EFFICIENT MUSIC IDENTIFICATION METHOD BASED ON STRAIGHT LINE DETECTION SCHEME

Jeong-soo Yu
ITBOB, Korea
yjs@ilv.co.kr

Jeong-hwan Lee
ITBOB, Korea
jh.lee@ilv.co.kr

Gyeongsik Bang
ITBOB, Korea
ks.bang@ilv.co.kr

ABSTRACT

Our music identification method is based on line detection in large similarity matrix image. To overcome the problem of the infeasible dimension of similarity matrix for large-scale database, we utilize vector quantization-based index structure to find local lines quickly, pruning most calculation of similarity matrix value.

1. INTRODUCTION

This document shows an efficient audio fingerprinting scheme based on line detection on similarity matrix. Music identification is a problem of finding similar section between query audio and audios in the database. Generally, similar sections follow a linear model. Therefore, the identification problem is to find a straight line in the similarity matrix image. The two challenges of this line detection problem are that the similarity matrix is very large for large-scale database and the lines are usually very faint due to severe additive noise. In order to solve this challenge, our identification method first selects high similarity candidate points through fast approximate search, and then finds a straight line by gradually obtaining local images around the candidate points.

2. IDENTIFICATION METHOD

As stated above, the basic concept of finding a straight line for music identification is illustrated in Fig. 1. An audio is firstly converted into a spectrogram, and sequence of features $r_0, r_1, r_2, \ldots, r_i \in \mathbb{R}^p$ are extracted from each vertical line of the spectrogram. A dot of similarity matrix is the similarity value of two feature vectors. The more similar the two

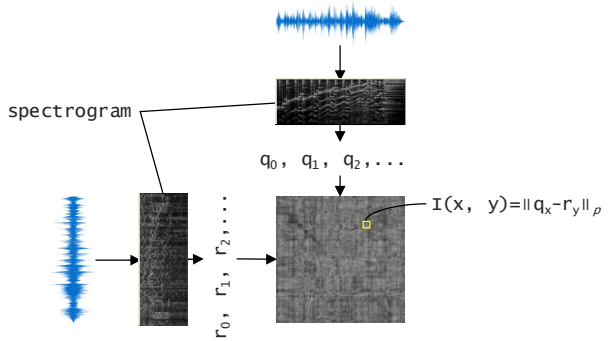

Figure 1. Example illustration of similarity matrix image of a pair of similar audio section.

features, the darker the dot value, and a dark straight line exists if the speed difference between the two audios is constant. If the speed of two audio is the same (snippet), the line is an exact $\pi/4$ diagonal.

The dimension of the similarity matrix is the product of the query audio length and the total sum of the lengths of the audios in the database. Therefore, it is not possible to obtain the entire matrix for a large-scale database. Our approach to solving this problem is as follows. First, candidate points are constructed by quickly searching for features in the database similar to each feature of query audio through vector-quantization based indexing. The indexing method is based on H. Jegou *at al.* [1] and we've improved the method to suit audio identification problem. Then the local straight line is detected while gradually calculating the matrix values near each candidate point, thus effectively pruning most matrix value calculation. Finally, the matching lines are obtained by merging the local lines and verified through the Hough-transform based line detection method [2]. Fig. 2 illustrates this identification process.

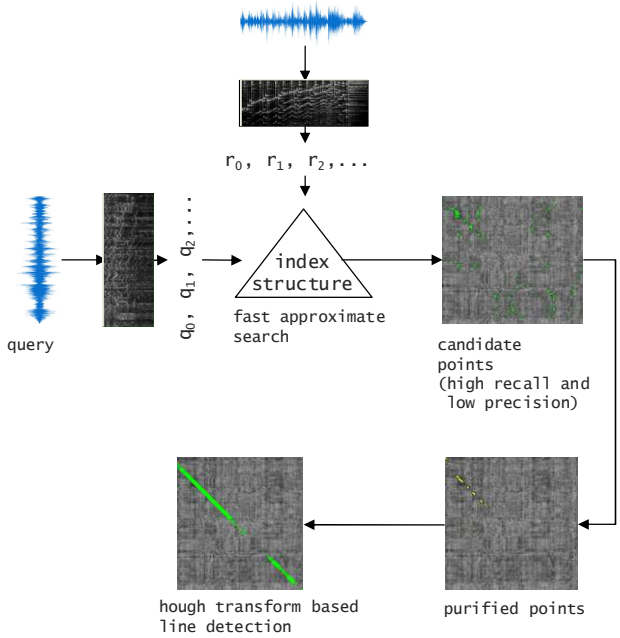

Figure 2. An illustration of identification process with approximate search and local line detection.

3. EFFICIENCY FEATURES

3.1 Space Cost

The bit rate of DNA files extracted from audio is about 16.5 KB per 1 minutes and the memory usage of the index structure is about 10GB per 10,000 songs.

3.2 Computation Cost

Extraction time of DNA files is about 20 seconds per 10,000 songs on Xeon silver 4114 2.2GHz and the cost of decoding MP3 is about 90%. Search time is about 2.5 seconds for 100 seconds query to 10,000 songs in database on the same machine.

4. CONCULUSION

We developed an efficient music identification method which is based on efficient line detection via vector quantization indexing. Our index method makes it possible to construct candidate points effectively through fast approximate search. We also are developing a more improved parallel search algorithm using GPGPU for fast music identification system of large-scale database.

5. REFERENCES

[1] H. Jegou, M. Douze, and C. Schmid: "Hamming embedding and weak geometric consistency for large scale image search," *Proceedings of the 10th ECCV*, pp. 304–317, 2008.

[2] Rau, J.Y. and L.C. Chen: "Fast straight lines detection using Hough Transform with principal axis analysis," *Journal of Photogrammetry and Remote Sensing*, Vol. 8, No.1, pp. 15–34, 2003.