

# Das spezifische Lautheits-Zeitmuster als gehörgerechte Klangdarstellung

Christian Sabrautzky<sup>1</sup>, Robert Mores<sup>2</sup>

<sup>1</sup> Email: christian@sabrautzky.com

<sup>2</sup> Fakultät Design Medien und Information, HAW Hamburg, Deutschland, Email: mores@mt.haw-hamburg.de

## Einleitung

Bei der Berechnung der spezifischen Lautheit werden sowohl die spektralen als auch die zeitlichen Eigenschaften des Gehörs berücksichtigt. Der zeitliche Verlauf der spezifischen Lautheit stellt eine geeignete Basis für eine gehörgerechte Darstellung von Klängen dar.

Für die Ermittlung der Lautheit eines akustischen Signals existieren verschiedene Modelle, welche die Gesamtlautheit durch spektrale Integration der spezifischen Lautheiten ermitteln. Zu den etablierten Modellen gehören die von Zwicker und die von Moore beschriebenen Verfahren.

In dieser Arbeit wird das Modell nach Moore implementiert und erweitert, um zu einer gehörgerechteren Vorverarbeitung zu kommen. Die zeitlichen Aspekte werden bereits für die spezifische Lautheit berücksichtigt und optimiert, um eine wahrnehmungsnahere Darstellung des Signals zu ermöglichen.

## Lautheit dynamischer Klänge

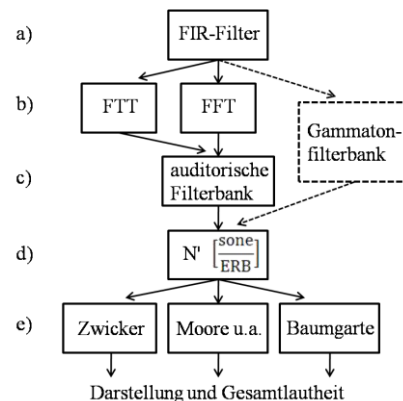
Das Lautheitsmodell von Moore und Glasberg wurde zunächst als Überarbeitung von Zwickers Verfahren veröffentlicht [1]. Es unterscheidet sich insbesondere hinsichtlich der Ermittlung des Erregungsmusters und der Berechnung der spezifischen Lautheit. Die Erregung wird, nach Filterung mit der Übertragungsfunktion von Außen- und Mittelohr, mittels auditorischer Filter bestimmt. Die äquivalente Bandbreite der Filter (equivalent rectangular bandwidth, ERB) ist dabei stets kleiner als die Frequenzgruppenbreite der Bark-Skala, vor allem bei sehr tiefen und hohen Frequenzen. Für jedes Band wird die Erregung in spezifische Lautheit umgerechnet. Die Gesamtlautheit ergibt sich durch Integration der Teillautheiten.

Zwicker u. a. beschreiben Verfahren zur Nachbildung der zeitlichen Verarbeitung im Gehör anhand von analogen Schaltungen. Eine aktuelle Implementierung von Zwickers Lautheitsmodell für dynamische Klänge wird von Chalupper und Fastl beschrieben [2].

Auch für das Modell von Moore und Glasberg existiert eine Variante für dynamische Signale [3]. Im Anschluss an die spektrale Analyse mittels schneller Fouriertransformation (FFT), wird für jede Millisekunde das Erregungsmuster anhand der auditorischen Filterbank bestimmt und in die spezifische Lautheit überführt. Die zeitliche Verarbeitung der Gesamtlautheit fasst Nachverdeckung und zeitliche Integration in einer automatischen Verstärkungssteuerung zusammen, in dem für steigende und fallende Signalabschnitte unterschiedliche Zeitkonstanten gewählt wurden. Abhängigkeit von der Dauer und dem Pegel des Signals werden nicht berücksichtigt.

## Erweiterung und Realisierung des Verfahrens

Das vorliegende Verfahren basiert auf dem Lautheitsmodell für zeitvariable Klänge nach Moore und Glasberg [3]. Für die Darstellung des spezifischen Lautheits-Zeitmusters als gehörgerechtes Spektrogramm sind einige Anpassungen und Erweiterungen vorgenommen.



**Abbildung 1:** Blockdiagramm des erweiterten Verfahrens mit den optionalen Komponenten; a) Außen- und Mittelohr-Filterung, b) spektrale Analyse, c) Erregungsmuster, d) Lautheitskompression, e) zeitliche Verarbeitung.

## Spektrale Analyse

In [3] werden zur spektralen Analyse sechs FFT unterschiedlicher Blocklängen verwendet, um die Zeit- und Frequenzauflösung des Gehörs in verschiedenen spektralen Bereichen anzunähern. Hier kommt es zu systematischen Fehlern bei der Kombination der Teilspektren. Aufgrund der unterschiedlichen Fensterbandbreiten kann es dazu kommen, dass Komponenten im Übergangsbereich zweimal im Gesamtspektrum enthalten sind oder verloren gehen.

Als Alternative zur FFT wurde eine Fourier-t-Transformation (FTT) zweiter Ordnung implementiert [4]. Entgegen der festen Fensterbreite bei der FFT kommt hier ein spezielles, frequenzabhängiges Analysefenster zum Einsatz. Um optimal mit der auditorischen Filterbank abgestimmt zu sein, wurde die Fensterfunktion an die ERB der Filter angepasst.

Optional kann ein Filtern im Zeitbereich mit einer Gammatonfilterbank erfolgen [5]. Vorteil dieser Variante ist der direkte Übergang vom Zeitsignal zum Erregungsmuster mit einer sehr guten Zeitauflösung. Sie eignet sich dennoch nicht für die Ermittlung der spezifischen Lautheit dynamischer Klänge, da es keine einfache Möglichkeit zur pegelabhängigen Steuerung der Filterflanken gibt.

## Zeitliche Verarbeitung

Da die hier im Vordergrund stehende Größe die spezifische Lautheit ist, wird die zeitliche Verarbeitung in jedem Filterband und nicht erst nach der spektralen Integration durch-

geführt. Da das Modell von Moore und Glasberg zudem ein stark vereinfachtes Verfahren zur Nachbildung der temporalen Effekte enthält, wurden zum Vergleich unterschiedliche Modelle der zeitlichen Verarbeitung eingesetzt.

Daher wurde als Alternative ein von Zwicker beschriebener, nichtlinearer Tiefpass herangezogen, der die Abhängigkeit der Nachverdeckung von der Signaldauer berücksichtigt. Eine digitale Implementierung ist in [6] beschrieben. Diese wurde mit leichten Änderungen übernommen.

Alternativ hierzu wurde eine Methode implementiert, die neben der Abhängigkeit von der Dauer auch den Einfluss des Signalpegels auf den Verlauf der Nachverdeckung nachbilden soll. Dies geschieht durch zwei zu addierende Exponentialfunktionen, deren Zeitkonstanten und Gewichtung von Dauer und Pegel gesteuert werden. Das zugrundeliegende RC-Netzwerk, das mit Hilfe rekursiver Filter umgesetzt wurde, wird von Baumgarte beschrieben [7].

Die zeitliche Verdeckung wird durch einen einfachen Tiefpass 1. Ordnung realisiert. Je nach Literaturquelle variiert die Zeitkonstante dieses Filters zwischen ca. 20 ms und 50 ms. Vergleichende Messungen zeigen die beste Übereinstimmung mit Ergebnissen aus veröffentlichten Studien für eine Zeitkonstante um 20 ms, was einer Grenzfrequenz von ca. 8 Hz entspricht.

## Verifikation

Aufgrund der vorgenommenen Änderungen mussten die Konstanten in den Gleichungen zur Berechnung der spezifischen Lautheit angepasst werden, sodass ein kalibriertes, statisches Signal mit einer Frequenz von 1 kHz und einem Schallpegel von 40 dB eine Lautheit von 1 sone erzeugt.

Weitere Tests mit reinen sowie komplexen Tönen unterschiedlicher Frequenz, Pegel und Dauer zeigen für die Gesamtlautheit und die Teillautheiten der einzelnen Bändern gute Übereinstimmungen mit den Ergebnissen von Moore.

Zu Abweichungen kam es erwartungsgemäß im zeitlichen Verlauf der Lautheit durch Verwendung der alternativen Verfahren für die Nachbildung der zeitlichen Verarbeitung. Die mittels des nichtlinearen Tiefpasses nach Zwicker errechneten Werte zeigen die beste Übereinstimmung mit Literaturwerten. Das Verfahren nach Baumgarte führt in der Regel zu einem relativ langsamen Abklingen der Lautheit, welches die in der Literatur angegebenen Grenzwerte überschreitet.

Problematisch ist die Abbildung amplitudenmodulierter Signale. Für Signale mit niedriger Modulationsrate weichen alle Verfahren von den aus der Literatur bekannten Daten ab. Allerdings unterscheiden sich hier auch die Ergebnisse verschiedener Studien. Eine mögliche Näherung stellt der in [3] beschriebene langfristige Lautheitseindruck („long-term loudness“) dar.

## Diskussion

Das vorgestellte Verfahren ermittelt die spezifische Lautheit und ihren zeitlichen Verlauf nach anerkannten psychoakustischen Modellen. Bisher sind noch keine eigenen Hör-

versuche durchgeführt worden und das Modell zunächst anhand von veröffentlichten Forschungsdaten verifiziert.

Diese Verifikation ist durch Widersprüche und Unschärfen innerhalb sowie zwischen den Studien erschwert. Zudem sind die in Modellen verwendeten Werte gemittelt und basieren häufig auf Schätzungen oder Überlegungen, die nicht weiter belegt sind. Anderweitig bekannte psychoakustische Eigenschaften des Gehörs werden nur zum Teil beachtet.

Daher ist die tatsächliche Genauigkeit der Ergebnisse des vorgestellten Verfahrens abschließend nur schwer einzuschätzen. Ein direkter Vergleich mit den Ergebnissen eines anderen dynamischen Lautheitsmodells wie [2] wäre wünschenswert. Die ermittelten Werte stellen lediglich Näherungen dar, von denen die persönliche Wahrnehmung teils erheblich abweichen kann.

## Zusammenfassung

Die interaktive Darstellung des Zeitmusters der spezifischen Lautheit ermöglicht die Betrachtung der für den Lautheitseindruck relevanten Klanganteile einfach strukturierter Klänge. Bei zeitlich komplexen Klängen wie dem Vibrato einer Geige hingegen bestehen noch erhebliche Abweichung zwischen der Abbildung und dem tatsächlichen Höreindruck. Hier ist eine weitere Optimierung des Verfahrens oder eine weitere Verarbeitung der spezifischen Lautheit notwendig.

## Dank

Das BMBF fördert das Projekt „Wissenschaftliche Begleitung der klang-ästhetischen Weiterentwicklung der Violine“.

## Literatur

- [1] Moore, B. C., & Glasberg, B. R.: A Revision of Zwicker's Loudness Model. *Acta Acustica united with Acustica*, 82 (1996), 335-345
- [2] Chalupper, J., & Fastl, H.: Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners. *Acta Acustica united with Acustica*, 88 (2002), 378-386
- [3] Glasberg, B. R., & Moore, B. C.: A Model of Loudness Applicable to Time-Varying Sounds. *Journal of the Audio Engineering Society*, 50 (5) (2002), 331-342
- [4] Schlang, M., & Mummert, M.: Die Bedeutung der Fensterfunktion für die Fourier-t Transformation als gehörgerechte Spektralanalyse. *Fortschritte der Akustik* (1990), 1043-1046
- [5] Hohmann, V.: Frequency analysis and synthesis using a Gammatone filterbank. *Acta Acustica united with Acustica*, 88 (2002), 433-442
- [6] Widmann, U., Lippold, R., & Fastl, H.: Ein Computerprogramm zur Simulation der Nachverdeckung für Anwendungen in akustischen Meßsystemen. *Fortschritte der Akustik* (1998), 96-97
- [7] Baumgarte, F.: Ein psychophysiologisches Gehörmodell zur Nachbildung von Wahrnehmungsschwellen für die Audiocodierung. Universität Hannover. (2000)