

Klassifikation von Meeressäugern mit Mitteln der Spracherkennung

Dissertation

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

(Dr.-Ing.)

der Technischen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Roman Kreimeyer

Kiel 2015

Tag der Disputation: 18.03.2016

- | | |
|----------------------|------------------------------------|
| 1. Berichterstatter: | Prof. Dr.-Ing. habil. Ulrich Heute |
| 2. Berichterstatter: | Prof. Dr.-Ing. Peter Vary |

Arbeiten über Digitale Signalverarbeitung

Band 43

Roman Kreimeyer

**Klassifikation von Meeressäugern
mit Mitteln der Spracherkennung**

Shaker Verlag
Aachen 2016

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Zugl.: Kiel, Univ., Diss., 2016

Copyright Shaker Verlag 2016

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-4568-0

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen
Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9
Internet: www.shaker.de • E-Mail: info@shaker.de

Zusammenfassung

Neben der Verschmutzung der Meere durch Müll, Abwässer oder die Schifffahrt nimmt auch die sog. akustische Verschmutzung der Meere, z.B. durch Schiffslärm, akustische Kommunikation, Echolote oder SONAR-Systeme immer weiter zu. Hiervon sind Meeressäuger mit ihrer starken Anpassung an den Lebensraum Meer besonders betroffen. Das Projekt *Protection of Marine Mammals* der Europäischen Verteidigungsagentur soll Meeressäuger besser vor den Auswirkungen aktiven SONARs bei Beibehaltung der Einsatzfähigkeit solcher Systeme schützen. Hierfür ist eine Klassifikation der Meeressäuger im Einflussgebiet des SONARs notwendig, um entsprechende Schutzmaßnahmen planen und ausführen zu können. Diese Arbeit greift für diese Aufgabe Methoden der Spracherkennung auf, um Synergieeffekte aus mehreren Forschungsgebieten zu nutzen. Ausgangsbasis für die Klassifikation ist eine im Laufe des Projektes erstellte Datenbank mit Meeressäuger-Lauten. Die für eine erfolgreiche Klassifikation notwendigen Merkmale zur Parameterisierung der Meeressäuger-Laute werden aus den zur Verfügung stehenden Signalen gewonnen. Neben klassischen Merkmalen aus verschiedenen Signaldomänen werden weitere Merkmale mittels zweier eigenständiger Detektoren gewonnen. Diese bestehen zum Einen aus einer Detektion von Klicklauten, der eine Rekonstruktion bzw. Schätzung der Klicksequenzen mittels Clustering folgt, und zum Anderen aus einem Detektor für tonale Laute.

Die bestimmten Merkmale werden einer Transformation unterzogen und anhand eines Qualitätskriteriums in ihrem möglichem Beitrag zur Klassifikation verglichen. Anschließend wird basierend auf einer Grundparametrierung mithilfe einer Kreuzvalidierung untersucht, welche Anzahl transformierter Merkmale für ein gutes Klassifikationsergebnis notwendig sind.

Zur Bewertung des Klassifikationsverhaltens werden drei der vorgestellten Klassifikationsmethoden verglichen. Zu den Auswahlkriterien zählen neben der Klassifikationsgenauigkeit auch die benötigte Verarbeitungszeit um die Implementierung als Echtzeitsystem zu ermöglichen.

Abschließend werden die Ergebnisse der Kreuzvalidierung mit dem Detektions- und Klassifikationsverhalten während der Forschungsfahrt Sirena 11 eines an der Forschungsanstalt der Bundeswehr für Wasserschall und Geophysik vorhandenen Klassifikationssystems verglichen. Des Weiteren wird an einigen Aufnahmen der Forschungsfahrten Sirena10, Memphis und Sirena11 das Verhalten des vorgestellten Systems dargestellt.

Inhaltsverzeichnis

Abkürzungen	xi
Notation	xiii
1 Einleitung	1
1.1 Motivation	2
1.2 Projektbeschreibung	4
1.2.1 Arbeitspakete	4
1.3 Unterwasserschall	5
1.3.1 SONAR-Gleichungen	6
1.4 Zusammenfassung	12
2 Meeressäuger und ihre Laute	15
2.1 Definition Meeressäuger	15
2.2 Laute	16
2.2.1 Pulsartige Laute	16
2.2.2 Tonale Laute	21
2.2.3 Robben	23
2.2.4 Direktivität	24
2.3 Gehör	26
2.4 Einteilung	28
2.5 Zusammenfassung	29
3 Sprechererkennung und Klassifikation	31
3.1 Definition und Anwendung	31
3.2 Detektion	32
3.3 Merkmalsextraktion	33
3.4 Merkmalstransformation	33

3.5	Klassifikation	34
3.6	Zusammenfassung	35
4	Datenbank	37
4.1	Übersicht	37
4.2	Sensoren	38
4.2.1	Bojen und feste Verankerungen	38
4.2.2	Geschleppte Systeme	38
4.3	Datenquelle	40
4.3.1	Forschungsfahrten	40
4.3.2	Projektpartner	41
4.3.3	Andere Länder (PoMM request for data)	41
4.4	Anforderungen	43
4.4.1	Überlappende Laute	43
4.4.2	Internationale Zusammenarbeit	44
4.5	Ereignisbasierte Struktur	44
4.6	Labelling-Prozess	46
4.7	Inhalt	48
4.7.1	Test- und Trainings-Datensatz	49
4.7.2	Gruppen	49
4.7.3	Laute	50
4.7.4	Qualität der gelabelten Daten	51
4.8	Zusammenfassung	52
5	Rahmenwerk	53
5.1	Übersicht	53
5.2	Datenbankanbindung	53
5.2.1	Initialisierung	53
5.2.2	Funktionen zur Datenbankabfrage	54
5.2.3	Funktionen zur Schlüsselkonvertierung	55
5.3	Zusammenfassung	57
6	Graphische Benutzeroberfläche	59
6.1	Initialisierung	59
6.2	Konfiguration	60
6.3	Modifikation der Konfiguration	60

6.4	Bedienung	61
6.4.1	Umstrukturierung der Daten	61
6.4.2	Merkmalsbestimmung	61
6.4.3	Merkmalstransformation	61
6.4.4	Training/Validierung	62
6.5	Visualisierung/Editierung	62
6.5.1	Datenbankübersicht	62
6.5.2	Darstellung des Klassifikationsergebnisses	62
6.6	Zusammenfassung	64
7	Vorverarbeitung	65
7.1	Grundlegende Vorverarbeitung	65
7.2	Abstratanpassung	66
7.3	Rahmenbildung	66
7.4	Detektion	67
7.4.1	Initiale Geräuschschätzung	68
7.4.2	SNR-Berechnung	68
7.4.3	Entscheidung	69
7.4.4	Aktualisierung der Geräuschschätzung	69
7.5	Geräuschunterdrückung	70
7.5.1	Spektrale Subtraktion	70
7.5.2	Wiener-Filter	70
7.6	Zusammenfassung	71
8	Merkmalsextraktion	73
8.1	Zeitbereichsmerkmale	73
8.2	Frequenzbereichsmerkmale	75
8.3	Cepstraldomänenmerkmale	77
8.4	Detektionsbasierte Merkmale	78
8.4.1	Klicklaute	79
8.4.2	Tonale Laute	83
8.5	Merkmalsselektion / Merkmalstransformation	86
8.6	Zusammenfassung	88
9	Klassifikationsansatz	89
9.1	Klassifikationsmethoden	89

9.1.1	Klassifikationsbaum	89
9.1.2	Gaußsche Mischverteilungsmodelle	90
9.1.3	Nächste-Nachbarn-Klassifikation	92
9.2	Aufteilung der Klassifikation	93
9.3	Zusammenfassung	93
10	Ergebnisse	95
10.1	Merkmalsselektion	95
10.2	Gegenüberstellung der Klassifikationsmethoden	101
10.2.1	Klassifikationsbaum-Klassifikation	103
10.2.2	Nächste-Nachbarn-Klassifikation	104
10.2.3	Klassifikation mit Gaußschen Mischmodellen	105
10.2.4	Verarbeitungszeit	106
10.2.5	Auswahl	109
10.3	Nähere Betrachtung GMM	110
10.3.1	Einfluss der Parameteranzahl	110
10.3.2	Einfluss der Komponentenzahl	111
10.4	Auswirkungen der automatischen Detektion	112
10.5	Vergleich mit dem PASS-System	114
10.6	Klassifikationsverhalten mit Daten aus Forschungsfahrten	118
10.6.1	Memphis	118
10.6.2	Sirena 2010	119
10.6.3	Sirena 2011	121
10.7	Zusammenfassung	124
11	Zusammenfassung und Ausblick	125
11.1	Zusammenfassung	125
11.2	Ausblick	127
A	Zusätzliche Informationen zum Abschnitt 4 Datenbank	131
B	Zusätzliche Informationen zum Abschnitt 6 Graphische Benutzeroberfläche	137
	Literaturverzeichnis	139

Abkürzungen

ACAI	<i>Advanced Customizable Audio Interface</i> , Graphische Benutzeroberfläche für den Audio-Teil der PoMM Datenbank
AKF	Autokorrelationsfolge, Autokorrelationsfunktion
AM	Amplitudenmodulation
ASCH	<i>American Standard Code for Information Interchange</i> , Zeichenkodierungstabelle
CART	<i>Classification And Regression Tree</i> , Klassifikationsbaum
CIBRA	<i>Centro Interdisciplinare di Bioacustica e Ricerche Ambientali</i> , Interdisziplinäres Zentrum für Bioakustik der Universität Pavia,
CPAM	<i>Compact Passive Acoustic Monitoring</i> , Pass. Sonarsystem des NURC
DCL	<i>Detection Localization Classification</i> , Detektion Lokalisierung Klassifikation
DE	emphGermany, Deutschland
EDA	<i>European Defence Agency</i> , Europäische Verteidigungsagentur
FM	Frequenzmodulation
FN	<i>False Negative</i> , Anzahl der falsch negativen Klassifikationen
FP	<i>False Positive</i> , Anzahl der falsch positiven Klassifikationen
FWG	Forschungsanstalt der Bundeswehr für Wasserschall und Geophysik
GMM	<i>Gaussian Mixture Model</i> , Gaußsches Mischverteilungs Modell
GUI	<i>Graphical User Interface</i> , Graphische Benutzeroberfläche
ICI	<i>Inter Click interval</i> , Interklickintervall
INI	<i>Initialization</i> , Initialisierung
INR	<i>Input to Noise Ratio</i> , Eingangs-Rausch-Verhältnis
IT	<i>Italy</i> , Italien
LDA	<i>Lineare Diskriminanzanalyse</i>
LOFAR	<i>Low Frequency Analysis And Recording</i>
LPC	<i>Linear Predictive Coding</i>
N	Negativität, Anzahl der falschen Klassifikationen
NF	<i>Niederfrequenz</i> , Frequenzbereich
NO	<i>Norway</i> , Norwegen
NURC	<i>Nato Undersea Research Centre</i> , NATO Forschungszentrum La Spezia

P	Positivität, Anzahl der korrekten Klassifikationen
PAM	<i>Passive Acoustic Monitoring</i> , Passive Akustische Überwachung
PASS	<i>Passive Acoustic Software Sonar</i> , Programmbezeichnung der Firma Kaon
PALAOA	<i>PerenniAL Acoustic Observatory in the Antarctic Ocean</i> , Akustisches Observatorium im Südpolarmeer
PoMM	<i>Protection of Marine Mammals</i> , Projektbezeichnung
PDA	<i>Pitch Detection Algorithm</i> , Algorithmus zur Schätzung der Grundfrequenz
PDF	<i>Probability Density Function</i> , Wahrscheinlichkeitsdichtefunktion
RMS	<i>Root Mean Square</i> , Quadratisches Mittel
ROC	<i>Receiver Operating Characteristic</i> , Operationscharakteristik eines Empfängers/Klassifizierers
SNR	<i>Signal to Noise Ratio</i> , Signal-Rausch-Verhältnis
SONAR	<i>Sound Navigation and Ranging</i> , Schallortungsverfahren
SVP	<i>Sound Velocity Profile</i> , Schallgeschwindigkeitsprofil
TN	<i>True Negative</i> , Anzahl der richtig negativen Klassifikationen
TP	<i>True Positive</i> , Anzahl der richtig positiven Klassifikationen
UK	<i>United Kingdom</i> , Großbritannien
ULA	<i>Uniform Linear Array</i> , Gleichverteilte Linear Anordnung
WE	<i>Work Element</i> , Projektabschnitt
WP	<i>Work Package</i> , Arbeitspaket
WMS	<i>Web Mapping Service</i> , Netzbasierte Kartendarstellung

Notation

Konventionen und Operatoren

$s(t)$	zeitkontinuierliches Signal
$s(k)$	Folge von Abtastwerten, zeitdiskretes Signal
$\mathbf{x}$	Vektor mit Elementen $x_i, i \in \mathbb{N}$
$\mathbf{x}^T$	transponierter Vektor $\mathbf{x}$
$\mathbf{X}$	Matrize $\mathbf{X}$
$\arg\{\}$	Argument (Phase) eines komplexen Ausdrucks / Zahl
$\exp\{\}$	Exponentialfunktion
$\max\{\}$	Maximalwert
$\min\{\}$	Minimalwert
$\text{diag}\{\mathbf{X}\}$	Diagonalelemente der Matrix $\mathbf{X}$ (Gemäß Matlab-Notation)