

# Adaptive Lautheitsmessung mit dynamischen Frequenzgruppenbreiten

Ludwig Kollenz<sup>1</sup>

<sup>1</sup> *Institut für Musikwissenschaft, Universität Wien, E-Mail: a0325926@unet.univie.ac.at*

## Einleitung

Schon 1957 gab es Arbeiten zur Lautstärkewahrnehmung: Zwicker, Flottorp und Stevens [1] erforschten die Auswertung von Lautstärke durch Frequenzgruppen. Zwicker und Feldtkeller erstellten 1967 eine *willkürliche* Reihung aus 24 Frequenzgruppen[2] (In der Arbeit von 1957 waren noch andere Mittenfrequenzen für die Frequenzgruppen angegeben). Die Tabelle von 1967 wird seither immer wieder zitiert [3] [4] und eine ähnliche Einteilung ist auch bei der Lautstärkenmessung nach Zwicker, dem Terzpegeldiagramm [5] in Verwendung. In diesem Verfahren wird der Frequenzbereich in Terzbänder unterteilt, die oberhalb von 280Hz Frequenzgruppenbreiten entsprechen, unterhalb zusammengefasst werden, um den Frequenzgruppen zu entsprechen. Nach Ermittlung der Pegel aller Frequenzgruppen wird die Verdeckung berechnet und eine Gesamtlautheit errechnet. In der DIN 45631 von 1967 sind diverse Diagramme zur händischen Berechnung für frontale und diffuse Schallfelder abgebildet, die in den Pegelteilungen verschieden sind. Die Frequenzgruppen weisen stets dieselben Grenzfrequenzen auf.

Sein Verfahren verglich Zwicker 1966 [6] mit denen von Stevens und Niese, wobei letztere zu hohe bzw. zu niedrige Ergebnisse im Vergleich zur subjektiv bewerteten Lautstärke lieferten.

Die neuste Fassung der DIN 45631 ist von 2010; einen guten Überblick bietet das Buch von Möser [7]. Heutzutage wird natürlich nicht mehr händisch gearbeitet, sondern es liegt eine CD-ROM mit einem Basic-Programm bei. Zwickers Verfahren wurde seit 1967 kaum verbessert [7]. Dies ist verwunderlich, denn es wurde bereits darauf hingewiesen, dass Frequenzgruppenbreiten veränderlich sind, sich also dynamisch und eben nicht fest verhalten. Daher müsste eine adäquate Lautheitsmessung diesem Umstand Rechnung tragen [8]. Der Autor des vorliegenden Artikels möchte eine Verbesserung des Terzpegeldiagrammes vorschlagen: Die Implementierung dynamischer Frequenzgruppenbreiten.

## Ziele

Ziel ist es ein Verfahren zur Lautheitsmessung zu entwickeln, welches sich auf Basis des Terzpegelverfahrens nach Zwicker besser dem menschlichem Empfinden für Lautstärke annähert. Die adaptiven Frequenzgruppenbreiten sind der erste Schritt, weitere Verbesserungen werden im Ausblick angesprochen. Das Verfahren soll auch vor allem durch Beachtung des zeitlichen Verhaltens des Gehörs bessere Werte als bisherige Ansätze liefern.

## Methode

Vom Autor wurde zur Simulation ein Script in Scilab verfasst, welches alle 10ms dynamisch Frequenzgruppen aufbaut, die Verdeckung berechnet und dann innerhalb der Frequenzgruppen die Lautstärke auswertet, um schließlich einen finalen Wert zu liefern.

Zunächst wird das Eingangssignal (derzeit nur mono) in Blöcke von 10ms Dauer zerlegt und mit einer FFT in den Bildbereich transformiert. Alle Komponenten des Amplitudenfrequenzgangs die unterhalb der absoluten Hörschwelle liegen werden entfernt. Von den Pegeln der verbleibenden Komponenten wird der Pegel der Hörschwelle bei dieser Frequenz abgezogen. Ein Peak-Finding-Algorithmus ermittelt die Frequenzen der Teiltöne, welche für die Berechnung der Breite der Frequenzgruppen nötig sind.

Die Ermittlung der Frequenzgruppen erfolgt indem zunächst die Frequenzgruppe des stärksten Teiltones gebildet wird. Alle weiteren Teiltöne innerhalb der Grenzen werden für die Bildung weiterer Frequenzgruppen von dem Algorithmus nicht mehr berücksichtigt. Danach wird die Frequenzgruppe des nächst-schwächeren Teiltones gebildet. Sollten zwei Frequenzgruppenbreiten überlappen wird die des schwächeren Teiltones zugunsten der stärkeren verkürzt. Das basiert auf der Wanderwellentheorie [9], eine stärkere Resonanz auf der Basilarmembran müsste eine schwächere überlagern und das Gehirn müsste es entsprechend auswerten.

Nach Bildung der Frequenzgruppenbreiten wird deren Pegel aus dem umfassten Spektralbereich ermittelt. Die Bildung dieses Pegels ist Grundlage für die Verdeckung nach Terhardt ([7] bzw. [4]).

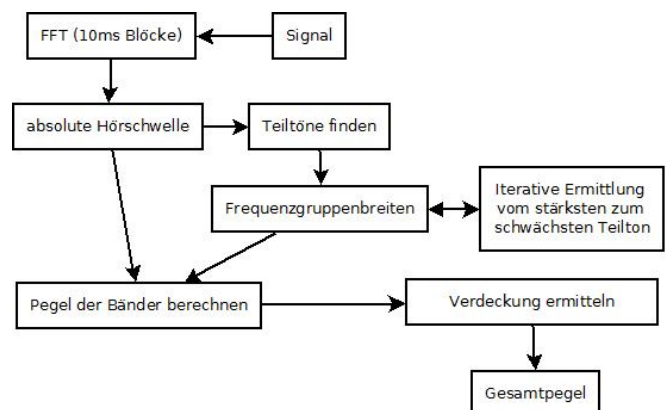


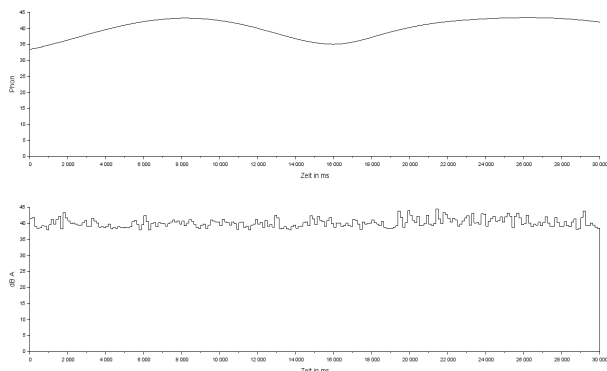
Abbildung 1: Flussdiagramm des Algorithmus.

Abschließend wird durch Pegeladdition die Gesamtlautstärke in Phon ausgegeben. Problematisch ist

immer noch die Einheit in dB/Bark, da die Frequenzgruppenbreiten hier variabel sind. Die Ergebnisse des Programms sind daher noch nicht als valide zu betrachten. Die Ausgabe in Sone, also die Berechnung der Gesamtlautheit, ist noch nicht möglich, da auf die Einheit mel/bark (Tonheit  $z$ ) zurückgegriffen wird, was bei einem voll dynamischen System in der Form nicht zulässig ist.

## Ergebnisse

Der Algorithmus wurde an einer Aufnahme des Donauinselfests 2014 getestet. Der Vorverstärker der Soundkarte wurde zusammen mit dem Messmikrofon auf 114 dB SPL = -6dB FS kalibriert. Analysiert wurden versuchs halber 30 Sekunden. Im Vergleich zu einer A-Bewertung dieser Aufnahme deren Mittelwert bei etwa 40dB A lag bzw. bei etwa 60dB SPL (das Spektrum zeigt vornehmlich tiefe Frequenzen, die entsprechend stark abgeschwächt werden) liefert der adaptive Algorithmus basierend auf Zwickers Methode durchschnittlich 39 Phon, weist allerdings je nach Spektrum teilweise starke Abweichungen zu dB A auf.



**Abbildung 2:** Oben der adaptive Algorithmus; unten dB A fast(125ms). Die obere Kurve ist deutlich glatter durch die 10ms Analysedauer.

## Ausblick

Die Forschung soll sich auf die Ausbildung der Frequenzgruppen konzentrieren, auch die zeitliche Entwicklung der Hörschwelle gilt es noch zu implementieren. Der derzeitige Stand der Software ist nur ein Prototyp bzw. eine Alpha-Version. Das Programm soll im nächsten Schritt um die Nachverdeckung erweitert werden sowie um die Isophonen verschiedener Lautstärken. Damit lässt sich die Anhebung der Hörschwelle simulieren. Dazu muss das Programm um ein „Gedächtnis“ erweitert werden, um ab einem gewissen anliegendem Lautheitspegel auf eine höhere Hörschwelle umzuschalten und bei verringertem Pegel die Schwelle langsam abzusenken.

Die Auflösung der FFT ist derzeit ein weiteres Problem, da bei 192kHz und 10ms Dauer der FFT die Auflösung  $\Delta f$  (=Abstand zweier Spektrallinien) 100Hz beträgt. Das Programm soll in einer späteren Version auf eine Waveletanalyse umgestellt werden. Da sie bei tiefen eine gute und bei hohen Frequenzen eine geringere Auflösung bietet, ähnelt sie dem menschlichen Gehör [10].

Eine weitere Aufgabe wird es sein das Programm für rauschartige Signale zu optimieren. Derzeit werden

ausgeprägte Teiltöne benötigt, die im Rauschen fehlen. Der Algorithmus würde keine oder enorm viele Teiltöne finden, was zu falschen Ergebnissen führt. Um zwischen tonalen oder rauschartigem Charakter zu unterscheiden, ist die Verteilung des Spektrums zu schätzen [4]. Bei Detektion eines Rauschsignals könnte das Programm eine feste Einteilung der Frequenzgruppenbreiten nutzen.

Sobald alle Features implementiert sind und die Ergebnisse als reliabel angesehen werden können, gilt es vergleichende Studien zwischen der hier vorgeschlagenen Erweiterung und dem Verfahren nach der DIN 45631 durchzuführen. Gibt es keine signifikanten Unterschiede kann festgestellt werden, dass die Variabilität der Frequenzgruppen unerheblich ist. Sollten aber Abweichungen entstehen, gilt es die Forschung weiter zu verfolgen. Dann wären Hörversuche aufschlussreich, um zu bestimmen, wie gut die gemessenen Werte mit der Lautheitswahrnehmung von Probanden korrelieren.

Auch der Unterschied zur ERB-Skala mit ihren 40 Bändern [7] wäre aufschlussreich, deren Herangehensweise einer Überabtastung gleicht. Der Unterschied zwischen Bark, ERB und dem dynamischen Verfahren müsste sich vornehmlich in der Verdeckung zeigen und auch darin, dass der adaptive Algorithmus unterschiedliche Breiten der Frequenzgruppen zulässt, die Auswertung also verschieden ist.

Abschließend bleibt zu bemerken, dass hier unter Umständen viel Forschung wartet.

## Literatur

- [1] Zwicker, E.; Flottorp, G.; Stevens, S.: Critical Band Width in Loudness Summation. JASA Vol. 29 (1957), No. 5, S. 548 – 557.
- [2] Zwicker, E.; Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. Hirzel, Stuttgart, 1967.
- [3] Fastl, H.; Zwicker, E.: Psychoacoustics. Springer, Berlin u.a., 2007.
- [4] Zölzer, U.: Digitale Audiosignalverarbeitung, Teubner, Stuttgart u.a., 2008.
- [5] DIN 45631 1967
- [6] Zwicker, E.: Lautstärkeberechnungsverfahren im Vergleich. in Acustica, Vol. 17 (1966), S. 278 – 284.
- [7] Möser, M. (Hrsg.): Messtechnik der Akustik. Springer, Berlin u.a., 2010.
- [8] Reuter, Ch.: Der Einschwingvorgang nichtperkussiver Musikinstrumente. Peter Lang, Berlin u.a., 1995.
- [9] Keidel, W. D. (Hrsg.): Physiologie des Gehörs. Thieme, Stuttgart, 1975.
- [10] Grünigen, D. Ch. von: Digitale Signalverarbeitung. Fotorotar, Egg, 2008.