

Berechnungsmodelle der akustischen Rauigkeit

Hannes Löschke¹, Sébastien Barre²

¹ *Institut für Musikinstrumentenbau, 08267 Zwota, Deutschland, Email: hannes_loeschke@yahoo.de*

² *Sinus Messtechnik GmbH, 04347 Leipzig, Deutschland, Email: sb@sinusmess.de*

Einleitung

Psychoakustische Rauigkeit ist eine gut wahrnehmbare und verhältnismäßig leicht zu erklärende Empfindungsgröße. Es sind verschiedene Modelle zur Vorhersage dieser Empfindung bekannt und bereits in kommerziell erhältlichen Messgeräten implementiert, dennoch findet sie nur wenig Akzeptanz in der breiten Öffentlichkeit. Die Modellbildung stützt sich auf zwei unterschiedliche Faktoren bei der Wahrnehmung von Rauigkeit. Zum Einen lässt sich leicht ein systematischer Zusammenhang zwischen der empfundenen Rauigkeit und der vorhandenen Modulation aufzeigen, zum Anderen kann man die Rauigkeit mit der sensorischen Konsonanz in Verbindung bringen. Daraus ergeben sich zwei gänzlich unterschiedliche Modellansätze mit unterschiedlichen daraus entstehenden Problemen bei der Umsetzung.

Modulationsgrad

Die Modelle, auf der Basis von Modulationsanalysen gehen im Grunde auf das Modell von Aures zurück. Wichtige Weiterentwicklungen stellen das Modell von Daniel und Weber, sowie das Verfahren von Sottek dar. Dabei wird das zu analysierende Signal in den Frequenzbereich transformiert und in 1 bark breite Frequenzbänder zerlegt. Innerhalb dieser Bänder wird die Amplitude der Modulationshüllkurve bestimmt und ein dem Modulationsgrad entsprechendes Maß aus dem Gleich- und dem Wechselanteil der Modulationshüllkurve ermittelt.

Bei der Zusammenfassung der so in jedem Band ermittelten spezifischen Rauigkeit ist zu beachten, dass korrelierte Modulation in unterschiedlichen Frequenzbereichen eine größere Rauigkeitsempfindung hervorrufen, als unkorrelierte. Dieser Effekt tritt jedoch bei wachsendem Frequenzabstand in den Hintergrund, so dass in den Modellen nur der Korrelationsgrad direkt benachbarter Bänder bei der Summierung berücksichtigt wird. Die spezifischen Rauigkeiten werden mit dem Korrelationsgrad multipliziert, wodurch bei vollständiger Korrelation die Rauigkeit maximal wird und sie bei dekorrelierter Modulation verschwindet. Dies trägt der Beobachtung Rechnung, dass weißes Rauschen als sehr wenig rau empfunden wird, jedoch in jedem Frequenzband starke zufällige Modulationen zu messen sind. Erst wenn das Rauschsignal insgesamt noch amplitudenmoduliert wird, kann Rauigkeit wahrgenommen werden, da in allen Bändern eine korrelierte Modulation hinzugefügt wird.

Problem bei der Berechnung der Modelle entstehen an mehreren Stellen. Zunächst wird das Zeitsignal mittels FFT in den Frequenzbereich transformiert. Da diese Transformation nur über einen endlichen Zeitabschnitt

ausgeführt werden kann, kommt die Transformation zugleich einer Faltung mit einer Zeitfensterfunktion gleich. Durch diese Faltung wird dem Spektrum die Charakteristik des Fensters aufgeprägt und es entstehen neben der tatsächlich vorhandenen Frequenzlinie weitere Linien, die je nach Fensterfunktion mit der Entfernung unterschiedlich schnell abklingen. Diese Artefakte der Fensterung rufen auch in weit entfernten Bändern Störungen hervor, die in großen Modulationsgraden resultieren und zudem eine sehr große Korrelation aufweisen. Daraus resultiert bei ungünstiger Fensterwahl eine unrealistisch hohe Rauigkeit. Bei üblichen Signalpegeln sind Fenster mit einer Seitenlinienunterdrückung von mindestens 60dB und möglichst schnell weiter abklingendem Spektrum notwendig um diesen Fehler zu vermeiden.

Ein weiteres Problem ist die Verwendung des Modulationsgrades an sich. Durch die Normierung auf den Gleichanteil der Modulationshüllkurve kann auch bei kaum noch wahrnehmbarer Signalamplitude ein maximaler Modulationsgrad und damit eine sehr hohe Rauigkeit berechnet werden. Somit werden möglicherweise auch Rauschkomponenten mit einer großen berechneten Rauigkeit bewertet, die in der Wahrnehmung keine weitere Rolle spielen. Somit ist es sinnvoll, bei der Zusammenfassung der spezifischen Rauigkeiten den Pegel der jeweiligen Komponente einzubeziehen und zu schwache Signale auszublenden.

Analyse tonaler Komponenten

Ein weiterer Ansatz zur Berechnung der Rauigkeit basiert auf Beobachtungen der sensorischen Konsonanz. Die Rauigkeit wird als der Konsonanz umgekehrt proportional angenommen und nimmt für zwei Töne mit steigendem Frequenzabstand zunächst sehr schnell zu und nach Überschreiten eines Maximums wieder langsam ab. Sind sie tonalen Komponenten eines Spektrums bekannt, lässt sich für jede Kombination der Frequenzabstand und daraus der jeweilige Beitrag zur Gesamtrauigkeit ermitteln. Für zwei reine Töne bildet das Modell das selbe Verhalten nach, wie der Ansatz über die Modulationsgrade. Bei Tonkomplexen stellen sich jedoch aufgrund der unterschiedlichen Frequenzauflösung Unterschiede heraus. Vor allem die sichere Erkennung von tonalen Komponenten stellt bei diesem Verfahren die größte Herausforderung dar. Gerade bei modulierten und damit ständig schwankenden Signalen hängt die Erkennung von Peaks im Spektrum von der Frequenzauflösung und der gewählten Fensterfunktion ab. Diese Probleme treten im Grunde genau so bei der Modellierung der empfundenen Tonalität auf.

Vereinfachter Vorschlag zur Rauigkeitsberechnung

Beide Modellansätze gehen die Verarbeitung von modulierten Signalen im Menschen sehr theoretisch an. Die Modell nach Aures orientieren sich dabei noch grundsätzlich an der Informationsverarbeitung im Gehör, benötigen jedoch komplexe Verarbeitungsschritte, wie die Bildung von Korrelationsfunktionen. Die Analyse tonaler Komponenten erklärt sehr gut die Effekte, die bei wenigen, stark ausgeprägten Tönen zu beobachten sind. Sie setzt jedoch noch wesentlich komplexere Verarbeitungsschritte voraus und scheitert an komplizierteren Signalen. Da Rauigkeit gerade auftritt, wenn Einzeltöne nicht mehr durch das Gehör getrennt werden können, erscheint der zugrundeliegende Ansatz wenig gehörgerecht.

Da es in der Natur nur wenige Beispiele für periodisch modulierte Geräusche gibt, die essentiell für das Überleben sind, stellt sich die Frage, wozu ein derart komplexes Verarbeitungssystem entstanden sein soll, wenn es sich auf bereit vorhandene, notwendige Verarbeitungsstufen zurückführen lässt. Am Institut für Musikinstrumentenbau (IfM) wird seit einiger Zeit der Ansatz verfolgt, dass die grundlegende Verarbeitung des Schalls zur Bildung eines Rauigkeitsmaßes, ähnlich der Lautheit, bereits sehr früh in der Verarbeitungskette des Gehörs erfolgt. Demnach wird versucht den Vorgang auf möglichst einfache Verarbeitungsschritte zurückzuführen, die im Idealfall als Nebenprodukt der Lautheitsbildung entstehen.

Der erste Schritt ist die Bildung der Erregungsfunktion mit Hilfe einer gehörgerechten Filterbank. Die Anzahl der Einzelfilter sollte dabei einen Kompromiss zwischen dem Frequenzauflösungsvermögen des Gehörs und der für die Bildung der Rauigkeit notwendigen Präzision darstellen, um das Modell nicht unnötig zu verkomplizieren. Bei Versuchen am IfM hat eine Steigerung der Kanalzahl über 120 keine wesentliche Veränderung mehr gezeigt. Bei der Umwandlung des Schallsignals in Nervenimpulse wird es bereits in der Gehörschnecke gleichgerichtet und schließlich demoduliert. Bereits am Hörnerv steht damit ein Signal an, das nur noch der Hüllkurve des entsprechend bandpassgefilterten Schalls entspricht.

Die phasenrichtige Überlagerung gleichzeitig dargebotener modulierter Signale innerhalb einer Frequenzgruppe lässt sich an dieser Stelle durch eine seitliche Kopplung der Informationskanäle erklären. Im Modell wird dies durch eine gleitende Mittelung über ein Fenster von der Breite einer Frequenzgruppe erreicht. Gegenphasige Hüllkurvenschwankungen sind somit in der Lage, sich gegenseitig auszulöschen, gleichphasige verstärken sich. Modulierte Signale, die nicht der selben Frequenzgruppe zugehören, gehen mit ihren Einzelrauigkeiten additiv in die wahrgenommene Gesamtrauigkeit ein. Die Extraktion der Signalschwankungen aus dem Kanalsignal wird im Modell zunächst durch die Beseitigung des Gleichanteils, der ja der wahrgenommenen Lautheit entspricht, und eine anschließende Effektivwertbildung durchgeführt. Durch diesen Schritt wird die Phaseninformation der Hüllkurvenschwankung entfernt, die nachfolgende

Summation über alle Kanäle erfolgt demnach phasenunabhängig.

Beurteilung des Modells

Das so gewonnene Rauigkeitsmaß zeigt noch einige Schwächen bei der Darstellung von Rauschsignalen, wobei hier diskutiert werden kann, inwiefern weißes Rauschen tatsächlich als nicht rau wahrgenommen wird. Man könnte diesen Effekt auch durch eine Überlagerung der empfundenen Rauigkeit und der wahrgenommenen Bandbreite des Signals auf einer höheren Verarbeitungsebene erklären.

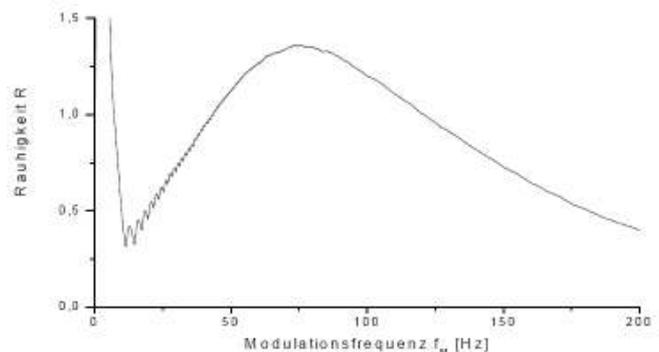


Abbildung 1: Berechnete Rauigkeit zweier überlagerter Sinustöne mit wachsendem Frequenzabstand.

Die Antwort des Modells auf ein einfaches moduliertes Signal ist in Abb. 1 dargestellt. Das Signal entstand durch Überlagerung zweier Sinustöne mit ständig wachsendem Frequenzabstand. Die ursprüngliche Zeitachse wurde durch resultierende Modulationsfrequenz ersetzt. Der ungewöhnlich große Wert zu Beginn des Signals ist auf abruptes Einschalten zurückzuführen, auf unmodulierte Sinustöne reagiert das Modell erwartungsgemäß nicht. Mehrere Versuche, das beschriebene Modell auf musikalische Schalle anzuwenden, verliefen bisher vielversprechend. Dennoch bedarf dieser Ansatz noch ausführlicher Prüfung.

Literatur

- [1] AURES, W.: Ein Berechnungsverfahren der Rauigkeit. *Acustica* 58 (1985), S. 268-280.
- [2] DANIEL, P.; WEBER, R.: Psychoacoustical roughness: Implementation of an optimized model. *Acustica/acta acustica* 83 (1997), S. 113-123
- [3] LÖSCHKE, H.: Zur Differenzierbarkeit von Musikinstrumenten. *Fortschr. d. Akust. DAGA '06*. Braunschweig 2006