

Instantanfrequenzgesteuerte Multikanal-Dynamikkompression für Hörgeräte

Thomas Bisitz, Volker Hohmann
 Medizinische Physik, Universität Oldenburg

Einleitung

Wir stellen ein vereinfachtes Cochlea-Signalverarbeitungsmodell vor, das einen pathologischen Lautheitsanstieg (Rekrutment) ausgleichen soll. Die als nahezu instane Dynamikkompression verstandenen aktiven Verstärkungs-Prozesse in der gesunden Cochlea werden mit Hilfe einer Gammaton-Filterbank nachempfunden. An deren Ausgang wird in jedem Kanal anhand der Instantanfrequenz entschieden, ob dieser Kanal aktiv verstärkt und komprimiert wird oder nicht. Im Gegensatz zu herkömmlichen Dynamikkompressionsverfahren ist eine gehörgerechte Zeit- und Frequenzauflösung realisiert. Bei bisherigen instantanen Verfahren auftauchende Artefakte unterbleiben. Das implementierte Signalverarbeitungsmodell konnte direkt genutzt werden, um Messungen mit Schwerhörenden durchzuführen. Mit Anpassung an den individuellen Hörverlust wurden insbesondere Auswirkungen auf die Sprachverständlichkeit sowie auf subjektive Qualitätsbeurteilungen untersucht.

Recruitment

Sowohl physiologische Basilmembranmessungen [1] als auch psychoakustische Messungen [2] lassen auf einen möglichen Verlauf der effektiven Kompressionskennlinie der Cochleaverarbeitung schließen. Die aktiven Verstärkungsprozesse für niedrige Schwingungsamplituden bzw. niedrige Eingangspegel nehmen bei Verlust äußerer Haarzellen bzw. Hörverlust ab oder fallen ganz aus. Wird dem folgend eine Kennlinie in der Cochlea wie in Abb. 1 (links) angenommen [2], ergeben sich die Kennlinien zum Ausgleich eines Hörverlustes wie in Abb. 1 (rechts) dargestellt.

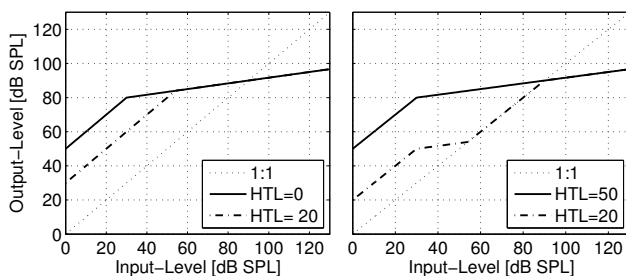


Abbildung 1: links: Angenommene Kompressionskennlinie in der gesunden Cochlea (HTL = 0) und bei 20dB Hörverlust (HTL = 20); rechts: Kennlinien im Hörgerät zum Ausgleich (HTL = 20, 50)

Algorithmus

Eine Gammatonfilterbank [3] (GTFB) mit Resynthesemöglichkeit ermöglicht durch Betragsbildung der komplexen Ausgangssignale der Filterbankkanäle eine instantane Berechnung der Einhüllenden in jedem Kanal. Die Mittenfrequenzen der Filter sind auf der ERB-Skala verteilt, pro ERB sind hier zwei Filter gewählt worden.

Instantanfrequenzsteuerung: Zur Realisierung einer instantanen Dynamikkompression wird in jedem Kanal der GTFB eine gain-Kennlinie entsprechend Abb. 1 (rechts) basierend auf der instantanen Einhüllenden angewendet (Abb. 2). Um on- und off-frequency-Signale in einem Kanal unterscheiden zu können, wird in jedem Kanal die Differenz der Instantanfrequenz (IF) zur Mittenfrequenz des Kanals berücksichtigt. In einem Filterbankkanal liegt ein Signal $s(t)$ mit einer wohldefinierten Amplitude $a(t)$ und Phase $\rho(t)$ vor. Es ergibt sich dann $IF = d\rho(t)/dt$. Mit zunehmender Abweichung der IF von der jeweiligen Mittenfrequenz des Kanals wird die Verstärkung und Kompression zunehmend zurückgenommen. Für Abweichungen $> 2\text{ ERB}$ („off-frequency“) wird so eine lineare Kennlinie ohne Verstärkung erreicht. Für einen definierten Ruhezustand des Systems wird auf dem Kontrollpfad ein Rauschen addiert, was für Signale mit kleinem Pegel überwiegt und somit die Instantanfrequenz bestimmt. Dazu wird für jeden Kanal ein schmalbandiges Rauschen mit einer Mittenfrequenz erzeugt, die unterhalb der Mittenfrequenz des Kanals liegt. So wird implizit ein Noisegate erreicht, da diese off-frequency Signale nicht verstärkt werden.

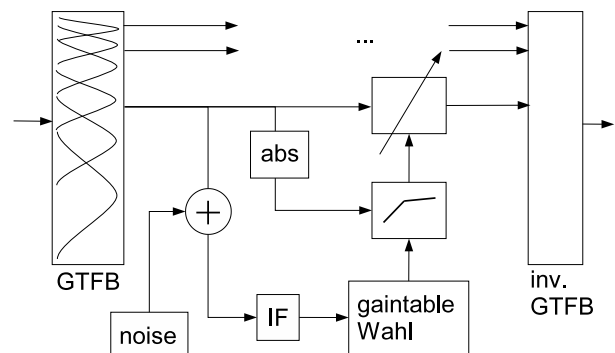


Abbildung 2: Implementiertes Dynamikkompressionsverfahren mit Instantanfrequenzsteuerung und Pfadrauschen

Messungen

Das instantane Kompressionsverfahren wurde mit einer angepassten linearen Verarbeitung verglichen. Bei dieser linearen Verstärkung wurde die mittlere Verstärkung angewandt, die ein sprachähnliches Rauschen bei der Kompression hervorruft, um eine möglichst gleiche Lautheit und ein möglichst gleiches Spektrum zu erreichen. Für die sechs schwerhörenden Versuchspersonen wurde für jeden GTFB-Kanal die gain-Kennlinie aus einem für die nötigen Mittenfrequenzen interpolierten Audiogramm an den individuellen Hörverlust angepasst. Zur Messung der Sprachverständlichkeit wurde in einem adaptiven Messverfahren (Oldenburger Satztest [4]) der Signalrauschabstand bestimmt, bei dem die Hälfte der Wörter verstanden wird (SRT, Speech Reception Threshold). Die Messung fand im unmodulierten sowie modulierten Störgeräusch statt. Die Signale wurden mit einer auf 1dB begrenzten SNR-Auflösung vorverarbeitet. Zur subjektiven Beurteilung wurde in Paarvergleichsmessungen nach Klangqualität, der Höranstrengung (bzw. des leichten Verstehens) bei Sprache und der Wahrnehmbarkeit aller Instrumente bei Musik gefragt. Dafür wurden Sprachsignale mit und ohne moduliertem und unmoduliertem Störgeräusch sowie Aufnahmen von Musik genutzt. In einer anderen Messung wurden ein herkömmliches Kompressionsverfahren und eine lineare Verarbeitung (half-gain-rule) verglichen.

Ergebnisse und Diskussion

In Tab. 1 sind die ermittelten SRT für die erste Messung (herkömmliches Kompressionsverfahren) dargestellt. Der Vergleich der instantanen Kompression mit der angepassten linearen Verarbeitung findet sich in Tab. 2. Die Anzahl der Gewinne im Paarvergleich zeigt Tab. 3.

VP:	bg	fe	bh	goe	bu	gre
unmod						
lin	-2,9	-5,2	-3,8	-4,5	-2,9	-5,1
komp	-3,1	-5,7	-3,6	-4,0	-1,5	-4,7
Δ	0,2	0,5	-0,2	-0,5	-1,4	-0,4
Mittel: -0,3						
mod						
lin	-8,7	-15,6	-8,6	-12,5	-6	-13,6
komp	-8,0	-11,9	-7,2	-10,4	-3,2	-13,6
Δ	-0,7	-3,7	-1,4	-2,1	-2,8	0,0
Mittel: -1,8						

Tabelle 1: Für die Versuchspersonen (VP) ermittelte SRT (dB SNR) und Verbesserung durch herkömmliche Kompression (komp) im Vergleich zu linearer Verarbeitung (lin) (mit half-gain-rule) für (un)moduliertes Störgeräusch.

Die Tests mit sechs Versuchspersonen lassen Unterschiede zwischen den Versuchspersonen erkennen, die nur durch weitere Tests mit diesen und anderen Versuchspersonen erklärbar werden und dann zu größerer Aussagekraft führen.

Eine Verschlechterung der Sprachverständlichkeit in unmoduliertem Störgeräusch ist durch Kompression zu er-

VP:	bg	fe	bh	goe	ph	gre
unmod						
lin2	-2,1	-5,9	-2,6	-4,2	-0,5	-5,9
instkmp	-1,2	-4,9	-2,1	-3,2	-1,6	-4,9
Δ	-0,9	-1,0	-0,5	-1	-2,1	-1
Mittel: -1,1						
mod						
lin2	-7	-13,1	-8,4	-9,8	1,7	-8,2
instkmp	-3,7	-15,7	-7,3	-13	2,6	-12,9
Δ	-3,3	2,6	-1,1	3,2	-0,9	4,7
Mittel: 0,9						

Tabelle 2: Ermittelte SRT (dB SNR) und Verbesserung durch instantane Kompression (instkmp) im Vergleich zu angepasster linearer Verarbeitung (lin2) für (un)moduliertes Störgeräusch.

	herk komp/half gain	inst komp/lin
Gewinne:	46 / 29	118 / 106

Tabelle 3: Anzahl Gewinne der Paarvergleichsmessungen für einen herkömmlichen Kompressor gegenüber einer linearen Verarb. (half gain) und für die inst. Kompression mit angepasster lin. Verarb.

warten, sie liegt bei dem instantanen Verfahren jedoch kaum über der bei der verwendeten herkömmlichen Kompression. In moduliertem Störgeräusch gibt es drei von sechs Versuchspersonen, die eine deutliche Verbesserung der Sprachverständlichkeit bei instantaner Kompression zeigen. Zu einer genaueren Erklärung sind weitere Messungen sowie eine genauere Klassifizierung der individuellen Hörverluste über ein Audiogramm hinaus nötig, z.B. mit Hilfe einer kategorialen Lautheitsskalierung. Bei der Qualitätsbewertung wird das Verfahren insgesamt gleich gut angenommen wie die lineare Verarbeitung und von keiner Versuchsperson abgelehnt.

Der Ansatz einer instantanfrequenz-gesteuerten Multikanal-Dynamikkompression erscheint daher als vielversprechend, um mit gehörgerechter Zeit- und Frequenzauflösung einen pathologischen Lautheitsanstieg auszugleichen. Auch eine Kombination mit Störgeräuschunterdrückungs-Algorithmen bietet weitere Entwicklungsmöglichkeiten.

Literatur

- [1] Ruggero, M.A. and Rich, N.C., 1991. „Furosemide alters organ of Corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane,“ *J. Neurosci.*, 11: 1057-1067.
- [2] Plack, J., Drga, V., Lopez-Poveda, E.A., 2004. „Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss,“ *J. Acoust. Soc. Am.*, vol 115, no. 4, pp. 1684-1694
- [3] Hohmann, V., 2002. „Frequency analysis and synthesis using a Gammatone filterbank,“ *Acustica / acta acustica*, vol 88, no. 3, pp. 433-443
- [4] Wagener, K., Kühnel, V., Kollmeier, B., 1999. „Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests,“ *Zeitschrift für Audiologie/Audiological Acoustics*, 38(1), p. 4-15.