

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/311848940>

A study of Shazam's Audio Recognition

Presentation · December 2016

DOI: 10.13140/RG.2.2.21768.21766

CITATIONS

0

READS

3,620

1 author:



Guendalina Palmirotta

University of Luxembourg

5 PUBLICATIONS 0 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Lower bounds for kissing numbers [View project](#)



Construction of flexible polyhedra [View project](#)

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam
Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)
Fingerprinting and
Hash functions
Anchor point and
target zone
Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

A study of Shazam's Audio Recognition

Seminar in Data Science

PALMIROTTA Guendalina

University of Luxembourg,
Faculty of Science, Technology and Communication,
Master in Mathematics
Winter semester 2016-2017

December 21, 2016

Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

1 Get in touch with Shazam

- Motivation

2 The magic behind the Shazam algorithm

- From digital sound to spectrogram
 - Spectrogram (Sonic visualization)
- Fingerprinting and Hash functions
 - Anchor point and target zone
- Matching of a song
 - Scatterplot and Histogram

3 Conclusion

- Roboustness, Speedness and Noise resistanceness

4 A mini Shazam on R

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)
Fingerprinting and
Hash functions
Anchor point and
target zone
Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

1 Get in touch with Shazam

• Motivation

2 The magic behind the Shazam algorithm

- From digital sound to spectrogram
 - Spectrogram (Sonic visualization)
- Fingerprinting and Hash functions
 - Anchor point and target zone
- Matching of a song
 - Scatterplot and Histogram

3 Conclusion

- Roboustness, Speedness and Noise resistanceness

4 A mini Shazam on R

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Robustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Motivation – Introduction

- 1 Capture the music for a few seconds (5-15s)
- 2 Identification of the song
- 3 Display the information (name, artist, album)



Goal

Recognize our unknown song in a short time using Shazam music application

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Difficulty and constraints

Develop an algorithm that is able to:

- Capture by a little microphone a short sample of music
- Often with mixed heavy ambient noise
- Quick identificaton over a large database of music \Rightarrow 2M tracks



Keywords:

Roboustness, Noise resistanceness and Speedness

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

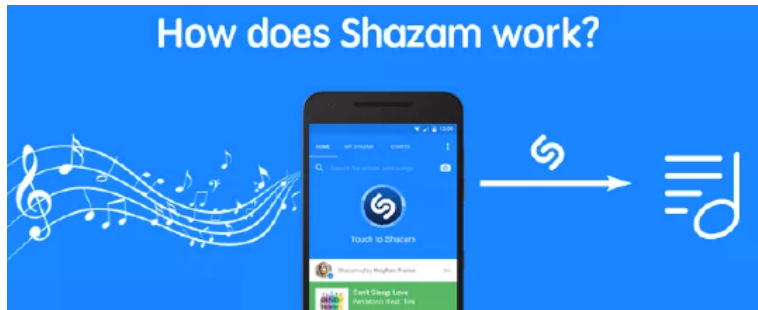
Matching of a song
Scatterplot and
Histogram

Conclusion

Robustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Seems to be magic, but...



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

1 Get in touch with Shazam

- Motivation

2 The magic behind the Shazam algorithm

- From digital sound to spectrogram
 - Spectrogram (Sonic visualization)
- Fingerprinting and Hash functions
 - Anchor point and target zone
- Matching of a song
 - Scatterplot and Histogram

3 Conclusion

- Roboustness, Speedness and Noise resistanceness

4 A mini Shazam on R

Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind the
Shazam algorithm

From digital sound to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

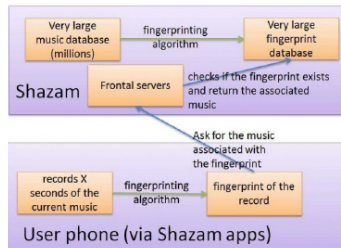
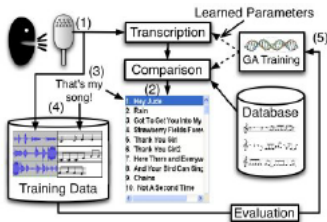
Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Overview

Shazam two sides : 'Client' side and the 'Server' side



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

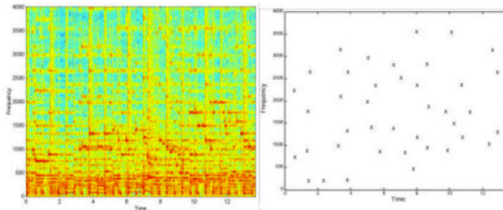
Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Step by step – Continuous signal to discrete signal

Step 1

The song must be transformed in a time-frequency graph, that we call spectrogram. Then we do a kind of filtration, we get a constellation map and keep only the 'important points'.



How to get a spectrogram?

Goal: From digital sound to frequency

Discrete Fourier Transform (DFT)

$$X(n) = \sum_{k=0}^{N-1} x[k] e^{-\frac{2\pi i kn}{N}},$$

- N size of (Hamming) window
- $X(n)$ the n -th bin of frequencies
- $x[k]$ the k -th sample of the audio signal

⇒ Use Fast Fourier Transform (FFT) instead of the DFT

Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

**Spectrogram (Sonic
visualization)**

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

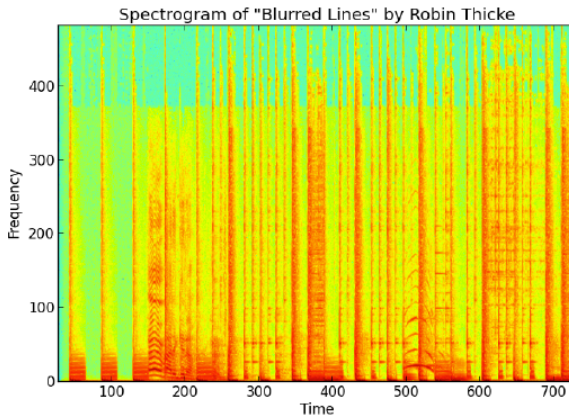
Conclusion

Robustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Spectrogram filtering

Combination of sinewaves at multiple frequencies



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind the Shazam
algorithm

From digital sound to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

Conclusion

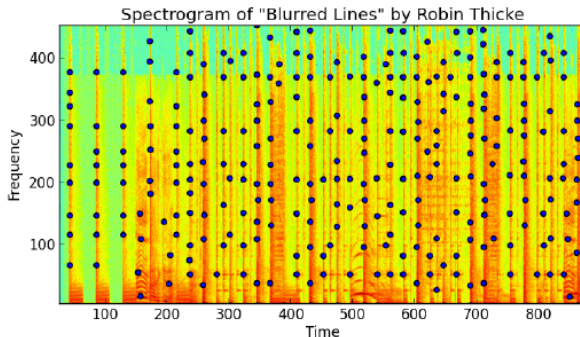
Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Filtration : Consellation map

What are the important nodes to consider?

- Depends on the coefficients of the bins
- Depends on the number of bands of the strongest time-frequency point



Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Step by step – Fingerprinting and Hash functions

Step 2

We code the song in a unique acoustic fingerprints and store it in a hash tag table.



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind the
Shazam algorithm

From digital sound to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

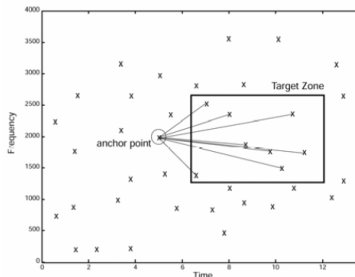
A mini Shazam
on R

How to store?

Use anchor point with their corresponding zone, called target zone

Idea

To look for multiple points at the same time instead of comparing each point one by one.



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

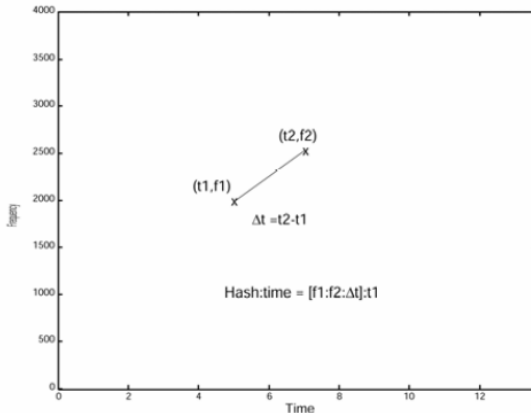
Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Hash details

Each time an anchor point lies inside the target zone \rightarrow a hash is created.



Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram
Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Searching and Scoring the fingerprints

Fingerprint

Unknown song: $[(f_1, f_2, \Delta t), t_1] \rightarrow [t_1]$

Update a new song: $[(f_1, f_2, \Delta t), t_1, ID] \rightarrow [t_1, ID]$

- (t_1, f_1) time-frequency at which the anchor point is located,
- (t_2, f_2) time-frequency at which the point in the target zone is positioned,
- $\Delta t = t_2 - t_1$ the time difference between t_1 and t_2 ,
- ID of the song (name, artist, album, of the song).

Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with Shazam

Motivation

The magic behind the Shazam algorithm

From digital sound to spectrogram

Spectrogram (Sonic visualization)

Fingerprinting and Hash functions

Anchor point and target zone

Matching of a song

Scatterplot and Histogram

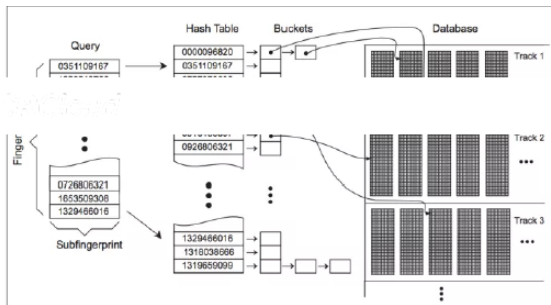
Conclusion

Robustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Hash function and table

⇒ Store in a hash tag table in the database



- Hash tag table: $\text{Hash}_{\text{table}} = [\text{Hash}(1), \dots, \text{Hash}(n)]$
- Bucket is a specific location in database

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Step by step – Matching of a song

Step 3

For the matching factor we use a scatter graph and the corresponding histogram graph.

Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind the
Shazam algorithm

From digital sound to
spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

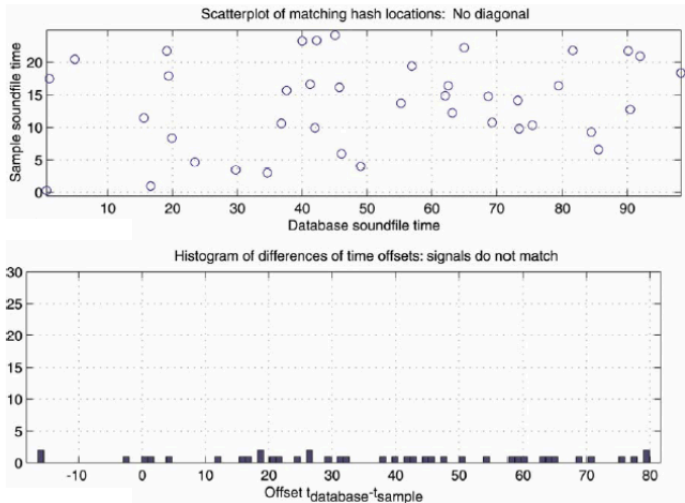
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Scatterplot and Histogram of no matching



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

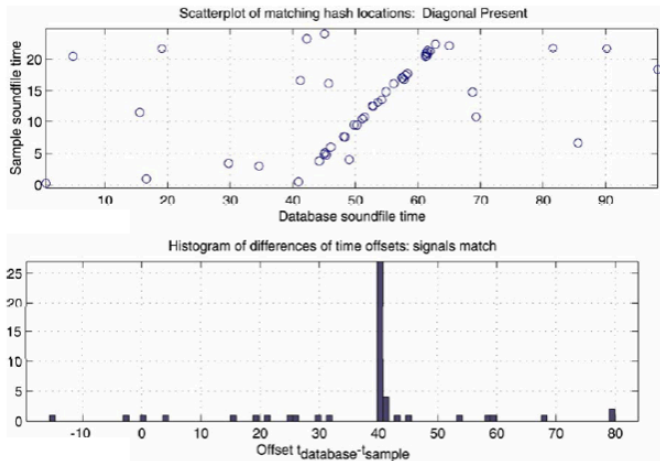
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Scatterplot and Histogram of matching



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

1 Get in touch with Shazam

- Motivation

2 The magic behind the Shazam algorithm

- From digital sound to spectrogram
 - Spectrogram (Sonic visualization)
- Fingerprinting and Hash functions
 - Anchor point and target zone
- Matching of a song
 - Scatterplot and Histogram

3 Conclusion

- Roboustness, Speedness and Noise resistanceness

4 A mini Shazam on R

Shazam's Audio
Recognition

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind the
Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

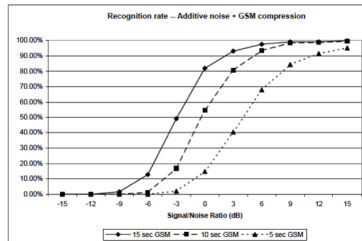
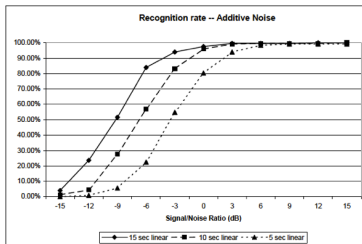
Scatterplot and
Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam
on R

Roboustness, Speedness and Noise resistanceness



- Identify the music with a rate of 90% of correctness in a short time with noise!

**Shazam's Audio
Recognition**

PALMIROTTA
Guendalina

Get in touch with
Shazam

Motivation

The magic behind
the Shazam
algorithm

From digital sound
to spectrogram

Spectrogram (Sonic
visualization)

Fingerprinting and
Hash functions

Anchor point and
target zone

Matching of a song

Scatterplot and
Histogram

Conclusion

**Roboustness,
Speedness and
Noise
resistanceness**

A mini Shazam
on R

Thank you for your attention!



Shazam's Audio Recognition

PALMIROTTA
Guendalina

Get in touch with Shazam

Motivation

The magic behind the Shazam algorithm

From digital sound to spectrogram

Spectrogram (Sonic visualization)

Fingerprinting and Hash functions

Anchor point and target zone

Matching of a song
Scatterplot and Histogram

Conclusion

Roboustness,
Speedness and
Noise
resistanceness

A mini Shazam on R

- 1 Get in touch with Shazam
 - Motivation
- 2 The magic behind the Shazam algorithm
 - From digital sound to spectrogram
 - Spectrogram (Sonic visualization)
 - Fingerprinting and Hash functions
 - Anchor point and target zone
 - Matching of a song
 - Scatterplot and Histogram
- 3 Conclusion
 - Roboustness, Speedness and Noise resistanceness
- 4 A mini Shazam on R