

Zur Wahrnehmung überlagerter Signalreflexionen

Karl Roman, Hans-Joachim Maempel, Stefan Weinzierl

TU Berlin, Fachgebiet Audiokommunikation

karl.roman@googlemail.com, hans-joachim-maempel@tu-berlin.de, stefan.weinzierl@tu-berlin.de

Einleitung

In analogen und digitalen Übertragungsstrecken kann es bei Reflexionen an schallharten Begrenzungsflächen, Verwendung mehrerer Mikrofone, Einsatz mehrerer Lautsprecher in der Beschallungstechnik oder systembedingten Latenzen in digitalen Prozessoren zu einer Überlagerung von identischen, zueinander zeitversetzten Signalen kommen. Durch solche in der Systemtheorie als nichtrekursive Kammfilter bezeichneten Systeme treten in Abhängigkeit von Pegel- und Zeitdifferenz der beiden Signale häufig unerwünschte klangliche Veränderungen auf. In diesem Zusammenhang sollte durch einen Hörversuch ermittelt werden,

1. ab welcher Pegeldifferenz zwischen verzögertem und nicht verzögertem Signal eine klangliche Veränderung eben wahrnehmbar ist und
2. welche perzeptiven klanglichen Merkmale für die Detektion des Kammfilters maßgeblich sind.

Unterschiedsschwellen

Zur Ermittlung des ebenmerklichen Unterschieds zwischen einem Originalsignal und dem durch seine verzögerte Überlagerung gebildeten Kammfilter-Signal wurde ein Hörversuch nach dem 3AFC-Verfahren durchgeführt. Dabei wurde die Stärke des Kammfilters, gegeben durch die Pegeldifferenz zwischen Original- und verzögertem Signal, gemäß einer Best-PEST-Prozedur [1] adaptiv nachgeführt. Als Audioinhalte wurden ein männlicher Sprecher, ein Klavier und ein Snare drum roll mit einer annähernd rosafarbenen spektralen Leistungsdichte verwendet. In einem Vorversuch wurde eine maximale Sensitivität der Vpn für eine Präsentationsdauer der Stimuli von 3 s ermittelt [2]. Für alle Stimuli wurde mittels Matlab eine Lautheitsanpassung nach Zwicker vorgenommen. Am Hörversuch nahmen 30 Vpn ($\bar{\Delta} 27$ Jahre) mit musikalischer Vorbildung teil, die in einer Trainingsphase Gelegenheit hatten, sich mit den auditiven Cues vertraut zu machen.

Abb. 1 zeigt die im Hörversuch ermittelten durchschnittlichen und minimalen Unterschiedsschwellen für Verzögerungszeiten zwischen 0,1 und 60 ms in 16 Schritten. Die Mittelwert-Kurven für Piano und Snare drum weisen jeweils ein Minimum bei $\Delta t = 0,8$ ms auf. Im Falle der Snare drum genügt hier den Versuchspersonen (Vpn) durchschnittlich eine Pegeldifferenz von 18 dB zur Erkennung des kammgefilterten Signals, im Einzelfall kann die Differenz sogar 27 dB betragen (Minimalwert bei $\Delta t = 0,5$ ms). Dies entspricht einer Überhöhung der Peaks im Frequenzgang des Kammfilters von 1,0 bzw. 0,4 dB. Für Sprache besitzt die Mittelwert-Kurve kein Minimum, sondern einen weitgehend monoton abnehmenden Verlauf, der oberhalb von 10 ms steil abfällt. Die Mittelwerte liegen in der gleichen Größenordnung wie

die von Kuhl [3] für weißes Rauschen und von Anazawa et al. [4] für verschiedene Musikschnitte ermittelten Werte. Im vorliegenden Versuch dürfen aufgrund des Forced-Choice-Verfahrens mit zum Teil mehrfacher Wiederholung von Trials darüber hinaus auch alle individuellen Messwerte, das heißt auch die Minimalwerte, als reliable Schwellen der sensorischen Erkennbarkeit betrachtet werden.

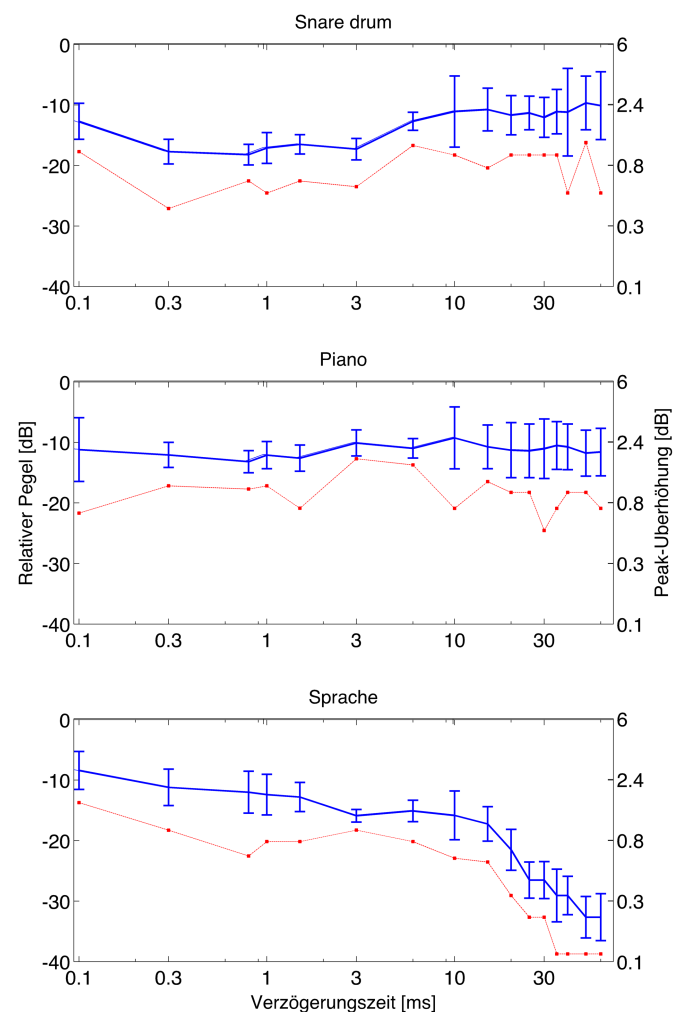


Abb. 1: Ebenmerkliche Unterschiede für die Detektion kammfilterartiger Verzerrungen (Mittelwerte $\pm$ Standardabweichung, Minimalwerte)

Klangmerkmale

Die an der Ermittlung der Unterschiedsschwellen beteiligten Vpn wurden gebeten anzugeben, anhand welcher Wahrnehmungsqualitäten ihnen die Detektion des gefilterten Signals möglich war. Insgesamt wurden vier Merkmale genannt:

- Veränderte Klangfarbe
- Zusätzliche Tonhöhe
- Veränderter Raumeindruck
- Getrennte Hörereignisse (Echo)

Um die relative Bedeutsamkeit der einzelnen Merkmale für die Wahrnehmbarkeit von Kammfiltereffekten zu ermitteln, wurde sie in einem zweiten Hörversuch auf einer vierstufigen Skala zwischen den Polen „unhörbar“ (entsprechend „unbedeutend“) und „sehr bedeutend“ abgefragt. Präsentiert wurden dieselben Audiosignale und Verzögerungszeiten wie bei der Ermittlung der Schwellwerte. Weil für diesen nicht kriterienfreien Rating-Versuch mit Blick auf eine Mittelwertbildung die Daten möglichst vieler Vpn erforderlich sind, die den Kammfiltereffekt zuverlässig detektieren, wurden drei leicht erkennbare Pegeldifferenzen getestet: 0 dB (gleicher Pegel von Signal und verzögertem Signal), die sich aus dem ersten Versuch für die jeweilige Verzögerungszeit ergebende mittlere Pegeldifferenz plus Standardabweichung sowie der dazwischenliegende Wert. Eine 3AFC-Prozedur stellte sicher, dass die Ausprägungen der Merkmale nur dann eingestuft werden konnten, wenn das gefilterte Signal zuverlässig erkannt wurde.

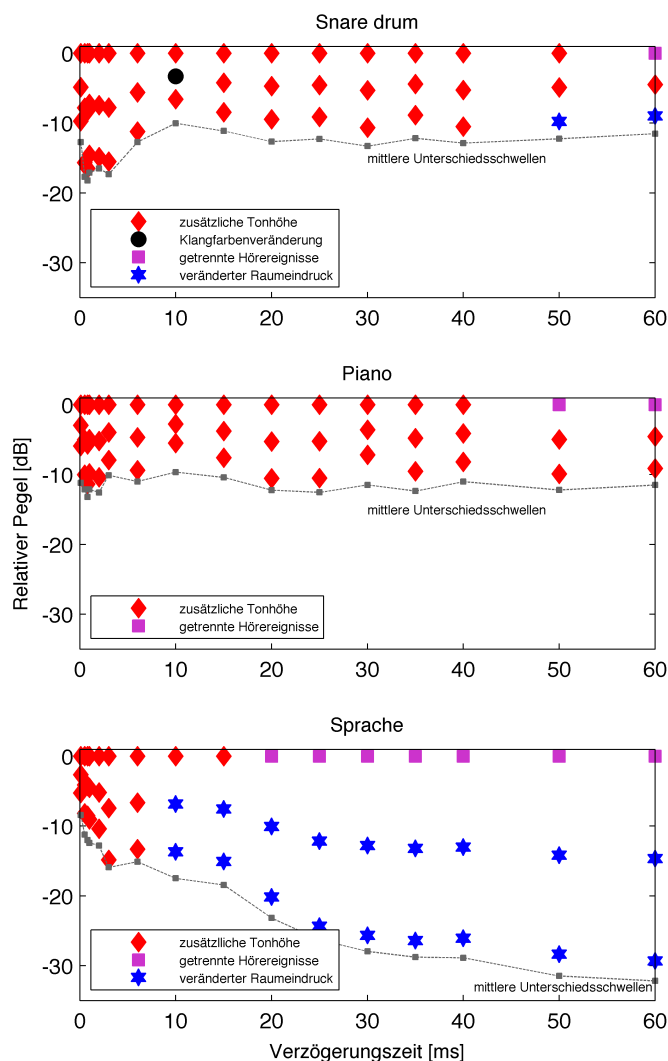


Abb. 2: Dominierende Klangmerkmale bei der Detektion von Kammfiltereffekten

Die Ergebnisse für 12 Vpn sind Abb. 2 zu entnehmen. Dargestellt sind nur die Merkmale, die bei der jeweiligen Zeit-/Pegeldifferenz als am bedeutsamsten eingestuft wurden, also die dominierenden perceptiven Cues. Im Falle der Audioinhalte Piano und Snare drum dominiert bei fast allen Verzögerungszeiten und Pegeldifferenzen durchweg die Empfindung einer zusätzlichen Tonhöhe (repetition pitch, [6]), die durch die Periodizität des Kammfilter-Frequenzgangs hervorgerufen wird. Erst ab einer Verzögerungszeit von 50 ms treten zum Teil die Wahrnehmung getrennter Hörereignisse und ein veränderter Raumeindruck in den Vordergrund. Im Falle von Sprache erfolgt die Erkennung des Kammfiltereffekts schon ab einer Verzögerungszeit von 10 ms überwiegend anhand des Raumeindrucks sowie ab 20 ms (für $\Delta L = 0$ dB) anhand getrennter Hörereignisse.

Zusammenfassung

Es wurden ebenmerkliche Unterschiede zwischen Original- und kammgefiltertem Signal für verschiedene Verzögerungszeiten und drei repräsentative Typen von Audiosignalen ermittelt. Während die Mittelwerte von 30 Vpn mit früheren Untersuchungen konform sind, weisen einzelne Vpn im vorliegenden Versuch zuverlässig eine deutlich höhere Detektionsleistung auf. Unterhalb der Echoschwelle sind im Vpn-Mittel Pegeldifferenzen von 13 dB (Piano), 17 dB (Sprache) bzw. 18 dB (Snare drum) eben wahrnehmbar, individuell sogar Pegeldifferenzen von 24,5 dB (Piano), 22,5 dB (Sprache) bzw. 27 dB (Snare drum). Maßgeblich für die Detektion des Kammfiltereffekts ist unterhalb einer Verzögerungszeit t_0 die Wahrnehmung einer Klangfarbenänderung bzw. einer zusätzlichen Tonhöhe. Obwohl im Hörversuch beide Merkmale als ähnlich bedeutsam eingestuft werden, dominiert die Tonhöhe in den allermeisten Fällen leicht (Abb. 2). Oberhalb von t_0 ist vor allem die Wahrnehmung eines geänderten Raumeindrucks oder eines Echos für die Detektion relevant. Je nach Pegeldifferenz liegt t_0 für die Inhalte Piano und Snare drum zwischen 50 und 60 ms, für Sprache zwischen 10 und 20 ms.

Literatur

- [1] Pentland, A.: „Maximum likelihood estimation: The best PEST.“ *Perception & Psychophysics* 28 (1980). S. 377-379.
- [2] Brunner, S.; Maempel, H.-J.; Weinzierl, S.: „On the audibility of comb-filter distortions. 122nd AES Convention, Vienna 2007, Preprint 7047.“
- [3] Kuhl, W.: „Unterschiedliche Bedingungen beim Hören in einem Raum und bei elektroakustischen Übertragern.“ *Rundfunktechnische Mitteilungen* 13/5 (1969). S. 205-208.
- [4] Anzawa, T.; Takahashi, Y.; Clegg, A. H.: „Digital Time-Coherent Recording Technique.“ *83th AES Convention, New York 1987. Preprint 2493.*
- [5] Haas, H.: „The Influence of a Single Echo on the Audibility of Speech.“ *Journal of the Audio Engineering Society* 20/2 (1972). S. 146-159.
- [6] Bilsen, F. A.; Ritsma, R. J.: „Repetition Pitch and Its Implication for Hearing Theory.“ *Acoustica* 22 (1969/70). S. 64-73.