



**Technische Universität Berlin**

Fakultät I - Geisteswissenschaften  
Fachgebiet Audiokommunikation  
Audiokommunikation und -technologie M.Sc.

# **Self-Organizing Maps for Sound Corpus Organization**

MASTER'S THESIS

**Vorgelegt von:** Jonas Margraf  
**Matrikelnummer:** 372625  
**E-Mail:** jonasmargraf@me.com

**Erstgutachter:** Prof. Dr. Stefan Weinzierl  
**Zweitgutachter:** Dr. Diemo Schwarz  
**Datum:** February 7, 2019



## Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den February 7, 2019

.....  
Jonas Margraf

**Abstract** An english abstract.

**Zusammenfassung** Die Zusammenfassung auch auf Deutsch.

## **Acknowledgements**

This is where the thank yous go.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation and Problem Description . . . . .	1
1.2	Aims and Objectives . . . . .	1
1.3	Previous Work . . . . .	1
<b>2</b>	<b>Background</b>	<b>2</b>
2.1	Audio Feature Extraction . . . . .	2
2.1.1	Fundamentals . . . . .	2
2.1.2	Audio Pre-Processing . . . . .	2
2.1.3	Time-Domain Features . . . . .	2
2.1.3.1	Root Mean Square (RMS) . . . . .	2
2.1.3.2	Zero-Crossing Rate (ZCR) . . . . .	2
2.1.4	Frequency-Domain Features . . . . .	2
2.1.4.1	Spectral Centroid . . . . .	2
2.1.4.2	Spectral Flatness . . . . .	2
2.1.4.3	Spectral Kurtosis . . . . .	2
2.1.4.4	Spectral Skewness . . . . .	2
2.1.4.5	Spectral Slope . . . . .	2
2.1.4.6	Spectral Spread . . . . .	2
2.1.4.7	Spectral Rolloff . . . . .	2
2.1.5	Perceptual Features . . . . .	2
2.1.5.1	Loudness . . . . .	2
2.2	Self-Organizing Map . . . . .	2
<b>3</b>	<b>Implementation</b>	<b>3</b>
3.1	Groundwork: CataRT Extension . . . . .	3
3.2	SOM Browser . . . . .	3
<b>4</b>	<b>Evaluation</b>	<b>4</b>
4.1	Measuring SOM-Induced Quantization . . . . .	4
4.2	Online Sound Similarity Survey . . . . .	4
4.3	Semistructured User Interviews . . . . .	4
<b>5</b>	<b>Results</b>	<b>5</b>
<b>6</b>	<b>Discussion</b>	<b>6</b>
6.1	Outlook . . . . .	6
	<b>Appendices</b>	<b>7</b>

<b>A</b>	<b>LaTeX Sources</b>	<b>7</b>
<b>B</b>	<b>Thesis Bibliography</b>	<b>7</b>
	<b>Glossary</b>	<b>I</b>
	<b>Acronyms</b>	<b>I</b>
	<b>List of Figures</b>	<b>II</b>
	<b>List of Listings</b>	<b>III</b>
	<b>List of Tables</b>	<b>IV</b>
	<b>Digital Resource</b>	<b>V</b>



## **1 Introduction**

This is the Introduction.

### **1.1 Motivation and Problem Description**

### **1.2 Aims and Objectives**

### **1.3 Previous Work**

## 2 Background

This is the Background section.

### 2.1 Audio Feature Extraction

#### 2.1.1 Fundamentals

#### 2.1.2 Audio Pre-Processing

#### 2.1.3 Time-Domain Features

**2.1.3.1 Root Mean Square (RMS)** Root Mean Square (RMS) goes here.

#### 2.1.3.2 Zero-Crossing Rate (ZCR)

#### 2.1.4 Frequency-Domain Features

##### 2.1.4.1 Spectral Centroid

##### 2.1.4.2 Spectral Flatness

##### 2.1.4.3 Spectral Kurtosis

##### 2.1.4.4 Spectral Skewness

##### 2.1.4.5 Spectral Slope

##### 2.1.4.6 Spectral Spread

##### 2.1.4.7 Spectral Rolloff

#### 2.1.5 Perceptual Features

##### 2.1.5.1 Loudness

### 2.2 Self-Organizing Map

Something about Self-Organizing Maps (SOMs) and also neurons have IDs.

## **3 Implementation**

This is the Implementation.

### **3.1 Groundwork: CataRT Extension**

### **3.2 SOM Browser**

## **4 Evaluation**

This is the Evaluation.

### **4.1 Measuring SOM-Induced Quantization**

### **4.2 Online Sound Similarity Survey**

### **4.3 Semistructured User Interviews**

## 5 Results

This is the Results section.

## 6 Discussion

This is the Discussion.

### 6.1 Outlook

# Appendices

## **A   LaTeX Sources**

The  $\text{\LaTeX}$  sources for this work can be found in XXX.

## **B   Thesis Bibliography**

The references used in this work can be found in XXX.

## **Glossary**

ID A name or number that identifies an object.

## **Acronyms**

RMS Root Mean Square.

SOM Self-Organizing Map.



## List of Figures

## List of Listings

## List of Tables

## Digital Resource

This page holds a data disk.