



Digitale Signalverarbeitung und Systemtheorie  
Christian-Albrechts-Universität zu Kiel

---

## Arbeiten über digitale Signalverarbeitung und Systemtheorie

Herausgegeben von Gerhard Schmidt

# Instrumentelle Analyse von Parkinson-Sprache

Sarah Christin Baasch

# **Instrumentelle Analyse von Parkinson-Sprache**

## **Dissertation**

zur Erlangung des akademischen Grades  
Doktor der Ingenieurwissenschaften  
(Dr.-Ing.)  
der Technischen Fakultät  
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

**Sarah Christin Baasch**

Kiel 2019

Tag der Einreichung: 04.02.2019  
Tag der Disputation: 02.05.2019

Berichterstatter: Prof. Dr.-Ing. Gerhard Schmidt  
Prof. Dr.-Ing. Sebastian Möller

Arbeiten über digitale Signalverarbeitung und Systemtheorie

Band 7

**Sarah Christin Baasch**

**Instrumentelle Analyse von Parkinson-Sprache**

Shaker Verlag  
Düren 2019

**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., 2019

Copyright Shaker Verlag 2019

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

Printed in Germany.

ISBN 978-3-8440-6775-0

ISSN 2197-7089

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren

Telefon: 02421 / 99 0 11 - 0 • Telefax: 02421 / 99 0 11 - 9

Internet: [www.shaker.de](http://www.shaker.de) • E-Mail: [info@shaker.de](mailto:info@shaker.de)

# Vorwort

Meine Arbeit an dieser Dissertation begann im November 2014 am Lehrstuhl für Digitale Signalverarbeitung und Systemtheorie (DSS) im Rahmen eines Forschungsprojektes mit der Deutschen Forschungsgemeinschaft (DFG) an der Christian-Albrechts-Universität zu Kiel. Ziel dieses Projektes ist es Zusammenhänge zwischen Sprachverständlichkeit, Sprechtraktverhalten und Gehirnaktivität bei sprechbehinderten Parkinson-Patienten zu untersuchen. Obwohl mir die Arbeit an diesem Projekt viel Freude bereitet hat, bin ich froh, mit der vorliegenden Arbeit mein großes Ziel der Promotion erreicht zu haben. Dies konnte ich jedoch nicht alleine schaffen, viele Menschen haben zu diesem Erfolg beigetragen.

Zuerst möchte ich mich dafür bei meinem Doktorvater Prof. Dr.-Ing. Gerhard Schmidt bedanken, für seine fortwährende Unterstützung und Hilfsbereitschaft. Durch seine offene, herzliche Art hat er für ein sehr kollegiales Miteinander in der Arbeitsgruppe gesorgt.

Ebenfalls möchte ich mich bei Prof. Dr.-Ing. Sebastian Möller für die Erstellung des Gutachtens sowie bei den zusätzlichen Mitgliedern der Prüfungskommission Prof. Dr.-Ing. Dipl.-Wirt. Ing. Pachnicke und Prof. Dr.-Ing. McCord bedanken.

Der DFG danke ich für die Finanzierung meines Projektes HE 4465/6-1. Des Weiteren geht ein großer Dank an Prof. Dr.-Ing. Ulrich Heute, der mir während des gesamten Projektes mit seinem umfangreichen Fachwissen durch Rat und Tat zur Seite stand.

Alle aktuellen und ehemaligen Mitarbeiter des DSS-Lehrstuhls sowie alle Beteiligten am oben genannten Forschungsprojekt werde ich stets in besonderer Erinnerung behalten. Sie haben für mich ein angenehmes und freundschaftliches Arbeitsumfeld geschaffen. Vielen Dank, dass ich so eine tolle Zeit mit euch erleben durfte!

Ich danke Tim Owe Wisch, Alexandra Reermann und Tobias Hübschen für das Korrekturlesen meiner Arbeit. Eure hilfreichen Anmerkungen haben es mir ermöglicht, meine Arbeit aus einem anderen Blickwinkel zu sehen, der einem als Autor fehlt.

Zu guter Letzt möchte ich mich bei meiner Familie als und bei Kevin bedanken. Sie haben mich immer nach Kräften, in allen Lebenslagen unterstützt und an mich geglaubt, auch wenn ich es nicht getan habe. Meine Eltern haben alles getan, um mir das Studium und die anschließende Promotion zu ermöglichen.



# Abstract

Parkinson's Disease is one of the most frequent neurodegenerative diseases worldwide. Besides motor disorders, patients affected by this disease mostly suffer from a speech disorder named dysarthria.

It will be treated by a speech therapist with a speech therapy, its success as well as the progress of the dysarthria shall be documented. Therefore, a multitude of different methods are available to do so, but all of them have one thing in common: they are not fully automatic. There is always a subjective component, where a rater or another person influences the process.

This work presents a system, named SINAS, for fully automatically rating the dysarthria. The system contains two main components: a recording tool and an analysis tool.

The first one gives the possibility to the speech therapist to guide the patient easy and with visual aid by HTML pages through different speech tasks. Thereby the recordings will be robust in level and independent of the position of the microphone.

In the analysis tool acoustic measures are calculated from the recordings, which are intended to evaluate the three clusters of symptoms of dysarthria. These measures form the entry of a neural network, which gives an NTID rating as a result. The NTID scale rates the intelligibility of the recording and therefore the dysarthria of the patient in 6 steps.

The validation of the tool is done by comparison of the results with a survey, where people rated the recordings of Parkinson patients according to the NTID scale, the mean value for each recording is then taken as a reference. As cost functions for evaluating the developed system the correlation, the mean absolute error, as well as the variance of the error are taken, on the basis of these functions the system will be optimized.

For further evaluation and to take into account the uncertainty of the raters, the epsilon insensitive RMSE is used to evaluate the performance of the system. This clearly shows the possibility of a fully automatic NTID rating of the patients with the presented SINAS system.

The developed tool can now form the basis for many applications to support the speech therapy of Parkinson patients.



# Kurzfassung

Morbus Parkinson ist eine der am weitesten verbreiteten neurodegenerativen Krankheiten weltweit. Neben motorischen Symptomen wird diese Krankheit häufig von einer Sprechstörung, Dysarthrie genannt, begleitet. Diese wird von einem Logopäden/einer Logopädin meist mit einer Sprachtherapie behandelt und in diesem Zuge der Erfolg sowie der Verlauf der Störung dokumentiert. Dazu werden in einer Logopädiesitzung aufgenommene Sprachsignale beurteilt. Hierzu gibt es eine Vielzahl verschiedener Verfahren, die alle eines gemeinsam haben: Keines ist vollständig objektiv, da nicht vollständig automatisch und es gibt immer eine subjektive Komponente bei der ein Bewerter oder eine andere Person Einfluss auf das Verfahren nimmt.

In dieser Arbeit wird ein System, SINAS, zur vollständig automatischen Bewertung der Dysarthrie vorgestellt. Das System besteht aus zwei Hauptkomponenten: einem Aufnahme-Tool und einem Analyse-Tool.

Erstgenanntes ermöglicht es dem Logopäden, den Patienten einfach, visuell durch verschiedene Sprechaufgaben zu leiten und dabei unabhängig von der Mikrofonposition pegelrobuste Aufnahmen zu erstellen.

Im Analyse-Tool werden akustische Merkmale aus den Sprachaufnahmen berechnet, welche drei Merkmalscluster der Dysarthrie beurteilen. Diese Merkmale bilden den Eingang eines Mustererkennters, welcher eine NTID-Stufe ausgibt. Die NTID-Skala bewertet dabei in 6 festen Stufen die Verständlichkeit des Gesprochenen und damit die Schwere der Dysarthrie des Patienten.

Validiert wird das Tool über den Vergleich mit einer durchgeführten Umfrage in welcher zuvor Sprachdateien von Parkinson-Patienten mit einer NTID-Stufe bewertet wurden. Die mittlere Wertung wird als Referenz verwendet. Als Kostenfunktionen für die Auswertung des Systems wurden die Korrelation, der mittlere Betragsfehler sowie dessen Varianz verwendet. Auf dieser Basis wurde das System optimiert.

Um darüber hinaus die Unsicherheit der wertenden Personen in der Umfrage zu berücksichtigen, wurde der sogenannte *epsilon insensitive RMSE* zur Auswertung des Systems herangezogen. Dieses Maß zeigt deutlich, dass es mit SINAS möglich ist, auf Basis der entwickelten Merkmale eine voll-automatische NTID-Einstufung der Patienten vorzunehmen.

Das entwickelte Tool kann als Basis für eine Vielzahl von Anwendungen zur Unterstützung der Sprachtherapie von Parkinson-Patienten dienen.



# Inhaltsverzeichnis

Abbildungsverzeichnis	vii
Tabellenverzeichnis	xi
Abkürzungen und Notation	xv
Symbolverzeichnisse	xix
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Aufbau der Dissertation . . . . .	2
<b>2 Grundlagen</b>	<b>5</b>
2.1 Der Sprechvorgang . . . . .	5
2.1.1 Anatomisch und physiologisch . . . . .	6
2.1.2 Akustisch . . . . .	8
2.1.3 Phonetisch . . . . .	9
2.1.4 Prosodisch . . . . .	12
2.2 Morbus Parkinson . . . . .	12
2.2.1 Motorische Störungen . . . . .	13
2.2.2 Dysarthrie . . . . .	15
2.3 Praxis der klinischen Dysarthriediagnostik . . . . .	19
2.3.1 Evaluierete Testverfahren . . . . .	19
2.3.2 Nichtevaluierete und nichtstandardisierte Untersuchungsverfahren . . . . .	23
2.3.3 Bewertungsskalen . . . . .	24
2.3.4 Selbsteinschätzung . . . . .	26
2.3.5 Akustische Beurteilung des Sprechers . . . . .	26
2.3.6 Apparative Untersuchungen . . . . .	27

2.3.7	Abgrenzung zur vorliegenden Arbeit . . . . .	28
<b>3</b>	<b>SINAS (Klinisches System zur instrumentellen Analyse von Sprechstörungen)</b>	<b>29</b>
3.1	Konzeptionierung . . . . .	30
3.1.1	Zielsetzung . . . . .	30
3.1.2	Technisches Konzept . . . . .	30
3.2	NTID-Umfrage . . . . .	37
3.3	Umsetzung der Sprecherlokalisierung . . . . .	40
3.3.1	Pegel-Algorithmus . . . . .	41
3.3.2	Kleusberg-Algorithmus . . . . .	42
3.3.3	Pegel-Korrektur . . . . .	44
3.4	Umsetzung der Merkmalsextraktion . . . . .	46
3.4.1	Untersuchung der prosodischen Auffälligkeiten . . . . .	46
3.4.2	Untersuchung der Artikulationsunschärfe . . . . .	51
3.4.3	Bewertung von Sprechtempo und Redefluss . . . . .	71
3.4.4	Auswertung der Merkmale . . . . .	75
3.5	Umsetzung des Mustererkenners . . . . .	76
3.5.1	GMM-Training . . . . .	76
3.5.2	Training des neuronalen Netzes . . . . .	80
3.6	Optionale Zusatzmodule . . . . .	83
3.6.1	Haltedauer und Pegelreduktion . . . . .	85
3.6.2	Jitter und Shimmer . . . . .	86
3.6.3	Silbenrate . . . . .	90
<b>4</b>	<b>Evaluierung und Optimierung des Gesamtsystems</b>	<b>93</b>
4.1	Optimierung des Mustererkenners . . . . .	94
4.1.1	Optimierung der Struktur . . . . .	94
4.1.2	Optimierung der Aktivierungsfunktion . . . . .	97
4.1.3	Optimierung der Regularisierung . . . . .	98
4.1.4	Auswahl des Trainingsalgorithmus . . . . .	101
4.2	Optimierung der Merkmale . . . . .	102
4.2.1	Optimierung der Parameteranzahl . . . . .	102
4.2.2	Optimierung der Parameterauswahl . . . . .	104
4.3	Vergleich mit Hörerwertungen . . . . .	107

---

<b>5</b>	<b>Schlussbetrachtung</b>	<b>111</b>
5.1	Zusammenfassung . . . . .	111
5.2	Fazit und Ausblick . . . . .	113
5.2.1	Weiterer Untersuchungsbedarf . . . . .	113
5.2.2	Mögliche Anwendungsbereiche . . . . .	114
<b>A</b>	<b>Signalverarbeitungsalgorithmen</b>	<b>117</b>
A.1	LPC-Analyse . . . . .	117
A.2	Levinson-Durbin-Algorithmus . . . . .	119
A.3	K-means-Algorithmus . . . . .	120
A.4	LoOP-Algorithmus . . . . .	122
A.5	DBSCAN-Methode . . . . .	123
A.6	LBG-Algorithmus . . . . .	124
A.7	Methode der kleinsten Quadrate . . . . .	126
A.8	EM-Algorithmus . . . . .	127
A.9	Lineare Diskriminanzanalyse . . . . .	129
A.10	Levenberg-Marquardt-Algorithmus . . . . .	131
A.11	Resilient-Backpropagation-Algorithmus . . . . .	133
A.12	Konjugierte Gradienten-Backpropagation mit Fletcher-Reeves-Updates . . . . .	135
<b>B</b>	<b>Ergebnisse der durchgeführten Simulationen</b>	<b>137</b>
B.1	Ergebnis der Korrelationsauswertung der Einzelmerkmale . . . . .	137
B.2	Ergebnisse der Strukturoptimierung des neuronalen Netzes . . . . .	144
B.3	Ergebnisse der Optimierung der Aktivierungsfunktion des neuronalen Netzes	155
B.4	Ergebnisse der Optimierung der Regularisierung des neuronalen Netzes . . . . .	162
	<b>Literaturverzeichnis</b>	<b>173</b>



# Abbildungsverzeichnis

1.1	Alterspyramide 1950, 2015 und 2100 [Min16]. . . . .	1
2.1	Vokaltrakt mit den beim Sprechen beteiligten Strukturen (in Anlehnung an [ND08]). . . . .	6
2.2	Akustisches Modell des Vokaltraktes mit Anregungsquelle und Vokaltraktfilter (in Anlehnung an [VHH98] und [VM06]). . . . .	8
2.3	Klassisches Röhrenmodell des Vokaltraktes. . . . .	9
2.4	Vokalviereck der Kardinalvokale mit Formantbereichen für die deutsche Sprache (in Anlehnung an [SSB13] und [Mac96]). . . . .	11
2.5	Typische Körperhaltung eines Parkinson-Patienten [Gow86]. . . . .	14
3.1	Strukturdiagramm des Gesamtsystems. . . . .	31
3.2	Strukturdiagramm des Aufnahme-Tools. . . . .	32
3.3	Ansicht für den behandelnden Mediziner im entwickelten Aufnahme-Tool. . . . .	33
3.4	Patienten-Ansicht des entwickelten Aufnahme-Tools. . . . .	33
3.5	Aufnahmeumgebung des Aufnahme-Tools. . . . .	35
3.6	Strukturdiagramm des Analyse-Tools. . . . .	36
3.7	Patientenblatt als Ausgabe des SINAS-Systems. . . . .	37
3.8	Präsentation der Umfrage für den Bewerter. . . . .	39
3.9	Frequenzgang und Gruppenlaufzeit des Bandpassfilters (in Anlehnung an [Jas15]) . . . . .	40
3.10	Im Pegel-Algorithmus verwendete Komponenten. Der Zeitindex wird aus Gründen der Übersichtlichkeit vernachlässigt (in Anlehnung an [JBS <sup>+</sup> 16]). . . . .	42
3.11	Im Kleusberg-Algorithmus verwendete Komponenten. Der Zeitindex wird aus Gründen der Übersichtlichkeit vernachlässigt (in Anlehnung an [JBS <sup>+</sup> 16]). . . . .	43
3.12	Testergebnisse des kombinierten Lokalisierungsalgorithmus für einen Soll-Abstand $d_{\text{soll}}$ von 25 cm (oben) und 50 cm (unten) und Normierung auf einen Refrenzabstand von $d_{\text{ref}} = 50$ cm (in Anlehnung an [JBS <sup>+</sup> 16]). . . . .	45
3.13	Frequenzantwort eines A-Gewichtungsfilters. . . . .	47
3.14	Ergebnis einer Sprachpegelanalyse eines Parkinson-Patienten. . . . .	48
3.15	Ergebnis einer Sprachgrundfrequenzanalyse eines Parkinson-Patienten. . . . .	50
3.16	Ergebnis der Analyse der spektralen Flachheit eines Parkinson-Patienten. . . . .	51

3.17	Kurzzeitspektrum (blau) und spektrale Einhüllende (orange) zur Bestimmung der Formantfrequenzen $F_1 = 592$ Hz und $F_2 = 1314$ Hz. . . . .	53
3.18	Spektrogramm eines Sprachsignals mit den geschätzten ersten beiden Formanten. . . . .	54
3.19	Klassisches Formantdreieck eines Patienten vor und nach einer Sprachtherapie. 54	
3.20	Ergebnis eines Codebuch-Trainings für Tonhalteaufgaben der 5 Kernvokale. 55	
3.21	Ergebnis der Analyse zur aufgespannten Codebuchfläche. Die Verschiebung der Codebuch-Vektoren zum Mittelpunkt ist deutlich zu erkennen. . . . .	56
3.22	Ergebnis einer Analyse von Formant-Daten mit dem LoOP-Algorithmus (in Anlehnung an [Pri16]). Ausreißer sind in rot eingekreist. . . . .	57
3.23	Klassifikation mit der <i>density based scan</i> Methode. Kernpunkte sind in grau, Randpunkte in blau und Rauschpunkte in orange gekennzeichnet. Es wurde eine minimale Anzahl von 4 Datenpunkten in der k-Nachbarschaft gewählt. Die schwarzen Kreise zeigen die k-Nachbarschaft an (in Anlehnung an [Pri16]). 58	
3.24	Ergebnis der Auswertung des Merkmals der Einhüllenden aller Formant-Paare. 59	
3.25	Graphische Veranschaulichung der Distanz der Codebuch-Vektoren. . . . .	60
3.26	Graphische Darstellung der mittleren Distanz zu den Codebuch-Vektoren. . . 61	
3.27	Mel-Filterbank mit 10 Dreiecksfiltern. . . . .	62
3.28	Strukturdiagramm der Mel-gefilterte Cepstral-Koeffizienten (engl. <i>Mel Frequency Cepstral Coefficients</i> , MFCC)-Berechnung. . . . .	63
3.29	Darstellung der Varianzen innerhalb der Cluster und über alle MFCCs. . . 64	
3.30	Darstellung des berechneten Flächenmaßes. . . . .	64
3.31	Ergebnis der Sprachaktivitätsdetektion (engl. <i>Voice Activity Detection</i> , VAD). 66	
3.32	Eingefrorenes Bild des Videos der Vokaltraktanalyse. . . . .	67
3.33	Algorithmus zur Berechnung des Modulationsspektrums. . . . .	73
3.34	Oktav-Filterbank mit Mittenfrequenzen zwischen 125 Hz und 4 kHz realisiert durch Butterworth-Bandpassfilter. . . . .	73
3.35	Darstellung des Modulationsspektrums für jedes der 6 Teilbänder sowie die Mittelung über alle Teilbänder. . . . .	74
3.36	Darstellung des Flächenmaßes zur Auswertung des Modulationsspektrums. 75	
3.37	Feed-Forward-Struktur eines neuronalen Netzes. . . . .	82
3.38	Ergebnisse der Korrelationsauswertung des neuronalen Netzes. . . . .	84
3.39	Einhüllende und Regressionsgerade für den gehaltenen Vokal /a:/. . . . .	86
3.40	Beispiel des Nachteils der Jitter-Analyse mit Maximumdetektion bei einem Parkinson-Patienten. . . . .	87
3.41	Ergebnis der korrelationsbasierten Jitter-Analyse eines Parkinson-Patienten. 90	
3.42	Tiefpass mit Grenzfrequenz von 6 Hz, zur Berechnung der Einhüllenden. . . 91	
3.43	Tiefpassgefiltertes Signal mit Markierungen für jede gesprochene Silbe. . . 91	
4.1	Aktivierungsfunktionen der Neuronen des neuronalen Netzes. . . . .	98
4.2	Simulationsergebnisse zur Optimierung der Merkmalsanzahl. . . . .	103

---

4.3	Histogramm der absoluten Häufigkeiten der <i>National Technical Institute for the Deaf</i> (NTID)-Wertungen. . . . .	107
4.4	Mittlere Betragsabweichung der NTID-Wertung vom Mittelwert sowie dessen Varianz. . . . .	108
A.1	Struktur eines Prädiktor-Fehler-Filters. . . . .	118



# Tabellenverzeichnis

2.1	Artikulationsorte nach Bildungsstellen unterteilt [SSB13]. . . . .	10
2.2	Beschreibung und Bezeichnung von Konsonanten [SSB13]. . . . .	10
2.3	Beispiel für den Aufbau eines Ensembles im MVP [ND08]. . . . .	23
2.4	Unified Parkinson's Disease Rating Scale - Skala zur Beurteilung der Sprache [Goe08]. . . . .	25
2.5	Ursprüngliche NTID-Skala [KHZ06]. . . . .	25
2.6	Modifizierte NTID-Skala nach Ramig [Ray92]. . . . .	25
3.1	Anzahl der Datensätze zum Training und zur Evaluierung des Mustererkenners mit ganzen NTID-Stufen. . . . .	77
3.2	Ergebnisse der GMM-Auswertung für ganze NTID-Stufen und unterschiedliche Parameteranzahlen. . . . .	78
3.3	Anzahl der Datensätze zum Training und zur Evaluierung des Mustererkenners mit halben NTID-Stufen. . . . .	79
3.4	Ergebnisse der GMM-Auswertung für halbe NTID-Stufen und unterschiedliche Parameteranzahlen. . . . .	79
3.5	Ergebnisse der GMM-Auswertung für halbe NTID-Stufen und LDA. . . . .	80
4.1	Übersicht über die verwendeten Netzwerkstrukturen in den versteckten Schichten. . . . .	94
4.2	Übersicht über die verwendeten Trainingsalgorithmen [Sar14]. . . . .	96
4.3	Ergebnisse der Simulation zur Optimierung der Netzstruktur. . . . .	97
4.4	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion. . . . .	99
4.5	Ergebnisse der Simulation zur Optimierung der Regularisierung. . . . .	100
4.6	Ergebnisse zur Auswahl des Trainingsalgorithmus. . . . .	101
4.7	Übersicht über die verwendeten Merkmale nach Optimierung der Parameteranzahl. . . . .	105
4.8	Übersicht über die verwendeten Merkmale nach Optimierung der Parameterauswahl. . . . .	106
B.1	Ergebnisse der Korrelationsauswertung der Einzelmerkmale. . . . .	143

B.2	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainbfg. . . . .	144
B.3	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus traingdx. . . . .	145
B.4	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainr. . . . .	146
B.5	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainb. . . . .	147
B.6	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainlm. . . . .	148
B.7	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainbr. . . . .	149
B.8	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus traingcf. . . . .	150
B.9	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus traingcb. . . . .	151
B.10	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus traingcp. . . . .	152
B.11	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus traingcg. . . . .	153
B.12	Ergebnisse der Simulation zur Strukturoptimierung für den Trainingsalgorithmus trainrp. . . . .	154
B.13	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainbfg. . . . .	155
B.14	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus traingdx. . . . .	155
B.15	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainr. . . . .	156
B.16	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainr mit angepassten Parametern. . . . .	156
B.17	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainb. . . . .	157
B.18	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainlm. . . . .	157
B.19	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus trainbr. . . . .	158
B.20	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus traingcf. . . . .	158
B.21	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus traingcb. . . . .	159

---

B.22	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus <code>traincgb</code> mit angepassten Parametern. . . . .	159
B.23	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus <code>traincgp</code> . . . . .	160
B.24	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus <code>trainscg</code> . . . . .	160
B.25	Ergebnisse der Simulation zur Optimierung der Aktivierungsfunktion für den Trainingsalgorithmus <code>trainrp</code> . . . . .	161
B.26	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainbfg</code> . . . . .	162
B.27	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>traingdx</code> . . . . .	163
B.28	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainr</code> mit angepassten Trainingsparametern. . . . .	164
B.29	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainb</code> . . . . .	165
B.30	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainlm</code> . . . . .	166
B.31	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>traincgf</code> . . . . .	167
B.32	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>traincgb</code> mit angepassten Trainingsparametern. . . . .	168
B.33	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>traincgp</code> . . . . .	169
B.34	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainscg</code> . . . . .	170
B.35	Ergebnisse der Simulation zur Optimierung der Regularisierung für den Trainingsalgorithmus <code>trainrp</code> . . . . .	171



# Abkürzungen

<b>AKF</b>	Autokorrelationsfunktion
<b>AMDNS</b>	Aachener Materialien zur Diagnostik neurogener Sprechstörungen
<b>ANSI</b>	<i>American National Standards Institute</i>
<b>BODYS</b>	Bogenhausener Dysarthrieskalen
<b>DBSCAN</b>	<i>Density Based Spatial Clustering of Applications with Noise</i>
<b>DCT</b>	Diskrete Kosinustransformation (engl. <i>discrete cosine transform</i> )
<b>DSS</b>	Digitale Signalverarbeitung und Systemtheorie
<b>EEG</b>	Elektroenzephalografie
<b>EM</b>	<i>Expectation-Maximization</i>
<b>EMG</b>	Elektromyografie
<b>EPG</b>	Elektropalatografie
<b>FFT</b>	schnelle Fourier-Transformation (engl. <i>Fast Fourier Transform</i> )
<b>GMM</b>	Gauß'sche Mischmodelle
<b>HTML</b>	<i>Hypertext Markup Language</i>
<b>IIR</b>	<i>Infinite Impulse Response</i>
<b>ITU</b>	<i>International Telecommunication Union</i>
<b>LBG</b>	<i>Linde, Buzo und Gray</i>
<b>LDA</b>	Lineare Diskriminanzanalyse (engl. <i>Linear Discriminant Analysis</i> )
<b>LDS</b>	Leistungsdichtespektrum (engl. <i>Power Spectral Density</i> )
<b>LOF</b>	lokaler Ausreißer-Faktor (engl. <i>Local Outlier Factor</i> )
<b>LoOP</b>	Lokale-Ausreißerwahrscheinlichkeiten (engl. <i>Local Outlier Probabilities</i> )
<b>LPC</b>	lineare Prädiktion (engl. <i>Linear Predictive Coding</i> )
<b>LSVT</b>	<i>Lee Silverman Voice Treatment</i>
<b>MDS</b>	<i>Movement Disorder Society</i>
<b>MFCC</b>	Mel-gefilterte Cepstral-Koeffizienten (engl. <i>Mel Frequency Cepstral Coefficients</i> )
<b>MODIAS</b>	„Modulares Diagnostik-System für Sprechstörungen“ (engl. <i>Modular Diagnostic System for Speech Disorders</i> )
<b>MVP</b>	Münchener Verständlichkeitsprofil
<b>NST</b>	<i>Nucleus subthalamicus</i>
<b>NTID</b>	<i>National Technical Institute for the Deaf</i>

<b>PLOF</b>	<i>Probabilistic Local Outlier Factor</i>
<b>RMSE</b>	Wurzel aus dem mittleren quadratischen Fehler (engl. <i>Root Mean Square Error</i> )
<b>SINAS</b>	klinisches System zur instrumentellen Analyse von Sprechstörungen
<b>SNR</b>	Signal-Rauschabstand (engl. <i>Signal-to-Noise Ratio</i> )
<b>TDoA</b>	<i>Time Difference of Arrival</i>
<b>UKSH</b>	Universitätsklinikum Schleswig-Holstein
<b>UNO</b>	Vereinte Nationen
<b>UNS</b>	Untersuchungsbogen neurologisch bedingter Sprech- und Stimmstörungen
<b>UPDRS</b>	<i>Unified Parkinson's Disease Rating Scale</i>
<b>VAD</b>	Sprachaktivitätsdetektion (engl. <i>Voice Activity Detection</i> )
<b>VHI</b>	<i>Voice Handicap Index</i>

# Notation

$x(n)$	Folge von Abtastwerten
$x(k, \eta)$	Endliches, gefensterter Segment des Signals $x(n)$ im $k$ -ten Rahmen
$X(k, \mu)$	Diskrete Fourier-Transformation des (endlichen) Signals $x(k, \eta)$
$x$	Skalar
$\mathbf{x}$	Vektor
$\mathbf{x}^T$	Transponierter Vektor
$\mathbf{X}$	Matrix
$\mathbf{X}^T$	Transponierte Matrix
$\mathbf{X}^{-1}$	Inverse Matrix
$\mathbb{X}$	Menge aller Datenpunkte $\mathbf{x}_i$
$m_x$	Mittelwert der Größe $x$
$\hat{x}$	Schätzwert bzw. Erwartungswert der Größe $x$
$\bar{x}$	Geglättetes Signal $x$
$\phi_{xx}(\kappa)$	Autokorrelationsfolge des Signals $x(n)$
$S_{xx}(\mu)$	Autoleistungsdichtespektrum des Signals $x(n)$
$f_s$	Abtastfrequenz
$h(\cdot)$	Impulsantwort
$H(\cdot)$	Übertragungsfunktion
$E\{\cdot\}$	Operation zur Berechnung des Erwartungswertes
FFT $\{\cdot\}$	Zeitdiskrete schnelle Fourier-Transformation
$\ \cdot\ _p$	$p$ -Vektornorm (Euklidische Norm für $p = 2$ )
$\max\{\cdot\}$	Maximalwertoperator
$\min\{\cdot\}$	Minimalwertoperator
$\operatorname{argmax}_{\kappa}\{\cdot\}$	Operation zur Bestimmung des Arguments des Maximums
$\operatorname{argmin}_{\kappa}\{\cdot\}$	Operation zur Bestimmung des Arguments des Minimums
$\sigma_x^2, \sigma_x$	Varianz/Standardabweichung der Größe $x$



# Symbolverzeichnis

## Verzeichnis lateinischer Symbole

$a_i$	Polynomparameter
$\mathbf{a}_i^\psi$	Neuronenausgangsvektor an der Schicht $\psi$ für Eingangsvektor $\mathbf{x}_i$
$AIC$	Akaikes Informationskriterium
$A(k, p)$	Röhrenquerschnitt an der Stelle $p$ im Rahmen $k$
$\mathbf{A}^\lambda$	Hesse-Matrix zu Iterationsschritt $\lambda$
$b_i$	Abstand in Metern vom Referenzmikrofon zum Mikrofon $i$
$b_q^\psi$	Bias an Neuron $q$ in der Schicht $\psi$
$\mathbf{b}^\psi$	Biasvektor in der Schicht $\psi$
$\hat{B}_\vartheta$	Maximalwert der Amplitude in der Pitchperiode $\vartheta$
$c$	Grenzwert des <i>Center Clipping</i> Verfahrens
$c_0$	Schallgeschwindigkeit in Luft
$c_i$	Filterkoeffizienten des Sprechtraktfiltermodells
$\mathbf{c}_i$	Codebuchvektor zu Cluster $i$
$ci_{95}(i)$	95% Konfidenzintervall für Datenpunkt $i$
$C$	Kalibrierungskonstante zur A-Gewichtung
$d$	Freiheitsgrade beim Mapping
$d_i(k)$	Distanz in Metern vom Mikrofon $i$ zum Referenzmikrofon, im Rahmen $k$
$d_{i,j}$	Euklid'sche Distanz vom Codebuchvektor $i$ zu Codebuchvektor $j$
$d_{\min}$	Minimal zu erreichende Distanz beim k-means Algorithmus
$d_{\text{prob}}(\Gamma, \mathbf{x}_i, \mathbb{S}_{\mathbf{x}_i})$	Probabilistische Distanz
$d_{\text{ref}}$	Festgelegter Sollabstand von der Quelle zum Referenzmikrofon
$d_{\text{soll}}$	Tatsächlicher Abstand von der Quelle zum Referenzmikrofon
$d(\mathbf{x}_i, \mathbf{c}_m)$	Euklid'sche Distanz vom Trainingsvektor $\mathbf{x}_i$ zum Codebuchvektor $\mathbf{c}_m$

$d_{\Delta,i}(k)$	Gangunterschied von Mikrofon $i$ zum Referenzmikrofon, im Rahmen $k$
$\tilde{d}(\lambda)$	Aktueller Wert der Kostenfunktion des k-means Algorithmus bei Iteration $\lambda$
$e(n)$	Fehlerfunktion
$\mathbf{e}_i$	Einheitsvektor vom Referenzmikrofon zum Mikrofon $i$ , Fehlervektor beim Training des neuronalen Netzes
$e_{\text{NTID}}$	Mittlerer Betragsfehler der NTID Schätzung
$\mathbf{e}_Q$	Einheitsvektor vom Referenzmikrofon zur Schallquelle $Q$
$E_{\min}^{(\lambda)}$	Minimale Fehlerleistung im Iterationsschritt $\lambda$
$E(\boldsymbol{\xi})$	Fehlerfunktion für das Training des neuronalen Netzes
$f_s$	Abtastfrequenz
$f_0$	Sprachgrundfrequenz
$f(k, p)$	Reflexionskoeffizienten
$F_1$	Erster Formant
$F_2$	Zweiter Formant
$F(\boldsymbol{\xi})$	Kostenfunktion für das Training des neuronalen Netzes
$g_{A,\text{avg}}$	Mittlerer Sprachpegel in dBA
$g_{A,\text{var}}$	Varianz des A-gewichteten Sprachpegels
$g_{\text{bil}}$	Spektrumbilanz
$g_{\text{FCB},\text{dm}1}$	Mittlere Distanz der Datenpunkte zum zugeordneten Codebuchvektor für den 1. Formanten
$g_{\text{FCB},\text{dm}2}$	Mittlere Distanz der Datenpunkte zum zugeordneten Codebuchvektor für den 2. Formanten
$g_{\text{FCB},\text{fl}}$	Aufgespannte Fläche der Codebuchvektoren
$g_{\text{FCB},\text{F}_1\text{v}}$	Varianz des 1. Formanten über alle Datenpunkte
$g_{\text{FCB},\text{F}_2\text{v}}$	Varianz des 2. Formanten über alle Datenpunkte
$g_{\text{FCB},\text{fn}}$	Differenz aus aufgespannter Fläche und Einhüllender
$g_{\text{FCB},\text{nv}}$	Fläche der Einhüllenden aller Formant-Paare
$g_{\text{FCB},\text{md}}$	Mittlere Distanz der Datenpunkte zum zugeordneten Codebuchvektor
$g_{\text{FCB},\text{wd}}$	Gewichtete Distanz der Codebuchvektoren zueinander
$g_{\text{t},\text{F}_1\text{v}}$	Zeitliche Varianz des 1. Formanten
$g_{\text{t},\text{F}_2\text{v}}$	Zeitliche Varianz des 2. Formanten
$g_{\text{t},\text{mv}}$	Mittlere zeitliche Varianz der Formanten
$g_i$	Merkmalsausprägungen
$\mathbf{g}_i$	Merkmalsvektor am Eingang der Mustererkennung

$g_{F_t, F_{1s}}$	Standardabweichung des 1. Formanten
$g_{F_t, F_{2s}}$	Standardabweichung des 2. Formanten
$g_{F_t, ms}$	Standardabweichung der Formanten
$\mathbf{g}_i$	Merkmalsvektor am Eingang der Mustererkennung
$g_m$	Gewicht der Gauß-Glocke $m$
$g_{M,A}$	Flächenmaß zur Modulationsanalyse
$g_{MFCC,dc}$	Mittlere Euklid'sche Distanz der Codebuchvektoren zueinander
$g_{MFCC,f}$	Alternatives Flächenmaß der MFCC-Analyse
$g_{MFCC,md}$	Mittlere Euklid'sche Distanz der Datenpunkte zum zugeordneten Codebuchvektor
$g_{P,avg}$	Mittlere Sprachgrundfrequenz
$g_{P,var}$	Varianz der Sprachgrundfrequenz
$g_{P,f0/f1}$	Leistungsverhältnis des Signals an der Grundfrequenz zur 1. Harmonischen
$g_{P,f0/f2}$	Leistungsverhältnis des Signals an der Grundfrequenz zur 2. Harmonischen
$g_{P,f1/f2}$	Leistungsverhältnis des Signals an der 1. Harmonischen zur 2. Harmonischen
$\tilde{g}_{SF}(k)$	Spektrale Flachheit des Signals im Rahmen $k$
$g_{SF}$	Mittlere Spektrale Flachheit des Signals
$g_{ST,pr}$	Anteil Pausenzeit vom Gesamtsignal
$g_{ST,sp}$	Verhältnis Sprachzeit zu Pausenzeit
$g_{ST,sr}$	Anteil Sprachzeit vom Gesamtsignal
$g_{T,dg}$	Differenz zwischen minimalem und maximalem Röhrenquerschnitt an der Glottis
$g_{T,dgmax}$	Differenz von maximaler und durchschnittlicher Querschnittsfläche an der Glottis
$g_{T,dgmin}$	Differenz von minimaler und durchschnittlicher Querschnittsfläche an der Glottis
$g_{T,dl}$	Differenz zwischen minimalem und maximalem Röhrenquerschnitt an den Lippen
$g_{T,dlmax}$	Differenz von maximaler und durchschnittlicher Querschnittsfläche an den Lippen
$g_{T,dlmax,gmin}$	Differenz zwischen minimalem Querschnitt an der Glottis und maximalem Querschnitt an den Lippen
$g_{T,dlmin}$	Differenz von minimaler und durchschnittlicher Querschnittsfläche an den Lippen
$g_{T,dlmin,gmax}$	Differenz zwischen maximalem Querschnitt an der Glottis und minimalem Querschnitt an den Lippen

---

$g_{T,dmax}$	Differenz zwischen maximalen Röhrenquerschnitten an den Lippen und der Glottis
$g_{T,dmin}$	Differenz zwischen minimalen Röhrenquerschnitten an den Lippen und der Glottis
$g_{T,dmk,rk}$	Differenz der mittleren Größen von mittlerer und hinterer Kavität
$g_{T,dvk,mk}$	Differenz der mittleren Größen von vorderer und mittlerer Kavität
$g_{T,gmax}$	Maximaler Röhrenquerschnitt an der Glottis
$g_{T,gmean}$	Durchschnittliche Querschnittfläche an der Glottis
$g_{T,gmin}$	Minimaler Röhrenquerschnitt an der Glottis
$g_{T,lmax}$	Maximaler Röhrenquerschnitt
$g_{T,lmean}$	Durchschnittlicher Querschnitt an den Lippen
$g_{T,lmin}$	Minimaler Röhrenquerschnitt an den Lippen
$g_{T,mf}$	Mittlere Größe der gemittelten vorderen Röhrenquerschnitte
$g_{T,mfk}$	Mittlere Größe der vorderen Kavität
$g_{T,mgk}$	Mittlere Größe der Gesamtkavität
$g_{T,mm}$	Mittlere Größe der gemittelten mittleren Röhrenquerschnitte
$g_{T,mmax}$	Mittlere Position des maximalen Röhrenquerschnittes
$g_{T,mmk}$	Mittlere Größe der mittleren Kavität
$g_{T,mr}$	Mittlere Größe der gemittelten hinteren Röhrenquerschnitte
$g_{T,mrk}$	Mittlere Größe der hinteren Kavität
$g_{T,pt}$	Varianz der Position der maximalen Querschnittsfläche
$g_{T,rg}$	Verhältnis von minimalem zu maximalem Röhrenquerschnitt an der Glottis
$g_{T,rgmax}$	Verhältnis von maximaler zu durchschnittlicher Querschnittsfläche an der Glottis
$g_{T,rgmin}$	Verhältnis von minimaler zu durchschnittlicher Querschnittsfläche an der Glottis
$g_{T,rl}$	Verhältnis von minimalem zu maximalem Röhrenquerschnitt an den Lippen
$g_{T,rlmax}$	Verhältnis von maximaler zu durchschnittlicher Querschnittsfläche an den Lippen
$g_{T,rlmax,gmin}$	Verhältnis von minimalem Querschnitt an der Glottis zu maximalem Querschnitt an den Lippen
$g_{T,rlmin}$	Verhältnis von minimaler zu durchschnittlicher Querschnittsfläche an den Lippen
$g_{T,rlmin,gmax}$	Verhältnis von maximalem Querschnitt an der Glottis zu minimalem Querschnitt an den Lippen
$g_{T,rmax}$	Verhältnis von maximalem Röhrenquerschnitt an den Lippen zu dem an der Glottis

$g_{T,\text{rmin}}$	Verhältnis von minimalem Röhrenquerschnitt an den Lippen zu dem an der Glottis
$g_{T,\text{rm,r}}$	Differenz der mittleren Größen von gemittelten mittleren und hinteren Röhrenquerschnitten
$g_{T,\text{rv,m}}$	Differenz der mittleren Größen von gemittelten vorderen und mittleren Röhrenquerschnitten
$g_{T,\text{vdmax}}$	Mittlere Änderung des Betrags der maximalen Querschnittsfläche
$g_{T,\text{vf}}$	Varianz der gemittelten vorderen Röhrenquerschnitte
$g_{T,\text{vfk}}$	Varianz der vorderen Kavität
$g_{T,\text{vi}}$	Varianz des Röhrenquerschnittes an Position $i$
$g_{T,\text{vm}}$	Varianz der gemittelten mittleren Röhrenquerschnitte
$g_{T,\text{vmax}}$	Mittlere Positionsverschiebung des Röhrensegmentes mit maximaler Querschnittsfläche
$g_{T,\text{vmk}}$	Varianz der mittleren Kavität
$g_{T,\text{vr}}$	Varianz der gemittelten hinteren Röhrenquerschnitte
$g_{T,\text{vrk}}$	Varianz der hinteren Kavität
$g_{T,\text{vs}}$	Varianz der Röhrenquerschnitte innerhalb eines Vokaltraktmodells gemittelt über das gesamte Signal für alle Röhrenabschnitte
$g_{T,\text{vs,f}}$	Varianz der Röhrenquerschnitte innerhalb eines Vokaltraktmodells gemittelt über das gesamte Signal für die vorderen Röhrenabschnitte
$g_{T,\text{vs,m}}$	Varianz der Röhrenquerschnitte innerhalb eines Vokaltraktmodells gemittelt über das gesamte Signal für die mittleren Röhrenabschnitte
$g_{T,\text{vs,r}}$	Varianz der Röhrenquerschnitte innerhalb eines Vokaltraktmodells gemittelt über das gesamte Signal für die hinteren Röhrenabschnitte
$g_{T,\text{vt}}$	Mittlere Varianz der Röhrensegmente
$\mathbf{g}^\lambda$	Gradient der Kostenfunktion zu Iterationsschritt $\lambda$
$h_A(\cdot)$	Koeffizienten des Filters zur A-Gewichtung
$h_i$	Filterkoeffizienten des Prädiktorfilters
$h_{\text{mel}}(\cdot)$	Koeffizienten der Mel-Filterbank
$h_w(\cdot)$	Koeffizienten der Fensterfunktion
$\hat{J}$	Geschätzter Jitter in %
$J(\boldsymbol{\xi})$	Jacobi-Matrix
$k$	Rahmenindex, Anzahl der Nachbarn in der $k$ -Nachbarschaft, Anzahl ausgewählter Eigenwerte bei LDA-Verfahren

$l$	Summenlaufindex, Klassenindex bei LDA-Verfahren
$l_{\text{free}}$	Anzahl der zu schätzenden Parameter im GMM-Modell
$L$	Anzahl zu unterscheidender Klassen beim LDA-Verfahren
$\hat{L}$	Wert der Log-Likelihood Funktion
$LoOP$	Lokale-Ausreißerwahrscheinlichkeiten
$m$	Clusterindex
$m_{\text{ges}}$	Mittelwert aller Datenpunkte $\mathbf{x}_i$
$m_l$	Mittelwert aller Datenpunkte einer Klasse $l$
$m_{\text{NTID}}$	Mittlere NTID-Wertung
$m_{\widehat{\text{NTID}}}$	Mittlere NTID-Wertung der Mustererkennung bzw. des Mappings
$\text{minPts}$	Minimale Anzahl von Datenpunkten in der $k$ -Nachbarschaft
$m_{\text{pmax}}$	Mittlere Position des maximalen Röhrenquerschnittes
$m_{\text{se}}$	Mittlerer quadratischer Fehler
$m_{\text{sereg}}$	Modifizierte Kostenfunktion zum mittleren quadratischen Fehler mit Regularisierung
$M$	Polynomgrad
$M_i$	Mikrofon $i$
$n$	Diskreter Zeitindex
$n\text{PLOF}(\Gamma, \mathbb{S}\mathbf{x}_i)$	Normalisierter <i>Probabilistic Local Outlier Factor</i>
$\mathbf{n}_{\text{L},i}^{\vartheta}$	Hilfsvektoren des langen Suchintervalls der Jitter- und Shimmer-Analyse
$n_{\text{min}}^{\vartheta}$	Zeitpunkt des Amplitudenminimums der Pitchperiode $\vartheta$
$\mathbf{n}_{\text{S},i}^{\vartheta}$	Hilfsvektoren des kurzen Suchintervalls der Jitter- und Shimmer-Analyse
$N_{\text{act}}$	Anzahl der aktiven Signalabschnitte
$N_{\text{CB}}$	Anzahl der Codebuchvektoren
$N_{\text{data}}$	Gesamtanzahl der Datenpunkte
$N_{\text{frame}}$	Anzahl der Abtastwerte in einem Rahmen $k$
$N_{\text{FFT}}$	Länge der FFT
$N_{\text{G}}$	Anzahl der Glottispulse
$N_{\text{GMM}}$	Anzahl der Gauß-Glocken
$N_{\text{k}}$	Anzahl der Rahmen in einem Sprachsignal
$N_l$	Anzahl der Datenpunkte in einer Klasse $l$
$N_{\text{L}}$	Anzahl von Abtastwerten im langen Suchintervall der Jitter- und Shimmer-Analyse
$N_{\text{S}}$	Anzahl von Abtastwerten in $o$ Pitchperioden
$NTID$	NTID-Wertung

$\widehat{NTID}$	Geschätzte NTID-Wertung durch Mustererkennung/Mapping
$N_W$	Anzahl der Gewichte des neuronalen Netzes
$\mathcal{N}(\mathbf{x}_i   \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m)$	Gaußverteilung des Clusters $m$
$o$	Anzahl der beobachteten Pitchperioden in einem Suchintervall der Jitter- und Shimmer-Analyse
$p$	Index der Röhrenquerschnitte
$p(g_i)$	Polynomwert für Merkmalsausprägung $g_i$
$\bar{p}_i(k)$	Mittlerer Schallpegel an Mikrofon $i$ im Rahmen $k$
$P$	Anzahl der Röhrenquerschnitte
$PLOF(\Gamma, \mathbb{S}\mathbf{x}_i)$	<i>Probabilistic Local Outlier Factor</i>
$P_\vartheta$	Pitchperiodenzeit der Periode $\vartheta$
$Q$	Schallquelle
$r$	Index der Kosinustransformation
$rmse$	Wurzel aus dem mittleren quadratischen Fehler
$rmse^*$	Epsilon insensitiver RMSE
$rmse^*_{\text{mod}}$	Modifizierter epsilon insensitiver RMSE
$R$	Länge eines Rahmens in Sekunden
$R_{NTID}$	Pearson-Korrelationskoeffizient der tatsächlichen mit der geschätzten NTID-Wertung
$\mathbf{R}_{xx}$	Autokorrelationsmatrix
$s(k, \eta)$	Segment des Sprachsignals, zum Zeitpunkt $k \cdot N_{\text{frame}} + \eta$ , mit $0 \leq \eta \leq N_{\text{frame}} - 1$
$s(n)$	Diskretes Sprachsignal
$s_{\text{cl}}(n)$	Sprachsignal nach <i>Center-Clipping</i> -Verfahren
$\tilde{s}_{\text{dB,env}}(n_{\text{max,act}})$	Einhüllende des Sprachsignals in dB
$\tilde{s}_{\text{env}}(n_{\text{max,act}})$	Einhüllende des Sprachsignals
$s_j$	Hilfsgröße für die Methode der kleinsten Quadrate
$\mathbf{s}_{i,j}(k, \eta)$	Datenpunkt der Menge $\mathbb{S}\mathbf{x}_i$
$s_{\text{norm}}(k, \eta)$	Abstandsnormiertes Sprachsignal
$\hat{S}$	Geschätzter Shimmer in dB
$S(k, \mu)$	Kurzzeit-Spektrum des Sprachsignals in Teilband $\mu$ , im Rahmen $k$
$\hat{S}_{\text{bb}}(k, \mu)$	Geschätztes Leistungsdichtespektrum des Hintergrundrauschens in Teilband $\mu$ , im Rahmen $k$
$\hat{S}_{\text{dBA}}(k)$	Geschätzter, mittlerer A-gewichteter Sprachpegel in dBA im Rahmen $k$

$\tilde{\mathbf{S}}_i^{\text{Nnn}}$	Sensitivitätsmatrix
$S_{\text{inter}}$	Interklassen-Varianz
$S_{\text{intra}}$	Intraklassen-Varianz
$S_{\text{mel}}(k, \rho)$	Mel-gefiltertes Spektrum in Mel-Teilband $\rho$ , im Rahmen $k$
$S_{\text{mfcc}}(k, r)$	MFCC-Spektrum im Rahmen $k$
$\widehat{SNR}(k)$	Geschätzter Signal-Rauschabstand im Rahmen $k$
$\hat{S}_{\text{ss}}(k, \mu)$	Geschätztes Leistungsdichtespektrum des Sprachsignals in Teilband $\mu$ , im Rahmen $k$
$\mathbb{S}_{\mathbf{x}_i}$	$k$ -Nachbarschaft des Datenpunktes $\mathbf{x}_i$
$\hat{t}_0$	Geschätzte Dauer einer mittleren Pitchperiode
$t_{\text{hold}}$	Haltedauer für gehaltene Vokale
$\hat{\mathbf{t}}_i$	Netzwerkausgangsvektor
$t_{\Delta, i}(k)$	Laufzeitunterschied Mikrophon $i$ zum Referenzmikrophon im Rahmen $k$
$T_{\text{VAD, min}}$	Amplitudengrenzwert der Sprachaktivitätsdetektion
$T_{\text{SNR}}$	SNR-Grenzwert für die Sprachaktivitätsdetektion
$v(k, \eta)$	Zwischengröße zur aktiv/passiv-Klassifizierung eines Rahmens $k$
$\mathbf{v}_l$	Eigenvektor der Klasse $l$
$v(n)$	Allgemeines Anregungssignal des Filters
$V(k)$	Verstärkungsfaktor im Rahmen $k$
$VAD(k)$	Ergebnis der Sprachaktivitätsdetektion im Rahmen $k$
$\widetilde{VAD}(k)$	Inverses Signal zur Sprachaktivitätsdetektion im Rahmen $k$
$v_i$	Hilfsgröße für die Methode der kleinsten Quadrate
$v_{L, i}^{\vartheta}(j)$	Betrachteter, gefensterter Signalabschnitt des langen Suchintervalls der Jitter- und Shimmer-Analyse
$v_{S, i}^{\vartheta}(j)$	Betrachteter, gefensterter Signalabschnitt des kurzen Suchintervalls der Jitter- und Shimmer-Analyse
$\mathbf{V}$	Matrix der Eigenvektoren
$w$	Gewicht im neuronalen Netz
$w_{\text{P}}$	Gewichtung des Pegelalgorithmus
$w(r)$	Gewichtungsfunktion der Diskreten Kosinustransformation
$w_{\text{K}}$	Gewichtung des Kleusberg-Algorithmus
$w_{\text{q}, j}^{\psi}$	Gewicht des neuronalen Netzes von Neuron $j$ zu Neuron $q$ in der Schicht $\psi$
$\mathbf{W}$	Transformationsmatrix des LDA-Verfahrens
$\mathbb{W}$	Menge der Gewichte im neuronalen Netz
$\mathbf{W}^{\psi}$	Gewichtsmatrix des neuronalen Netzes in der Schicht $\psi$

$\mathbf{x}_i$	Trainingsdatenvektor
$x(n)$	Allgemeines Signal am Filterausgang
$\mathbf{X}$	Matrix aller Trainingsdaten
$\mathbb{X}$	Menge aller Trainingsdaten
$\mathbf{Y}$	Matrix der transformierten Daten beim LDA-Verfahren
$z_i(l)$	Parameter, der die Zugehörigkeit eines Datenpunktes $l$ zum Cluster $i$ angibt
$z_m$	Latente, nicht beobachtbare Variable
$\mathbf{z}^\lambda$	Suchrichtung des Trainingsalgorithmus des neuronalen Netzes

## Verzeichnis griechischer Symbole

$\alpha^\lambda$	Lernrate zu Iterationsschritt $\lambda$
$\beta$	Glättungskonstante für ein IIR-Filter
$\beta^\lambda$	Gewichtungsparameter zum Update der Suchrichtung in Iterationsschritt $\lambda$
$\gamma$	Regularisierungsparameter
$\gamma_{i,m}^\lambda$	Hilfsvariable für den k-means-Algorithmus, welche die Zugehörigkeit zu einem Cluster anzeigt
$\gamma(z_m, i)$	Bedingte Wahrscheinlichkeit von $z_m = 1$ gegeben $\mathbf{x}_i$
$\Gamma$	Signifikanz
$\delta_{\text{dec}}$	Parameter für eine Verstärkung der Gewichtskorrektur
$\delta_{\text{inc}}$	Parameter für eine Abschwächung der Gewichtskorrektur
$\delta_L$	Langes Suchintervall der Jitter- und Shimmer-Analyse
$\delta_S$	Kurzes Suchintervall der Jitter- und Shimmer-Analyse
$\Delta_{\text{dek}}$	Steigungsparameter bei Anstieg des Geräuschpegels
$\Delta \mathbf{g}^\lambda$	Korrektur des Gradienten der Kostenfunktion im Iterationsschritt $\lambda$
$\Delta_{\text{ink}}$	Steigungsparameter bei Verringerung des Geräuschpegels
$\Delta_{q,j}$	Hilfsparameter zur Korrektur der Gewichte beim Training des neuronalen Netzes
$\Delta w_{q,j}^\lambda$	Korrektur des Gewichtes von Neuron $j$ zu Neuron $q$ im Iterationsschritt $\lambda$

---

$\epsilon$	Kleiner Wert mit $\epsilon > 0$
$\tilde{\epsilon}_s(k)$	Energie des A-gewichteten Signals im Rahmen $k$
$\vartheta$	Index der Pitchperioden
$\kappa$	Zeitliche Verschiebung in Abtastwerten
$\kappa_{\max}$	Zeitliche Verschiebung des ersten lokalen Maximums der AKF
$\kappa_{S,\max}$	Zeitliche Verschiebung des ersten lokalen Maximums der AKF im kurzen Suchintervall der Jitter- und Shimmer-Analyse
$\lambda$	Iterationsindex
$\lambda_{\max}$	Maximal festgelegte Anzahl an Iterationen
$\mu$	Teilbandindex
$\boldsymbol{\mu}_m$	Mittelwertvektor der Gauß-Glocke $m$
$\eta$	Zeitindex innerhalb eines Rahmens $k$
$\boldsymbol{\xi}$	Vektor der Netzwerkparameter
$\phi_{ss}(k, \kappa)$	Autokorrelationsfunktion des Sprachsignals in Abhängigkeit der Zeitverschiebung $\kappa$ , im Rahmen $k$
$\phi_{\max}(k)$	Maximalwert der AKF im Grundfrequenzbereich, im Rahmen $k$
$\phi_{ss,\text{norm}}(k, \kappa)$	Normierte AKF des Sprachsignals zur Zeitverschiebung $\kappa$ , im Rahmen $k$
$\boldsymbol{\phi}_{xx}$	Autokorrelationsvektor
$\rho$	Index der Mel-Bänder
$\sigma(n)$	Amplitude des Anregungssignals
$\sigma_{\text{NTID}}(n)$	Standardabweichung der NTID-Schätzung
$\sigma_u(n)$	Amplitude stimmloser Signalanteile des Anregungssignals
$\sigma_v(n)$	Amplitude stimmhafter Signalanteile des Anregungssignals
$\Sigma_m$	Kovarianzmatrix von Gauß-Glocke $m$
$\psi$	Index der Schichten im neuronalen Netz