Identification and classification of defects in vinyl disc records

Vaclav MOLDAN
Dept. of Radioelectronics, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
moldavac@fel.cvut.cz

Abstract. This paper presents current research focusing on identification and classification of defects on vinyl discs. To identify defects, identification method using wavelet transform model is adopted from literature. The current project’s goal is to classify the identified defects by use of KNN classification algorithm. This method is now able to successfully identify 57% of the records. These results are expected to be further improved by performing larger number of tests and after testing various modifications of the algorithm.

Keywords
Identification, classification, vinyl disc, impulse noise, wavelet transform, KNN, MATLAB

1. Introduction
To ensure the best possible sound quality to the customer, vinyl disc manufacturers need to have a control mechanism to check the vinyl disc and detect possible defects. Moreover, it is also useful to be able to identify what kind of defect is present in the record and based on that also identify what was the cause of it and during which phase of the manufacturing process was it created.

One of the possible methods is to have a dedicated person to listen to the created records and check for impulse noise and other attributes of the defect present in the recording. This approach has many disadvantages, since it is time consuming and can’t be very reliable, because human hearing tends to get tired after prolonged hours of use.

2. Suggested classification algorithm
In bachelor thesis [3] have been described and tested various algorithms to identify impulse noise in vinyl disc records. Best results were achieved with use of wavelet transform [9].

The goal of the current project is to extend the existing algorithm and implement a way to identify type of the defect of the disc. Since K-nearest neighbors algorithm has been successfully implemented in other audio records identification projects [4,5], it was decided to test to use this algorithm also for purpose of this project. The basic implementation of the KNN algorithm is adopted from [4].

Block diagram of the suggested algorithm can be seen in fig. 1.

3. KNN Classification algorithm
K-nearest neighbors [6] is a type of instance-based machine learning algorithm. At first, algorithm is trained from training data set. The training data are put into classes, to which the unknown test data is later also classified. During the classification itself, distance is measured between the training data set and the test data. The test data class is determined as the class of the most common element in the group of the K measured minimum distances.
4. Vinyl disc defects

The process of vinyl records manufacturing consists of several phases, and during each of these phases defects might be caused. These defects mainly result in impulse noise present in the records. Impulse noise is characterized by clicks and crackling present in the audio signal. Frequency spectrum of the impulse noise is usually flat, all spectral components are present with similar power [8].

The defects resulting in impulse noise can be further categorized. Table 1 describes some of the types of defects, which were provided by GZ Media, including audio samples for purpose of testing of the detection algorithm.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click, pop</td>
<td>General defect, resulting in clicks and pops of various intensity in the audio signal. The cause might be electrostatic click, dirt in the groove, overcut, jump (skip), scratch or even error in the master record.</td>
</tr>
<tr>
<td>Crackle, crackling</td>
<td>Sequence of clicks with various intensities. The cause might be defects from galvanoplasty or from pressing.</td>
</tr>
<tr>
<td>Scratch</td>
<td>There is a visible line across the vinyl disc. In case the scratch crosses the bottom of the groove, it causes click in each turn of the disc. If the scratch doesn't go deep enough to damage the bottom of the groove, it can be only optical defect.</td>
</tr>
<tr>
<td>Scratched matrix</td>
<td>In case of scratched matrix is the defect present on every vinyl disc created from the matrix. This leads to other defects, mainly clicks in each turn of the disc, single click and crackling. This type of defect is caused mainly by incorrect manipulation with the matrix.</td>
</tr>
<tr>
<td>Non-fills, un-fill</td>
<td>This defect results in crackling in one channel (right channel). Non-fills are placed on the lee side of the edge of the matrix. In most cases, the defect crosses more grooves, therefore the crackling can be heard on each turn. The cause is the pressing, when the mass doesn't surround the matrix enough. This defect usually appears near the end of the side. We can see the non-fills as clusters of white points.</td>
</tr>
<tr>
<td>Stitching, separation damage</td>
<td>This type of defect is created during separation of manufacturing tools, which can result in cutting in incorrect spots on the disc.</td>
</tr>
<tr>
<td>Stitching</td>
<td>Other type of stitching can be created during pressing of the disc, if the disc hits the matrix.</td>
</tr>
<tr>
<td>Horns, burrs</td>
<td>Horns (burrs) are created when cutting into the lacquer and can have several causes: attributes of the lacquer, shape of the cutting noise, temperature and speed of the cutting stylus. This type of defect causes inerasable smudge on the disc.</td>
</tr>
<tr>
<td>Dimple (optical defect) thump (acoustic defect)</td>
<td>Dimple or thump is a type of defect which is created already during preparation for the pressing. If dirt is not properly cleaned from the matrix, it causes defect of the matrix which is then present also on every disc created from the matrix. Thump is characterized by low frequency sound.</td>
</tr>
</tbody>
</table>

Tab. 1. Examples of vinyl disc defect types ([7], modified)
In detection algorithms which were dealing with audio record identification and were using KNN algorithm, signal is divided into short segments. Then value of certain parameter is calculated for each segment and based on the value of the parameter, most similar records are identified using KNN algorithm.

The task in this project differs from the other audio classification tasks, because we need to classify the defect of the vinyl record, which is present in only some parts of the signal. Therefore, segmentation as used in [4, 5] is not expected to have positive result, because then the classification would be more focused on the content of the whole record then on the impulse noise.

One option which I have decided to implement and test, is to separate the impulse noise from the audio record. The output of wavelet transform is a vector carrying information whether each sample contains defect or not. To separate short segments of signal containing the impulse noise, I have identified clusters of samples containing the error. I have set the maximum distance of samples with error within one cluster to 30.

For further computation of parameters for classification with KNN algorithm, it is useful to have all segments (in this case clusters of samples containing the impulse noise) of the same length. To solve this problem, I added vector containing zeros to the end of each separated segment so that the segment will now be extended to the length of 30 samples.

To easily characterize the audio record, existing parameters from MIR toolbox [2] were used. Parameters for initial testing were chosen according to the results from previous projects [4, 5].

Algorithm is at first trained for the certain parameter - the value of the parameter is computed for each error cluster of each record. Then during the classification process the value of the parameter is computed for the unknown test signal and the most similar record is found with KNN algorithm.

For initial test of the algorithm, the same records were used as training data set and as test data set. 8 records were provided with known type of defect from GZ Media. One of these records was not detected as containing defect, which provides 7 records for testing.

Test was performed with some of the parameters from MIR Toolbox. Results of the test can be seen in the table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Records successfully identified (out of 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brightness</td>
<td>4</td>
</tr>
<tr>
<td>Flux</td>
<td>4</td>
</tr>
<tr>
<td>RMS</td>
<td>2</td>
</tr>
<tr>
<td>Roll off</td>
<td>4</td>
</tr>
<tr>
<td>Tempo</td>
<td>4</td>
</tr>
</tbody>
</table>

Tab.2: Defects classification results

The results of the test are still not sufficient and are not giving satisfying result. Larger number of tests might
be necessary until results will be sufficient. The problem is currently mainly in KNN classification algorithm itself, since even if the value of the computed parameter value is the same for the classified record and record for which was the algorithm trained, KNN is struggling to match the records.

Another possible approach might be to take in account physical attributes of each defect type, such as the distance between each occurrence of the error (between each cluster of samples containing error) or periodic repetition of the occurrences for some types of defects.

6. Conclusion and future work

This paper presents current research project which focuses on classification of defect of vinyl disc records. With use of existing algorithms from other projects from FEE CTU, we can identify and separate segments containing defects from the audio signal.

For further classification of the defects, parameters from MIR Toolbox and KNN classification algorithm were used. The highest achieved success rate was 57 %. To receive better results, further testing with modifications of the existing algorithm might be needed.

Acknowledgements

Research described in the paper was supervised by Ing. Frantisek Rund Ph.D., FEE CTU in Prague and supported by the Grant Agency of the Czech Technical University in Prague, grant No SGS17/190/OHK3/3T/13.

Audio samples and description of vinyl disc defect types were kindly provided by GZ Media, a.s.

References

[1] BOLEK, M., Defect detection at mechanical sound records, FEE CTU, 2016
[7] List of vinyl degradations, GZmedia, internal material, 2015

About Authors...

Vaclav MOLDAN was born in Prague, Czech Republic in 1994. In 2017 he received his bachelor’s degree from Communication, Multimedia and Electronics program at Faculty of Electrical Engineering (FEE) Czech Technical University (CTU) in Prague. His bachelor’s thesis was focusing on audio records identification.

He is currently studying master’s program Electronics and Communication at the same faculty.