

Modellierung der Lautstärkeschwankung für Normal- und Schwerhörige

Josef Chalupper

Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München

Einleitung

Nach dem Modell der Schwankungsstärke nach Fastl [1] kann für Normalhörende die Schwankungsstärke synthetischer Schalle aus dem Pegelhub ΔL des Mithörschwellen-Zeitmusters vorhergesagt werden. Ergebnisse von Hörversuchen mit Normal- und Schwerhörigen, bei denen die "Lautstärkeschwankung" und der Pegelhub ΔL des Mithörschwellen-Zeitmusters gemessen wurden, zeigen, daß Innenohrschwerhörige trotz eines deutlich kleineren ΔL dieselbe Lautstärkeschwankung wahrnehmen können wie Normalhörende [2]. Da dieses Verhalten durch die steileren Lautheitsfunktionen der Schwerhörigen ("Recruitment") erklärt werden kann, wurde ein erweitertes Modell der Schwankungsstärke vorgestellt:

$$F \sim \frac{\int \Delta L(z) \cdot r(z) dz}{\frac{f_{\text{mod}}}{4 \text{ Hz}} + \frac{4 \text{ Hz}}{f_{\text{mod}}}} \quad (1)$$

wobei $r(z)$ die (zum Normalhörenden) relative Steigung der spezifischen Lautheitsfunktionen ist. Die Lautstärkeschwankung zeigt im Gegensatz zur Schwankungsstärke keine Bandpaßabhängigkeit von der Modulationsfrequenz. Durch Weglassen des Nenners in (1) erhält man somit ein Maß für die Lautstärkeschwankung.

Da es möglich ist, den Pegelhub ΔL mithilfe des Lautheitsmodells nach Zwicker [1] zu bestimmen, wird im Folgenden ein Verfahren vorgestellt, das aus einem Zeitsignal die Lautstärkeschwankung für Normal- und Schwerhörige berechnen und außerdem frühere Untersuchungen zur Schwankungsstärke nachvollziehen kann.

Modellierung der Lautstärkeschwankung

Die Berechnung der Lautstärkeschwankung wird in zwei Schritten durchgeführt. Zuerst werden mit einem modifizierten Lautheitsmodell (s. Abb.1) die Kanallautheiten berechnet, aus denen danach über den Pegelhub ΔL die Lautstärkeschwankung (s. Abb.2) abgeleitet wird.

I. Modifiziertes Lautheitsmodell nach Zwicker

Das Freifeldübertragungsmaß berücksichtigt die Schallübertragung durch Außen- und Mittelohr. Die Frequenzgruppenfilterbank wurde in der vorliegenden Implementierung in Form einer Fourier-t-Transformation (FTT) [3] [4] realisiert. Als Analysefrequenzen wurden die Frequenzgruppenmittelfrequenzen gewählt. Die mit einem Zeitfenster vierter Ordnung realisierten Analysefilter haben eine 3dB-Bandbreite von 0,887 Bark. Bei diesen Parametereinstellungen ist die FTT-Analyse identisch mit einer Gammatone-Filterbank, deren Filter eine äquivalente rechteckige Breite von 1 Bark haben [5].

Die zeitlichen Hüllkurven der Bandpaßsignale werden mit einem gehörgerechten Zeitfenster [6] extrahiert. Die Form dieses Fensters ist dem FTT-Zeitfenster sehr ähnlich, seine äquivalente rechteckige Dauer wurde jedoch - im Gegensatz zum FTT-Fenster - für alle Analysefrequenzen auf 8 ms festgelegt. Der Vorschub beträgt 2 ms.

Im nächsten Schritt werden die Erregungspegel in Kernlautheiten transformiert [Anmerkung: $L_E = 10 \log(E/E_0)$ dB]:

$$N'(z) = N_0 \left(\frac{E_{\text{THQ}}(z)}{s(z)E_0} \right)^{0,23} \left[\left(1 - s(z) + s(z) \frac{E(z)}{E_{\text{THQ}}(z)} \right)^{0,23} - 1 \right] \quad (2)$$

Um das Lautheitsmodell nach Zwicker auf Schwerhörige zu erweitern, wird gemäß dem Zwei-Komponenten-Ansatz nach Moore der Hörverlust in einen dämpfenden und einen expansiven Anteil aufgeteilt [7]. Der dämpfende Anteil wird vom Erregungspegel abgezogen, während der expansive Anteil zu E_{THQ} in (2) addiert wird. Die Aufteilung des Hörverlusts wird anhand einer Hörfeldskalierung mit der Methode nach Levenberg-Marquardt erreicht [8].

Die Nachverdeckung wird mit einem nichtlinearen Tiefpaß modelliert. Im Gegensatz zu einer früheren Implementierung [9] wurde zusätzlich die Pegelabhängigkeit der Nachverdeckung berücksichtigt.

Bei der folgenden Berechnung der Flankenverdeckung gemäß DIN 45631 findet ein Übergang auf eine Auflösung von 0,1 Bark statt.

Die Kanallautheiten erhält man, indem man das resultierende spezifische Lautheits-Tonheitsmuster barkweise über der Tonheit aufsummiert und mit einem Tiefpaß, der die zeitliche Lautheitsintegration nachempfiehlt, bewertet. Wird über die gesamte Bark-Skala spektral integriert, erhält man die zeitabhängige Lautheit.

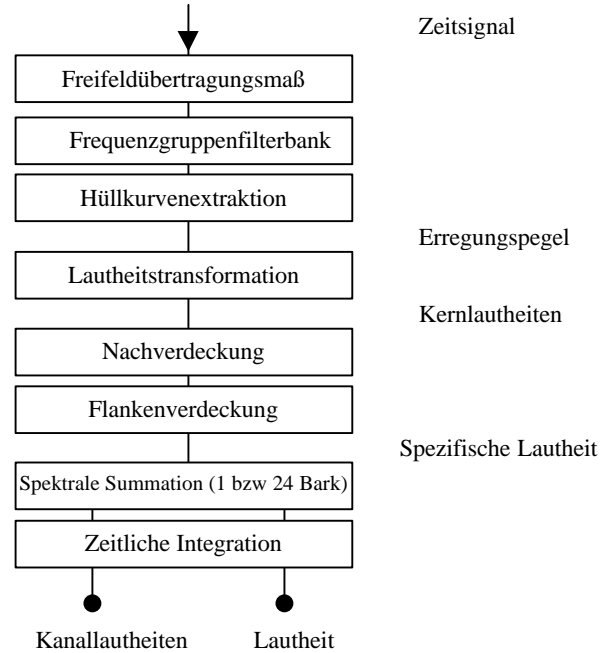


Abb.1: modifiziertes Lautheitsmodell

II. Modell der Lautstärkeschwankung

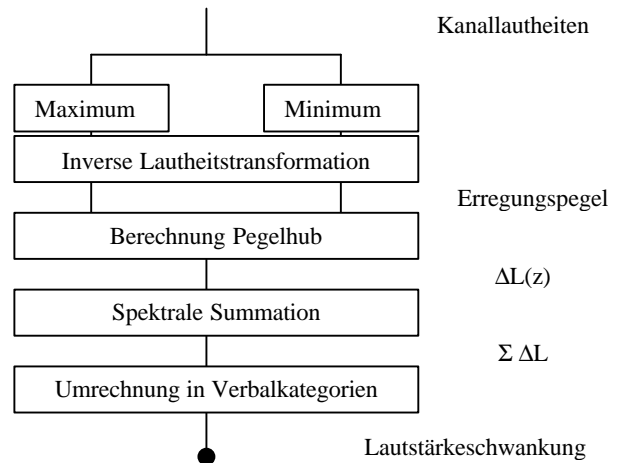


Abb. 2 Modell der Lautstärkeschwankung

In jeder der 24 Frequenzgruppen werden zunächst minimale und maximale Kanallautheit bestimmt. Durch Auflösung von (2) nach L_E erhält man die inverse Lautheitstransformation, mit deren Hilfe minimale und maximale Kanallautheiten in Erregungspegel umgerechnet werden.

Anstatt hierbei die spezifischen Lautheitsfunktionen der Schwerhörigen zugrunde zu legen und danach wie in (1) $\Delta L(z)$ mit $r(z)$ zu multiplizieren, werden auch für Schwerhörige die inversen Lautheitsfunktionen von Normalhörenden benutzt, weshalb auf die Multiplikation mit $r(z)$ verzichtet werden kann. Dies führt zum gleichen Resultat wie in (1) und unterstreicht außerdem die Vorstellung, daß die Lautstärkeschwankung bei Normal- und Schwerhörigen in gleicher Weise zustandekommt.

Im Block "Berechnung Pegelhub" werden zunächst die Differenzen zwischen maximalen und minimalen Erregungspegeln mit den Korrelationsfaktoren zwischen benachbarten Kanälen gewichtet, um zu verhindern, daß bei stationären breitbandigen Rauschen die durch die Frequenzgruppenfilter verursachten Eigenfluktuationen zu hohen berechneten Lautstärkeschwankungen führen. Nach einer Begrenzung auf 30 dB werden die Pegelhub $\Delta L(z)$ gemäß (1) über der Tonheit zu $\Sigma \Delta L$ integriert. Soll nun die Schwankungsstärke berechnet werden, so muß noch deren bandpaßartige Abhängigkeit von der Modulationsfrequenz berücksichtigt werden. Zur Modellierung der Lautstärkeschwankung wird $\Sigma \Delta L$ über einen funktionalen Zusammenhang in Verbalkategorien umgerechnet.

Evaluierung

I. Lautstärkeschwankung bei amplitudenmodulierten Rauschen

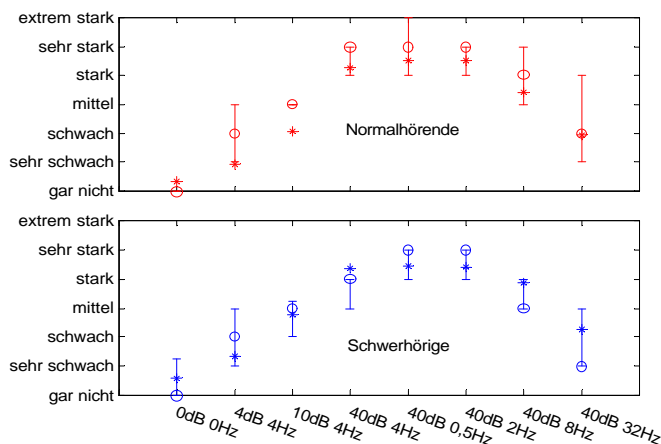


Abb. 3: Berechnete(*) und gemessene(o) Lautstärkeschwankung von AM-GLR für angegebene Modulationstiefen (dB) und -frequenzen(Hz)

Abb. 3 zeigt Mediane (o) und Interquartile der Lautstärkeschwankung für "Gleichmäßig Lautes Rauschen" (GLR), das mit verschiedenen Modulationsfrequenzen und -tiefen amplitudenmoduliert wurde [2]. Bei gleicher Modulationstiefe wird für Normal- und Schwerhörige annähernd dieselbe Lautstärkeschwankung gemessen. Die Funktion zur Umrechnung von $\Sigma \Delta L$ in Verbalkategorien wurde aus den Werten der Normalhörenden abgeleitet und unverändert auch auf die schwerhörigen Probanden angewendet. Die berechneten Mediane (*) stimmen sowohl für Normal- als auch Schwerhörige sehr gut mit den gemessenen überein.

II. "Modulation Matching" bei "pegelmodulierten" Sinustönen

Im Gegensatz zu obigem Hörversuch wurde bei Experimenten mit Sinustönen, deren Pegel sinusförmig moduliert wurde, festgestellt, daß Schwerhörige in der Regel eine geringere Modulationstiefe d als Normalhörende benötigen, um Sinustöne als gleich stark schwankend wahrzunehmen [10]. Da diese Messungen aufgrund fehlender Hörfeldskalierungen mit dem vorliegenden Modell nicht nachvollzogen werden können, werden hier Berechnungen für einen typischen Altersschwerhörigen vorgestellt. Die Steigung seiner Lauthheitsfunktion bei 1000 Hz ist ungefähr doppelt so groß wie bei Normalhörenden, weshalb ein pegelmodulierter 1 kHz - Sinuston bei halber Modulationstiefe die gleiche Lautstärkeschwankung wie bei Normalhörenden hervorrufen sollte. In Abb.4 (links) werden Modellvorhersagen für Normalhörende (Quadrate) verglichen mit Berechnungen für den Schwerhörigen bei gleicher (Rauten) und halber (Dreiecke) Modulationstiefe. Zwar wird vom Modell bei gleicher Modulationstiefe stets eine größere Lautstärkeschwankung für den Schwerhörigen als für Normalhörende vorhergesagt, aber die Berechnungen bei halber Modulationstiefe liegen durchweg unter den Werten der Normalhörenden. In Abb. 4 (rechts) sind die jeweiligen $\Delta L(z)$ für den Sinuston mit $f_{mod}=4\text{Hz}$ und $d=12\text{ dB}$ dargestellt. Im Bereich der Kernerregung (8,5 Bark) wird für den Schwerhörigen bei halbem d nahezu das gleiche ΔL erreicht, bei 7,5 Bark dagegen 10dB weniger. Da die Lautstärkeschwankung durch spektrale Integration

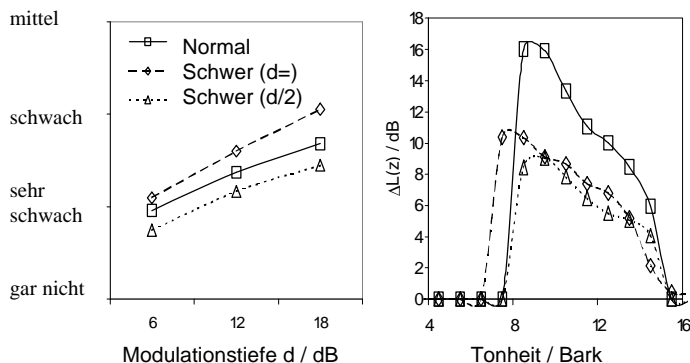


Abb. 4: berechnete Lautstärkeschwankung bei "Modulation Matching" (links) und $\Delta L(z)$ bei $d=12\text{ dB}$ (rechts). $f_{mod}=4\text{ Hz}$

von $\Delta L(z)$ berechnet wird, folgt daraus eine Unterschätzung der Lautstärkeschwankung von Sinustönen für Schwerhörige. Die Ursache hierfür ist, daß bei der Berechnung der Flankenverdeckung im Lauthheitsmodell nach Zwicker nur die obere Flanke berücksichtigt wird. Für pegelmodulierte Breitbandrauschen wird vom Modell bei halbem d nahezu gleiche Lautstärkeschwankung vorhergesagt.

III. Diskussion

Bei AM-Rauschen führt also gleiche Modulationstiefe zu gleicher Lautstärkeschwankung bei Normal- und Schwerhörigen, bei Pegelmodulation dagegen reicht eine geringere Modulationstiefe. Dies kann einerseits durch die angehobene Ruhehörschwelle und andererseits durch einen geringeren Einfluß der Nachverdeckung bei Pegelmodulation erklärt werden.

Für die Hörgeräteanpassung bedeutet dies, daß zur Restauration der Lautstärkeschwankung bei pegelmodulierten Schallen extrem kurze Ein- und Ausschwingzeiten benötigt werden, bei amplitudenmodulierten Schallen jedoch eine lineare Verstärkung - quasi unendlich lange Regelzeiten - ausreichend sein kann. Mithilfe des vorgestellten Modells können für einen bestimmten Schall und Hörverlust die benötigten Regelzeiten abgeschätzt werden.

Zusammenfassung und Ausblick

Basierend auf dem Lauthheitsmodell nach Zwicker wurde ein Verfahren vorgestellt, mit dem es möglich ist, für breitbandige synthetische Schalle die Lautstärkeschwankung sowohl für Normal- als auch Schwerhörige vorherzusagen.

Die Erweiterung der Gültigkeit auf Sinustöne erfordert eine verbesserte Berücksichtigung der Flankenverdeckung.

Für die Verwendung des Modells zur Berechnung von Kompressionsregelzeiten sollte außerdem die Gültigkeit für natürliche Schalle überprüft werden.

Literatur

- [1] Zwicker, E., Fastl, H., "Psychoacoustics. Facts and Models", 2nd edition. Springer, Heidelberg, New York, 1999.
- [2] Chalupper, J., "Loudness fluctuation and temporal masking in normal and hearing impaired listeners", J. Acoust. Soc. Am. 105 (2), 1023, 1999.
- [3] Terhardt, E., "Fourier transformation of time signals: conceptual revision", Acustica 57, 242-256, 1985.
- [4] Mummert, M., "Sprachcodierung durch Konturierung eines gehörangepaßten Spektrogramms und ihre Anwendung zur Datenreduktion", VDI Reihe 10, Nr. 522. Düsseldorf, VDI Verlag.
- [5] Chalupper, J., Fastl, H., "Simulation of Hearing Impairment based on the Fourier Time Transformation", Tagungsband ICASSP 2000 (in press).
- [6] Plack, J.P., Moore, B.C.J., "Temporal window shape as a function of frequency and level", J. Acoust. Soc. Am. 87 (5), 2178-2187, 1990.
- [7] Moore, B.C.J., Glasberg, B. R., Vickers, D.A., "Further evaluation of a model of loudness perception applied to cochlear hearing loss", J. Acoust. Soc. Am. 106 (2), 898-907, 1999.
- [8] Marquardt, D.W., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters", J. Soc. Appl. Math. 11 (2), 431-441, 1963.
- [9] Widmann, U., Lippold, R., Fastl, H., "Ein Computerprogramm zur Simulation der Nachverdeckung für Anwendungen in Akustischen Meßsystemen. DAGA 98, 96-97.
- [10] Moore, B.C.J., Wojtczak, M., Vickers, D.A., "Effect of loudness recruitment on the perception of amplitude modulation", J. Acoust. Soc. Am. 100 (1), 481-489, 1996.