

# COLLECTING GROUND TRUTH ANNOTATIONS FOR DRUM DETECTION IN POLYPHONIC MUSIC

Koen Tanghe<sup>1</sup>, Micheline Lesaffre<sup>1</sup>, Sven Degroeve<sup>2</sup>,  
Marc Leman<sup>1</sup>, Bernard De Baets<sup>2</sup> and Jean-Pierre Martens<sup>4</sup>

<sup>1</sup> IPEM, Department of Musicology, Ghent University, Blandijnberg 2, 9000 Ghent, Belgium  
Koen.Tanghe@UGent.be

<sup>2</sup> Department of Applied Mathematics, Biometrics and Process Control, Ghent University

<sup>4</sup> Department of Electronics and Information Systems, Ghent University

## ABSTRACT

In order to train and test algorithms that can automatically detect drum events in polyphonic music, ground truth data is needed. This paper describes a setup used for gathering manual annotations for 49 real-world music fragments containing different drum event types. Apart from the drum events, the beat was also annotated. The annotators were experienced drummers or percussionists. This paper is primarily aimed towards other drum detection researchers, but might also be of interest to others dealing with automatic music analysis, manual annotation and data gathering. Its purpose is threefold: providing annotation data for algorithm training and evaluation, describing a practical way of setting up a drum annotation task, and reporting issues that came up during the annotation sessions while at the same time providing some thoughts on important points that could be taken into account when setting up similar tasks in the future.

**Keywords:** drum detection, annotation, data gathering

## 1 INTRODUCTION

Drum events provide important clues about the rhythmical organisation of a musical piece. For many music genres nowadays, rhythmic structures have become (at least) equally important as melodic or tonal structures. In the same way that melody lines can be seen as a representation of one aspect of the musical content of a piece of music, drum sequences can be seen as another type of musical content representation, but more related to rhythm. This information could be used to allow people to search for a particular drum sequence or a typical drum pattern, for automatically classifying music pieces into different genres or subgenres, or for obtaining information about tempo and metrical structure.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2005 Queen Mary, University of London

Research on drum detection is relatively new (compared to melody-related research) and there is a real need for reference material to train and test drum detection algorithms. One of the interesting initiatives for which this type of reference material is important is the MIREX “contest” track [1] of the annual ISMIR conference [2]. The goal of this “contest” is to compare state-of-the-art algorithms and systems relevant for Music Information Retrieval (MIR), and ground truth data is needed in order to perform objective evaluations. Some efforts towards setting up realistic databases of reference music and annotations have already been made (e.g. the Real World Computing (RWC) Music Database [3]), and some people have also started developing specialized annotation tools to ease the cumbersome task of performing manual annotations (e.g. the Sound Onset Labelizer in [4] or the semi-automatic beat annotation tool in [5]). In contrast to the speech analysis research field, however, the music analysis field in general does not (yet?) have flexible music annotation tools or large sets of real-world annotated data that can serve as reference material and thus as a catalyst for faster and higher-quality algorithm development. We believe that research projects dedicated solely to data gathering and preparation will be needed in the near future. We hope that the results of our little annotation task may be a valuable contribution in that sense.

## 2 CONTEXT

In the context of the Musical Audio Mining (MAMI) project [6], research is being done on extracting drum events from polyphonic music. Algorithms have been designed and software has been developed to localize and label drum events, based on a model consisting of three main parts: onset detection, feature extraction and feature vector classification.

There are a few reasons why we need realistic ground truth data for this drum detection system. The first reason is that it contains a machine learning algorithm for the classification part that needs to be trained in a supervised way on data that represents the true task of drum detection before it can actually do anything useful. So, the labelled data is used to train the drum type classification models.

Another reason why we need the ground truth data is parameter optimization. The components of the system typically have a certain number of parameters that all have some influence on the performance of the overall system. In order to obtain a parameter combination that

leads to a good overall performance of the system, it is necessary to optimize the various parameters over a set of realistic music examples. This optimization process requires minimizing some cost function that indicates how bad the system performs. And the best way to obtain a “badness indication” is to compare the detected drum events with the “correct” drum events (ground truth data) and to calculate a quantitative error measure from the comparison results.

Finally, as already stated in the introduction, ground truth data is very useful for comparing various drum detection systems against each other in a systematic way. That gives us an objective indication of how well our trained and optimized system performs compared to other systems, which is interesting in se, but it also tells us (and the designers of the other systems) something about the strong and weak points of the used algorithms.

### 3 DATA AND PEOPLE

#### 3.1 Music

Our goal is to make the drum detection system work on “real music”. We are not interested in making it work for a few carefully selected test examples in a laboratory environment. Of course, working with special cases or selected examples can be useful for evaluating some specific aspects of the various parts of the system. Also, working with MIDI files can be an interesting first step when quickly trying out and comparing various algorithms. But the end goal is still drum detection on fully produced music (both recorded and sequenced) from different popular genres.

We collected a set of 52 music fragments digitally extracted from various commercial music CD’s. The fragments are 30 seconds long and in 16 bit 44100 Hz stereo PCM WAV format. In selecting the music, we have tried to make a compromise between diversity and annotator preferences. Having as much different styles and genres as possible is important to evaluate our algorithms for robustness and flexibility, but making sure that our annotators are familiar with the music is also important because it usually leads to more reliable annotations. Therefore, prior to actually setting up the annotation task, we had asked the annotation candidates to answer a short list of questions about (amongst others) the genres and styles of music they are most familiar with. Since most of our candidates also took part in a separate online inquiry aimed at recruiting a large group of subjects willing to participate in diverse annotation experiments [7], we have also taken into account the music they had then specified as being their “favourite”. A full list of the music we have used can be found at the end of this paper in Table 3.

Apart from this set of “real music” fragments, we also added three other fragments: a very simple, self-made reference file containing clear drums of various types and no music, and two recordings of MIDI files with drums and music. These fragments were meant as

“reference” material to get an idea about the annotation quality. Since we generated these audio fragments from a symbolic representation, it is possible to compare the manual annotations with the true events.

#### 3.2 Annotations

The goal of the annotation task was to come up with reliable ground truth data that represents the positions in the sound files *where* drum events occur, together with labels specifying *which* types of drum events are occurring at these locations. Unlike monophonic melody lines, drum events can overlap in time, so for a specific position, multiple drum events can occur at the same time. The 18 types of drum events we have considered are listed in Table 1. While we initially planned to use only 6 types of drum events (BD, SD, HH, CY, TM and “other”), feedback from the annotators during preliminary tests showed that some other drum types were considered as being important nonetheless. And since these other types would make no big difference in the annotation task itself, we decided to go for a more elaborated list of drum types. For our current algorithms, we can still reduce these full annotations by remapping or omitting some drum types to only a set of very basic drum types. For future versions that might be able to deal with more drum types, we can then use the full annotations. The annotations were stored in a MIDI file where a MIDI “note on” message per annotated drum event encodes both the position and drum type.

Since we are also interested in tempo detection for a later stage of our research, we asked the annotators to tap along with the beat of each music fragment. This information is again stored in a MIDI file where each beat is represented by a MIDI “note on” message with MIDI note number 76.

**Table 1.** Overview of the annotated drum types with their labels and MIDI note numbers.

Full name	Label	Note
bass drum	BD	36
snare drum	SD	40
open hi-hat	OH	46
closed hi-hat	CH	42
ride cymbal	RC	59
crash cymbal	CC	57
low tom	LT	45
mid tom	MT	47
high tom	HT	50
claps	CP	39
rim shot	RS	37
splash cymbal	SC	55
shaker	SH	70
tambourine	TB	54
wood block	WB	77
low conga	LC	64
high conga	HC	63
cow bell	CB	56
other drum	-D	75

### 3.3 Annotators

Annotation of the different fragments was performed by a team of 10 experienced male<sup>1</sup> drummers and percussionists varying in age from 23 to 57 (half of them were 25 or younger though). Most of the annotators were selected from the pool of participants in the online inquiry mentioned in section 3.1. We chose the ones who indicated that they play a percussion instrument and who thought of themselves as having a high level of musicality. Some were also recruited directly through other connections because they were known as good drummers. They were all volunteers but were nonetheless financially compensated for their time and efforts. We did not want to “use” students for the task because we think it requires a more than average familiarity with drum sounds and a lot of true motivation to perform the task rigorously.

As already mentioned, the annotators had been asked in advance to answer a short list of questions about the music styles/genres they are most familiar with. Other questions were related to their acquaintance with sequencer software/hardware, the (in their opinion) relative importance of various drum sound types and the appropriateness of different methods for entering a drum sequence into a computer. All these answers have been taken into account while setting up the annotation task.

Finally, all annotators signed a statement saying that they give us permission to freely use the gathered data for research purposes.

## 4 ANNOTATION METHOD

### 4.1 Setup

After a short evaluation of a few available candidate annotation tools, we decided to use a standard music production software package with multi-track audio and MIDI sequencing capabilities, namely Cakewalk Sonar<sup>2</sup>. A summary of the advantages and disadvantages of the considered programs is given in Table 2.

Each annotator was presented with a Sonar multi-track project consisting of 1 audio track for the music fragments, 1 MIDI drum track for the drum annotations, and 1 MIDI drum track for the beat annotations (see Figure 1). The music fragments were placed after each other on the audio track (usually between 4 and 8 fragments per annotation session) with a gap of 5 seconds of silence in between. The top half of the screen showed the three tracks and standard controls, and the bottom half showed a drum grid view of either the drum track or the beat track, depending on which one was selected. Navigation was done by the mouse and by special keyboard shortcuts (e.g. for jumping to the start of a fragment). For starting and stopping playback and re-

coding, the standard transport buttons (and the corresponding keyboard shortcuts) could be used. Snapping to a grid and quantization were turned off completely, and time was displayed in seconds (we didn’t want to force the tempo track of the project to match the tempo of the fragments to avoid biased timings).

**Table 2.** Advantages and disadvantages of a few candidate annotation tools

<b>Cakewalk Sonar</b>	
+	easy-to-use interface
+	allows you to listen to your annotations and also to a mix of the original and your annotations
+	many musicians already know how to use this type of program
+	possibility of using a MIDI input device
+	synchronization between audio and annotations (MIDI) is automatic
+	avoids annotation errors (no values or data entering in a difficult GUI as in Praat)
-	no possibility to see audio and annotations (audio and MIDI track) in the same edit window
<b>Praat</b>	
+	audio and annotations in the same window
+	easier to add complicated annotations (reliability scores etc...) compared to using MIDI CC’s
-	no possibility to listen to your annotations (!)
-	unknown by annotators, not focused on music
-	user interface is not so easy to use, editing is a bit difficult for music annotations
-	potential for making typing errors, when typing in the occurring drum types is needed
<b>Audacity</b>	
+	has potential for the future: open source C++ code, possibility to extend with own code
-	mark tracks are not fully worked out (almost no editing possibilities)

The MIDI tracks were connected to a virtual sound module, which means that the annotators could actually listen to their annotated events by playing back the annotation track through the sound module. Of course, this could also be done while the original music was playing back at the same time. Volume controls for each track could then be used to setup the right balance between both signals, and stereo panning could be used to playback the original sound in one channel (left) and the sound of the annotated events in the other (right). The audio track was also supplied with a filter that could be switched to low-pass, band-pass or high-pass mode in order to make it easier to listen to specific spectral regions of the music (e.g. low-pass filtering to ease the annotation of the bass drum events).

We also attached a “tweaked” MIDI keyboard to the annotation computer as an alternative to entering all the drum events one by one into the drum grid view. The MIDI keyboard was “tweaked” in the sense that we stucked drum labels on a range of white MIDI keys and inserted a strip of insulation material underneath these

<sup>1</sup> This was not a deliberate choice. We just didn’t find female drummers or percussionists who wanted to participate in the annotation task.

<sup>2</sup> Sonar is a registered trademark of Twelve Tone Systems, Inc.

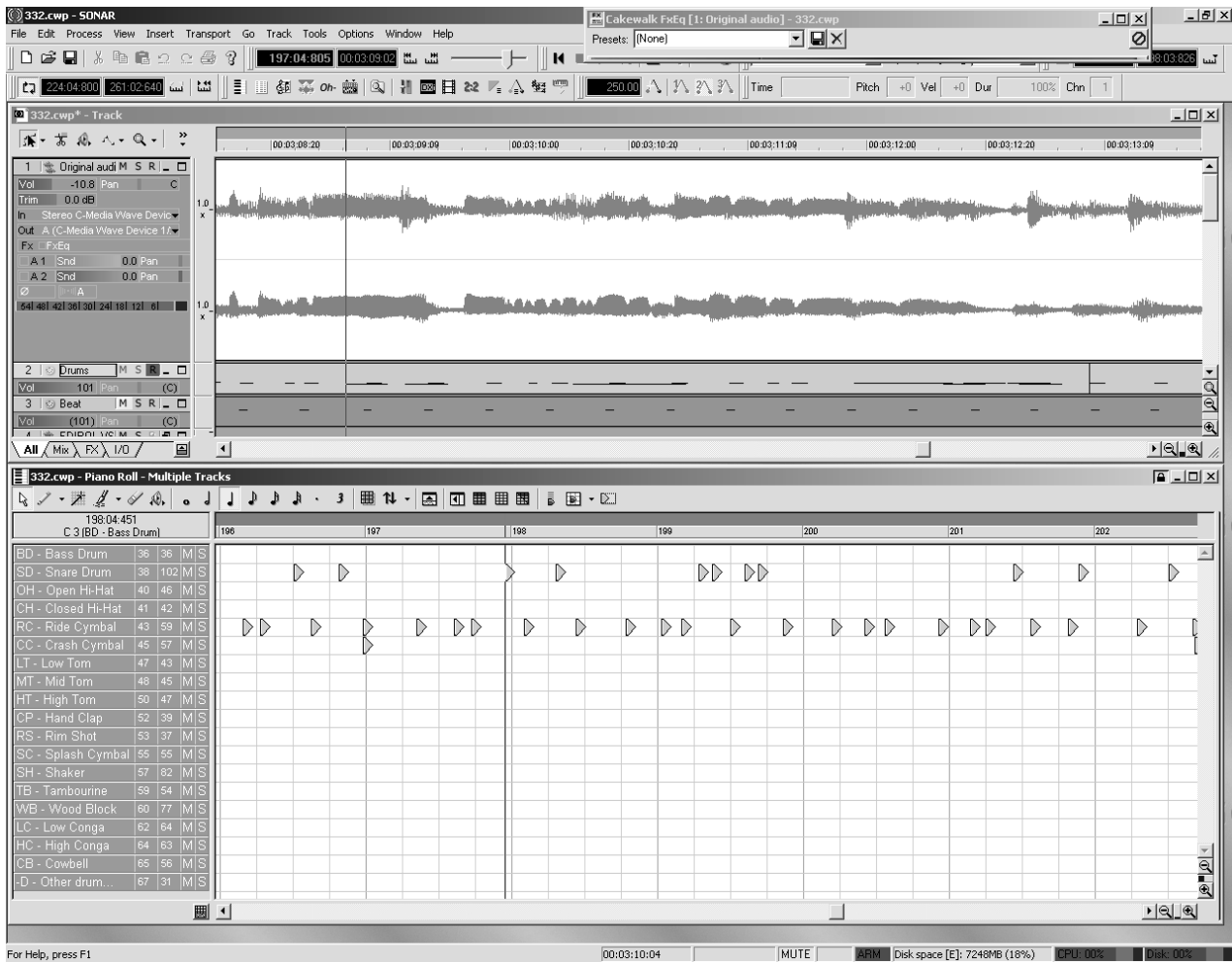


Figure 1. A multi-track sequencer drum and beat annotation setup in Cakewalk Sonar 2.2

keys to limit their downward movement and make it easier to play fast drum sequences.

Of course, a combination of visual editing and MIDI keyboard recording was possible too. Usually, the drummer would start the annotation task by recording a few drum parts played on the MIDI keyboard while listening to the original music. This recording process would then be repeated for a few other drum parts, until the most important drums were registered. Then, the drummer would start adjusting or extending the recorded annotations by moving, adding or deleting drum events in the drum grid view using the graphical interface. At any time it was possible to playback the annotated events together with the original music in order to verify the accuracy of the annotations.

As for the physical setup: the annotation task was set up in a standard office room with an old DX7 MIDI keyboard connected through a Midisport2x2 USB MIDI I/O device to a WindowsXP PC running Cakewalk Sonar 2.2. The used soundcard was a low-latency M-Audio Delta, the sound module for playing back the drums was an Edirol Virtual Sound Canvas DXi shipped with Sonar, and the annotators were wearing medium-quality headphones.

## 4.2 Guidelines

In order to make sure that all annotators fully understood the purpose of the task and how the offered setup could be used to complete it, we prepared a small document with guidelines that was verbally presented to them by one of the organizers and was left at their disposal in the annotation room as a reference.

At the one hand, this document served as a practical user guide for the annotation setup: it included a screenshot of the Sonar user interface with a description of the most important views and controls and how to interact with them. Of course, one of the organizers was always available in case there were questions or problems.

On the other hand and maybe more importantly, given the fact that we were working with different annotators, these guidelines also contained a few more semantics-related thoughts on how to annotate the music in a consistent way. For example, it was stressed that although the specific *sound* of the played back drum samples may differ from the ones found in the original music fragment, it is nonetheless very important to choose the correct drum *type*. Also, if the music frag-

ment contains a drum type that is not contained in the given list of drum types, and the annotator thinks it is nonetheless very important, he should annotate the drum as “other drum”. However, the annotators were asked to use the given list of drum types as much as possible. For example: if a sound fragment contains two types of crash cymbals, they should annotate both types with the same label for “crash cymbal”, instead of using “other drum” for one of the two types.

Another point we wanted to emphasize, is that both the timing (exact location) and the labelling (choice of the drum types) are equally important. We prefer to collect fewer annotations of good quality than more annotations that are hardly usable as ground truth.

Finally, we encouraged the annotators to write down any thoughts, remarks, suggestions and problems that came up during the annotation session.

## 5 RESULTS AND DISCUSSION

### 5.1 Revision of initial goals

Initially, we built up each annotation setup out of 3 reference fragments and 4 real music fragments, but we decided to omit one of the reference fragments so that there was more time left for the real fragments. An annotation session lasted roughly one morning or afternoon (about 4 hours, depending on the difficulty of the fragments). Annotators who came back for a second or third annotation session could immediately start with the real music fragments.

We also intended to have each fragment cross-annotated by three different annotators so that we could check the inter-subject consistency of the annotations, but since obtaining a single reliable annotation for each fragment was already hard enough to accomplish, this idea had to be dropped in the end. Only the self-made reference file and one of the two recordings of the MIDI reference files have been cross-annotated by 9 different annotators. At the time of writing, however, a detailed analysis of the inter-annotator agreement for these reference fragments has not been performed yet.

### 5.2 Difficult drum types and articulations

From the remarks and suggestions of the annotators, it was clear that some drum types and articulations are very difficult to annotate. Brushes for example have a typical “dragged” sound which is hard to annotate as a single percussive event. In this case most annotators chose to register the accents of the brush sounds. Snare rolls do consist of a series of discernable percussive onsets, but it’s very hard to annotate the many fast strokes accurately. The same is true for “flammed” drums (typically the snare drum) where two hits of the same drum type are deliberately played almost (but not quite) at the same time, leading to the sensation of a ghost note occurring slightly before a main note. The difficulties with hi-hats on the other hand have more to

do with the different ways in which the two metal discs of this instrument can be controlled while playing. After evaluating the answers to the questions we had asked the annotators in advance, we had already decided to include both closed and open hi-hats, but a few annotators reported that half-open hi-hats and the typical sound of an open hi-hat closed by the foot pedal should also be included in our list of drum types. One annotator reported that he found it difficult at some point to decide whether he should annotate a series of drum sounds as hi-hat or as shaker events, while another one had some doubts about a particular sequence where it was unclear to him whether the percussive events he was hearing were generated by a strummed rhythm guitar or by a hi-hat. A similar type of confusion was also reported for fragments that were heavily post-processed with audio effects. Reverb typically smears out sudden events in the audio and filtering can alter the original timbre of an instrument to a degree where it becomes very difficult to keep recognizing it. Also, dynamics processing on the bass frequencies of a song can sometimes make it hard to perceive the bass drum and the bass lead as separate entities.

### 5.3 Use of multi-track software

Apart from the recurring request to provide a count-in sound at the start of the fragments, most annotators got around surprisingly well with our sequencer-based setup, even those who were not really acquainted with multi-track sequencer software. There were nonetheless a few types of mistakes that were caused by non-optimal use of the software. There were a couple of double events, probably due to re-recording a drum sequence played on the keyboard without deleting the already registered events, and some events had been recorded on both the drum track and the beat track, which is typically caused by forgetting to switch on or off the “arm for recording” button for the appropriate MIDI track.

Only one annotator clearly had troubles understanding how to use the setup. Apart from the sporadic errors already mentioned above, he had somehow shifted all audio fragments in time which led to big synchronization problems with his annotated events. He had also clearly copied/pasted parts of his annotations which resulted in many incorrectly annotated events, and he did not annotate the beat at all. Furthermore, he had reported that “there was too much info on the screen for a beginner”, which kind of makes sense, but funnily enough he had also asked if it was possible to “remove some instruments from the audio mix so that it would be easier to hear the drum sounds one by one”, which is of course precisely what we are trying to do with our automatic drum detection research. This person also complained that the drum grid view did not have any subdivision in measures or beats, which was a deliberate choice to avoid biased timings. All of this shows that the clearly explained purpose of the annotation task was not fully understood. These annotations were totally

unusable as ground truth, so we had to ask another annotator to come back and annotate these files correctly.

#### 5.4 Use of a MIDI keyboard interface

Our setup with a tweaked MIDI keyboard seemed to work sufficiently well for most drummers, although some problems did exist. In particular, we had a request to use more than one single MIDI key for each drum event, as this would facilitate the annotation of fast drum sound sequences by alternately pressing the two MIDI keys. There were also a few remarks about it being too hard to annotate complex sequences using a MIDI keyboard. This came up when drummers wanted to play in more than 2 drum parts using the keyboard.

It may be useful at this point to mention that we had also done some experiments using an electronic drum kit with MIDI output, in order to stay as close as possible to the natural way of drumming, but this didn't work out as expected. Many drum parts can not be played on a standard electronic drum kit alone, and require a broader range of drum types. Also, it's harder (physically) to switch between live playing on the drums and visual editing on the computer, which is very important in order to be able to evaluate, correct and overdub the played drum sequences easily. We also noticed that when playing live on an electronic kit, the drummer tends to start playing along with the music without actually making sure to play exactly what is being played in the music, which is of course very important for this task. Apart from that, there were also technical reasons why we decided not to continue with an electronic drum kit: sometimes drum events were "double triggered" or there was cross-talk between the different drum brain sensors (e.g. a tom sensor being triggered slightly when a hard snare is played). At some point, we had to completely disconnect a foot pedal because, although not explicitly being played at all, it was still being triggered by the arm and body movements propagating through the legs of the drummer while playing. These technical issues introduced ghost notes, which required manual ad hoc checking and correction, so it turned out that our intentions to use an electronic drum kit to avoid these manual post-processing efforts afterwards were not very realistic. Together with the higher setup cost and the disadvantages already mentioned above, we eventually felt that using an electronic drum kit would not really help us much for this task.

#### 5.5 Evaluation and corrections

After all annotations had been gathered, we manually went through each and every one of them to check their quality and select the best ones. Occurring mistakes ranged from double events (usually due to overdubs) or single missing events, to spurious events caused by copy/paste actions without checking if they are really accurate at the pasted location. The beginning and end of a fragment were most susceptible to mistakes. At the beginning of a fragment timing was sometimes a bit

jittery because it is difficult to start with the correct tempo immediately (hence the suggestion of adding a count-in at the start), whereas near the end, annotators may start feeling overconfident that the same is being repeated all the time, which may make them miss subtle changes that do occur and actually indicate semantically interesting moments (like the end of a measure or a short pause). This is somehow related to the already mentioned danger of starting to play along instead of focussing on doing an accurate transcription. There were also a few cases where similar drum event types were annotated with different labels (like two different snare drums) or where incorrect drum types had been used in order to make a difference between different "other" drum types (like udu, tabla and djembe). This "trying to make a distinction by using another (inappropriate) label" mistake also occurred for a few pitched drum types (two woodblocks with a different pitch, or more than 3 types of toms). As for the beat, a recurring problem was the missing or jittery beats near the start of a fragment, which is related to the "settling time" we need in order to lock on to a periodicity in the beat-range. Most drummers handled this by extrapolating the beat towards the start of the fragment, but the first beats are usually a bit jittery nonetheless. The beat events also sometimes ended up on the drum events track instead of the beat track. This has to do with forgetting to toggle the recording status of the different tracks, and might be avoided in the future by just using a single MIDI track and a special note number for the beat, although this may introduce a bit more work afterwards to split up the track in drum and beat events. In any case, only the very obvious mistakes were corrected (like a clearly missed cymbal, an extra bass drum where there is none, excessive timing jitter, events on wrong tracks, ...) in order to keep the annotations as "true" as possible.

Finally, we should mention that we had to eliminate 3 fragments from the original 52 because the annotators (remember: experienced drummers and percussionists) considered them impossible or too time-consuming to annotate. One of these fragments was an electronic music fragment with dense layers of drums in a complex rhythmical structure; another one was a jazz fragment with hard to follow drum sequences with lots of in-between hits and lush cymbals and hi-hats, and the third one was a fragment in which the hi-hats were very difficult to hear and in which lots of lower dynamics drum rolls occurred.

## 6 CONCLUSIONS

In this paper, we presented a method for collecting ground truth data for drum detection in polyphonic music. This method has been used by a group of 10 experienced drummers and percussionists for locating and labelling 18 types of drum events in 49 polyphonic music fragments from different popular genres. The beat for these fragments has been annotated as well. The obtained data can be found on the MAMI web site un-

der “Test collections and annotation material” in the public section of the site [6].

Being able to see the events and the audio at the same time, together with the possibility to actually hear the annotated events proved to be essential for the annotation process. Although we have managed to use a standard multi-track software package for this annotation task, we do believe there is a need for easy-to-use and flexible tools dedicated to music annotation. Apart from the already mentioned aural and visual feedback, things like multi-layer and hierarchically structured annotations, connectivity to external user interfaces, slowdown of fast audio sequences, flexible input and output formats are all features that were found to be important for gathering reliable and well-structured music annotations in a fast and comfortable way.

Finally, the choice of audio fragments should be carefully considered. For this annotation task, we had decided to use “real music” from commercial CD’s, but the problem with that is that we cannot distribute the music itself along with the annotation data due to copyright restrictions: we can only provide a documentation file with all the info needed to obtain the exact same fragment we used for our annotations. In the future, we will consider using music released under a more open license like the Creative Commons “attribution, non-commercial, share alike” license [8].

### ACKNOWLEDGEMENTS

This work was done in the context of the “Musical Audio Mining” (MAMI) project, which is funded by the Flemish Institute for the Promotion of Scientific and Technological Research in Industry.

The authors wish to thank Liesbeth De Voogdt and Dirk Van Steelant for their help with the annotator guidelines, music selection and practical organization.

### REFERENCES

- [1] Music Information Retrieval Evaluation eXchange (MIREX), <http://www.music-ir.org/mirexwiki>
- [2] International Conference on Music Information Retrieval (ISMIR), <http://www.ismir.net>
- [3] Goto, M. “Development of the RWC Music Database”, Proceedings of the 18<sup>th</sup> International Congress on Acoustics (ICA 2004), pp. I-553-556, April 2004
- [4] Leveau P., Daudet L. and Richard G., “Methodology and Tools for the evaluation of automatic onset detection algorithms in music”, Proceedings of the 5<sup>th</sup> International Symposium on Music Information Retrieval (ISMIR 2004), pp. 72-75, Barcelona, Spain, 2004
- [5] Gouyon F., Wack N. and Dixon S., “An open source tool for semi-automatic rhythmic annotation”, Proceedings of the 7<sup>th</sup> International Conference on Digital Audio Effects (DAFx 2004), Naples, Italy, 2004
- [6] Musical Audio Mining (MAMI), Ghent University, Belgium, <http://www.ipem.ugent.be/MAMI>
- [7] Lesaffre M., Leman M., De Baets B. and Martens J.-P., “Methodological considerations concerning manual annotation of musical audio in function of algorithm development”, Proceedings of the 5<sup>th</sup> International Conference on Music Information Retrieval (ISMIR 2004), pp. 64-71, Barcelona, Spain, 2004
- [8] Creative Commons, <http://creativecommons.org>

**Table 3.** Overview of the music we used for the annotation task. For the fragments marked with an asterisk, no reliable annotations could be obtained. More detailed information (including CD identification numbers and start and end times of the fragments) can be found on the MAMI project web site.

Title	Performer
Achterbank	De nieuwe Snaar
Afro-Left	Leftfield
Ahmad's Blues	Ahmad Jamal Trio
Angel	Massive Attack
Bard Dance	Enya
Billie Jean	Michael Jackson
Boom Boom	John Lee Hooker
Both Sides of the Story	Phil Collins
Business	Eminem
Busted *	Johnny Cash
Caroline Hard-Core Ecstasy	Frank Zappa, Captain Beefheart, George Duke, Napoleon Murphy Brock, Bruce Fowler,...
Chicken Walk	The Jon Spencer Blues Explosion
Cold Water	Tom Waits

Dans La Spirale	Starflam
Dejamer ser mujer	Axelle Red
Dromen zijn bedrog	Marco Borsato
Have a cigar	Pink Floyd
Haw	16 Horsepower
Highway to Hell	AC/DC
In Bloom	Nirvana
Join Hands	Laurent Garnier
Jumbo	Underworld
Land of...	St Germain
Leave Home	Chemical Brothers
Links 2,3,4	Rammstein
Looking Through the Eye of a Pig	Cypress Hill
Malegria	Manu Chao
Marilou Reggae	Serge Gainsbourg
Market Daze	Nitin Sawhney
Meisjes	Raymond van het Groenewoud
Molten Universe	Kuyss
My World	Metallica
Nooit met krijt	Kadril
Poofter's Froth Wyoming Plans Ahead	Frank Zappa, Captain Beefheart, George Duke, Napoleon Murphy Brock, Bruce Fowler,...
Prison Shoe Romp	16 Horsepower
Miserlou	Dick Dale & His Del-Tones
Queremos Paz	Gotan Project
Rocket 88	The Jimmy Cotton Blues Quartet
Say What You Say	Eminem
Sink To The Bottom	Fountains of Wayne
Sorte	Gal Costa, Gaetano Velosa
St. Anger	Metallica
Stinkfist	Tool
Sunday Bloody Sunday	U2
The Box *	Orbital
The Time Is Now	Moloko
The Watcher	Dr. Dre
Triptico	Gotan Project
Waterloo	ABBA
We Speak *	Booker Little, Eric Dolphy, Julian Priester, Max Roach, ...
When It Sings	Elvis Costello
Yellow	Coldplay