Applications of duplicate detection in music archives: from metadata comparison to storage optimisation

The case of the Belgian Royal Museum for Central Africa

Joren Six, Federica Bressan, and Marc Leman

IPEM, Ghent University,
Miriam Makebaplein 1, Belgium
{joren.six,federica.bressan,marc.leman}@ugent.be

Abstract. This work focuses on applications of duplicate detection for managing digital music archives. It aims to make this mature music information retrieval (MIR) technology better known to archivists and provide clear suggestions on how this technology can be used in practice. More specifically applications are discussed to complement meta-data, to link or merge digital music archives, to improve listening experiences and to re-use segmentation data. To illustrate the effectiveness of the technology a case study is explored. The case study identifies duplicates in the archive of the Royal Museum for Central Africa, which mainly contains field recordings of Central Africa. Duplicate detection is done with an existing Open Source acoustic fingerprinter system. In the set, 2.5% of the recordings are duplicates. It is found that meta-data differs dramatically between original and duplicate showing that merging meta-data could improve the quality of descriptions. The case study also shows that duplicates can be identified even if recording speed is not the same for original and duplicate.

Keywords: MIR applications, documentation, collaboration, digital music archives

1 Introduction

Music Information Retrieval (MIR) technologies have a lot of untapped potential in the management of digital music archives. There seems to be several reasons for this. One is that MIR technologies are simply not well known to archivists. Another reason is that it is often unclear how MIR technology can be applied in a digital music archive setting. A third reason is that considerable effort is often needed to transform a potentially promising MIR research prototype into a working solution for archivists as end-users.

In this article we focus on duplicate detection. It is an MIR technology that has matured over the last two decades for which there is usable software available. The aim of the is article is to describe several applications for duplicate detection

and to encourage the communication about them to the archival community. Some of these applications might not be immediately obvious since duplicate detection is used indirectly to complement meta-data, link or merge archives, improve listening experiences and it has opportunities for segmentation. These applications are grounded in experience with working on the archive of the Royal Museum for Central Africa, a digitised audio archive of which the majority of tracks are field recordings from Central Africa.

2   Duplicate detection

The problem of duplicate detection is defined as follows:

> *How to design a system that is able to compare every audio fragment in a set with all other audio in the set to determine if the fragment is either unique or appears multiple times in the complete set. The comparison should be robust against various artefacts.*

The artefacts in the definition above include noise of various sources. This includes imperfections introduced during the analog-to-digital (A/D) conversion. Artefacts resulting from mechanical defects, such as clicks from gramophone discs or magnetic tape hum. Detecting duplicates should be possible when changes in volume, compression or dynamics are introduced as well.

There is a distinction to be made between *exact*, *near* and *far* duplicates [1]. Exact duplicates contain the exact same information, near duplicates are two tracks with minor differences e.g. a lossless and lossy version of the same audio. Far duplicates are less straightforward. A far duplicate can be an edit where parts are added to the audio – e.g. a radio versus an album edit with a solo added. Live versions or covers of the same song can also be regarded as a far duplicate. A song that samples an original could again be a far duplicate. In this work we focus on duplicates which contain the *same recorded material* from the original. This includes samples and edits but excludes live versions and covers.

The need for duplicate detection is there since, over time, it is almost inevitable that duplicates of the same recording end up in a digitised archive. For example, an original field recording is published on an LP, and both the LP as the original version get digitised and stored in the same lot. It is also not uncommon that an archive contains multiple copies of the same recording because the same live event was captured from two different angles (normally on the side of the parterre and from the orchestra pit), or because before the advent of digital technology, copies of degrading tapes were already being made on other tapes. Last but not least, the chance of duplicates grows exponentially when different archives or audio collections get connected or virtually merged, which is a desirable operation and one of the advantages introduced by the digital technology (see 2).

From a technical standpoint and using the terminology from [2] a duplicate detector needs to have the following requirements:

- It needs to be capable to mark duplicates without generating false positives or missing true positives. In other words **precision and recall** need to be acceptable.
- It should be capable to operate on large archives. It should be **efficient**. Efficient here means quick when resolving a query and efficient on storage and memory use when building an index.
- Duplicates should be marked as such even if there is noise or the speed is not kept constant. It should be **robust** against various modifications.
- Lookup for short audio fragments should be possible, the algorithm should be **granular**. A resolution of 20 seconds or less is beneficial.

Once such system is available, several applications are possible (in [1] many of these applications are described as well but, notably, the application of re-use of segmentation boundaries is missing).

Duplicate detection for complementing meta-data . Being aware of duplicates is useful to **check or complement meta-data**. If an item has richer meta-data than a duplicate, the meta-data of the duplicate can be integrated. With a duplicate detection technology conflicting meta-data between an original and a duplicate can be resolved or at least flagged. The problem of conflicting meta-data is especially prevalent in archives with ethnic music where often there are many different spellings of names, places and titles. Naming instruments systematically can also be very challenging.

Duplicate detection to improve the listening experience . When multiple recordings in sequence are marked as exact duplicates, meaning they contain the exact same digital information, this **indicates inefficient storage use**. If they do not contain exactly the same information it is possible that either the same analog carrier was accidentally digitised twice or there are effectively two analogue copies with the same content. To **improve the listening experience** the most qualitative digitised version can be returned if requested, or alternatively to assist philological research all the different versions (variants, witnesses of the archetype) can be returned.

Duplicate detection for segmentation . It potentially solves **segmentation** issues. When an LP is digitised as one long recording and the same material has already been segmented in an other digitisation effort, the segmentation boundaries can be reused. Also duplicate detection allows to identify when different segmentation boundaries are used. Perhaps an item was not segmented in one digitisation effort while a partial duplicate is split and has an extra meta-data item – e.g. an extra title. Duplicated detection allows re-use of segmentation boundaries or, at the bare minimum, indicate segmentation discrepancies.

Duplicate detection for merging archives . Technology makes it possible to **merge or link digital archives** from different sources – e.g. the creation of a single point of access to documentation from different institutions concerning a special subject; the implementation of the "virtual re-unification" of collections

and holdings from a single original location or creator now widely scattered [3, p.11]. More and more digital music archives 'islands' are bridged by efforts such as Europeana Sounds. Europeana Sounds is a European effort to standardise meta-data and link digital music archives. The EuropeanaConnect/DISMARC Audio Aggregation Platform provides this link and could definitely benefit from duplicate detection technology and provide a view on unique material.

If duplicates are found in one of these merged archives, all previous duplicate detection applications come into play as well. How similar is the meta-data between original and duplicate? How large is the difference in audio quality? Are both original and duplicate segmented similarly or is there a discrepancy?

2.1 Robustness to speed change

Duplicate detection robust to speed changes has an important added value. When playback (or recording) speed changes from analogue carriers, both tempo and pitch change accordingly. Most people are familiar with the effect of playing a 33 rpm LP at 45 rpm. But the problem with historic archives and analogue carriers is more subtle: the speed at which the tape gets digitised might not match the original recording speed, impacting the resulting pitch. Often it is impossible to predict with reasonable precision when the recording device was defective, inadequately operated, or when the portable recorder was slowly running out of battery.

So not only it is nearly impossible to make a good estimation of the original non-standard recording speed, but it might not be a constant speed at all, it could actually fluctuate 'around' a standard speed. This is also a problem with wax cylinders, where there are numerous speed indications but they are not systematically used – if indications are present at all. In the impossibility to solve this problem with exact precision, a viable approach, balancing out technical needs and philological requirements, is normally to transfer the audio information at standard speed with state-of-the-art perfectly calibrated machinery. The precision of the A/D transfer system in a way compensates for the uncertainty of the source materials. We still obtain potentially sped-up or slowed-down versions of the recording, but when the original context in which the recording was produced can be reconstructed, it is possible to add and subtract quantities from the digitised version because that is exactly known (and its parameters ought to be documented in the preservation meta-data).If the playback speed during transfer is tampered, adapted, guessed, anything that results in a non-standard behaviour in the attempt of matching the original recording speed, will do nothing but add uncertainty to uncertainty, imprecision to imprecision.

An additional reason to digitise historical audio recordings at standard speed and with state-of-the-art perfectly calibrated machinery, is that by doing so, the archive master [4] will preserve the information on the fluctuations of the original. If we are to "save history, not rewrite it" [5], then our desire to "improve" the quality of the recording during the process of A/D conversion should be held back. Noises and imperfections present in the source carrier bear witness to its history of transmission, and as such constitute part of the historical document.

Removing or altering any of these elements violates basic philological principles [6] that should be assumed in any act of digitisation which has the ambition to be culturally significant. The output of a process where sources have been altered (with good or bad intention, consciously or unconsciously, intentionally or unintentionally, or without documenting the interventions) is a *corpus* that is not authentic, unreliable and for all intents and purposes useless for scientific studies. Therefore, in the light of what has been said so far, the problem of speed fluctuation is structural and endemic in historical analogue sound archives, and cannot be easily dismissed. Hence the crucial importance of algorithms that treat this type of material to consider this problem and operate accordingly.

3 Acoustic Fingerprinting

Some possible applications of duplicate detection have been presented in the previous section, now we see how they can be put into practice. It is clear that naively comparing every audio fragment – e.g. every five seconds – with all other audio in an archive quickly becomes impractical, especially for medium-to-large size archives. Adding robustness to speed changes to this naive approach makes it downright impossible. An efficient alternative is needed and this is where *acoustic fingerprinting techniques* comes into play, a well researched MIR topic.

The aim of acoustic fingerprinting is to generate a small representation of an audio signal that can be used to reliably identify identical, or recognise similar, audio signals in a large set of reference audio. One of the main challenges is to design a system so that the reference database can grow to contain millions of entries. Over the years several efficient acoustic fingerprinting methods have been introduced [7,8,9,1]. These methods perform well, even with degraded audio quality and with industrial sized reference databases. However, these systems are not designed to handle duplicate detection when speed is changed between the original and duplicate. For this end, fingerprinting system robust against speed changes are desired.

Some fingerprinting systems have been developed that take pitch-shifts into account [10,11,12] without allowing time-scale modification. Others are designed to handle both pitch and time-scale modification [13,14]. The system by [13] employs an image processing algorithm on an auditory image to counter time-scale modification and pitch-shifts. Unfortunately, the system is computationally expensive, it iterates the whole database to find a match. The system by [14] allows extreme pitch-shifting and time-stretching, but has the same problem.

The ideas behind both [15,16] allow efficient duplicate detection robust to speed changes. The systems are built mainly with recognition of original tracks in DJ-sets in mind. Tracks used in DJ-sets are manipulated in various ways and often speed is changed as well. The problem translates almost directly to duplicate detection for archives. The respective research articles show that these systems are efficient and able to recognise audio with a ±30% speed change.

Only [15] seems directly applicable in practice since it is the only system for which there is runnable software and documentation available. It can be down-

loaded from `http://panako.be` and has been tested with datasets containing tens of thousands of tracks on a single computer. The output is data about duplicates: which items are present more than once, together with time offsets.

The idea behind Panako is relatively simple. Audio enters the system and is transformed into a spectral representation. In the spectral domain peaks are identified. Some heuristics are used to detect only salient, identifiable peaks and ignore spectral peaks in areas with equal energy – e.g. silent parts. Once peaks are identified, these are bundled to form triplets. Valid triplets only use peaks that are near both in frequency as in time. For performance reasons a peak is also only used in a limited number of triplets. These triplets are the fingerprints that are hashed and stored and ultimately queried for matches.
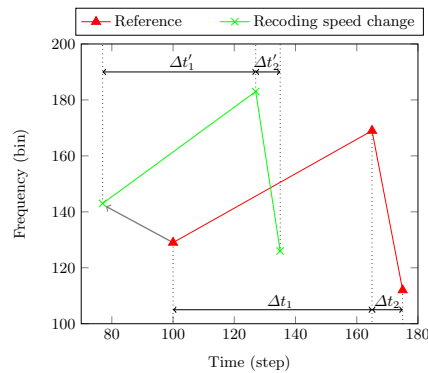

Fig. 1: The effect of speed modification on a fingerprint. It shows a single fingerprint extracted from reference audio (—▲—) and the same fingerprint extracted from audio after recording speed modification (—✕—).

Exact hashing makes lookup fast but needs to be done diligently to allow retrieval of audio with modified speed. A fingerprint together with a fingerprint extracted from the same audio but with modified speed can be seen in Figure 1. While absolute values regarding time change, ratios remain the same: $\frac{\Delta t_1}{\Delta t_2} = \frac{\Delta t_1'}{\Delta t_2'}$. The same holds true for the frequency ratios. This information is used in a hash. Next to the hash, the identifier of the audio is stored together with the start time of the first spectral peak.

Lookup follows a similar procedure: fingerprints are extracted and hashes are formed. Matching hashes from the database are returned and these lists are processed. If the list contains an audio identifier multiple times and the start times of the matching fingerprints align in time accounting for an optional linear scaling factor then a match is found. The linear time scaling factor is returned

together with the match. An implementation of this system was used in the case study.

4 The sound archive of the Royal Museum for Central Africa: a case study

The Royal Museum for Central Africa, Tervuren, Belgium preserves a large archive with field recordings mainly from Central Africa. The first recordings were made on wax cylinders in the late 19th century and later on all kinds of analogue carriers were used from various types of gramophone discs to sonofil. During a digitisation project called DEKKMMA (digitisation of the Ethnomusicological Sound Archive of the Royal Museum for Central Africa) [17] the recordings were digitised. Due to its history and size it is reasonable to expect that duplicates are present in the collection. In this case study we want to identify the duplicates, quantify the similarity in meta-data between duplicates and report the number of duplicates with modified speed. Here it is not the aim improve the data itself, this requires specialists with deep knowledge on the archive to resolve or explain (meta-data) conflicts: we mainly want to illustrate the practical use of duplicate detection.

With the Panako [15] fingerprints of 35,306 recordings of the archive were extracted. With the default parameters of Panako this resulted in an index of 65 million fingerprints for 10 million seconds of audio or 6.5 fingerprints per second. After indexing, each recording was split into pieces of 25 seconds with 5 seconds overlap, this means a granularity of 20 seconds. Each of those pieces (10,000,000 s / 20 s = 500,000 items) was compared with the index and resulted in a match with itself and potentially one or more duplicates. After filtering out identical matches, 4,940 fragments of 25 seconds were found to be duplicates. The duplicate fragments originated from 887 unique recordings. This means that 887 recordings (2.5%) were found to be (partial) duplicates. Thanks to the efficient algorithm, this whole process requires only modest computational power. It was performed on an Intel Core2 Quad CPU Q9650 @ 3.00GHz, with 8GB RAM, introduced in 2009.

Due to the nature of the collection, some duplicates were expected. In some cases the collection contains both the digitised version of a complete side of an analogue carrier as well as segmented recordings. Eighty duplicates could be potentially be explained in this way thanks to similarities in the recording identifier. In the collection recordings have an identifier that follows a scheme `collection name.year.collection identifier.subidentifier-track`. If a track identifier contains A or B it refers to a side of an analog carrier (cassette or gramophone disc). The duplicate pair `MR.1979.7.1-A1` and `MR.1979.7.1-A6` suggest that `A1` contains the complete side and A6 is track 6 on that side. The following duplicate pair suggests that the same side of a carrier has been digitised twice but stored with two identifiers: `MR.1974.23.3-A` and `MR.1974.23.3-B`. Unfortunately this means that one side is probably not digitised.

The 800 other duplicates do not have similar identifiers and lack a straightforward explanation. These duplicates must have been accumulated over the years. Potentially

duplicates entered in the form of analogue copies in donated collections. It is clear that some do not originate from the same analog carrier when listening to both versions. The supplementary material contains some examples. Next, we compare the meta-data difference between original and duplicate.

(a) Filing cabinet in the museum

(b) Main part of meta-data on file. Some fields use free, handwritten text (e.g. title) others a pre-defined list which are stamped (e.g. language)

Fig. 2: Meta-data on file

4.1 Differences in meta-data

Since the duplicates originate from the same recorded event, to original and duplicate should have identical or very similar meta-data describing their content. This is unfortunately not the case. In general, meta-data implementation depends on the history of an institution. In this case the older field-recordings are often made by priests or members of the military who did not follow a strict methodology to describe the musical audio and its context. Changes in geographical nomenclature over time, especially in Africa, is also a confounding factor [18]. There is also a large amount of vernacular names for musical instruments. The lamellophone for example is known as Kombi, Kembe, Ekembe, Ikembe Dikembe and Likembe [18] to name only a few variations. On top of that, the majority of the Niger-Congo languages are tonal (Yoruba, Igbo, Ashanti, Ewe) which further limits accurate, consistent description with a western alphabet. These factors, combined with human error in transcribing and digitising information, results in an accumulation of inaccuracies. Figure 2 shows the physical meta-data files. If there are enough duplicates in an archive, duplicate detection can serve as a window on the quality of meta-data in general.

Table 1 show the results of the meta-data analysis. For every duplicate a pair of meta-data elements is retrieved and compared. They are either empty, match exactly or differ. Some pairs match quite well but not exactly. It is clear that the title of the original *O ho yi yee yi yee* is very similar to the title of the duplicate *O ho yi yee yie yee*. To capture such similarities as well, a fuzzy string match algorithm based on SørensenDice coefficients is employed. When comparing the title of an original with a

Field	Empty	Different	Exact match	Fuzzy or exact match
Identifier	0.00%	100.00%	0.00%	0.00%
Year	20.83%	13.29%	65.88%	65.88%
People	21.17%	17.34%	61.49%	64.86%
Country	0.79%	3.15%	96.06%	96.06%
Province	55.52%	5.63%	38.85%	38.85%
Region	52.03%	12.16%	35.81%	37.95%
Place	33.45%	16.67%	49.89%	55.86%
Language	42.34%	8.45%	49.21%	55.74%
Functions	34.12%	25.34%	40.54%	40.54%
Title	42.23%	38.40%	19.37%	30.18%
Collector	10.59%	14.08%	75.34%	86.71%

Table 1: Comparison of pairs of meta-data fields for originals and duplicates. The field is either empty, different or exactly the same. Allowing fuzzy matching shows that fields are often similar but not exactly the same.

duplicate, only 19% match. If fuzzy matches are included 30% match. The table makes clear titles often differ while country is the most stable meta-data field. It also makes clear that the overall quality of the meta-data leaves much to improve. To correctly merge meta-data fields requires specialist knowledge - is it *yie* or *yi* - and individual inspection. This falls outside the scope of this case study.

Original title	Duplicate title
Warrior dance	Warriors dance
Amangbetu Olia	Amangbetu olya
Coming out of walekele	Walekele coming out
Nantoo	Yakubu Nantoo
O ho yi yee yi yee	O ho yi yee yie yee
Enjoy life	Gently enjoy life
Eshidi	Eshidi (man's name)
Green Sahel	The green Sahel
Ngolo kele	Ngolokole

Table 2: Pairs of titles that match only when using a fuzzy match algorithm.

4.2 Speed modifications

In our dataset only very few items with modified speed have been detected. For 98.8% of the identified duplicates the speed matches exactly between original and duplicate. For the remaining 12 identified duplicates speed is changed in a limited range, from -5% to +4%. These 12 pieces must have multiple analogue carriers in the archive. Perhaps copies were made with recording equipment that was not calibrated; or if the live event was captured from multiple angles, it is possible that the calibration of the original recorders was not consistent. There is a number of reasons why a digitised archive ends

up containing copies of the same content at slightly different speeds, but it is normally desirable that the cause for this depends on the attributes of the recordings *before* digitisation, and it is not introduced *during* the digitisation process. Our case study shows that duplicates can be successfully detected even when speed is modified. How this is done is explained in the following section.

5 De-duplication in practice

In this section, the practical functioning of Panako is described. The Panako acoustic fingerprinting suite is Java software and needs a recent Java Runtime. The Java Runtime and TarsosDSP[19] are the only dependencies for the Panako system, no other software needs to be installed. Java makes the application multi-platform and compatible with most software environments. It has a command-line interface, users are expected to have a basic understanding of their command line environment.
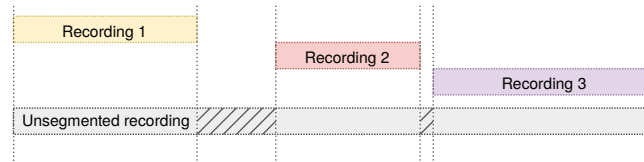

Fig. 3: Reuse of segmentation boundaries. The recording 1, 2 and 3 are found in a long unsegmented track. The segmentation boundaries (dotted lines) can be reused. Some parts in the unsegmented track remain unlabeled (the parts with diagonal lines).

Panako contains a deduplicate command which expects either a list of audio files or a text file that contains the full path of audio files separated by newlines. This text file approach is more practical on large archives. After running the deduplicate program a text file will contain the full path of duplicate files together with the time at which the duplicate audio was detected.

Several parameters need to be set for a successful de-duplication. The main parameters determine the granularity level, allowed modifications and performance levels. The granularity level determines the size of the audio fragments that are used for de-duplication. If this is set to 20 seconds instead of 10, then the number of queries is, obviously, halved. If speed is expected to be relatively stable, a parameter can be set to limit the allowed speed change. The performance can be modified by choosing the number of fingerprints that are extracted per second. The parameters determine several trade-offs between query speed, storage size, and retrieval performance. The default parameters should have the system perform reasonably effectively in most cases.

The indirect applications of linking meta-data is dependent on organization of the meta-data of the archive but has some common aspects. First, the audio identifiers of duplicates are arranged in original/duplicate pairs. Subsequently, the meta-data of these pairs is retrieved from the meta-data store (e.g. a relational database system). Finally, the meta-data element pairs are compared and resolved. The last step can use a combination of rules to automatically merge meta-data and manual intervention when

a meta-data conflict arises. The manual intervention requires analysis to determine the correct meta-data element for both original and duplicate.

Reuse of segmentation boundaries needs similar custom solutions. However, there are again some commonalities in reuse of boundaries. First, audio identifiers from the segmented set are identified within the unsegmented set resulting in a situation as in 3. The identified segment boundaries can subsequently be reused. Finally, segments are labeled. Since these tasks are very dependent on file formats, database types, meta-data formats and context in general it is hard to offer a general solutions. This means that while the duplicate detection system is relatively user friendly and ready to use, applying it still needs a software developer but not, and this is crucial, an MIR specialist.

6   Conclusions

In this paper we described possible applications of duplicate detection techniques and presented a practical solution for duplicate detection in an archive of digitised audio of African field recordings. More specifically applications were discussed to complement meta-data, to link or merge digital music archives, to improve listening experiences and to re-use segmentation data. In the case study on the archive of the Royal Museum of Central Africa we were able to show that duplicates can be successfully identified. We have shown that the meta-data in that archive differs significantly between original and duplicate. We have also shown that duplicate detection is robust to speed variations.

The archive used in the case study is probably very similar to many other archives of historic recordings and similar results can be expected. In the case study we have shown that the acoustic fingerprinting software Panako is mature enough for practical application in the field today. We have also given practical instructions on how to use the software. It should also be clear that all music archives can benefit from this technology and we encourage archives to experiment with duplicate detection since only modest computing power is needed even for large collections.

References

1. Orio, N.: Searching and classifying affinities in a web music collection. In: Italian Research Conference on Digital Libraries, Springer (2016) 59–70
2. Cano, P., Batlle, E., Kalker, T., Haitsma, J.: A review of audio fingerprinting. The Journal of VLSI Signal Processing **41** (2005) 271–284
3. IFLA - Audiovisual and Multimedia Section: Guidelines for digitization projects: for collections and holdings in the public domain, particularly those held by libraries and archives. Technical report, International Federation of Library Associations and Institutions (IFLA), Paris (France) (March 2002)
4. IASA-TC 04: Guidelines on the Production and Preservation of Digital Objects. IASA Technical Committee (2004)

5. Boston, G.: Safeguarding the Documentary Heritage. A guide to Standards, Recommended Practices and Reference Literature Related to the Preservation of Documents of all kinds. UNESCO (1998)

6. Bressan, F., Canazza, S., Vets, T., Leman, M.: Hermeneutic implications of cultural encoding: A reflection on audio recordings and interactive installation art. In Agosti, M., Bertini, M., Ferilli, S., Marinai, S., Orio, N., eds.: Digital Libraries and Multimedia Archives – Proceedings of the 12th Italian Research Conference on Digital Libraries (IRCDL 2016). Procedia - Computer Sciences, Elsevier (2017) 47–58

7. Wang, A.L.C.: An industrial-strength audio search algorithm. In: Proceedings of the 4th International Symposium on Music Information Retrieval (ISMIR 2003). (2003) 7–13

8. Haitsma, J., Kalker, T.: A highly robust audio fingerprinting system. In: Proceedings of the 3th International Symposium on Music Information Retrieval (ISMIR 2002). (2002)

9. Ellis, D., Whitman, B., Porter, A.: Echoprint - an open music identification service. In: Proceedings of the 12th International Symposium on Music Information Retrieval (ISMIR 2011). (2011)

10. Fenet, S., Richard, G., Grenier, Y.: A Scalable Audio Fingerprint Method with Robustness to Pitch-Shifting. In: Proceedings of the 12th International Symposium on Music Information Retrieval (ISMIR 2011). (2011) 121–126

11. Bellettini, C., Mazzini, G.: Reliable automatic recognition for pitch-shifted audio. In: Proceedings of 17th International Conference on Computer Communications and Networks (ICCCN 2008), IEEE (2008) 838–843

12. Ramona, M., Peeters, G.: AudioPrint: An efficient audio fingerprint system based on a novel cost-less synchronization scheme. In: Proceedings of the 2013 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2013). (2013) 818–822

13. Zhu, B., Li, W., Wang, Z., Xue, X.: A novel audio fingerprinting method robust to time scale modification and pitch shifting. In: Proceedings of the international conference on Multimedia (MM 2010), ACM (2010) 987–990

14. Malekesmaeili, M., Ward, R.K.: A local fingerprinting approach for audio copy detection. Computing Research Repository (CoRR) **abs/1304.0793** (2013)

15. Six, J., Leman, M.: Panako - A scalable acoustic fingerprinting system handling time-scale and pitch modification. In: Proceedings of the 15th ISMIR Conference (ISMIR 2014). (2014) 1–6

16. Sonnleitner, R., Widmer, G.: Quad-based Audio Fingerprinting Robust To Time And Frequency Scaling. In: Proceedings of the 17th International Conference on Digital Audio Effects (DAFx-14). (2014)

17. Cornelis, O., De Caluwe, R., Detr, G., Hallez, A., Leman, M., Matth, T., Moelants, D., Gansemans, J.: Digitisation of the ethnomusicological sound archive of the rmca. IASA Journal **26** (2005) 35–44

18. Cornelis, O., Lesaffre, M., Moelants, D., Leman, M.: Access to ethnic music: Advances and perspectives in content-based music information retrieval. Signal Processing **90**(4) (2010) 1008 – 1031 Special Section: Ethnic Music Audio Documents: From the Preservation to the Fruition.

19. Six, J., Cornelis, O., Leman, M.: TarsosDSP, a real-time audio processing framework in Java. In: Proceedings of the 53rd AES Conference (AES 53rd), The Audio Engineering Society (2014)