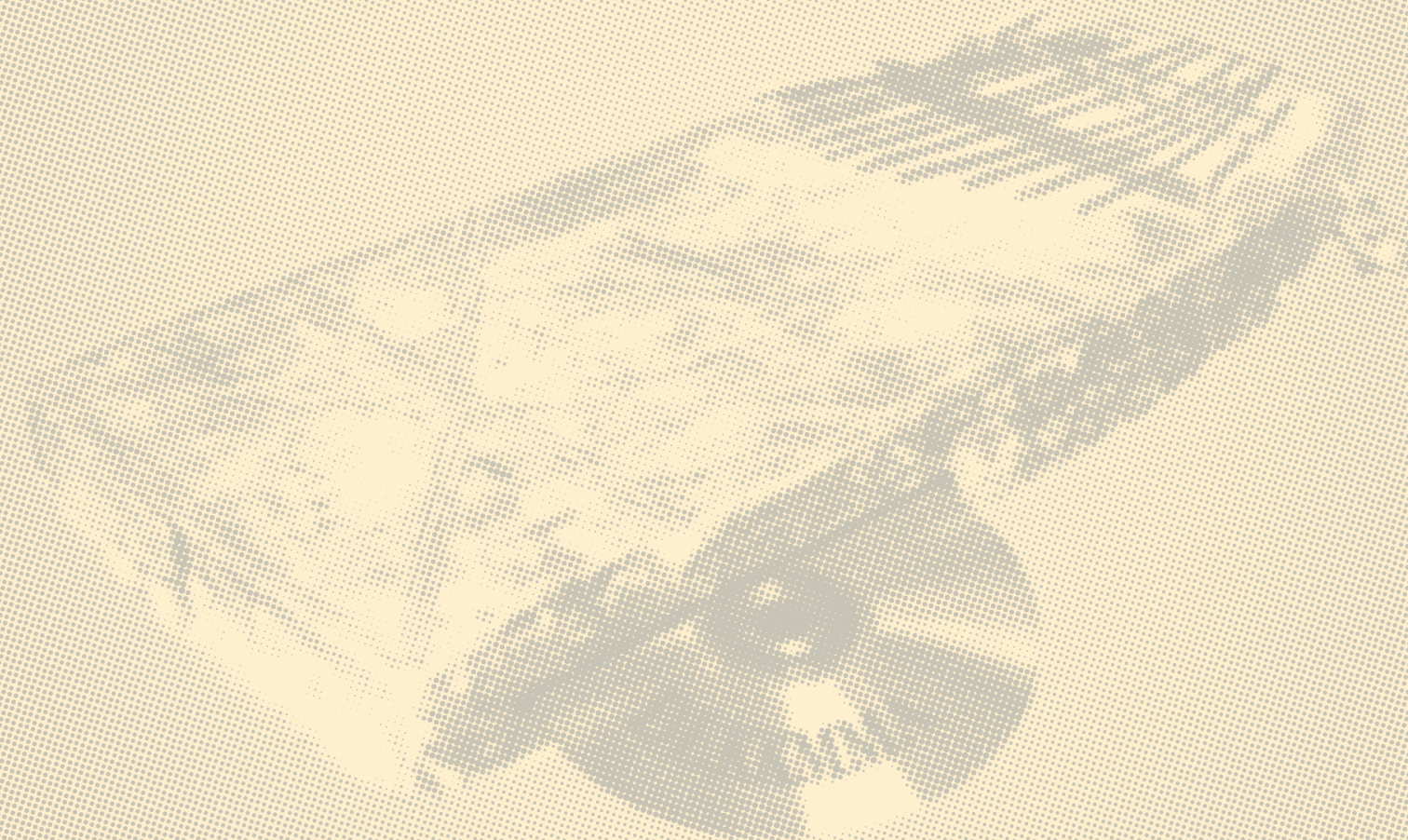


The Problems and Opportunities of Content-based Analysis and Description of Ethnic Music

Dirk Moelants, Olmo Cornelis & Marc Leman

Jos Gansemans

Rita De Caluwe, Guy De Tré, Tom Matthé & Axel Hallez



The Problems and Opportunities of Content-based Analysis and Description of Ethnic Music

● **Dirk Moelants, Olmo Cornelis & Marc Leman**
IPEM - Department of Musicology, University of Ghent, Belgium

Jos Gansemans

Department of Cultural Anthropology, Royal Museum of Central Africa, Belgium

Rita De Caluwe, Guy De Tré, Tom Matthé & Axel Hallez

TELIN, University of Ghent, Belgium

ABSTRACT

The Belgian Royal Museum for Central-Africa (RMCA) holds a large collection of ethnographic artifacts, including a sound archive with music recordings from the early 20th Century up to recently. The archive is one of the biggest and best-documented archives worldwide for the region of Central Africa. An on-going digitisation project is part of a strategy to conserve this archive and make it accessible to the public by (i) the digitisation of the data, and (ii) the application of music information retrieval techniques for the digitised data. While state-of-the-art research in music information retrieval aims to search and retrieve music on the basis of content description, most of the existing tools are designed for Western music collections, without any guarantee that these techniques can be applied to music from other cultures. African music, in particular, creates new challenges for content-based description and information retrieval. This paper describes some general problems regarding the content-based description of African and other non-Western music. It suggests an approach for describing pitch structures which will allow for the description of both Western music and non-Western music.

underlines the value and importance of preserving collections and making it easier to consult them by means of modern digital infrastructures and appropriate content management tools. Digital audio files can be

stored on a redundant storage system, which combines hard disks with a tape backup system for extra security. Nowadays, the technological infrastructure, based on digital broadband and mobile communication, makes it possible to access musical information quickly. Music can now be available from anywhere and at any time, by a finger click on the computer mouse, and there is no reason why this facility should be restricted to commercial music. The opportunities for storage, back-up, accessibility and documentation, make digital systems particularly well-suited to the preservation of cultural heritage. Unlike the playback of analogue audio, playbacks of digital audio do not harm the original carrier, and the number of copies cannot be exhausted. Digital meta-data, which typically consist of records with information about the origin and nature of the sound recording, and perhaps also the recording conditions, can be added to the musical audio files.

During the last decade, traditional meta-data descriptions which relate to the recording context have been complemented with another type of meta-description that is focused on the musical content. An objective content-description would typically focus on aspects such as timbre, pitch, melody, harmony, rhythm and tempo, while a subjective description would typically focus on factors related to movement (static, dynamic, slow, fast, etc.), emotional descriptors (gay, sad, etc.) or other semantic or corporeal descriptors (Lesaffre et al., 2004; Lesaffre, 2005; Leman, 2007).

So-called 'audio-mining' techniques aim to extract these content-based descriptions directly from the musical audio files. These techniques are based on low-level feature extraction and classification into higher-level descriptors that can be accessed by the human mind. For example, melody extraction from polyphonic audio is based on frequency analysis techniques that

work on small time frames (e.g. 40 ms). The task is to select the fundamental frequency component that is representative for the melody, and to discard the information from percussion and accompanying instruments. At a higher analysis level, the time frames will be concatenated to form pitch objects that can be represented as an electronic score (e.g. Paiva, 2006). The melody provides a level of representation which users can access. For example, they can sing a melody and the sung melody can then be compared with a database of melodies that have been extracted from the audio archive.

The audio-mining approach is a first step in what is generally called 'data-mining' or the automatic search for patterns in large volumes of data, using techniques of statistical correlation and categorisation. Indeed, users would typically tend to extract further information from such a database, such as information about similarities in a large collection of musical pieces. Similarities can then be represented graphically, on maps and visualisation schemas (van Gulik & Vignoli, 2005; Pampalk 2005). Music information retrieval thus aims to combine audio-mining techniques and data-mining techniques for the search and retrieval of music in a digital music library. This would allow new query techniques to be developed - like searching for similar pieces of music (including the possibility of 'query-by-example', where the user uploads his audio fragment) or the use of semantic descriptions (e.g. adjectives relating to emotions or gestures) that are automatically connected to the low-level features of the musical content (Leman et al., 2005). However, although the techniques for content-based description and retrieval look promising, it remains very difficult to get technological success rates that would allow practical, real-world applications. There is a huge semantic gap between music as digital encoded audio, and music as

Introduction

During the last decade, digitisation projects for cultural heritage have received increasing financial support from both public sources and private institutions. This

something meaningful. The step from encoded audio to meaning involves many aspects of human perception, understanding and a thorough knowledge of the social context in which the music information retrieval activities are taking place.

At this moment, the audio-mining techniques that have been developed are designed principally to extract information from Western music. Hence, the concepts that underlie audio-mining techniques are often based on Western music theory. Examples are the use of the chromatic (equidistant 12-tone) scale or the assumption of a regular division of the measure in ternary or binary units. There seems to be a general lack of knowledge about cultural heritage from beyond the Western world, and music that has a fundamentally different structure is not usually considered at all.

The construction of a database with 'traditional' meta-data for non-Western, and especially ethnic, music requires an approach that is entirely different from that used for Western popular and classical music. In this paper we describe some of the problems encountered during the process of digitising the music archives of the Belgian Royal Museum of Central-Africa (RMCA), and we propose an approach that would allow more suitable descriptions for music with structural and sociological characteristics that are very different from the Western standard. First, we will give a short description of the collections of the RMCA and the digitisation process. Then we will introduce some of the problems - and opportunities - we have encountered in dealing with

databases of non - Western music, and in the last section we will present a system that analyses and compares the pitch content of musical pieces as an illustration of how content-based descriptions of music can deal with specific aspects of non-Western music.

Digitisation of the audio collection of the Belgian Royal Museum of Central-Africa (RMCA)

With its 50,000 sound recordings (with a total of 3,000 hours of music), dating from the early 20th century up to the present day, the music archive of the RMCA (located in Tervuren near Brussels) is one of the biggest in the world for the region of Central Africa. The music archive is part of a larger collection of artifacts from Central Africa that includes musical instruments, masks, tools, animals, plants and many other objects. Conservation and access to this collection is of particular importance given the current political instability, and the general destruction of cultural heritage that is taking place in this region (Cornelis et al., 2005). To conserve this important cultural heritage, and make it accessible to the public, the Belgian government provided a grant for a project called DEKKMMA. The project aimed to (i) digitise the different types of sound recordings (wax cylinders, sonofil, vinyl recordings and magnetic tapes) each with their own specific problems, (ii) construct a database suitable for describing different types of music (Matthé et al., 2005), (iii) explore tools that will allow further analysis

and classification using audio-mining and data-mining techniques, (iv) develop content pertaining to the description of musical culture and musical instruments from different geographic regions, (v) process audio, photographic and video material from the archives and from the museum collections as background information for the audio files. (The results of this project can be accessed on the website <http://music.africamuseum.be/>. The reader will find metadata as well as sound excerpts and accompanying documentation such as descriptions of musical instruments.)

The project included the construction of the website, which allows different user groups to search and retrieve data related to the music archive. Three user groups were identified. The first and largest group are people who are just interested in African music, but do not have much knowledge of it. These users typically want to retrieve music using a rather vague and general labelling, such as 'drumming', 'trance music' or 'some song from Rwanda'. The second group consists of users from Central Africa. They often have a good knowledge of certain repertoires and functions of the music, and therefore they tend to ask very specific questions - such as for music played by a specific performer, music from one particular village, lyrics, genres, instruments - and they may well use local terminology. Finally, the third group of users consists of researchers who use the database for further study. This group would typically tend to ask questions related to the geographical spread of certain types of instrument, or the relative importance of certain rhythmic or pitch musical structures in different regions. In short, music information retrieval has to take into account user groups with a whole range of different interests. Research in music information retrieval aims to develop new tools and find proper ways of dealing with the interests of these different user groups. In practice, this often requires a thorough empirical analysis of the needs, as well as the subject backgrounds, of the users (Lesaffre et al., in press).

To overcome the lack of knowledge of the amateur user, extra search strategies have been implemented, such as searching by clicking on a map of Africa, or listing musical instruments in a tree according to Sachs' and von Hornbostel's musical instrument classification. For example, the tree structure allows the user to search for wind instruments, then for flutes, then for straight and transverse flutes and finally a list of vernacular instrument names will be given, including the names of

the countries where a particular instrument has been found. Interesting applications to be developed for the first group of users are query-by-example, or search by effective parameters. Query-by-example would allow users to provide an audio example and formulate a request to find music with a similar rhythm. Query by effective parameters would allow users to specify music using descriptors that pertain to particular emotions, musical effects, moods, or movement characteristics. Users can also add valuable information to the database. For this reason, a forum will be set up where users can discuss topics, improve the descriptions of content and, hopefully, fill some gaps. The forum will allow users to come into contact with researchers working on developing the database, and that may produce ideas for the invention of new search tools and possibly allow the database to be extended to include copies of recordings kept at other institutions or by private individuals. In DEKKMMA, some of these facilities have been implemented, while other parts are still at the development stage.

Integrating non-Western music into databases: problems and opportunities

The major task of the DEKKMMA project is to transform the (analogue) music archive into a digital music library, using modern tools of music information retrieval. However, music information retrieval research usually takes Western music and its musical characteristics and semantic descriptions as a standard, and develops tools following a series of assumptions based on Western cultural concepts. These apply to structural aspects (e.g. tonal key, assumption of octave equivalence, instrumentation), social organisation of the music (e.g. composers, performers, audience) and technical aspects (e.g. record company, release date). There is no guarantee that these concepts can be readily applied to non-Western music. Indeed, the production and appreciation of music in oral cultures may be completely different from the way Westerners see music, with traditions of learning by listening and practicing, passing skills from father to son or from master to pupil. Trying to incorporate descriptions of non-Western music in databases that are structured to cope with the demands of classifying Western music, often causes problems. These problems can be found both in descriptive meta-data and in descriptions of musical content. Imposing



Figure 1
Logo of the DEKKMMA-project, symbolising the digitisation of the African music archive

Western concepts on to non-Western music can lead to incorrect information going into the databases, or important information being excluded through the lack of suitable database fields.

Music Information Retrieval (MIR) applications that focus on popular and classical Western music will make the search and retrieval of such music easier, and therefore more people will have access to it. In contrast, music that occupies a more marginal position risks being excluded by this technology, and will thus become even less accessible. In this way, the combination of digitisation and commercial large-scale distribution tends to push ‘vulnerable’ music even further into oblivion. Music information retrieval research should therefore take into account an ethical code that aims to develop tools for all types of music, not just for Western music. This is a huge challenge for music research and it will require a change of approach in most centres that currently deal with music information retrieval. To bring people into contact with music they would normally never have heard of requires a reconsideration of the concepts that underlie musical practices in non-Western cultures.

Simply integrating more ethnic music, and non-Western music in general, into the existing databases and indexes will solve the above-mentioned problem entirely. Indeed, as already mentioned, musical structures, as well as the relative importance of structural elements within the musical experience, can be fundamentally different in different cultures. A straightforward example is the Western focus on pitch and fixed tuning, whereas in African music fixed tuning does not exist. Instead, a large number of different pitch scales can be observed, and often relative pitch (higher-lower) is more important than absolute pitch. A proposal for a method that deals with the description of pitch will be outlined in the next section.

Another difficulty with integrating ethnic music into digital music libraries concerns the organisation of the meta-data. In describing field recordings, it often happens that some information that is ‘compulsory’ in the description of Western music is lacking, while other information that seems irrelevant in the description of Western music turns out to be very important. For example, names of composers are usually not known. Performers could be named but music is often seen as performed by ‘the community’ and therefore, the names of the participants are not considered to be very important. On the other hand, the location and date of the recording are important because location can be a

crucial search field for retrieving music from oral cultures. Since names of performers and composers are usually not known, the music is primarily identified with the country, region, ethnicity or town where it is produced. In many existing databases, location is not even an existing meta-data field, due to its low relevance in Western music.

A further problem in the meta-data descriptions of field recordings is related to the lack of standardisation. This is due to the fact that these meta-data descriptions often have a historical origin, and have been collected by many different field researchers, often amateurs, who used many different recording techniques. As a result, not all recordings are equally well-documented. For some interesting old field recordings even the most basic information is lacking, and one needs extensive historical research in order to have even a rough idea of the time and place where the music might have been recorded. But even recordings made by professional ethno-musicologists are sometimes not completely documented. In some cases, the documentation has been partially lost, or the connection between the recordings and the documentation is no longer clear. Given the fact that knowledge about traditional music within oral cultures is vanishing under pressure of urbanisation and Westernisation, the correct identification of the music and its meta-data descriptions, as well as the definition of its authenticity (in the digital context) becomes increasingly important.

Finally, there is a problem of terminology. There can be different local names for the same concept, and different researchers can use different terms for them. At this moment, the American Folklore Society and the American Folklife Center at the Library of Congress are constructing an ‘Ethnographic Thesaurus’, a comprehensive, controlled list of subject terms to be used in describing ethnographic and ethnological research collections (cf. <http://www.etproject.org>). But even a standardised list cannot solve all the problems. Consider the example of the ‘thumb piano’ (lamellophone). This instrument type has very diverse names (see Table 1). A user looking for one of these names should also be directed to pieces in which one of the other terms is used. This requires an elaborate thesaurus and a specific approach in the construction of the database (Matthé *et al.*, 2006). To make it even more complicated, one name does not necessarily point to a specific sub-type: size, material, number of pitches and

Table 1
Variable denominations for the lamellophone.

Tshisaasj	Ngombi	Kadimba	Agidogo
Tshisaji	Kombi	Kalimbe	Alogu
Tshisanji	Kembe	Malimba	Ashwa
Chisanzi	Ekembe	Marimba	Dongo
Sanzi	Ikembe	Irimba	Isen
Sanza	Likembe	Ilimba	Kankobele
Sansa	Dikembe	Limba	Mambalat
Esanzo	Kidebe		Prempensua
Issanji	Ibekee		Ubo
Kisanzi	Gbelee		...
Kassandji			
Kassayo			

tuning can vary widely. Therefore it is desirable that the user should be able to refine his search by looking for more specific instrument characteristics, or for instruments with similar tuning.

Audio-mining and the pitch structures in African music

Research on content-based music information retrieval aims i) to define the search and retrieval of music in terms of musical content descriptors and ii) to develop automated content-description and retrieval methods. Rather than having to specify the name of the composer or the title of the song, the content-based approach would allow one to specify musical content using descriptors related to its nature such as ‘happy’, ‘sad’, ‘dynamic’, ‘harmonious’, or using corporeal descriptors which define particular movements that are captured by sensors (e.g. indicating tempo or expression), using graphical navigation in databases, and using search and retrieval by providing audio examples (for relevant publications, see www.ismir.net). As content-description is often based on subjective descriptions, knowledge of the user’s background is an important issue (Lesaffre et al., in press).

In this section, we focus on one single structural characteristic of musical content, namely pitch. Although pitch is closely related to the physical characteristics of music, and less to subjective factors that would depend on education, gender, and familiarity, there are many problems with this as a content descriptor. The major problem is that researchers tend to develop pitch extraction algorithms that are based on Western concepts and assumptions from music theory which cannot easily be applied to non-Western music. A

straightforward example of such an assumption is the concept of the octave-reduced pitch representation, with a categorisation based on the chromatic 12-tone scale.

Within the collection of Central African music of the RMCA, there are a wide variety of tunings and scales. Often, a melody is closely connected to the tones of the Bantu tone languages. This is the case in vocal music where the melody-line has to follow the speech-tones, and in instrumental genres that are based on verbal elements. A similar phenomenon is seen in the music of other tonal languages, such as Chinese. In Bantu languages there is a continuum between speech and music, which makes it possible to transfer messages from one village to another by using drum signals of different pitch. The use of speech tones that do not have a precisely determined pitch makes it more important to distinguish between high and low pitches rather than having specific harmonic relationships between pitches. Typically, in this region, people prefer instrumental sounds with a very broad, percussive spectrum, which in turn reduces the importance of ‘correct’ pitch intervals. However, other instruments with fixed tuning, like flutes, zither, (wooden) trumpets or thumb pianos, allow the study of possible fixed pitch relations and pitch scales.

To include information about pitch scales and pitch tunings in the DEKKMMA database, a pitch description approach was used that represents pitch scales without reference to Western notes or scales. In avoiding a priori pitch categories, this approach is based on a continuous representation of pitch. Thus, rather than a discrete pitch representation, we propose a continuous pitch representation and continuously-based associated retrieval mechanisms.

In the first stage, the music is analysed by a melody extractor (De Mulder et al., 2004). This melody extractor

Figure 2

Pitch analysis of Ingendo y'inka. The x-axis shows the pitch in cents (0 = a 55 Hz), the y-axis shows the number of occurrences for every 1 cent interval, smoothed by taking the average of 7 bins around the middle. The vertical lines show the position of the a's, dividing the whole into three octaves.

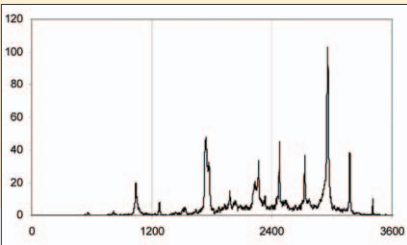


Figure 3

Octave-reduced pitch analysis of Ingendo y'inka. The x-axis shows the pitch with 0 = a, the y-axis shows the number of occurrences for every 1 cent interval, smoothed by taking the average of 7 bins around the middle. The vertical lines show the position of the a's, dividing the whole into three octaves.

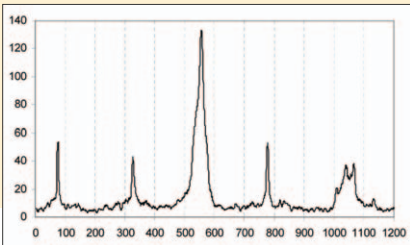


Table 2

Analysis of the most prominent peaks in Ingendo y'inka: the left column shows the pitch in cents above 55 Hz, the middle column shows the distance in cents between successive pitches and the right column shows the size of the octave relationships.

Pitch	interval	octave
1042	236	1225
1278	264	1199
1542	202	1185
1744	237	1214
1981	286	1198
2267	210	1140
2477	250	
2727	231	
2958	221	
3179	228	
3407		

was originally designed for the transcription of vocal queries. It was optimised for monophonic music and the normal voice range. Yet, testing the model on different types of music reveals that it can provide a straightforward image of the pitch distribution, even in music with a more complex texture. For complex polyphonic music with a dense texture, like Western symphonic music, the system is less successful, but other pitch detection systems could be used, following the same methodology. The melody extractor currently used gives a frequency for every time frame of 10 ms. In order to give an image of the scale, we transform these values to cent values (taking the low A, 55Hz as 0 cents). The cent scale divides every half tone into 100 subdivisions, which allows a very precise representation of the actual pitch. An additional advantage of the cent-scale is that the distances are the same in every octave and within every half-tone (this is not the case in a Hertz frequency representation, which has a somewhat logarithmic relationship to pitch). The representation of the pitch content of a piece is then given by plotting the number of occurrences of each cent value between 1 and 6000 cents. In this paper the method will be illustrated first by the analysis of one particular song, then seven other examples, some from the same region and some from completely different musical cultures, will be used to show how this methodology can be applied to a broad spectrum of types of music.

The sample song is called *Ingendo y'inka*. It is performed by a man singing with *ikembe* (thumb piano) accompaniment. The piece was recorded on 14 January 1973 by Jos Gansemans in the village Karengera, Cyangugu province, Rwanda. The singing is mainly in

parlando style, following the pitches of the instrument. The text describes the elegance of the local cows, because these animals are very importance in local society and are a sign of wealth. Extraction of the pitches in this song reveals that 11 notes occur, spread over several octaves (see figure 2, table 2). The *tessitura* used is quite low; within the octaves A to a' (110-440 Hz). Certain tones occur in the different octaves. A reduction to one octave of 1200 cents gives a better picture of the scale used, but at the same time it illustrates the danger of this octave reduction. Figure 3 shows the octave-reduced representation. The subdivisions on the X-axis (the discontinuous vertical lines) represent Western equal-tempered tuning: Every half tone has a distance of 100 cents, a whole tone measures 200 cents, a small third 300, a fifth 700, the octave 1200. These markers give a clear picture of the pitches of the song, and in this case put forward only 5 different notes that are frequently used. Looking at this pentatonic scale, 3 intervals around 226 cents and two around 266 cents (table 2) can be found.

The African tonality in this case, as measured, can be described as more or less equidistant pentatonic, although the small differences in interval size might be characteristic for this scale. The peak between 1000 and 1100 cents (g and g#) is much broader, and we can, in fact, see three small peaks in it (Figure 3). Although in most music the octaves and fifths relations are usually quite precise, due to the strong sense of consonance, we see that is not the case in this example. The highest octave is much (60 cents) too small, while the lowest is 25 cents too wide. These deviations are not uncommon in African music and they create a sought-after tension

Figure 4 : 73.9.3-6, flute

Octave-reduced pitch analysis of an umwirongi(flute) piece from Cyangugu province, Rwanda.

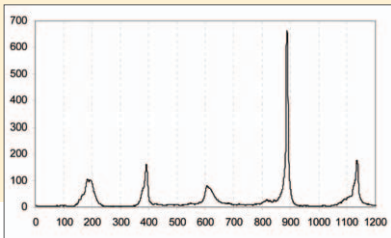


Figure 5 : 73.9.15-10, flute

Octave-reduced pitch analysis of an umwirongi(flute) piece from Ruhengeri province, Rwanda.

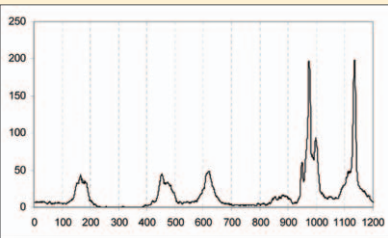
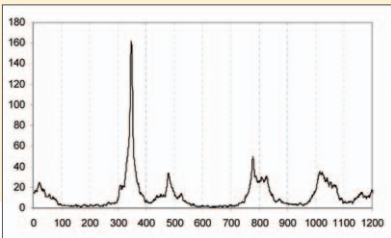


Figure 6 : 54.1.13-1, musical bow and singing

Octave-reduced pitch analysis of a Twa song with umuduri (musical bow) accompaniment from Kigali province, Rwanda.



instead of smooth sounding exact octaves. This shows that the octave-reduction has both an advantage and a disadvantage. It makes it much easier to illustrate the scale and to compare it with other examples, but it removes the subtle differences between octaves that are important in some styles of music.

For six other pieces, an octave-reduced graph is created in order to show the use of this method in characterising and comparing different tuning systems and scales. Thus every peak indicates the number of annotations made for every tone. Three examples (Figures 4-6) are taken from the collection of Rwandan music in the RMCA. The first two were also recorded in 1973 by Jos Gansemans, and they are both played on the *umwirongi* flute. The first (Figure 4) was recorded in the village Cyimbogo in the same Cyangugu province as *Ingendo y'inka*, the second (Figure 5) in the village Ndusu in Ruhengeri province. The third example (Figure 6) represents the pitch content of a song accompanied by the musical bow (*umuduri*) from the Twa people in the village of Nyanza in Kigali province; it was recorded in 1954 by the missionary Scohy-Stroobants.

All the examples point to pentatonic tunings. However, there is no standard. Figure 4 is very similar to figure 3, using a more or less equidistant pentatonic scale. If the pitch of figure 4 were lowered by 330 cents, the main peaks would fall together, and the two would be almost identical. Although the second piece was played on the same type of instrument, the tuning is clearly different, tending more towards irregular anhemitonic pentatonic scales. Another difference is that instead of having one main peak, this example has two large peaks of almost equal magnitude. The third example stands somewhere in between, and could be related to the others, even though the origin, instrumentation and recording date are very different.

Three additional examples illustrate how different musical styles and scales are represented. Figure 7 shows the pitch distribution of Mozart's piano sonata in G Major (KV 283). Figure 8 shows a piece of Korean classical orchestral court music (*Pohoja* performed by the Seoul National Orchestra of Classical Music), and Figure 9 shows a piece of Persian santur (zither) music (*Avaz* from *Bayate Esfahan* played by Mohamad Heydari).

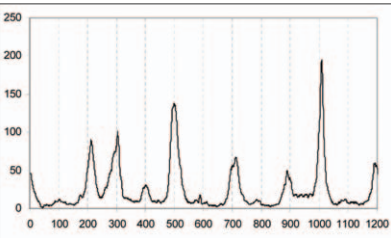


Figure 7 : Piano Sonata Mozart

Octave-reduced pitch analysis of W.A. Mozart's piano sonata in G-major (KV 283)

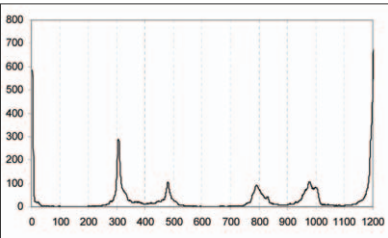


Figure 8 : Korea - Aag

Octave-reduced pitch analysis of "Pohoja", banquet music for the Korean court, performed by the "Seoul National Orchestra of Classical Music".

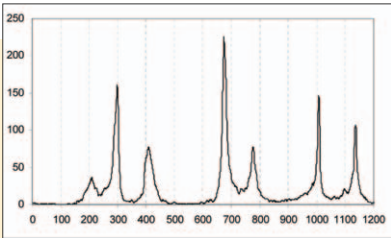


Figure 9 : Iran - Santur

Octave-reduced pitch analysis of *Avaz* from *Bayate Esfahan* performed by Mohamad Heydari on the Persian ceter santur.


These examples illustrate the difference between distinct Western and non-Western tone systems very well. The Western system uses a chromatic/diatonic scale in which interval sizes are multipliers of 100 cents, and there is a strict adherence to a 1200 cents octave. The Mozart piece clearly shows its G-major tonality and 100/200 cents based structure (in all examples 0 cents is A, 100 cents is A#, 200 cents is B”) so peaks appear at 1000/300/500 (representing notes G/C/D, which are the tonal grades I-IV-V in G-Major), and the peak at 200 cents (note B, III grade) reveals the major tonality.

In contrast, the Korean and the Persian pieces rely on well-defined tuning systems, but the graphs show characteristics that would not be represented idiomatically using Western tonal notation. The pitches used in the Korean court orchestra are closely related to the Western tuning system, but the music uses a pentatonic scale, just as in the African pieces described above, and the pitches do not correspond exactly to the divisions of the Western scale. The Persian example is clearly heptatonic, like most Western scales, but the distances between the pitches do not adhere to the standard patterns as the Persian modes use both ‘normal’ semi-tones and ‘enlarged’ intervals, like 3/4-tones. Comparisons of pitch scales can be based on continuous pitch representations using cross-correlation techniques. Although this is computationally intensive, it offers an appropriate method that avoids the danger of imposing Western pitch categories on to non-Western music. Further study may reveal that the cent-scale can be represented at a lower sampling rate to reduce the computational cost in search and retrieval. However, the concept of a continuous representation over different octaves would remain the core feature of this pitch representation schema.

Conclusion

Constructing applications for content-based descriptions of music that can deal with all the world’s musical traditions is difficult, but this seems very necessary to protect the world’s cultural heritage. It is important to bring together knowledge about the music from different cultures, and to make it accessible to a broad audience. Dealing with ethnic music reveals interpretational

problems related to musical practice, the semantic description of musical features, as well as the automated extracted musical content parameters. The fundamentally different use of pitch in African music compared to Western music illustrates the difficulty of applying existing Western discreet pitch categories to non-Western music. Other examples pertain to aspects of rhythm, timbre and articulation. In addition, one should be careful about adopting extra-musical descriptive characteristics, which typically relate to musical practice in its social/cultural context. The use of our Western semantic and perceptual framework is often inappropriate for accurate digitalisation and the subsequent development of a digital library for cultural heritage.

The DEKKMMA-project is an example of a digitalisation project that aims to develop methods for content-based description for all types of music. This was illustrated by the method for representing pitch scales, using a continuous representational schema that can be used in search and retrieval applications. 

REFERENCES

- Cornelis, O., De Caluwe, R., De Tré, G., Hallez, A., Leman, M., Matthé, T., Moelants, D. & Gansemans, J. 2005. Digitisation of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa (Belgium). pp. 35-43 in *IASA journal*, 26.
- De Mulder, T., Martens, J.P., Lesaffre, M., Leman, M., De Baets, B. & De Meyer, H. 2004. Recent improvements of an auditory model based front-end for the transcription of vocal queries. pp. 257-260 in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. (Montreal)
- Leman, M. 2007. *Embodied Music Cognition*. (Cambridge: MIT Press)
- Leman, M., Vermeulen, V., De Voogdt, L., Moelants, D. & Lesaffre, M. 2005. Prediction of Musical Affect Using a Combination of Acoustic Structural Cues. pp. 39-68 in *Journal of New Music Research* 34(1)
- Lesaffre, M. 2005. *Music Information Retrieval: Conceptual Framework, Annotation and User Behavior*. (Unpublished Ph D. thesis: Ghent University)
- Lesaffre, M. Leman, M. De Baets, B. Martens, J.-P. 2004. Methodological considerations concerning manual annotation of musical audio in function of algorithm development. pp. 64-71 in *Proceedings of the fifth International Conference on Music Information Retrieval*. (Barcelona)Lesaffre, M., De
- Voogdt, L., & Leman, M. (2007) “How potential users of music search and retrieval systems describe the semantic quality of music” (in press)
- Matthé, T., De Tré, G., Hallez, A., De Caluwe, R., Leman, M., Cornelis, O., Moelants, D. &Gansemans, J. 2005. A Framework for Flexible Querying and Mining of Musical Audio Archives. pp.1041-1045 in *Proceedings of the 16th International Conference on Database and Expert Systems Applications*. (Copenhagen)
- Matthé, T., De Caluwe, R., De Tré, G., Hallez, A. Verstraete, J., Leman, M., Cornelis, O. Moelants, D. & Gansemans, J. 2006. Similarity Between Multi-valued Thesaurus Attributes: Theory and Application in Multimedia Systems. pp. 331-342 *Lecture Notes in Computer Science 4027*.
- Paiva, R. P. 2006. *Melody Detection in Polyphonic Audio*. (Unpublished Ph.D thesis: University of Coimbra)
- Pampalk, E. 2004. A Matlab Toolbox to compute Music Similarity from Audio, pp. 254-257 in *Proceedings of the Fifth International Conference on Music Information Retrieval*. (Barcelona)
- Van Gulik, R., & Vignoli, F. 2005. Visual playlist generation on the artist map. pp.520-523 in *Proceedings of the Sixth International Conference on Music Information Retrieval*. (London)