

Fortschritt-Berichte VDI

VDI

Reihe 10

Informatik/
Kommunikation

Nr. 854

Dipl.-Ing. Stephan Preihs,
Hannover

Fehlerrobuste Audiocodierung mit geringer Latenz

The logo for the Institute for Information Processing (tnt) consists of the lowercase letters "tnt" in a bold, dark gray sans-serif font. The letters are slightly slanted to the right. The "t" and "n" are connected by a horizontal bar, and the "n" and "t" are connected by a vertical bar that also serves as the stem for the "t".

Institut für Informationsverarbeitung
www.tnt.uni-hannover.de

Fehlerrobuste Audiocodierung mit geringer Latenz

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades

Doktor-Ingenieur

(abgekürzt: Dr.-Ing.)

genehmigte

Dissertation

von

Dipl.-Ing. Stephan Preihs

geboren am 02. September 1983 in Hannover.

2016

Referent: Prof. Dr.-Ing. J. Ostermann
Korreferent: Prof. Dr.-Ing. G. Schuller
Vorsitzender: Prof. Dr.-Ing. B. Rosenhahn
Tag der Promotion: 06.09.2016

Fortschritt-Berichte VDI

Reihe 10

Informatik/
Kommunikation

Dipl.-Ing. Stephan Preihs,
Hannover

Nr. 854

Fehlerrobuste
Audiocodierung
mit geringer Latenz



Institut für Informationsverarbeitung
www.tnt.uni-hannover.de

Preihs, Stephan

Fehlerrobuste Audiocodierung mit geringer Latenz

Fortschr.-Ber. VDI Reihe 10 Nr. 854. Düsseldorf: VDI Verlag 2017.

128 Seiten, 36 Bilder, 4 Tabellen.

ISBN 978-3-18-385410-3, ISSN 0178-9627,

€ 48,00/VDI-Mitgliederpreis € 43,20.

Für die Dokumentation: Audiocodierung mit geringer Latenz – fehlerrobuste Subband-ADPCM
– Filterbankdesign – dynamische Bit-Allokation – Audiocodec Optimierung

In der vorliegenden Arbeit wird das Ziel der Entwicklung eines Verfahrens zur fehlerrobusten Audiocodierung mit geringer Latenz verfolgt. Hierzu wird der erfolgversprechende Ansatz einer Subband-ADPCM um Maßnahmen zum Erreichen einer algorithmischen Fehlerrobustheit ergänzt, die Codierstrukturen hinsichtlich einer Anpassung an wechselnde Signaleigenschaften erweitert und diese bezüglich ihrer Parameter global optimiert. Die Ermittlung der Leistungsfähigkeit des entwickelten Codierverfahrens und der Vergleich mit dem Stand der Technik erfolgt auf Basis des PEAQ-Algorithmus. Die Ergebnisse werden zusätzlich durch einen Hörversuch mit im kritischen Hören geübten Probanden validiert. Gemäß der in dieser Arbeit erfolgten Untersuchungen ist es mit dem entwickelten fehlerrobusten Audiocodec möglich, bei moderaten Bitraten von bis zu 200 kbit/s und einer algorithmischen Latenz von nur 0,5 ms eine Audioqualität zu erreichen, die selbst erhöhten Ansprüchen gerecht wird.

Bibliographische Information der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter <http://dnb.ddb.de> abrufbar.

Bibliographic information published by the Deutsche Bibliothek

(German National Library)

The Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliographie
(German National Bibliography); detailed bibliographic data is available via Internet at
<http://dnb.ddb.de>.

© VDI Verlag GmbH · Düsseldorf 2017

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe (Fotokopie, Mikrokopie), der Speicherung in Datenverarbeitungsanlagen, im Internet und das der Übersetzung, vorbehalten.

Als Manuskript gedruckt. Printed in Germany.

ISSN 0178-9627

ISBN 978-3-18-385410-3

Vorwort

Die vorliegende Arbeit entstand während meiner Zeit als Wissenschaftlicher Mitarbeiter am Institut für Informationsverarbeitung (TNT) der Gottfried Wilhelm Leibniz Universität Hannover.

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr.-Ing. Jörn Ostermann, der durch meine Einstellung am Institut, die Betreuung der Dissertation und die Unterstützung bei der Bearbeitung von Projekten sowie durch angeregte Diskussionen und kritisches Hinterfragen maßgeblich zum Gelingen dieser Arbeit beigetragen hat. Die Vielzahl an Entfaltungsmöglichkeiten, die über das rein wissenschaftliche Arbeiten hinausgingen sowie das stets entgegengebrachte Vertrauen und Verständnis wusste ich jederzeit sehr zu schätzen.

Mein Dank gilt auch Herrn Prof. Dr.-Ing. Gerald Schuller für die Annahme des Korreferates und das damit verbundene Anfertigen eines Gutachtens. Des Weiteren möchte ich Herrn Prof. Dr.-Ing. Bodo Rosenhahn für die Übernahme des Prüfungsvorsitzes, das Leiten des Promotionsverfahrens sowie für den unkomplizierten Umgang bei der Betreuung von Vorlesungen, Laboren oder institutsinternen Angelegenheiten danken.

Ein großer Dank geht an die Kollegen des TNT. Die durchweg gute Atmosphäre am Institut mit stets geöffneten Türen und Ohren und die vielfältigen Möglichkeiten des fachlichen und privaten Austausches haben mich immer motiviert. Besonders erwähnen möchte ich dabei die „Freunde aus dem eNIFE“ Björn Scheuermann, Florian Baumann, Arne Ehlers, Kai Cordes und Matthias Reso. Den aufeinander folgenden Bürokollegen Gerard Pons-Moll und besonders Thomas Krause möchte ich für die Bereicherung des Büroalltages durch angeregte Diskussionen, Hilfe bei fachlichen Fragestellungen und die aufmunternden Worte in schwierigen Zeiten danken.

Mein aufrichtiger Dank gebührt auch den Kolleginnen und Kollegen aus dem Sekretariat und der Administration des TNT namentlich Pia Bank, Hilke Brodersen und Doris Jaspers-Göring sowie Martin Pahl und Thomas Wehberg. Die hervorragende Unterstützung bei der Koordination von Institutsangelegenheiten, Planung und Durchführung von Dienstreisen und Bewältigung von vertraglichen Belangen hat mir immer sehr geholfen. Zusätzlich danke ich Herrn Matthias Schuh für die Unterstützung bei technischen Angelegenheiten und der Beschaffung von Hardware sowie den dabei herrschenden ungezwungenen Umgang.

Bedanken möchte ich mich auch bei den zahlreichen Studenten, die im Rahmen von Abschlussarbeiten und Tätigkeiten als hilfswissenschaftliche Kräfte durch ihre engagierte und zuverlässige Zusammenarbeit viele Beiträge zu meinen Forschungsaktivitäten geleistet haben. Mein Dank gilt zudem den Probanden des Hörversuches für ihre Bereitschaft zur Teilnahme sowie die Zeit und Geduld bei dessen Durchführung. Herrn Nikolaus Meine danke ich zusätzlich für die Bereitstellung des Quellcodes für die Benutzeroberfläche zur Realisierung der Hörversuche.

Ein großer Dank gilt meinen Eltern Ruth und Wolfgang Preihs. Ohne die Unterstützung im beruflichen und privaten Werdegang und die Rückendeckung in allen Lebenslagen wäre mein Studium und die Promotion nicht denkbar gewesen.

Mein spezieller Dank gilt meiner Frau Katha, die mir immer den Rücken frei hält und mit ihrer Geduld, dem vielen guten Zureden sowie dem akribischen Korrekturlesen der Arbeit deren Fertigstellung überhaupt erst ermöglicht hat.

für Oma Henny (1916-2016)

Inhaltsverzeichnis

Abkürzungen und Formelzeichen	VI
Kurzfassung/Abstract	XII
1 Einleitung	1
1.1 Stand der Technik: Audiodcodierung mit geringer Latenz	4
1.2 Ansatz ADPCM-basierter Subband-Codierung	7
1.3 Ziele, Beiträge und Aufbau der Arbeit	8
2 Filterbänke mit geringer Latenz	10
2.1 Filterbänke in der Audiodcodierung	10
2.2 <i>M</i> -Kanal Filterbankentwurfsverfahren	13
2.3 Vergleich verschiedener Filterbankentwürfe	21
2.4 Globale Optimierung von Entwurfsparametern	26
3 Fehlerrobuste ADPCM-Codierung der Subband-Signale	30
3.1 Grundlagen zur ADPCM	30
3.2 Fehlerrobuste adaptive Prädiktion	35
3.3 Fehlerrobuste adaptive Optimalquantisierung	45
3.4 Maßnahme zur Verbesserung der Subband-Codierung	64
4 Audiocodec Optimierung	72
4.1 Audiocodec Bewertungsgrößen	72
4.2 Methode zur globalen Parameteroptimierung	75
4.3 Ergebnisse von globalen Optimierungen der Audiocodec Parameter .	79
5 Evaluierung	83
5.1 Objektive Ergebnisse	83
5.2 Ergebnisse subjektiver Tests	88
5.3 Einfluss von Kanalübertragungsfehlern	94
6 Zusammenfassung und Ausblick	98
A Anhang: Tabellarische Ergebnisse	102
A.1 Ergebnisse der objektiven Evaluierung	102
A.2 Ergebnisse der subjektiven Evaluierung	105
Literaturverzeichnis	108

Abkürzungen und Formelzeichen

Abkürzungen

3D	<u>3</u> - <u>Dimensional</u>
AAC	<u>Advanced</u> <u>Audio</u> <u>Coding</u>
AD/DA	<u>Analog</u> - <u>Digital</u> / <u>Digital</u> - <u>Analog</u>
ADPCM	<u>Adaptive</u> <u>Differential</u> <u>Pulse</u> <u>Code</u> <u>Modulation</u>
AFB	<u>Analyse</u> <u>filter</u> <u>bank</u>
AKF	<u>Autokorrelations</u> <u>folge</u>
BER	<u>Bit</u> <u>Error</u> <u>Rate</u>
BP	<u>Band</u> <u>pass</u>
CELP	<u>Code</u> <u>Excited</u> <u>Linear</u> <u>Prediction</u>
CELT	<u>Constrained</u> <u>Energy</u> <u>Lapped</u> <u>Transform</u>
dB	<u>dezi</u> - <u>Bel</u>
dBA	<u>dynamische</u> <u>Bit</u> - <u>Allokation</u>
De-MUX	<u>De</u> - <u>Multiplexer</u>
DI	<u>Distortion</u> <u>Index</u>
DSP	<u>Digital</u> <u>Signal</u> <u>Processor</u>
ELD	<u>Enhanced</u> <u>Low</u> <u>Delay</u>
EVS	<u>Enhanced</u> <u>Voice</u> <u>Service</u>
FFT	<u>Fast</u> <u>Fourier</u> <u>Transform</u>
FIR	<u>Finite</u> <u>Impulse</u> <u>Response</u>
FLAC	<u>Free</u> <u>Lossless</u> <u>Audio</u> <u>Codec</u>
GAL	<u>Gradient</u> <u>Adaptive</u> <u>Lattice</u>
GSM	<u>Global</u> <u>System</u> for <u>Mobile</u> <u>Communication</u>
HP	<u>Hoch</u> <u>pass</u>
IETF	<u>Internet</u> <u>Engineering</u> <u>Task</u> <u>Force</u>
IIR	<u>Infinite</u> <u>Impulse</u> <u>Response</u>
ISO	<u>International</u> <u>Standardization</u> <u>Organization</u>
LSB	<u>Least</u> <u>Significant</u> <u>Bit</u>
MDCT	<u>Modifizierte</u> <u>Diskrete</u> <u>Cosinus</u> <u>Transformation</u>
MOV	<u>Model</u> <u>Output</u> <u>Variable</u>
MP3	<u>MPEG</u> - <u>I</u> <u>Audio</u> <u>Layer</u> <u>3</u>
MPEG	<u>Moving</u> <u>Pictures</u> <u>Expert</u> <u>Group</u>
MSB	<u>Most</u> <u>Significant</u> <u>Bit</u>
MUX	<u>Multiplexer</u>

NMP	Network <u>Music</u> <u>Performance</u>
NMR	Noise to <u>Mask</u> <u>Ratio</u>
ODG	Objective <u>Difference</u> <u>Grade</u>
PEAQ	Perceptual <u>Evaluation</u> of <u>Audio</u> <u>Quality</u>
PQMF	Pseudo <u>Quadrature</u> <u>Mirror</u> <u>Filter</u>
QMF	Quadrature <u>Mirror</u> <u>Filter</u>
segSNR	segmental <u>Signal</u> to <u>Noise</u> <u>Ratio</u>
SFB	Synthesefilter <u>bank</u>
SLQ	Sphärisch <u>Logarithmische</u> <u>Quantisierung</u>
SNR	Signal to <u>Noise</u> <u>Ratio</u>
SNR _r	Signal-to-reconstruction <u>Noise</u> <u>Ratio</u>
TP	Tiefpass
ULD	Ultra <u>Low</u> <u>Delay</u>

Formelzeichen

Allgemeine Formelzeichen, Parameter und Bezeichner

A	A-Law Parameter
a_i	Prädiktionskoeffizient des Direktstruktur-Prädiktionsfilters
α	Leakage-Faktor im Lattice-Filter bzw. GAL-Algorithmus
$\alpha(n)$	Zeitkonstante der Pegelschätzung in der dynamischen Bit-Allokation
\mathbf{b}	Vektor aus Ergebnissen der dynamischen Bit-Allokation
$b_m(n)$	Rückwärtsprädiktionsfehler in der m -ten Stufe des Lattice-Filters
β	Leakage-Faktor des GAL-Algorithmus
C_k	k -tes Codewort bzw. k -ter Repräsentativwert des Quantisierers
d_k	k -te Entscheidungsschwelle bzw. Maximal- und Minimalwert des Quantisierers
Δ_k	Breite der k -ten Stufe des skalaren Quantisierers
$e(n)$	Prädiktionsfehlersignal
$\tilde{e}(n)$	Signal des rekonstruierten Prädiktionsfehlers
$e_{\text{norm}}(n)$	Normalisiertes Prädiktionsfehlersignal
$\tilde{e}_{\text{norm}}(n)$	Rekonstruiertes normalisiertes Prädiktionsfehlersignal
$F(z)$	Übertragungsfunktion eines Direktstruktur-Prädiktionsfehlerfilters
$F_{\text{Alias.,max.}}$	Maximale Aliasingverzerrung einer Filterbank
$F_{\text{Ampl.,max.}}$	Maximale Amplitudenverzerrung einer Filterbank
$f_m(n)$	Vorwärtsprädiktionsfehler in der m -ten Stufe des Lattice-Filters
$G_i(z)$	Synthesefilter Übertragungsfunktion des i -ten Bandes
$g_i(n)$	Impulsantwort des i -ten Synthesefilters einer Filterbank
$G_{S,\text{min.}}$	Miminale Sperrdämpfung der Synthesefilter einer Filterbank
$G_{S,\text{mtl.}}$	Mittlere Sperrdämpfung der Synthesefilter einer Filterbank

G_p	Prädiktionsgewinn
$H_i(z)$	Analysefilter Übertragungsfunktion des i -ten Bandes
$h_i(n)$	Impulsantwort des i -ten Analysefilters einer Filterbank
$H_{i,S,\text{mtl.}}$	Mittlere Sperrdämpfung des i -ten Analysefilters einer Filterbank
$H_{S,\text{min.}}$	Minimale Sperrdämpfung der Analysefilter einer Filterbank
$H_{S,\text{mtl.}}$	Mittlere Sperrdämpfung der Analysefilter einer Filterbank
i, j, k, l	Allgemeine Laufvariablen
J	Blockgröße der dynamischen Bit-Allokation
K_C	Anzahl der Codewörter und Repräsentativwerte des Quantisierers
$K(\cdot)$	Operator zum Berechnen des Kostenwertes der globalen Parameteroptimierung
κ	Parameter der Pegelreduktion in der dynamischen Bit-Allokation
k_m	Lattice-Koeffizient in der m -ten Stufe des Lattice-Filters
L	Länge des Eingangsblockes zur Lattice-Prädiktoradaption
$\lambda(n)$	Parameter für die Hüllkurvenschätzung zur Signalnormalisierung
λ_{AT}	Parameter für die Hüllkurvenschätzung im „Ansprech-Fall“
λ_{RT}	Parameter für die Hüllkurvenschätzung im „Rücklauf-Fall“
M	Anzahl der Frequenzbänder des Subband-Codecs bzw. der Filterbank sowie sich daraus ebenfalls ergebender Unterabtastungs-, Überabtastungs- und Kompensationsfaktor
m	Stufenindex des Lattice-Filters
$\downarrow M$	Unterabtastung mit dem Faktor M
$\uparrow M$	Überabtastung mit dem Faktor M
$\hat{\mu}$	Grundschriftweite im GAL-Algorithmus
N	Anzahl der Koeffizienten der Subband-Filter einer Filterbank
$P(z)$	Übertragungsfunktion eines Direktstruktur-Prädiktionsfilters
$p(n)$	Prädiktionssignal
$\mathbf{P}_a(z)$	Polyphasen-Matrix der Analysefilterbank beim Entwurfsverfahren nach Schuller
$p_{\Delta_n}(n)$	Zeitlich verschobene Impulsantwort des Prototypfilters beim Entwurfsverfahren nach Keiler/Nguyen
$p_l(n)$	Impulsantwort des linearphasigen Prototypfilters beim Entwurfsverfahren nach Keiler/Nguyen
$p_{\text{opt.}}(n)$	Numerisch optimierte Impulsantwort des Prototypfilters beim Entwurfsverfahren nach Keiler/Nguyen
$\mathbf{P}_s(z)$	Polyphasen-Matrix der Synthesefilterbank beim Entwurfsverfahren nach Schuller
ϕ	Parameter der Pegelreduktion in der dynamischen Bit-Allokation
p	Ordnung des Lattice- oder Direktstruktur-Filters
$Q(\cdot)$	Operator zum Durchführen der skalaren Quantisierung

$\hat{q}(n)$	Quantisierungsindex
$\hat{q}_i(m)$	Subband-Quantisierungsindex des i -ten Bandes
$Q^{-1}(\cdot)$	Operator zum Abbilden des Quantisierungsindexes auf den Repräsentativwert
$Q_a(\cdot)$	Operator der adaptiven skalaren Quantisierung
$Q_a^{-1}(\cdot)$	Operator zum Abbilden des Quantisierungsindexes auf den Repräsentativwert und zur anschließenden Denormalisierung
$r_k(n)$	k -ter Eintrag für die Lookup-Tabellen des adaptiven Quantisierers
$r(\cdot)$	Operator zum Zugriff auf die Einträge der Lookup-Tabelle des adaptiven Quantisierers
$r_{xx}(k)$	Autokorrelationsfolge
r_v	Periodenlänge der Prädiktionsfehlernormalisierungsfaktorübertragung bei der Breitband-Codierung
ρ	Parameter der Pegelreduktion in der dynamischen Bit-Allokation
$S(z)$	Übertragungsfunktion eines Direktstruktur-Synthesefilters
s	Verzögerungsparameter beim Filterbankentwurf nach Keiler/Nguyen
$s_i(n)$	Geschätzter Pegel des i -ten Subbandes in der dynamischen Bit-Allokation
$\sigma_m^2(n)$	(Geschätzte) Fehlerenergie der m -ten Stufe im Lattice-Filter
σ_{\min}^2	Begrenzung des Normalisierungsfaktors im GAL-Algorithmus
T	Anzahl der Signale eines Testsets für die globale Parameteroptimierung
τ	Gruppenlaufzeit einer Filterbank in Samples
$u(n)$	Gefensteretes Eingangssignal
$v(n)$	Normalisierungssignal der adaptiven Quantisierung
v_{\min}	Minimalwert für das Normalisierungssignal der adaptiven Quantisierung
$w(n)$	Fensterfunktion
ω	Normierte Kreisfrequenz
$\omega_{S,i}$	Als Sperrbereich definierter Teil der normierten Kreisfrequenz
X	Pegelreduktion in der Iterationsschleife der Bit-Allokation
χ	Parametervektor bei der globalen Parameteroptimierung
$x(n)$	Eingangssignal des Subband-Codecs bzw. der ADPCM oder exemplarisches Eingangssignal des Prädiktions- bzw. Lattice-Filters
\mathbf{x}	Allgemeiner Signal- oder Filterkoeffizientenvektor
\mathbf{x}_0	Initialer Filterkoeffizientenvektor beim Entwurfsverfahren nach Schuller
$x_i(m)$	Unterabgetastetes Subband-Signal des i -ten Bandes
$x_i(n)$	Subband-Signal des i -ten Bandes

$\tilde{x}(n)$	Ausgangssignal des Subband-Codecs bzw. der ADPCM sowie Rekonstruktionssignal im Encoder
$\tilde{x}_i(m)$	Rekonstruiertes unterabgetastetes Subband-Signal
ζ	Leakage-Faktor zum Erreichen der Fehlerrobustheit in der adaptiven Quantisierung

Formelzeichen des Filterbankentwurfsverfahrens nach Nayebi

A_0	Matrix mit Prototypen-Koeffizienten der Analysefilterbank
A_k	Matrix mit den Koeffizienten der Analysefilterbank in der k -ten Iteration
A	Matrix mit den Koeffizienten der Analysefilterbank
B	Matrix mit verschobener Antidiagonalmatrix zum Einstellen der gewünschten Gruppenlaufzeit des Filterbankentwurfs
δ_{A_k}	Matrix zur Modifikation der Analysefilterbank-Koeffizienten in der k -ten Iteration
E	Wert der Nayebi-Kostenfunktion
E_{global}	Wert der Kostenfunktion zur globalen Optimierung der Nayebi-Gewichtsparameter
$E_{\min.}$	Schwelle für die Nayebi-Kostenfunktion zum Abbruch der Optimierung
ϵ_j	Allgemeiner Fehler der Übertragungsfunktion der Subband-Filter
ϵ_{pb}	Fehler der Übertragungsfunktion eines Subband-Filters im Durchlassbereich
ϵ_r	Rekonstruktionsfehler der Filterbank
ϵ_{sb}	Fehler der Übertragungsfunktion eines Subband-Filters im Sperrbereich
S	Matrix mit Koeffizienten der Synthesefilterbank
S_{LS}	Matrix mit Koeffizienten einer Least Squares-Lösung für die Synthesefilterbank
w_j	Gewichtsfaktor der Nayebi-Kostenfunktion

Formelzeichen der adaptiven Vor- und Nachfilterung

$A_k(z)$	Übertragungsfunktion des k -ten Filters zur Allpasstransformation
G_k	Grenzfrequenz des k -ten Prototypfilters
$g_k(n)$	Verstärkungsfaktor des k -ten Filters der Vorfilterkaskade
$g_k^{-1}(n)$	Verstärkungsfaktor des k -ten Filters der Nachfilterkaskade
$\check{H}_{\text{bpk}}(z)$	Übertragungsfunktion des k -ten Bandpasses zur Leistungsschätzung
$\check{H}_k(z)$	Übertragungsfunktion des k -ten Bandes der Vorfilterkaskade
$\check{H}_{k,\text{prot.}}(z)$	Übertragungsfunktion des k -ten Prototypfilters
$\check{H}_{\text{pre}}(z)$	Resultierende Übertragungsfunktion der Vorfilterkaskade
K	Anzahl der Subbänder der Vor- und Nachfilterkaskade

λ_p	Glättungsparameter des rekursiven Filters zur Leistungsschätzung
ν	Parameter zum Abbilden auf den Verstärkungsfaktor
$\omega_{c,k}$	Mittenfrequenz des k -ten Bandes
$\omega_{l,k}$	Untere Grenzfrequenz des k -ten Bandes
$\omega_{u,k}$	Obere Grenzfrequenz des k -ten Bandes
$\omega_{w,k}$	Breite des k -ten Bandes
$\bar{p}(n)$	Mittelwert der geschätzten Leistungen
p_{\min}	Minimalwert für die geschätzte Leistung
$p_k(n)$	Geschätzte Leistung des k -ten Bandes
$\tilde{x}_{j,\text{bpk}}(n)$	Rekonstruiertes überabgetastetes bandpassgefiltertes Subband-Signal
$y(n)$	Eingangssignal der Vorfilterkaskade
$\hat{y}(n)$	Ausgangssignal der Nachfilterkaskade

Formelzeichen der Sphärisch Logarithmischen Quantisierung

D	Zahl der verwendeten Dimensionen
Δ	Kantenlänge der Quantisierungszellen
$e_{\text{ADPCM}}(n)$	Fehlersignal der ADPCM-Codierung
η	Zahl der Durchläufe bei der Radiusschätzung
N_φ	Winkelindex
N_r	Radiusindex
N_{SLQ}	SLQ-Index
R	Resultierende Anzahl an Bits/Sample
r	Radius des Vektors
\hat{r}	Quantisierter Radius
\tilde{r}	Geschätzter Radius
\hat{x}	Vektor mit quantisierten Signalwerten

Kurzfassung

Fehlerrobuste Audiocodierung mit geringer Latenz hat eine Vielzahl von Anwendungsfeldern. So kann diese durch Einsatz in drahtlosen Mikrofon- und Kopfhörersystemen, Audiostreaming-Applikationen und modernen Hörhilfen für effizientere Übertragungsstrecken, erweiterte Funktionsumfänge und kostengünstigere Produkte sorgen. Allerdings eignen sich bestehende standardisierte, quelloffene oder in der Literatur zu findende Verfahren aufgrund zu hoher Latenzen, zu geringer Audioqualität oder nicht vorliegender Fehlerrobustheit nur eingeschränkt für die Verwendung in solchen Systemen. Aus diesem Grund wird bei vielen praktischen Anwendungen auf proprietäre Lösungen zurückgegriffen.

In der vorliegenden Arbeit wird daher das Ziel der Entwicklung eines Verfahrens zur fehlerrobusten Audiocodierung mit geringer Latenz verfolgt, das diese Nachteile überwindet. Hierzu wird der erfolgversprechende Ansatz einer Subband-ADPCM um Maßnahmen zum Erreichen einer algorithmischen Fehlerrobustheit ergänzt, die Codierstrukturen hinsichtlich einer Anpassung an wechselnde Signaleigenschaften erweitert und diese bezüglich ihrer Parameter global optimiert. Die Ermittlung der Leistungsfähigkeit des entwickelten Codierverfahrens und der Vergleich mit dem Stand der Technik erfolgt auf Basis des PEAQ-Algorithmus. Die Ergebnisse werden zusätzlich durch einen Hörversuch mit im kritischen Hören geübten Probanden validiert.

Gemäß der in dieser Arbeit erfolgten Untersuchungen ist es mit dem entwickelten Audiocodec möglich, bei moderaten Bitraten von bis zu 200 kbit/s und einer algorithmischen Latenz von nur 0,5 ms eine Audioqualität zu erreichen, die selbst erhöhten Ansprüchen gerecht wird, was durch die Ergebnisse des Hörversuches Bestätigung fand. Hierbei wird aufgrund der globalen Parameteroptimierung für das SQAM-Testset eine PEAQ-Bewertung ohne signifikante Ausreißer erreicht. Die Codierrobustheit bei auftretenden Kanalübertragungsfehlern ist dabei algorithmisch sichergestellt, wobei für geringe Fehlerraten eine Abschätzung der Audioqualität zu Ergebnissen führt, die mit denen von aktuellen fehlerrobusten Verfahren aus der Literatur vergleichbar sind.

Stichworte: Audiocodierung mit geringer Latenz, fehlerrobuste Subband-ADPCM, Filterbankdesign, dynamische Bit-Allokation, Audiocodec Optimierung.

Abstract

Error robust low delay audio coding has numerous applications. When utilized in technologies like wireless microphones and headsets, audiostreaming solutions or modern hearing aids it can lead to a more efficient transmission, enhanced functionalities and inexpensive products. Unfortunately most of the well known standardized, open source and published approaches are of limited usability in such scenarios because of too high delays, a not sufficiently high audio quality or a missing robustness against transmission errors. Thus for practical solutions in many cases proprietary codecs are used.

This thesis therefore targets at developing an audio coding scheme which can overcome the drawbacks of existing methods. For this, the promising approach of a subband-ADPCM is extended by methods for achieving a robustness against transmission errors and the codec structures are enhanced by an adaption to changing signal characteristics as well as globally optimized regarding their parameters. The evaluation of the codec performance and comparison to the state of the art is done by means of the PEAQ-algorithm. The tendencies of this evaluation are verified by a listening test with subjects that are trained in critical listening.

According to the results obtained with the research conducted in this work, the developed audio codec, with its algorithmic delay of only 0,5 ms, provides an audio quality that even meets higher demands at a bitrate of up to 200 kbit/s which was confirmed by the results of the listening test. Due to the applied global parameter optimization, a PEAQ-evaluation of the codec leads to results without significant outliers. The robustness against channel transmission errors is algorithmically ensured and for low error rates an audio quality is achieved that is comparable to the results of error robust state of the art approaches.

Keywords: Low delay audio coding, error robust subband-ADPCM, filter bank design, dynamic bit-allocation, audio codec optimization.

