International Journal of Engineering, Science and Mathematics
Vol.8 Issue 4, Apri2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

CONTENT RECOGNITION USING AUDIO FINGER PRINTING

G.SELVAKUMAR*
S.NIKIL NANDHA**
ADARSH.T.ARAVIND***
V.HARISH****

ABSTRACT

Sound fingerprinting is a procedure that utilizes PCs to break down little clasps of music chronicles to respond to a typical inquiry that individuals who tune in to music frequently solicit: "What is the name of that melody I hear?". Sound fingerprinting advances have as of late pulled in consideration since they permit the checking of sound autonomously of its organization and without the need of meta-information or watermark installing. Sound fingerprinting calculation make utilization of typically short sounds of three to thirty seconds long to make a sound unique finger impression. This sound unique mark is contrasted with a database of realized sound fingerprints to distinguish the first sound source. In the wake of getting distinguished it gives the metadata of the sound therefore. This work is finished with the assistance of Application Program Interface (ACR Cloud). The sound fingerprints of the sections don't really need to be of high caliber to be a match. Twists and obstruction of the first flag makes coordinating of the fingerprints less dependable, however (to a limited degree) it will even now be conspicuous. This can be utilized to look through a fragment of a sound document from a great many sound accumulations inside seconds with the assistance of ACR cloud API.

Copyright © 2019 International Journals of Multidisciplinary Research Academy. *All rights reserved.*

KEYWORDS:

Audio fingerprinting;
Watermark embedding;
ACR Cloud;
Content Recognition.

*Associate professor, Sri Shakthi Institute Of Engineering and Technology, Coimbatore
**UG Program, Sri Shakthi Institute Of Engineering and Technology, Coimbatore
***UG Program, Sri Shakthi Institute Of Engineering and Technology, Coimbatore
****UG Program, Sri Shakthi Institute Of Engineering and Technology, Coimbatore

1. INTRODUCTION

To develop an application that records audio content of a video using microphone and recognises the video id and its metadata using audio fingerprinting and indexing. The audio content of the video will be recorded for few seconds and the result can be displayed to the user.

People spend about 2 hours average per day on watching video content. Most of the time people are unaware about the video metadata. An average of 2 billion downloads occur for content recognition applications. Content recognition needs to be improvised and new innovations must be brought, such that they provide ease of access. Our application takes a simple problem statement of video recognition and identification that solves most of the real time complications on identifying meta data and provides ease of access for information gain.

Audio fingerprinting is used to take a short sample of an unknown audio recording and retrieve meta information about the recording. It does this by converting the data-rich audio signal into a series of short numerical values (fingerprints) that aim to uniquely identify a musical recording. Audio fingerprinting systems keep large databases of fingerprints for millions of known audio recordings. To identify an unknown audio recording query, the query's fingerprint is generated and compared to the reference database to find recordings that have identical or similar fingerprints. An audio fingerprinting system should be able to recognize audios of songs in the same way that a person can. If a person can recognize a song from a short clip of audio comes then an ideal computer system should also be able to recognize the song.

So as to coordinate an inquiry to a chronicle in a reference database, a fingerprinting calculation must produce indistinguishable fingerprints for the reference recording and question. The hashes must be indistinguishable notwithstanding when the question is extremely short, or when given a chronicle that sounds like a human, yet has an alternate sound flag. In the event that the account has been made, for instance, with a receiver in a boisterous room then the fingerprinting calculation ought to almost certainly separate the music in the chronicle from any extra clamor recorded by the amplifier. Reference databases of sound ought to contain unique mark codes for the whole term of sound. This is on the grounds that it is additionally valuable to distinguish a melody when just a bit of the tune is given as an inquiry. This portion could be recorded from any point in the melody, particularly if the account is made in an open spot from music that is being played

over a PA framework. Then again, unique mark query frameworks don't have to store a duplicate of the sound that is utilized to make the unique mark.

Fingerprints can be created and submitted to a database by any individual with a duplicate of the sound. While business fingerprinting frameworks can get fingerprints specifically from the music merchants, it is likewise conceivable to get fingerprints for uncommon music, no longer in production music, or music discharged through free marks straightforwardly from individuals who have duplicates of the sound. Sound watermarking is the demonstration of embeddings shrouded data into a sound stream, that can't be recognized by the human ear. It tends to be utilized for a large number of indistinguishable uses from fingerprinting, for instance, finding the copyright status of a melody, or giving metadata to a given account. Extra metadata can be incorporated into a watermark, including both data about the work or craftsman, or extra data (e.g., news communicate from a radio station alongside a melody). Sound watermarking procedures can store information in a sound stream at rates of up to 150kbps, giving enough space to incorporate metadata about the right now playing tune, or even work of art. The fundamental focus of the calculation for sound unique mark extraction is the capacity to institutionalize the substance of music/sound. A decent sound unique finger impression calculation extraction can speak to the music/sound from a mind boggling space of waveform into another component based area that can bolster looking at the likeness and contrast of two tunes/sound from each specific piece of its sound unique finger impression.

The principle focus of the calculation for sound unique finger impression extraction is the capacity to institutionalize the substance of music/sound. A decent sound unique mark calculation extraction can speak to the music/sound from an intricate area of waveform into another component based space that can bolster contrasting the likeness and distinction of two melodies/sound from each specific piece of its sound unique mark. In this usage, we use panako as the sound unique mark extraction calculation for the recognition of the mutilated sound from the first sound record. We acquired the benefits of the panako calculation for separating the sound unique mark and concentrating on proposing another shrewd stockpiling of fingerprints for the PC utilizing various GPGPUs and pleasing the parallel closest issue calculation that can support to deal with various questions in the meantime. Sound unique mark is a computerized vector that was extricated from the sound/melody waveform and ready to institutionalize the substance of the sound/tune source. Sound unique finger impression can without much of a stretch help for contrasting the similitudes and contrasts of tunes. Moreover, utilizing sound unique finger

impression for putting away can diminish the measure of unique sound/tune with the standard structure. In our framework's database, rather than capacity the genuine waveforms of the melodies we considered capacity of the sound fingerprints and its metainformation for each tune/track and changes over it into a recurrence area motion from which more data can be separated. Confining and cover decides what number of tests to think about while figuring a change of a period area flag. Each edge has a window connected to it to aid estimations, and the casings are handled in covering pieces from the time-arrangement flag. The element extraction process takes the flag that has been changed over into the recurrence area and chooses striking highlights that are utilized to describe the sound. At last, when the highlights have been picked and extricated from the flag, they should be changed over into a unique mark portrayal that can be put away in a database and contrasted with obscure question signals. For a full fingerprinting procedure, the system does not finish at the fingerprinting calculation (Figure 1). When the unique mark has been created, it must be put away in a reference database. The numerical portrayal of the unique finger impression is generally too awkward to be in any way utilized as an identifier, so a littler remarkable identifier is utilized. This could be as straightforward as the craftsman and tune name, or a short one of a kind string. A fingerprinting framework will give a query administration. This query task ought to almost certainly take an obscure info inquiry and match the question's unique finger impression with a unique finger impression that is in the reference database, restoring the identifier of the tune that the inquiry best matches, and alternatively the area in the melody that the question originates from. On the off chance that the tune being turned upward isn't in the reference database, the fingerprinting framework should report that the tune isn't in the database, instead of giving a wrong answer.

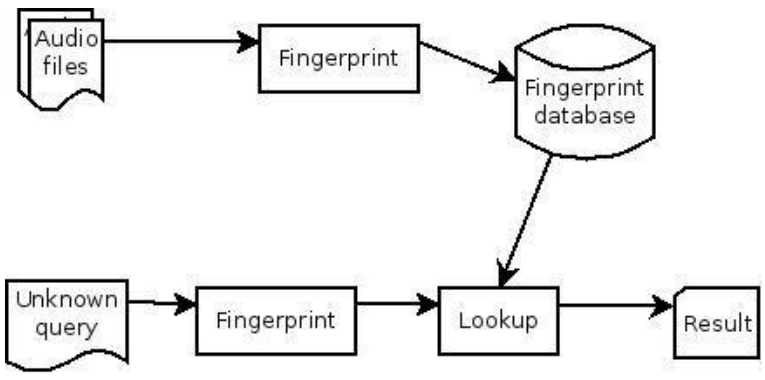


Figure 1. Audio fingerprinting process

EXISTING SYSTEM

'Shazam' is a free application that enables clients to make sense of the name of an appealing melody being played on the radio, TV or other spot. Clients just hold their iPhone gadget up to the speaker playing the music and the Shazam application will endeavor to distinguish the tune's collection, craftsman and tune title. Shazam works by investigating the caught sound and looking for a match dependent on an acoustic unique mark in a database of in excess of 11 million melodies. A spectrogram of the sound of a violin. The objective zone of a melody examined by Shazam. The client labels a tune for 10 seconds and the application makes a sound unique mark. The working of the shazam application is depicted in the figure.

For a fingerprinting calculation we need a decent recurrence goals (like 10.7Hz) to diminish range spillage and have a smart thought of the most critical notes played inside the melody. In the meantime, we have to diminish the calculation time beyond what many would consider possible and along these lines utilize the most minimal conceivable window measure.

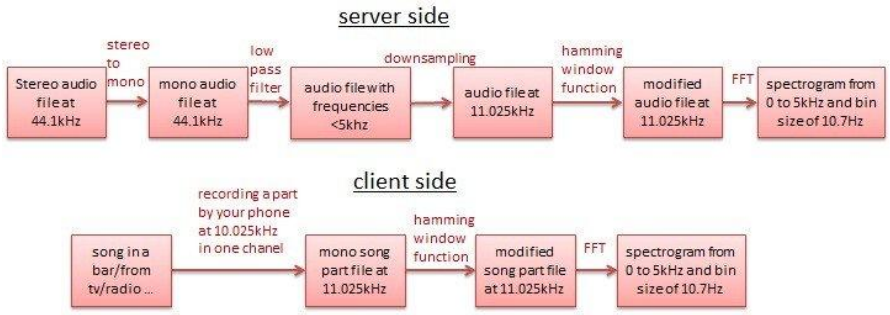


Figure 2. background process of shazam

On the server side (Shazam), the 44.1khz inspected sound (from CD, MP3 or whatever sound organization) needs to go from stereo to mono. It is conceivable to do that by taking the normal of the left speaker and the correct one. Before downsampling, at that point it have to channel the frequencies above 5kHz to abstain from associating. At that point, the sound can be downsampled at 11.025kHz.

On the customer side (telephone), the testing rate of the receiver that records the sound should be at 11.025 kHz. Shazam distinguishes melodies dependent on a sound unique finger impression dependent on a period recurrence chart called a spectrogram. It utilizes a cell phone or PC's worked in mouthpiece to accumulate a concise example of sound being played. Shazam stores an inventory of sound fingerprints in a database. The client labels a melody for 10 seconds and the application makes a sound unique mark. Shazam works by

breaking down the caught sound and looking for a match dependent on an acoustic unique finger impression in a database of in excess of 11 million tunes. On the off chance that it finds a match, it sends data, for example, the craftsman, tune title, and collection back to the client. A few executions of Shazam join important connects to administrations, for example, iTunes, Spotify, YouTube, or Groove Music. In the event that Shazam can't discover a match, it restores a "melody not known" exchange.

2. RESEARCH METHOD

Here we used a fingerprinting algorithm called ‘panako’to generate fingerprints from the fragment of the audio. User have to play the audio content of a video from any sources and the meta data about the audio content will be displayed on the screen. This is specially build for the TV series and also the custom audios that user can make and upload.By extracting the audio fingerprints (digital features) from the inputted audio fragment and then comparing to existing audio fingerprints in database, the detailed information of that audio fragment can be retrieved. This searching method does not only avoid common problems of traditional keyword-based search, such as lack of tags or wrong tags, but also allows users to perform queries without inputting any keywords.

SYSTEM ARCHITECTURE:

The figure shows the proposed framework design and here any YouTube or other sound or video sources can be taken as Recordings gathering. The unique mark of these recorded accumulations will be saved money on the database. It tends to be access by utilizing API. Furthermore, the custom recordings are recorded and transferred in the database further the fingerprints will be produced.

Sound fingerprinting, likewise named as sound hashing, has been outstanding as a ground-breaking method to perform sound recognizable proof and synchronization. It essentially includes two noteworthy advances: unique mark (sound example) plan and coordinating pursuit. While the initial step concerns the deduction of a strong and smaller sound mark, the second step as a rule requires learning about database and speedy inquiry calculations. In spite of the fact that this strategy gives a wide scope of true applications, to the best of the creators' information, an extensive study of existing calculations seemed over eight years prior. Along these lines, in this paper, we present a more cutting-edge audit and, for underscoring on the sound flag preparing viewpoint, we center our best in class study on the unique mark configuration venture for which different sound highlights and their tractable measurable models are examined.

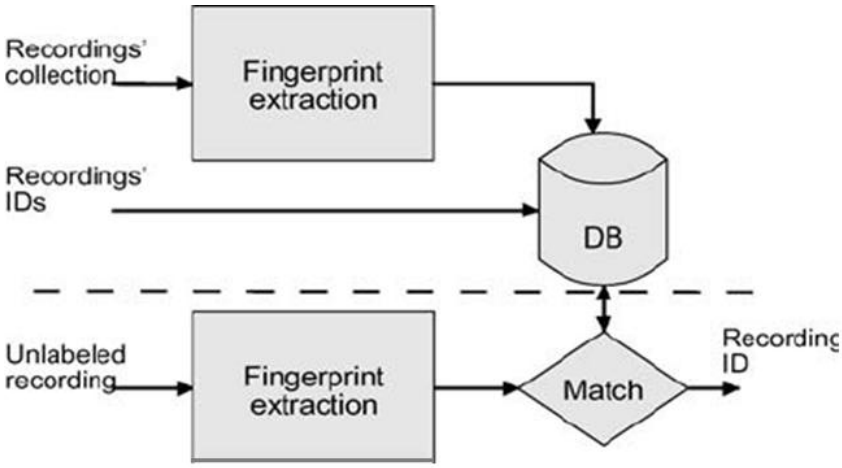


Figure 3. System Architecture

ACR LIBRARY IMPLEMENTATION

An audio fingerprint is a condensed digital summary, deterministically generated from an audio signal, that can be used to identify an audio sample or quickly locate similar items in an audio database. The follow picture gives us an intuitive understanding of audio finger print, we can take the black lines and points as finger prints.

ACR technology helps audiences easily retrieve information about the content they watched. For smart TVs and applications with ACR technology embedded the audience can check the name of the song which is played or descriptions of the movie they watched. In addition to that, the identified video and music content can be linked to internet content providers for on-demand viewing, third parties for additional background information, or complementary media.

The fingerprinting algorithm processes the signal of audios and extracts digital features called fingerprints for each audio. Fingerprints are very discriminative so the system can use them to identify the audio they belong to. Fingerprints are also robust which means they can resist the environment noise and this make it possible to identify recorded audios in rather noisy environments. If the system finds matched fingerprints of the query snippet, it can determine the most like audio in database and gives the position of the snippet in the source audio.

ACR BUCKET

Many individuals need to let their applications perceive their own sound and video substance with Audio fingerprinting advances to improve client experience and promoting viability. Typically,these substances incorporate music, Advertisement, network shows,

films and so on. Individuals can add substance they need to perceive into their own cans by transferring records specifically or sending just fingerprints through comfort APIs.

We utilized this ACR basin as a database for the capacity of custom sounds and its fingerprints. At the point when the calculation endeavors to perceive the sound it attempts to coordinate with both the ACR libraries and can.

ACR APPLICATION PROGRAM INTERFACE

Sound Fingerprinting (additionally called Acoustic Fingerprinting) is the sort of most stabled, compelling calculation of ACR and has been broadly utilized in numerous applications. A sound unique mark is a consolidated advanced synopsis, deterministically produced from a sound flag, that can be utilized to recognize a sound example or rapidly find comparable things in a sound database. The pursue picture gives us an instinctive comprehension of sound unique finger impression, we can take the dark lines and focuses as fingerprints.

Pragmatic employments of sound finger printing incorporate recognizing tunes, tunes, tunes, or commercials; audio effect library the board; and video document distinguishing proof. Media distinguishing proof utilizing acoustic fingerprints can be utilized to screen the utilization of explicit melodic works and exhibitions on radio communicate, records, CDs and shared systems. This recognizable proof has been utilized in copyright consistence, authorizing, and other adaptation plans.

A vigorous acoustic unique mark calculation must consider the perceptual attributes of the sound. In the event that two documents sound alike to the human ear, their acoustic fingerprints should coordinate, regardless of whether their twofold portrayals are very unique. Acoustic fingerprints are not bitwise fingerprints, which must be delicate to any little changes in the information. Acoustic fingerprints are increasingly undifferentiated from human fingerprints where little varieties that are immaterial to the highlights the unique finger impression utilizes are endured. One can envision the instance of a spread human unique finger impression which can precisely be coordinated to another unique mark test in a reference database; acoustic fingerprints work correspondingly. Perceptual qualities frequently misused by sound fingerprints incorporate normal zero intersection rate, assessed rhythm, normal range, unearthly levelness, conspicuous tones over a lot of recurrence groups, and data transfer capacity.

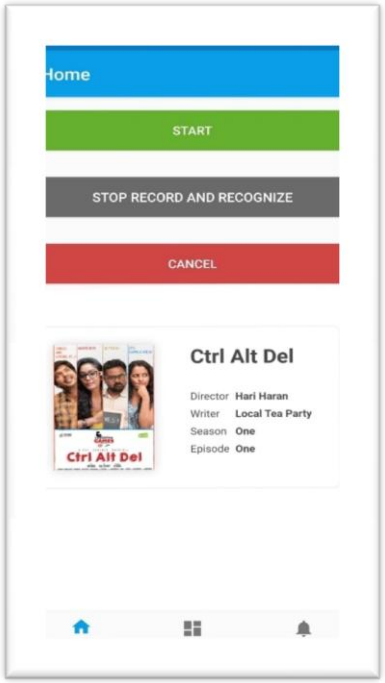
Most sound pressure methods (AAC, MP3, WMA, Vorbis) will roll out radical improvements to the twofold encoding of a sound record, without fundamentally influencing the manner in which it is seen by the human ear. A vigorous acoustic unique mark will enable an account to be distinguished after it has experienced such pressure, regardless of whether the sound quality has been diminished altogether. For use in radio communicate checking, acoustic fingerprints ought to likewise be unfeeling to simple transmission antiques. Then again, a great acoustic unique mark calculation must almost certainly distinguish a specific ace account among every one of the preparations of a craftsman or gathering. For use as proof in an official courtroom, an acoustic unique finger impression strategy must be legal in its exactness.

3. RESULTS AND ANALYSIS

As a result of the successful implementation of this algorithm we implemented an android application which recognize both local as well as custom audios with the help of Automatic content Recognition API. Within five to ten seconds this application converts the audio file in to fingerprints and match it with both the library collection and the bucket database which is used for the storage of custom audio fingerprints. The custom content as well as its fingerprint need to be uploaded in the bucket database for the matching purpose. After the audio gets converted to fingerprints the API matches it with the millions of fingerprints in the cloud. Which results to the successful recognition.

3.1 RECOGNITION OF CUSTOM AUDIOS

Here an audio which is uploaded in the ACR Bucket has recognized successfully. Which shows the successful recognition of custom audios. The local audios are recognized with the help of ACR API and custom vidos are recognized with the help of ACR Bucket. Here the main advantage is that it needs only a fragment of the audio to recognize it and not the entire audio.



Screenshot of uploaded audio recognition.

Here the figure represents the successful recognition of an audio which is uploaded in the ACR Bucket database. Any local videos can be identified with the help of ACR API and custom videos can be identified by ACR Bucket.

4. CONCLUSIONAND FUTURE SCOPE

A new mobile interface has been created with local and custom audio recognition. API for the content recognition is available in the ACR Cloud. With the help of this API we created a mobile interface that user can recognize the audio files including music, Tv shows and any other custom videos that are uploaded in the ACR Bucket. This interface helps the user to find the meta data about an audio file from any of the sources. One of the main advantages that this Interface provides is, the audio fingerprints of the segments do not necessarily have to be of high quality to be a match.Distortions and interference of the original signal makes matching of the fingerprints less reliable, but (to a certain extent) which takes only less than 10 seconds to recognize it. This provides the name of the audio which described in the Bucket along with its other description including image information.Within five to ten seconds this application converts the audio file in to fingerprints and match it with both the library collection and the bucket database which is used for the storage of custom audio fingerprints. The custom content as well as its fingerprint need to be uploaded in the bucket database for the matching purpose. After the

audio gets converted to fingerprints the API matches it with the millions of fingerprints in the cloud. Which results to the successful recognition.

This is an initial step in our research plan in which we are designing an Internet of Things system for hearing impaird who can communicate with the real world and respond to the various audio signals like door bell, telephone call, microwave oven and other alerts in kitchen and similar sounds where they have to respond. Our research will yield a wearable system through which the user can recognize the atmospheric sound and respond accordingly.

REFERENCES

[1] P. Cano, E. Batlle, T. Kalker, and J. Haitsma, "A review of audio fingerprinting," Journal of VLSI Signal Processing Systems, vo1.41, no.3, pp. 271-284, Nov. 2005.

[2] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney "Content-based music information retrieval: current directions and future challenges," Proceedings of the IEEE, vol. 96, no. 4, pp. 668-696, Apr. 2008.

[3] R. van Zwol, S. Rüger, M. Sanderson and Y. Mass: "Multimedia information retrieval: new challenges in audio visual search," WorkshopReport of ACM SIGIR Forum, 41(2), pp 77–82, Dec. 2007.

[4] W. Son, H. Cho, K. Yoon, and S. Lee, "Sub-fingerprint masking for a robust audio fingerprinting system in a real-noise environment for portable consumer devices," IEEE Transactions on ConsumerElectronics, vol. 56, no. 1, pp. 156-160, Feb. 2010.

[5] J. Haitsma and T. Kalker, "A Highly Robust Audio Fingerprinting System With an Efficient Search Strategy," J. New Music Research, vol. 32, no. 2, pp. 211–221, 2003.

[6] J. Haitsma and T. Kalker, "A Highly Robust Audio Fingerprinting System," Proc. Int. Symp. on Music Information Retrieval, pp. 107–115, 2002.

[7] M.L. Miller, M.C. Rodriguez, and I.J. Cox, "Audio Fingerprinting: Nearest Neighbor Search in High Dimensional Binary Spaces," J. VLSI Signal Processing, vol. 41, pp. 285–291, 2005.

[8] J. Oostveen, T. Kalker, and J. Haitsma "Feature Extraction and a Database Strategy for Video Fingerprinting, " Proc. Int. Conf. on Visual Information Systems, vol. LNCS 2314, pp. 117–128, 2002.

[9] P. Cano, E. Batlle, T. Kalker, and J. Haitsma, "A review of audio fingerprinting," Journal of VLSI Signal Processing Systems, vo1.41, no.3, pp. 271-284, Nov. 2005.

[10] M. Riley, E. Heinen, and J. Ghosh, "A Text-Retrieval Approach to Content-Based Audio Retrieval," in Proc. The 1st ACM International Conference on Multimedia Information Retrieval, Vancouver, Canada, pp. 105-112, Oct. 2008.

[11] M. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney "Content-based music information retrieval: current directions and future challenges," Proceedings of the IEEE, vol. 96, no. 4, pp. 668-696, Apr. 2008.

[12] R. van Zwol, S. Rüger, M. Sanderson and Y. Mass: "Multimedia information retrieval: new challenges in audio visual search," Workshop Report of ACM SIGIR Forum, 41(2), pp 77–82, Dec. 2007. 20.Y. Qian, H. Dou and Y. Feng, "A novel algorithm for audio information retrieval based on audio fingerprint," in Proc. 2010 International Conference on Information Networking and Automation, Kunming, China, pp. 266 - 270, Oct. 2010.