AUDIO FORENSICS MEETS MUSIC INFORMATION RETRIEVAL - A TOOLBOX FOR INSPECTION OF MUSIC PLAGIARISM

Christian Dittmar, Kay F. Hildebrand, Daniel Gaertner, Manuel Winges, Florian Müller*, Patrick Aichroth*

Fraunhofer IDMT, Department Metadata, Ilmenau, Germany
*University of Münster, Department of Information Systems, Münster, Germany

ABSTRACT

This paper presents a toolbox that has been developed in order to facilitate the inspection of suspected music plagiarism cases. The basic concept is the use of techniques from Music Information Retrieval for semi-automatic inspection of original and suspect song. Basic types of music plagiarism are discussed. Several signal processing approaches suitable to reveal these types are introduced. They are intended to be used under supervision of a human expert. Evaluation of the proposed methods in a non-supervised scenario is not within the scope of this paper.

Index Terms— Music Information Retrieval, Music Plagiarism, Audio Forensics

1. INTRODUCTION

Music plagiarism, i.e. the use of another work while presenting it as one's own original music, has always been a topic of public interest making headlines now and then. One recent example of so-called *sampling plagiarism* has been the case of the German Rap artist Bushido [1], whose producer used music excerpts from songs of the bands Dark Sanctuary and Dimmu Borgir without authorization or attribution. A prominent example for *melody plagiarism* was George Harrison's song *My Sweet Lord*, released 1970. The music label Bright Tunes Music sued Harrison for unauthorized usage of the song melody of *He's So Fine* by the Chiffons, released in 1962. The lawsuit lasted for more than ten years, finally finding that Harrison indeed imitated the melody, even though the responsible judge believed that he did so unintentionally [2].

Today, with huge public music databases and services such as YouTube[1], SoundCloud[2] or Spotify[3], there are endless opportunities not only for musical inspiration, but also for unintentional and intentional plagiarism. Thus, there is a need for approaches and tools to efficiently and transparently measure indications for plagiarism, thus helping to sift out the qualified cases, and to lower the costs associated with settling disputes.

Typically, when music plagiarism cases are brought to court, independent music experts, often musicologists, are asked to analyze the similarities between two songs, and the judges rely on their opinion. We believe that, in order to support such analysis, specialized software can be provided to analyze musical features of the suspected music recordings. Similarities can be identified by applying well-described pattern matching algorithms from the Music Information Retrieval (MIR) literature. Moreover, such software can display similarities in a way that experts can not only use to evaluate their importance, but also to visualize and explain it to an untrained audience.

This paper is organized as follows: Sec. 2 describes the types and intricacies of music plagiarism. Sec. 3 explains the proposed plagiarism analysis toolbox by outlining the signal processing approaches to inspect each plagiarism type. Sec. 4 concludes this work and gives future directions. It is important to note that a formal evaluation of the proposed algorithms is omitted in this paper. As the build-up of comprehensive test corpora is still under work, this remains a subject for future work.

2. TYPES OF MUSIC PLAGIARISM

A clear and precise definition of the term "plagiarism" is difficult to derive. However, the notion of intellectual property is known since ancient times. A poet named Martial called another poet "kidnapper" (lat. plagiarius), because he presented Martial's poems without permission and claimed that these were his own. This incident is perceived as the first mentioning of author's rights, even though an established copyright was unknown [3]. Authorship became more important with the invention of letterpress printing in the late 15^{th} century. In the realm of music, it became common practice to credit the composer for his sheet music since the 16^{th} century.

2.1. Sampling Plagiarism

The term sampling describes the re-use of recorded sounds or music excerpts in another song [4]. The samples are often

[1] www.youtube.com
[2] www.soundcloud.com
[3] www.spotify.com

manipulated in pitch or tempo to fit the rhythm and tonality of the new song. It is very common to mix additional instruments to the sample, such as additional vocals or drums. The most common use of samples is to crop an excerpt of one or more bars and loop them. More elaborate forms of sampling include rearrangement and post-processing of the respective sample beyond recognition. Sampling has strongly influenced popular music culture. Thus, there exist websites[4], where sampling cases are collected by a community of music aficionados. Due to the fact that sampling is basically the use of "a song in a song" it is related to the task of cover song detection [5]. Cover song detection is commonly approached by chroma features, as described in [6], [7] and [8]. A more recent approach is presented in [4], where a well-known audio fingerprinting algorithm [9] is modified in order to retrieve samples inside songs, based on spectral peak signatures.

2.2. Rhythm Plagiarism

A prominent example for rhythm plagiarism is the so-called "Amen Break". It originates from the 1969 Funk recording *Amen Brother* by The Winstons and is considered one of the most widely used drum loops in the history of Rap and Electronic music. Some of such extraordinary beats are protected by the law. But it is often difficult to judge, whether two songs share the same rhythm. There is no definition of which instrument is playing the rhythm. Commonly, the drums make up the beat. But a guitar can also be a dominant rhythmical instrument. In general, rhythm is formed by periodical pattern of accents in the amplitude envelopes of different frequency bands. Rhythm plagiarism has been scarcely covered in the literature but it is closely related to rhythm similarity estimation [10]. Paulus and Klapuri took the melody as a reference for rhythm [11]. They transformed the melody into rhythmical strings which are easier to compare along structural dimensions. Others extracted rhythmical features such as the beat spectrum or tempo in order to measure rhythmical similarity [12].

2.3. Melody Plagiarism

Copied melodies are less obvious than the previously explained plagiarism types. A melodic motive is considered to be identical, even if it is transposed to another key, slowed down, sped up or interpreted with different rhythmic accentuation. Thus, melody plagiarism is a gray area, where it is hard to discern copying from citation. However, MIR techniques [13], [14] are suited for inspection of such cases. In the MIR literature, a closely related task is Query-by-Humming (QbH). QbH can be used to retrieve songs from a database by letting the user hum or sing the respective melody [15]. Melody plagiarism inspection can be done with basically the same approach, since means to identify and evaluate melodic similarity are required. The main difference is, that QbH searches across extensive databases while plagiarism detection concentrates on one single comparison, which has to be more precise.

3. PLAGIARISM ANALYZER APPLICATION

We introduce the plagiarism analyzer application, which is developed in the scope of the REWIND[5] project. It features a graphical user interface and allows to import two music excerpts for analysis and comparison.

3.1. Sampling Plagiarism Inspection

As described in Sec. 2.1, sampling plagiarism occurs in different characteristics. In this paper, we consider the most common and simple approach of music excerpts that are resampled and looped as basis for the plagiarism song.

3.1.1. Brute Force Approach

The most straightforward approach to detect and inspect sampling plagiarism is to compare a time-frequency representation of both music excerpts [5]. We compute the magnitude spectrogram by means of Short-Term Fourier Transform (STFT) with an approximate hop-size of and block-size of 90ms. We convert each spectral frame to a constant-Q representation by means of re-sampling to a logarithmically spaced frequency axis, yielding the spectrograms of original X_o and suspected plagiarism X_s respectively. A number of hypotheses f for the applied re-sampling factor is derived by computing the pair-wise ratio of the strongest periodicities in the energy envelope of X_o and X_s. In order to retrieve the occurrences of X_o inside X_s, it is re-sampled both in time and frequency according to each entry in f, yielding $\tilde{X}_o$. Each $\tilde{X}_o$ is shifted frame-wise along all frames of X_s and the accumulated, absolute difference d is computed between all corresponding time-frequency tiles. Assuming only re-sampling and looping were applied, periodic minima will occur in d. These correspond to the point, where an optimal matching can be found. At this point, it is also possible to subtract the energy of $\tilde{X}_o$ from X_s, perform inverse STFT and auralize the result.

3.1.2. Decomposition Approach

The alternative approach is depicted in Fig. 1 and is based on decomposition of both X_o and X_s by means of Non-Negative Matrix Factorization (NMF) [16] and the modifications proposed in [17]. NMF is suited to factorize a spectrogram according to

$$X \approx B \cdot G \tag{1}$$

[4]www.whosampled.com, www.the-breaks.com

[5]www.rewindproject.eu

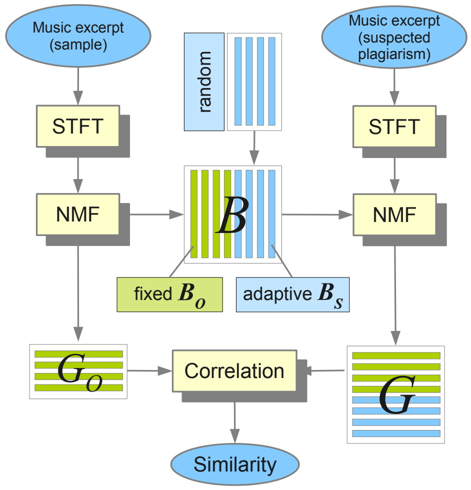

Fig. 1. Algorithm for inspecting sampling plagiarism

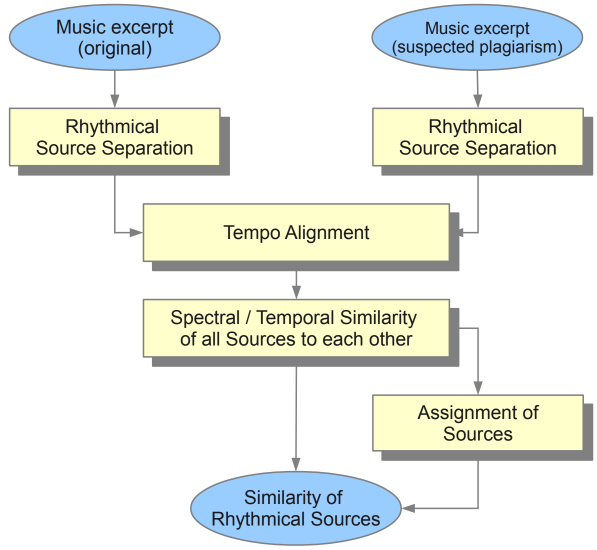

Fig. 2. Algorithm for inspecting rhythm plagiarism

The base spectra B represent characteristics of the used sounds. The amplitude envelopes G represent the time-varying gains associated to each spectrum in B. The plagiarism can be interpreted as a mixture of the known sample and unknown additional mixed sources. We create two sets of base spectra: one to model the original sample B_o and another to model the additional sounds B_s. The B_o are initialized by a preceding NMF of X_o. During the following NMF, these vectors stay fixed and will not be updated. The B_s are initialized with random values and are adapted with every NMF iteration. At the end of NMF the overall G are supposed to contain separated amplitude envelopes to model the sample and to model the additional sounds. Of course, this only works if the sample and the plagiarism have the same pitch and tempo, which is not guaranteed. Therefore, all possible variations in pitch are checked against by just shifting the B_o along the logarithmically spaced frequency axis, which equals re-sampling in the time domain. The above described process is repeated for all variations and scored via the reconstruction error.

3.2. Rhythm Plagiarism Inspection

As described in Sec. 2.2, rhythm plagiarism either poses a special case of sampling, or a complete re-arrangement. In this paper, we assume, that the original rhythm may have undergone a number of manipulations, such as time stretching, pitch shifting, re-sampling or even shuffling of individual beats. An overview of our approach to inspect rhythm plagiarism is depicted in Fig. 2.The single processing steps will be explained in the following sections.

3.2.1. Rhythmical Source Separation

First all rhythmical components of both X_o and X_s are again extracted by means of NMF. We follow the principle approach to compute the NMF with large number of components, and cluster these later on. From both the B_o,B_s and G_o,G_s features can be extracted that indicate an assignment to a certain instrument. We use a measure for periodicity [18] and further remove all components that show a low percussiveness [19]. Afterwards, a clustering of the components is necessary, since NMF often splits one instrument into several components. The assignment of components to each other is based on evaluating the correlation between the amplitude envelopes. For sake of brevity, further reading is referred to [20]. An example for the clustered G_o is depicted in Fig. 3. This visualization is also presented in the plagiarism analyzer application for visual inspection by the user.

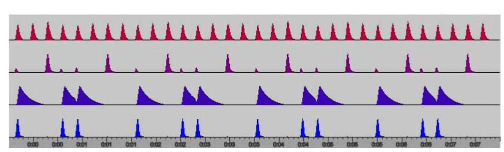

Fig. 3. Extracted amplitude envelopes of a drum-loop

3.2.2. Tempo Alignment

In order to compare the extracted sources, the tempi of the sequences have to be aligned to each other. Therefore, the G_o and G_s are transformed to logarithmically re-sampled autocorrelation functions (Log-Lag ACF) as described in [21]. On the log-lag axis, the lag shift between the Log-Lag ACF of the songs is retrieved by means of cross correlation. The shift corresponds to a re-sampling factor with which the tempo difference of the sequences can be compensated. Details of the method are described in [22].

3.2.3. *Similarity of Sources*

Every extracted source from the original is compared to the extracted ones from the suspected plagiarism. We take Pearson's correlation coefficient [23] of all possible permutations of G_o and G_s as well as B_o and B_s as similarity measure. Temporal and spectral shifts are accounted for by the use of normalized cross correlation. The best correlation values indicate that a pair of components can be assigned to each other. All distances are accumulated to an overall similarity measure by means of averaging across the pairs.

3.3. Melody Plagiarism Inspection

As explained in Sec. 2.3, melody plagiarism can be very hard to detect automatically. Thus, our approach is an automatic melody transcription as described in [24] yielding discrete MIDI note objects. Since automatic music transcription is still not mature enough to cope with any style and complexity of music recordings, we built a piano-roll view into the plagiarism analyzer that allows the user to transpose, stretch, move, merge and split notes as in conventional MIDI sequencers. Once a satisfactory transcription of the melody in original song and suspected plagiarism is available, the comparison is conducted by means of melodic similarity measurement.

3.3.1. *Pitch Vector Similarity*

The first approach is a local alignment similarity measure, utilizing Euclidean distances as described in [25]. The algorithm splits both melodies into smaller time windows of duration w, each of them represented by a vector with l pitch values. Rests are overwritten by extending the previous note. Each note of a melody has its own time window, starting with the onset of the note and ending with the note active w seconds later. The course of the pitch information within the time w is sub-sampled with sampling interval $\frac{w}{l}$. We subtract the mean-value from each of the sub-sampled vectors in order to guarantee invariance with respect to the musical key. Secondly, the duration of the windows is varied according to $w_s = \mu \cdot w_o$ where w_o is the window size of the original. A multiplier of $\mu = 1$ models the case where original and vector share the same tempo, a multiplier of $\mu < 1$ indicates that the suspect melody is played faster and a multiplier of $\mu > 1$ models the case where the suspect melody is played slower than the original.

3.3.2. *Sequence Alignment*

The second approach relies on the Smith-Waterman algorithm [26] to find a local alignment between symbol-sequences as described in [27]. The algorithm tries to identify subsequences of symbols, which encode intervals between consecutive notes in the MIDI transcription. On execution, each of these melody fragments is compared to the entire suspect sequence. The resulting scores are ordered descending and presented via the graphical user interface. Fig.4 shows an example of this visualization, the upper melody poses the original and the lower the suspect plagiarism. In this case the suspected melody is played faster, which is indicated by the relative length of both note sequences.

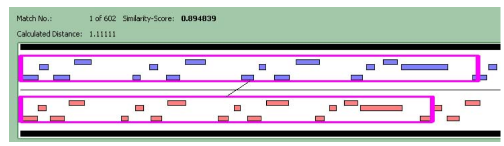

Fig. 4. Visualization of melodic similarity

4. CONCLUSIONS AND FUTURE WORK

We presented a signal processing toolbox for music plagiarism inspection. It combines several techniques from the MIR literature in order to allow semi-automatic analysis of suspected music plagiarism cases. A formal evaluation of the described methods has been omitted and will be subject to future work. Furthermore, other aspects of music plagiarism, such as the re-use of functional chord progressions are subjects for further research. It is planned to make a basic version of the plagiarism analyzer software toolbox freely available to the public.

5. ACKNOWLEDGMENTS

This research has been partially funded by the EU project REWIND. The project acknowledges the financial support of the Future and Emerging Technologies (FET) Programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 268478.

6. REFERENCES

[1] Billboard.biz, "Court Rules Bushido Plagiarized French Act," *Billboard.biz*, 2010, last accessed 2012–03–09.

[2] Music Copyright Infringement Resource, "Bright Tunes Music v. Harrisongs Music 420 F. Supp. 177 (s.d.n.y. 1976)," *UCLA - School of Law*, 1976, last accessed on 2011-09-24.

[3] G. Schricker, A. Dietz, C. Rudloff, D. Thum, and S. Hamadeh, *Urheberrecht*, Beck, 2006.

[4] J. Van Balen, Martín Haro, and J. Serrà, "Automatic identification of samples in hip hop music," M.S. thesis, London, UK, 2012.

[5] J. S. Downie, M. Bay, A. F. Ehmann, and M. C. Jones, "Audio cover song identification: Mirex 2006-2007 results and analyses," in *ISMIR Systems Evaluation Laboratory University of Illinois at Urbana-Champaign*, 2008.

[6] J. Serrà, E. Gómez, and P. Herrera, "Audio cover song identification and similarity: background, approaches, evaluation, and beyond," in *Advances in Music Information Retrieval*, 2010.

[7] Emanuele Di Buccio, Nicola Montecchio, and Nicola Orio, "A scalable cover identification engine," in *Proceedings of the international conference on Multimedia*, New York, NY, USA, 2010, MM '10, pp. 1143–1146, ACM.

[8] Emanuele Di Buccio, Nicola Montecchio, and Nicola Orio, "Falcon: Fast lucene-based cover song identification," in *Proceedings of the international conference on Multimedia*, New York, NY, USA, 2010, MM '10, pp. 1477–1480, ACM.

[9] A. L.-C. Wang, "An industrial strength audio search algorithm," in *Proc. of the 4th International Conference on Music Information Retrieval (ISMIR)*, Baltimore, USA, 2003.

[10] A. Kapur, M. Benning, and G. Tzanetakis, "Query-by-beat-boxing: Music retrieval for the DJ," in *Proc. of the 5th International Conference on Music Information Retrieval (ISMIR)*, 2004.

[11] J. Paulus and A. Klapuri, "Measuring the similarity of rhythmic patterns," in *3rd International Conference on Music Information Retrieval (ISMIR)*, 2002.

[12] J. Foote, M. Cooper, and U. Nam, "Audio retrieval by rhythmic similarity," in *3rd International Conference on Music Information Retrieval (ISMIR)*, 2002.

[13] R. Typke, F. Wiering, and R. C. Veltkamp, "A survey of music information retrieval systems," in *Proc. of the 6th International Conference on Music Information Retrieval (ISMIR)*, 2005.

[14] M. Müller, D. Ellis, A. Klapuri, and G. Richard, "Signal processing for music analysis," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 6, oct. 2011.

[15] M. Ryynänen, *Singing Transcription*, Springer, 2006.

[16] D. D. Lee and H. Sebastian Seung, "Algorithms for non-negative matrix factorization," in *In NIPS*. 2001, MIT Press.

[17] P. Smaragdis, "From learning music to learning to separate," in *Invited paper at Forum Acusticum*, Budapest, Hungary, 2005.

[18] D. Gärtner, "Tempo estimation from urban music using non-negative matrix factorization," in *Proc. of the AES 42nd International Conference on Semantic Audio*, Ilmenau, Germany, 2011.

[19] C. Uhle, C. Dittmar, and T. Sporer, "Extraction of drum tracks from polyphonic music using independent subspace analysis," in *Proc. of the 4th International Symposium on Independent Component Analysis and Blind Signal Separation*, 2003.

[20] C. Dittmar and C. Uhle, "Further steps towards drum transcription of polyphonic music," in *Proceedings Audio Engineering Society 116th Convention (AES)*, Berlin, Germany, 2004.

[21] C. Dittmar, M. Gruhne, and D. Gärtner, "Preprocessing methods for rhythmic mid-level features," in *Proceedings of International Conference on Acoustics NAG/DAGA*, Rotterdam, The Netherlands, 2009.

[22] T. Völkel, J. Abeßer, C. Dittmar, and H. Großmann, "Automatic genre classification on latin music using characteristic rhythmic patterns.," in *Proceedings of the Audio Mostly Conference*, Pitae, Sweden, 2010.

[23] J. L. Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician*, February 1998.

[24] K. Dressler, "Pitch estimation by the pair-wise evaluation of spectral peaks," in *Proc. of the AES 42nd International Conference on Semantic Audio*, Ilmenau, Germany, 2011.

[25] M. Ryynänen and A. Klapuri, "Query by humming of midi and audio using locality sensitive hashing," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2008.

[26] D. Gusfield, *Algorithms on Strings, Trees and Sequences*, Cambridge University Press, 1st edition, April 1997.

[27] M. Dittrich, "Untersuchung und Optimierung eines musikalisch fundierten Suchverfahrens zur Melodieerkennung," M.S. thesis, Ilmenau University of Technology, 2003.