

Hörmodell zur Detektion von Tonhaltigkeit in transienten Signalen

Peter Volk¹, Arne Oetjen², Steven van de Par³

¹ Leopold Kostal GmbH & Co. KG, 58513 Lüdenscheid, Deutschland, Email: p.volk@kostal.com

² Acoustics Group, Exzellenzcluster "Hearing4all", Carl-von-Ossietzky Universität, Oldenburg, Deutschland

³ Acoustics Group, Exzellenzcluster "Hearing4all", Carl-von-Ossietzky Universität, Oldenburg, Deutschland

Einleitung

Die akustische Optimierung von technischen Produkten ist heutzutage ein Standardprozess in der Komponenten- und Bauteilentwicklung bei vielen Herstellern. Es gilt nicht nur den Schalldruckpegel zu reduzieren, sondern auch den Klang subjektiv zu verbessern. Um die erhöhte Störwirkung tonaler Komponenten in einem Geräusch zu analysieren wird in der DIN 45681:2005 ein FFT-basierter Algorithmus vorgestellt. Der Algorithmus zeigt jedoch deutliche Schwächen, wenn das Signal zeitliche Änderungen, insbesondere in der Frequenz, aufweist, da die verwendeten Blocklängen zu hoch sind. Auch eine modifizierte Variante nach [4] mit kürzeren Blocklängen ($\tau = 100$ ms) detektiert zeitlich veränderliche tonale Komponenten nur sehr unzureichend. Die Ergebnisse einer solchen Analyse nach DIN 45691:2005 ist in der unteren Grafik von Abbildung 1 dargestellt.

Motivation

Die Norm findet vor allem in der Industrie wie zum Beispiel bei Fahrzeugherstellern und deren Zulieferern Anwendung. Die tonalen Elemente der einzelnen Bauteile sind unter anderem drehzahlabhängig und somit zeitlich variabel. Ein Turbolader erzeugt in einem Zeitbereich von ca. 1,4 s hochfrequente tonale Flanken, welche einen Frequenzbereich von ca. 16 kHz abdecken. Eine korrekte Analyse der Tonhaltigkeit mit den etablierten FFT-basierten Algorithmen ist aufgrund der unausreichenden zeitlichen Auflösung nicht möglich. In Zukunft werden, aufgrund der immer leiser werdenden Verbrennungsmotoren, psychoakustische Parameter wie die Tonhaltigkeit an Bedeutung gewinnen und in die Kaufentscheidung von Fahrzeugen einfließen. Demnach ist eine korrekte Detektion von tonalen Elementen unerlässlich.

Aufbau des Hörmodells zur Detektion von Tonhaltigkeit

Da das menschliche Gehör jedoch in der Lage ist solche tonalen Komponenten zu detektieren, wurde ein Hörmodell entwickelt, welche auf physiologischen und psychoakustischen Grundlagen Tonhaltigkeit detektieren kann. Für die Umsetzung wurde Auditory-Modelling-Toolbox [6] verwendet.

Das Front-End des Modells beinhaltet das Envelope-Power-Spectrum-Model nach [1] (siehe Abbildung 2). Um eine bessere Frequenzauflösung zu erlangen, wurde die Anzahl der auditorischen Kanäle um den Faktor Fünf erhöht. Die Mittenfrequenzen umfassen einen

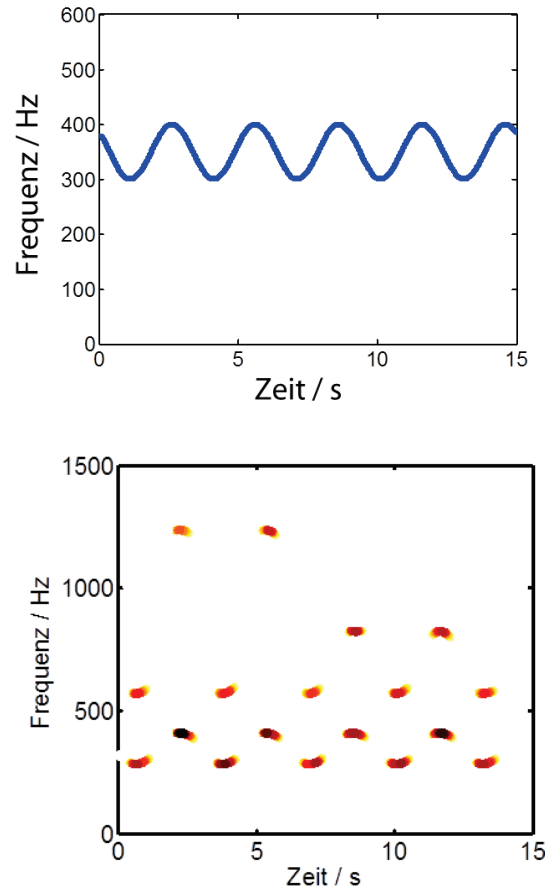


Abbildung 1: Obere Grafik: Grundfrequenz in der Amplitude über die Zeit konstanten harmonischen Tonkomplexes mit periodisch veränderlicher Frequenz. Untere Grafik: Ergebnis der Tonhaltigkeitsanalyse nach [4]. Tonale Komponenten werden nur an den Zeitpunkten detektiert, an denen der Gradient gering ist.

Frequenzbereich von 0,1 bis 20 kHz. Zudem wurde der Bandbreitenfaktor, welcher in der Filterbank nach [2] zu wählen ist auf 75 % eingestellt, was zu einer verbesserten Frequenzselektivität führt. Die Einhüllende jedes auditorischen Kanals wird über die Hilbert-Transformation berechnet. Die Modulationsfilterbank beinhaltet für jeden Kanal lediglich einen 0 Hz-Filter. In [3] wurde gezeigt, dass das Modulationsspektrum von bandbegrenztem Rauschen mit steigender Bandbreite ebenfalls breiter wird. Die Leistung des Rauschens fällt zu höheren Modulationsfrequenzen ab und besitzt eine Dreiecksform. Unmodulierte Sinustöne hingegen besitzen somit ausschließlich einen Peak in der 0 Hz Komponente. Während sich die gesamte Modulationsleistung eines Sinustones in die-

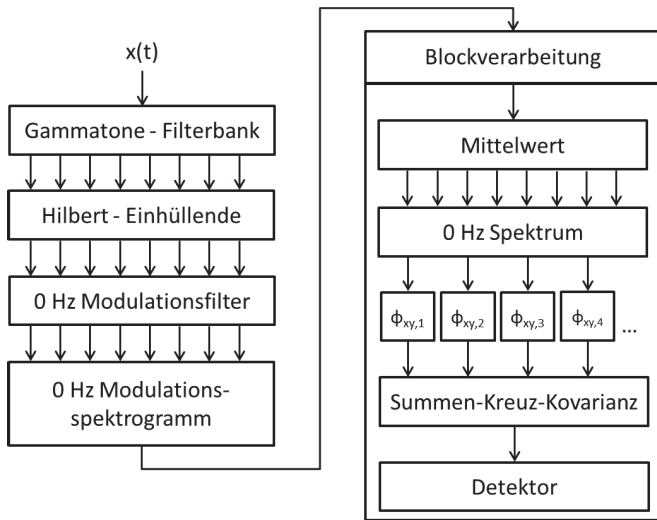


Abbildung 2: Schematische Darstellung der Verarbeitungsstufen des Hörmodells, Front-End: EPSM nach [1], Back-End: Detektion der Tonhaltigkeit über eine Kreuz-Kovarianzberechnung, Blocklänge: 50 ms, Overlap: 25 %

ser 0 Hz Komponente befindet, teilt sich die Leistung von Rauschen auf den Modulationsfrequenzbereich auf. Sinusartige Signale müssten demnach bei 0 Hz sehr hohe und Rauschen niedrigere Peaks ausbilden. Das Ziel ist es, mit dem Hörmodell Töne über den 0 Hz Modulationsfilter zu detektieren.

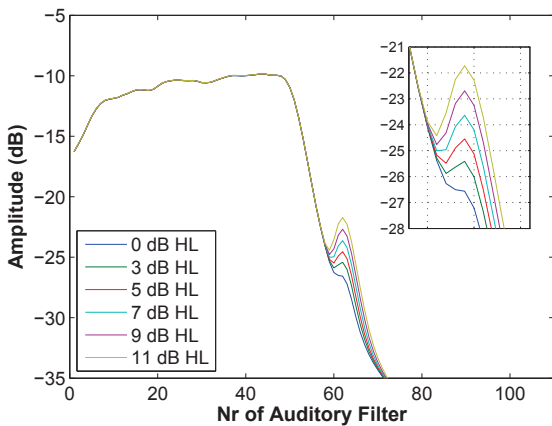


Abbildung 3: 0 Hz Spektrum eines bandbegrenzten Rosa-Rauschens und 1 kHz Sinustones. Die obere Grenzfrequenz des Rauschens liegt 2 ERB unterhalb der Tonfrequenz. Bei ansteigendem Tonpegel nimmt auch die Ausprägtheit des Tones im Erregungsmuster zu.

In Abbildung 3 ist das 0 Hz Spektrum eines zeitlich unveränderlichen Sinustones in bandbegrenztem Rosa-Rauschen dargestellt. Zuvor wurden über einen Hörversuch die Detektionsschwelle ermittelt. Somit ergeben sich Pegel in dB HL. Es ist zu erkennen, dass die Ausprägtheit des Tones mit steigendem Tonpegel zunimmt. Dies ist auch bei schmalbandigem Low-Noise-Noise, welches tonal wahrgenommen wird, zu beobachten. Für die Detektion von Tonhaltigkeit wird eine Kreuz-Kovarianzberechnung durchgeführt. Dafür werden zuvor einzeln unverrauschte Sinusoide, welche als maxi-

mal tonhaltig angenommen werden, in Oktavabständen eingelesen und das Erregungsmuster abgespeichert. Dann wird die Kreuz-Kovarianz aus diesen und dem aktuellem Erregungsmuster berechnet. Dies geschieht für jedes abgespeicherte Erregungsmuster einzeln. Anschließend werden die berechneten Kurven zu einer Summen-Kreuz-Kovarianzfunktion (SKKF) aufaddiert. Die Amplitude der SKKF beschreibt also die Tonhaltigkeit. Der Lag wird in auditorischen Filtern angegeben und kann in die entsprechende Frequenz umgerechnet werden. Eine Interpolation um den Faktor Zehn steigert die Frequenzauflösung. Die SKKF des Signals aus 3 ist in Abbildung 4 dargestellt.

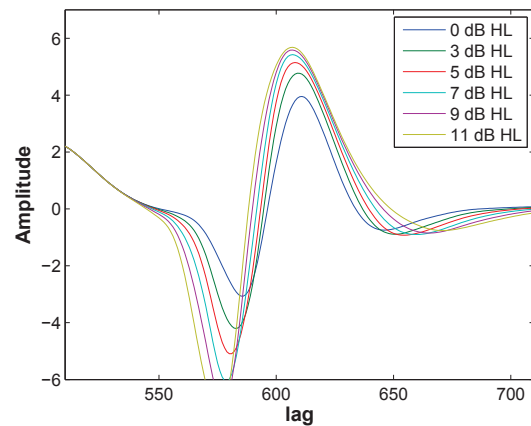


Abbildung 4: Summen-Kreuz-Kovarianzfunktion des Signals aus Abbildung 3 Die Amplitude steigt näherungsweise logarithmisch an, d.h. die Tonhaltigkeit steigt bei hohen Tonpegeln weniger stark an.

Bei erhöhtem Tonpegel, nimmt auch die Amplitude der SKKF zu. Der Anstieg erfolgt jedoch nicht mehr linear, sondern annähernd logarithmisch. Bei steigender Ausprägtheit gleicht der Ton im Erregungsmuster immer stärker einem unverrauschten Sinuston. Die Tonhaltigkeit steigt demnach zunächst stark an. Der Anstieg der Tonhaltigkeit flacht anschließend mit zunehmenden Tonpegel immer weiter ab. Eigenfluktuationen in niederfrequenten auditorischen Kanälen verursachen kleinere tonale Ausprägungen im Erregungsmuster. Folgedessen wurde bei einem empirisch ermittelten Wert eine Schwelle definiert, die die SKKF überschreiten muss, damit das Element als Ton klassifiziert wird. Da es sich bei diesem Modell um ein Hörmodell handelt, wird an dieser Stelle auf eine Umrechnung in tonalen Überschuss in dB verzichtet. Das auditorische System liefert keine numerischen Werte in dB, sondern es findet lediglich eine subjektive Einschätzung auf einer nominalen Skala über die Ausprägtheit des Perzeptes statt.

Problematik bei der Anwendung auf zeitlich veränderliche Signale - 0 Hz Modulationsfilter vs. zeitliche Auflösung

Bei der Analyse zeitlich veränderlicher Signale spielen vor allem die Einschwingzeiten der verwendeten Filter eine wichtige Rolle. Sind diese zu lang, so ist das Fil-

ter zu träge um kurze Zeitsignale korrekt erfassen zu können. Das 0 Hz Modulationsfilter ist ein Butterworth-Filter zweiter Ordnung und besitzt eine Grenzfrequenz von 2,5 Hz. Die hohe Flankensteilheit zieht ebenfalls eine höhere Einschwingzeit mit sich, was in der oberen Grafik von Abbildung 5 gezeigt ist. Erst nach ca. 500 ms ist die Einschwingphase vollständig beendet.

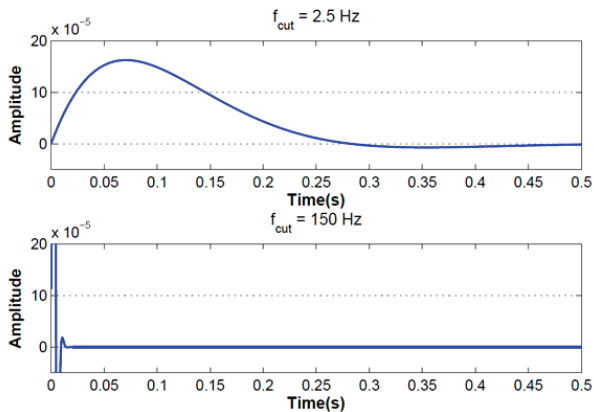


Abbildung 5: Vergleich zwischen der Einschwingzeit des original 0 Hz Modulationsfilters (obere Grafik) und des Modulationstiefpasses (untere Grafik) des Hörmodells

Demnach wurde die Grenzfrequenz auf 150 Hz nach dem Modell des leaky integratoräus [7] angehoben. Somit ist ein schneller Einschwingvorgang gewährleistet, welcher in der unteren Grafik von Abbildung 5 veranschaulicht ist. Durch diese Modifikation lassen sich nun auch transiente Signale auf die Tonhaltigkeit analysieren (siehe Abbildung 6). Es sei jedoch erwähnt, dass es sich hierbei nicht mehr um ein 0 Hz-Filter handelt, sondern um einen Modulationstiefpass, welcher hinter jeden auditorischen Kanal geschaltet wurde.

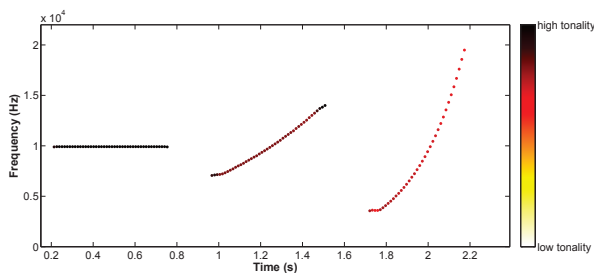


Abbildung 6: Tonhaltigkeitsspektrogramm des SKKF-Hörmodells, Unverrauschte logarithmische Sinus-Sweeps bei $f_c = 10 \text{ kHz}$, Länge: 0.5 s

Vergleich der Berechnungen aus DIN 45681:2005 und SKKF-Hörmodell

In Abbildung 1 wurde bereits gezeigt, dass die Detektion von zeitlich veränderlichen tonalen Elementen lediglich an den Zeitpunkten funktioniert, an denen der Gradient der Frequenzänderung gering ist. In der obere Grafik von Abbildung 1 ist das Spektrogramm des selben Signals dargestellt. Die untere Grafik zeigt den Output des Hörmodells. Es ist deutlich zu erkennen, dass alle

Harmonischen des verrauschten Tonkomplexes durchgehend detektiert werden. Es sind jedoch auch die bereits erwähnten Artefakte, hervorgerufen durch Eigenfluktuationen in niederfrequenten auditorischen Kanälen, zu vernehmen.

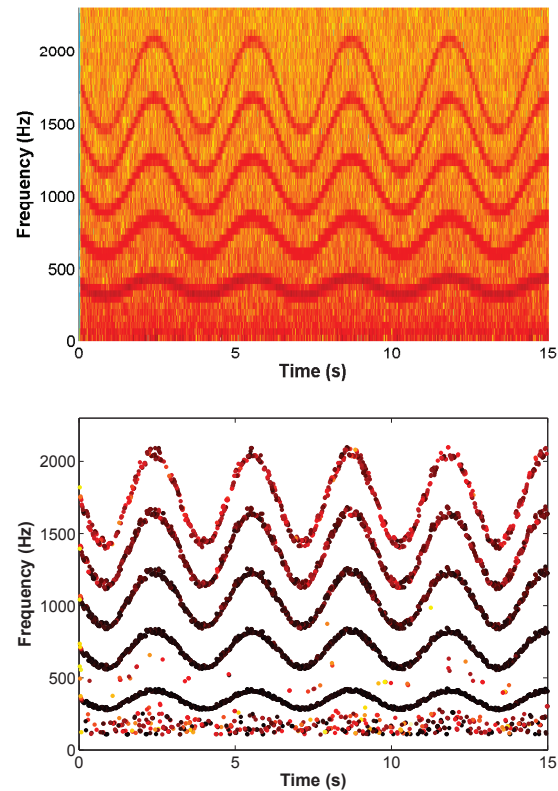


Abbildung 7: Obere Grafik: Spektrogramm des Soundbeispiels aus Abbildung 1. Untere Grafik: Tonhaltigkeitsspektrogramm des selben Soundbeispiels, berechnet mit dem SKKF-Hörmodells.

Anwendung des Hörmodells auf reale Signale

Es konnte bereits gezeigt werden, dass das vorgestellte Hörmodell transiente synthetische tonale Elemente detektiert kann. Das Modell soll jedoch vor allem reale Soundsignale hinsichtlich ihrer Tonhaltigkeit analysieren können. In Abbildung 8 sind die Tonhaltigkeitsspektrogramme von Soundbeispielen aus dem DEGA Fachausschuss für Fahrzeugakustik dargestellt. Die obere Grafik zeigt die Tonhaltigkeitsanalyse eines Turboladers, dessen tonale Flanke bei ca. 6 kHz startet und innerhalb von 1,4 s in den nicht hörbaren Bereich ansteigt. Eine niedrigere Ordnung des Turboladers kann ebenfalls detektiert werden. Die hochfrequenten waagerechten Verläufe spiegeln womöglich Reifenresonanzen wider.

Die untere Grafik zeigt die tonalen Komponenten eines stufenloses CVT-Getriebes, dessen Drehzahl über die Zeit leicht abnimmt. Die zwei Ordnungen, welche bei 2 und 4 kHz starten, können sehr gut abgebildet werden.

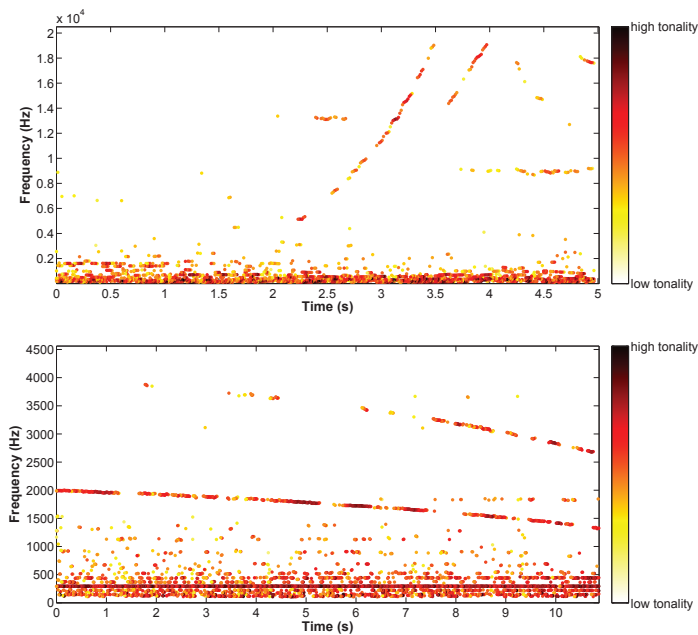


Abbildung 8: Obere Grafik: Tonalitätsspektrogramm des Soundbeispiels von einem Turbolader. Untere Grafik: Tonalitätsspektrogramm eines stufenlosen CVT-Getriebes.

Zusammenfassung und Ausblick

Für die akustische Bewertung und Optimierung von technischen Produkten werden vor allem Algorithmen benötigt, welche zum Beispiel Tonalität korrekt detektieren können. Der Algorithmus der aktuellen DIN 45681:2005 zeigt deutliche Schwächen wenn das Signale zeitliche Änderungen aufweist.

Das vorgestellte Hörmodell detektiert transiente tonale Elemente in Signalen anhand der Ausprägtheit im Erregungsmuster. Der ursprüngliche Ansatz Töne über die 0 Hz Modulation zu detektieren konnte aufgrund der Trägheit des 0 Hz Filters nicht umgesetzt werden. Stattdessen wurde die Grenzfrequenz des Filters auf 150 Hz nach Vorbild des „leaky integrators“ angehoben, was zu einer wesentlich kürzeren Einschwingphase führt. Die Tonalität wird über eine Kreuz-Kovarianzberechnung aus aktuellem Erregungsmuster und dem Erregungsmuster von zuvor abgespeichertem unverrauschten Sinusoiden berechnet. Es konnte gezeigt werden, dass die Performance des Modells im Vergleich zur DIN 45681:2005 wesentlich höher ist. Bei realen Anwendungen wie zum Beispiel die Tonalitätsanalyse von Turboladern können die in der Frequenz schnell ansteigenden tonalen Flanken nahezu durchgehend detektiert werden, ohne dass ein Tracking-Algorithmus implementiert wurde.

Literatur

- [1] Ewert, S. D., and Dau, T.: Characterizing frequency selectivity for envelope fluctuations, *J. Acoust. Soc. Am.* 108, 2000
- [2] Hohmann, V.: Frequency analysis and synthesis using a Gammatone filterbank, *Acta Acustica*, S. 433-

442, 2002

- [3] Lawson, Uhlenbeck, Threshold Signals, Dover Publications Inc., 1950
- [4] Oetjen, A., van de Par, S., Weber R., Letens, U.: Verbesserte Tonalitätsberechnung für instationäre Geräusche, DAGA 2014
- [5] Plomb, Mimpen: The Ear as a frequency analyzer II, *J. Acoust. Soc. Am.* 47, 1968
- [6] Søndergaard, Majdak: The Auditory Modeling Toolbox, *The Technology of Binaural Listening*, 2013
- [7] Viemeister: Temporal modulation transfer functions based on modulation threshold, *J. Acoust. Soc. Am.* 66, 1979
- [8] Zerbs, Martner, Wilholt: Subjektive und objektive Bewertung von zeitabhängiger Tonalität, DAGA 2010