Melodiediskrimination für Kammfilterrauschen: zeitliche oder spektrale Kodierung?

Hagen Wierstorf, Stefan Uppenkamp
Medizinische Physik, Institut für Physik, Carl von Ossietzky Universität, 26111 Oldenburg
Email: hagen.wierstorf@uni-oldenburg.de

Einleitung

Es gibt eine obere und eine untere Grenze für die Wahrnehmbarkeit des Grundtones eines harmonischen Tonkomplexes als Tonhöhe. Diese Grenzen wiederum hängen stark von den im Spektrum enthaltenen Frequenzen des Tonkomplexes ab. Pressnitzer *et al.* [1] haben eine Methode vorgestellt, mit der sie die untere Grenze in einem Melodietest bestimmt haben (*lower limit of melodic pitch*, LLMP). In der vorliegenden Studie wurde untersucht, in weit sich das LLMP eines Kammfilterrauschens (*iterated rippled noise*, IRN) ähnlich der eines harmonisches Tonkomplexes verhält.

Methode

Stimuli: Es kamen IRN-Stimuli mit 16 Iterationen und Grundfrequenzen F_G von 13,75 Hz bis 1661,22 Hz (entsprechen Verzögerungszeiten im *delay-and-add*-Prozess von 0,6 ms bis 72.73 ms) in Halbtonschritten zum Einsatz. Sie waren jeweils 300 ms lang und begannen bzw. endeten mit 6 ms langen Hanning-Flanken. Zur Untersuchung der Abhängigkeit vom Frequenzgehalt des IRN wurde eine untere Frequenz F_U und im zweiten Versuchsteil zusätzlich eine obere Frequenz F_O eingeführt. Im Spektralbereich wurden die so bandbegrenzten IRN-Stimuli unterhalb von der Frequenz F_U und oberhalb von F_O mit einem Rauschen mit gleichen RMS-Wert aufgefüllt. Anschließend wurde der Stimulus mit einem Butterworth-Filter vierter Ordnung zwischen 200 Hz und 8 kHz bandpassgefiltert.

Aus den IRN-Stimuli wurden Melodien mit vier Noten gebildet. Dazu wurde die Grundfrequenz F_G der tiefsten in der Melodie vorkommenden Note zur Klassifizierung der Melodie verwendet. Aus ihr und den bis zu drei Halbtonschritten darüber liegenden Noten wurde zufällig eine Melodie gebildet (vgl. [1]). Der zeitliche Abstand zwischen den einzelnen Noten betrug 20 ms.

Darbietung: Den Versuchspersonen wurde zuerst eine Melodie dargeboten und nach einer Pause von 660 ms die gleiche Melodie erneut, nur das sich nun eine der vier Noten um einen Halbtonschritt nach oben oder unten geändert hatte. Es galt also für die Versuchspersonen, die sich ändernde Note in einem Zwei-Interval 4AFC-Paradigma zu detektieren. Dieses geschah in Abhängigkeit von F_U und F_G. Dabei wurde zunächst F_U konstant gehalten und F_G wurde mit einem adaptiven 1up-2down-Algorithmus variiert, um das LLMP zu finden. Für F_U wurden dabei die Werte 250, 500, 1000, 1500, 2000, 3000, 4000 und 5000 Hz gewählt.

Anschließend ist dann für ein konstantes F_G von 82,4 Hz die psychometrische Funktion in Abhängigkeit von F_U und F_O ermittelt worden. Für F_U wurden die Werte 100, 200, 400, 600, 800, 1000, 1200 und 1500 Hz gewählt. F_O wurde so variiert, dass es immer $F_U + X$ betrug, wobei X die Werte 100, 300, 500, 700, 1500 oder 3100 Hz annehmen konnte.

Bei der adaptiven Messung wurden für jedes F_U zwei Werte für F_G ermittelt und der Mittelwert gebildet. Bei zu großem Unterschied der Einzelmessungen folgten ein bis zwei weitere Wiederholungen. Für die psychometrische Funktion wurden für jeden Punkt 30 Wiederholungen durchgeführt. Bei zu großer Standardabweichung wurden auch hier weitere Messungen durchgeführt (maximal 30 weitere je Punkt).

Versuchsdurchführung: Die Stimuli wurden an einem PC mit einer Abtastfrequenz von 44100 Hz erzeugt und über eine Soundkarte, D/A-Wandler und Kopfhörerverstärker diotisch über einen Kopfhörer HDA200 in einer doppelwandigen Hörkabine dargeboten. An dem Versuch nahmen sechs Probanden teil (3 w, 3 m, mittl. Alter 24,8 Jahre). Alle Probanden hatten eine normale Hörschwelle und bekamen eine Aufwandsentschädigung.

Ergebnisse

In Abbildung 1 sind die über alle sechs Versuchspersonen gemittelten Ergebnisse der adaptiven Messung aufgetragen. Für $F_U = 5000$ Hz konvergierte die adaptive Prozedur bei keiner der Versuchspersonen, auch wenn sie angaben, dass sie noch eine Tonhöhe wahrnahmen. Sie waren jedoch nicht mehr in der Lage, die beiden Melodien voneinander zu unterscheiden.

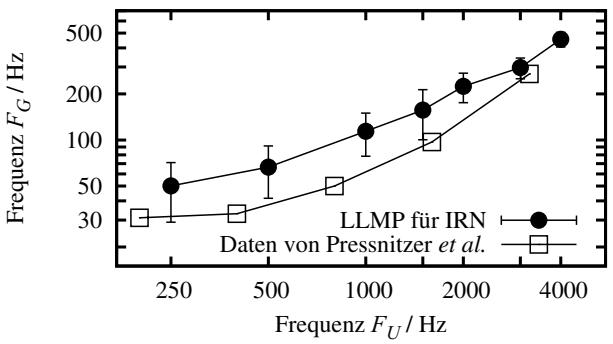

Abbildung 1: Ausgefüllte Kreise: LLMP der IRN-Stimuli als Mittelwert über alle sechs Versuchspersonen in Abhängigkeit von der Frequenz F_U. Unterhalb von F_U wurde das Signal mit Rauschen gleicher spektraler Leistungsdichte aufgefüllt. Offene Quadrate: Daten von Pressnitzer *et al.* für das LLMP von harmonischen Tonkomplexen. [1]

Für die anderen Frequenzen von F_U stieg die LLMP monoton von 50 Hz für $F_U = 250$ Hz bis auf 450 Hz für $F_U = 4000$ Hz an. Zusätzlich zu den eigenen Ergebnissen sind auch die Mittelwerte des LLMP für harmonische Tonkomplexe von Pressnitzer *et al.* [1] eingezeichnet. In ihrer Studie wurden die Tonkomplexe im Zeitbereich als periodische Pulsfolgen generiert.

In Abbildung 2 sind die psychometrischen Funktionen für ein F_G von 82,4 Hz und unterschiedliche spektrale Breiten M des IRN als Mittelwerte über alle Versuchspersonen eingetragen. Auf der x-Achse ist die unterste noch im IRN enthaltene Harmonische aufgetragen, diese ergibt sich mit $N = \frac{F_U}{F_G} + 1$. Die unterschiedlichen spektralen Breiten sind durch die Anzahl M an Harmonischen gegeben, die im IRN zwischen F_U und F_O enthalten sind. Es sind zwei Plateaus zu erkennen, eines bei 2-6 Harmonischen mit einer Richtig-Antwort-Wahrscheinlichkeit von ungefähr 80% und eines bei 12-19 Harmonischen mit einer Richtig-Antwort-Wahrscheinlichkeit von 35%, die somit nur knapp über der Ratewahrscheinlichkeit von 25% liegt. Die Abhängigkeit der Richtig-Antwort-Wahrscheinlichkeit von der spektralen Breite M ist äußert gering und wenn überhaupt, nur für große N zu beobachten.

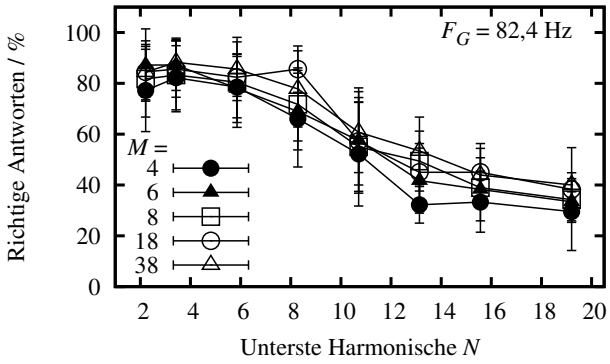

Abbildung 2: Gemittelte psychometrische Funktionen (6 Probanden) der IRN-Stimuli mit einer Grundfrequenz F_G von 82,4 Hz in Abhängigkeit von der untersten Harmonischen, die noch im IRN enthalten ist ($N = \frac{F_U}{F_G} + 1$). Die einzelnen psychometrischen Funktionen unterscheiden sich in der spektralen Breite des IRN, angegeben durch die Anzahl M der enthaltenen harmonischen Komponenten.

Diskussion

Bei der Wahrnehmung des Grundtons eines harmonischen Tonkomplexes ist bekannt, dass es zu einer stärkeren Tonhöhenempfindung kommt, wenn in ihm Harmonische vorhanden sind, die einzeln in den auditorischen Filtern aufgelöst werden können [2]. Definitionen auf Grund psychoakustischer Messungen der Auflösbarkeit von einzelnen Harmonischen in einem Tonkomplex lassen sich etwa bei Moore und Ohgushi [3] oder bei Shackleton und Carlyon [4] finden.

Zur Erklärung des LLMP liegt es daher nahe, sich dieses im Zusammenhang mit der Grenze der Auflösbarkeit von harmonischen Komponenten anzusehen. In Abbildung 3 sind noch einmal die Ergebnisse unserer Studie und die von Pressnitzer *et al.* mit einer veränderten Ordinate aufgetragen. Zusätzlich sind die Grenzen der Auflösbarkeit von einzelnen Harmonischen, wie sie sich aus den oben genannten Studien [3, 4] ergeben eingezeichnet. Es ist zu erkennen, dass die Grenze für das LLMP des IRN (ausgefüllte Symbole) genau in den Bereich des Übergangs zwischen aufgelösten und unaufgelösten Harmonischen fällt. Dieses ist jedoch nicht der Fall für die Daten von Pressnitzer *et al.* [1] (offene Symbole).

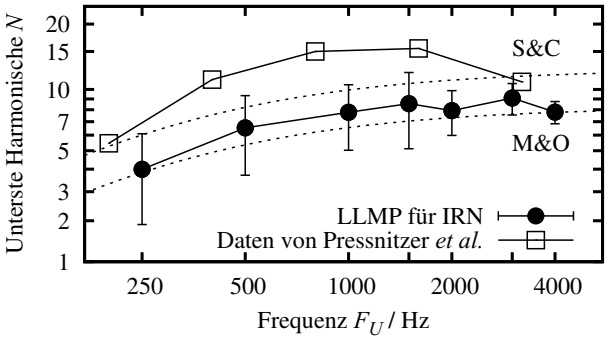

Abbildung 3: LLMP für IRN-Stimuli und für harmonische Tonkomplexe aus Abb. 1. Auf der Ordinate ist jetzt die unterste Harmonische N dargestellt, die noch im IRN bzw. im Tonkomplex enthalten ist. Die beiden gestrichelten Linien S&C und M&O stellen den Bereich für den Übergang zwischen aufgelösten und unaufgelösten Harmonischen nach [3, 4] dar.

Wird allerdings nochmals die psychometrische Funktion in Abbildung 2 betrachtet, lässt sich auch dieser Unterschied erklären. Die adaptive 1up-2down-Prozedur konvergiert bei dem 71%-Punkt der psychometrischen Funktion und dem dazugehörigem N. Dieser Punkt befindet sich in dem Übergangsbereich zwischen den beiden Plateaus. Somit ist das adaptiv gemessene N von der absoluten Lage der psychometrischen Funktion abhängig und würde sich zu höheren Werten verschieben, wenn die Richtig-Antwort-Wahrscheinlichkeit für aufgelöste Harmonische nahezu 100% betragen würde, was für den Versuch von Pressnitzer *et al.* für harmonische Tonkomplexe zu erwarten wäre.

Die Ergebnisse dieser Studie zeigen, dass sich die Grenzen der Tonhöhenwahrnehmung bei Kammfilterrauschen analog zu der bei harmonischen Tonkomplexen verhalten und beide mit der Grenze der Auflösbarkeit einzelner Harmonischer zusammenhängen. Es ist daher zu vermuten, dass auch im Kammfilterrauschen einzelne Harmonische aufgelöst werden können.

Literatur

[1] Pressnitzer D, Patterson RD, Krumbholz K (2001), J Acoust Soc Am **109** 2074-2084

[2] Houtsma AJM, Smurzynski J (1990), J Acoust Soc Am **87** 304-310

[3] Moore BCJ, Ohgushi K (1993), J Acoust Soc Am **93** 452-461

[4] Shackleton TM, Carlyon RP (1994), J Acoust Soc Am **95** 3529-3540