of the 9th Conference on Interdisciplinary Musicology (CIM14)*, Berlin, Germany, Dec. 2014, pp. 382–387.

[20] A. R. Tindale, A. Kapur, G. Tzanetakis, P. Driessen, and A. Schloss, "A comparison of sensor strategies for capturing percussive gestures," in *Proc. of the International Conference on New Interfaces for Musical Expression (NIME '05)*, Vancouver, Canada, 2005, pp. 200–203.

[21] UNESCO, "Representative list of the intangible cultural heritage of humanity – Candombe and its socio–cultural space: a community practice," 2009, http://www.unesco.org/culture/ich/RL/00182.

[22] R. Van Rooyen and G. Tzanetakis, "Pragmatic drum motion capture system," in *Proc. of the International Conference on New Interfaces for Musical Expression (NIME '15)*, Lousiana, USA, 2015, pp. ??–??

[23] A. Volk, F. Wiering, and P. van Kranenburg, "Unfolding the potential of Computational Musicology," in

Proc. of the International Conference on Informatics and Semiotics in Organisations, Leeuwarden, Netherlands, Nov. 2011, pp. 137–144.