

Bewertung der Gehörrelevanz von Partialtonzeitstrukturen in Klaviertönen

M. N. Valenzuela

Fachgebiet Akustische Kommunikation, Technische Universität München, Jetzt: Institut für Mechanik, UniBw München

Seit den Untersuchungen von Risset [4] mit Trompetentönen weiß man, daß für die synthetische Nachbildung von Naturinstrumenten der zeitliche Pegelverlauf der einzelnen Partialtöne bekannt sein muß. Deshalb werden bei der digitalen Klangerzeugung mittels additiver Synthese neben den Frequenzen auch die zeitabhängigen Amplitudenwerte der Partialtöne berücksichtigt und abgespeichert [1]. Vor allem das Abspeichern des Amplitudenverlaufs jedes Partialtons erfordert bei der authentischen Nachbildung von Klaviertönen einen hohen Speicherbedarf. Mit dem nachfolgend beschriebenen Bewertungsverfahren, das die Gehörrelevanz der Pegeländerungen im Ausklingen der Partialtöne ermittelt, kann dieser hohe Speicherbedarf um über die Hälfte reduziert werden. Den komplizierten Pegelverlauf eines Partialtons, dessen Pegeländerungen im Ausklingen als gehörirrelevant bewertet werden, ersetzt das Verfahren durch einen charakteristischen Pegelverlauf, für den ein wesentlich geringerer Speicherbedarf notwendig ist.

1. Charakterisierung der Pegelverläufe von Partialtönen

Das typische Ausklingverhalten eines Klaviertons ist durch zwei verschiedene schnelle Abklingphasen gekennzeichnet, die als Sofort- und Nachklang bezeichnet werden [2]. Während der ersten Abklingphase fällt der Schalldruckpegel recht schnell ab, in der direkt anschließenden zweiten Phase sinkt der Pegel deutlich langsamer ab. Auf diese Weise ergibt sich im Ausklingverhalten eines Klaviertons ein mehr oder weniger deutlich ausgeprägter Knick. Darüber hinaus können modulationsartige Feinstrukturen im Ausklingverhalten auftreten, die dem Klavierton einen belebten Charakter verleihen [3]. Die zeitlichen Pegelverläufe der Partialtöne eines Klaviertons lassen sich demnach durch einen "knicklinienförmigen" Pegelabfall und einer überlagerten modulationsartigen Feinstruktur kennzeichnen. In Abb. 1 sind beispielhaft zwei typische Pegelverläufe von Partialtönen dargestellt.

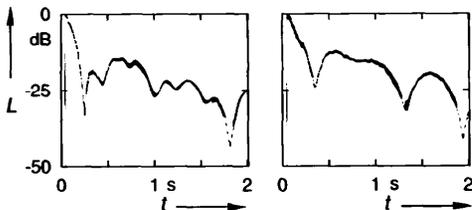


Abb. 1: Zwei typische Pegelverläufe im Ausklingverhalten von Partialtönen in Klaviertönen.

2. Bewertungsverfahren

Die Gehörrelevanz der modulationsartigen Feinstrukturen im Ausklingverhalten von Partialtönen wird mit Hilfe der Amplitudenmodulationsschwellen des Gehörs bewertet. Dazu interpretiert das Verfahren die Zeitstruktur im Pegelverlauf eines beliebigen Partialtons als modulierte "Knicklinie", wobei gleichzeitig die übrigen Partialtöne als Maskierer betrachtet werden. Da sich die Modulationsschwellen bei gleichzeitiger Darbietung eines verdeckenden Störgeräusches verändern [9], ist zwischen

ungestörten und gestörten Modulationsschwellen zu unterscheiden. Das Bewertungsverfahren läßt sich grob in drei Schritte einteilen.

Bestimmung der Pegeländerungen. Im ersten Schritt wird der modulationsartige Anteil eines Partialtons durch die auftretenden Pegeländerungen gekennzeichnet, wobei nur diejenigen Pegeländerung bestimmt werden, die nicht schon durch den Pegelabfall der Knicklinie selbst erfaßt werden. Demnach werden nur solche Pegeländerungen ermittelt, die sich zwischen lokalen Minima der Partialtonhüllkurve und den auf sie folgenden Maxima ergeben. Die Änderungsgeschwindigkeit der Pegeländerungen bleibt unberücksichtigt. Da das Gehör nur langsame Modulationsfrequenzen bis etwa 30 Hz [5] als Amplitudenschwankungen wahrnehmen kann, werden die Hüllkurven der Partialtöne zuvor mit einem entsprechenden Tiefpaß geglättet. Zu jeder ermittelten Pegeländerung ΔL_μ im Pegelverlauf eines Partialtons kann der korrespondierende Modulationsgrad m_μ zu

$$m_\mu = \frac{10^{\frac{\Delta L_\mu}{20\text{dB}} - 1}}{10^{\frac{\Delta L_\mu}{20\text{dB}} + 1}} \quad (1)$$

berechnet werden. Auf diese Weise wird der modulationsartige Anteil eines Partialtons durch die berechneten Modulationsgrade m_μ repräsentiert.

Um den verdeckenden Einfluß der übrigen Partialtöne auf den untersuchten Partialton zu ermitteln, wird der Pegelüberschuß $LX(t)$ nach [6] für jeden Zeitpunkt des aktuellen Partialtons berechnet. Für die Berechnung der Erregungsverteilungen wird die in [7] vorgeschlagene Näherung für die Mithörschwelle von Sinustönen verwendet.

Bewertung der Pegeländerungen. Im zweiten Schritt wird die Gehörrelevanz der ermittelten Pegeländerungen bestimmt, wobei vereinfachend angenommen wird, daß die Pegeländerungen immer mit einer Modulationsfrequenz von 4 Hz erfolgen. Da das Gehör im Bereich dieser Modulationsfrequenz am empfindlichsten gegenüber Pegeländerungen ist, erhält man schlimmstenfalls eine zu hohe Empfindlichkeit bei anderen Änderungsgeschwindigkeiten.

Die modulationsartige Feinstruktur eines Partialtons wird nur dann als gehörirrelevant bewertet, wenn keine der bestimmten Pegeländerungen ΔL_μ im Ausklingverhalten wahrgenommen werden kann. Dementsprechend darf keiner der korrespondierenden Modulationsgrade m_μ oberhalb der jeweiligen Modulationsschwelle $m_{S,\mu}$ liegen. Da zwischen ungestörten und gestörten Modulationsschwellen zu unterscheiden ist, muß für die Bestimmung der jeweiligen Modulationsschwelle $m_{S,\mu}$ neben der Frequenz f und dem Pegel L des Partialtons auch der Pegelüberschuß LX an der Stelle der Pegeländerung ΔL_μ berücksichtigt werden. Nach Zwicker [9] bleiben die Modulationsschwellen vom maskierenden Störgeräusch unbeeinflusst, wenn der Pegel des amplitudenmodulierten Sinustons den vom Maskierer hervorgerufenen Erregungspegel um mehr als etwa 20 bis 30 dB überschreitet. Für das Verfahren wurde die Grenze zwischen gestörten und ungestörten Modulationsschwellen bei einem Pegelüberschuß von 20 dB festgesetzt, um in jedem Fall

eine ausreichend hohe Empfindlichkeit gegenüber Pegeländerungen zu gewährleisten. Für die Modulationsschwelle $m_{S,\mu}$, die bei der Pegeländerung ΔL_μ gerade eben wahrgenommen werden kann, gilt also

$$m_{S,\mu}(f, L, LX) = \begin{cases} 1; & LX \leq 0\text{dB} \\ m_{SLX}(f, L, LX); & 0 < LX/\text{dB} < 20 \\ m_S(f, L); & LX \geq 20\text{dB} \end{cases} \quad (2)$$

wobei m_{SLX} die gestörte und m_S die ungestörte Modulationsschwelle bezeichnet.

Für die Nachbildung der ungestörten Modulationsschwellen m_S von amplitudenmodulierten Sinustönen (Modulationsfrequenz 4 Hz) beliebiger Frequenz f und Pegel L wird die Näherung

$$m_S(f, L) = \frac{3,9 \text{ dB}}{[L - L_{HS}(f) + L_{HS}(f = 1\text{kHz})]^{1,2}} \quad (3)$$

vorgeschlagen, wobei L_{HS} die Ruhhörschwelle bezeichnet. Diese kann mit der von Terhardt [6] angegebenen Formel berechnet werden. Für Pegel $L \leq L_{HS}$ wird $m_S = 1$ gesetzt. In Abb. 2 sind links die mit dieser Modellierung berechneten, ungestörten Modulationsschwellen m_S in der Hörfläche dargestellt. Rechts ist ein Schnitt bei der Frequenz 1 kHz dargestellt und zum Vergleich sind zusätzlich die von Zwicker und Feldtkeller [8, S.95] angegebenen Werte gestrichelt eingetragen.

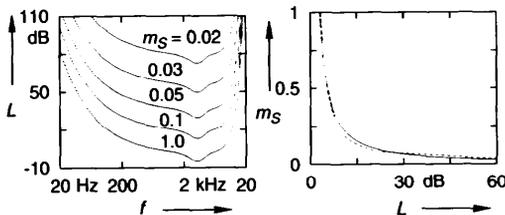


Abb. 2: Links: Mit der Näherung (3) berechnete ungestörte Modulationsschwellen, dargestellt in der Hörfläche. Parameter ist der jeweils eben wahrnehmbare Modulationsgrad m_S . Rechts: Vergleich zwischen berechneten (durchgezogen) und gemessenen (gestrichelt) [8, S.95] ungestörten Modulationsschwellen eines 1 kHz-Tons.

Im Bereich zwischen der vollständigen Verdeckung des modulierten Sinustons und der oberen Einflußgrenze des Maskierers, nämlich rund 20 bis 30 dB oberhalb seines Erregungspegels, geht die Modulationsschwelle sehr schnell von großen Werten, beginnend bei einem Modulationsgrad von 100%, in die ungestörte Modulationsschwelle über [9]. Dieser Übergang kann mit ausreichender Genauigkeit durch eine lineare Interpolation zwischen den beiden genannten Eckdaten beschrieben werden. Die Nachbildung der gestörten Modulationsschwellen m_{SLX} von amplitudenmodulierten Sinustönen mit beliebigem Pegel L und Frequenz f lautet demnach in Abhängigkeit des Pegelüberschuß LX für $0 < LX/\text{dB} < 20$

$$m_{SLX}(f, L, LX) = 1 - \frac{LX}{20\text{dB}} [1 - m_S(f, L)] \quad (4)$$

Bearbeitung der Pegeländerungen. Wird die modulationsartige Feinstruktur eines Partialtons als gehörrelevant bewertet, bleibt die gesamte Hüllkurve des Partialtons unverändert erhal-

ten. Wird die Feinstruktur dagegen als gehörirrelevant bewertet, ersetzt das Verfahren die Hüllkurve des Partialtons durch die entsprechende unmodulierte Knicklinie, die durch ein iteratives Anpassungsverfahren ermittelt wird.

3. Auditive Überprüfung und Diskussion

Die Brauchbarkeit des Bewertungsverfahrens wurde mit auditiven Vergleichen zwischen bearbeiteten und unbearbeiteten Klaviertönen untersucht. Dazu wurden jeweils 3 Töne (C_2, C_4, C_6) vier verschiedener Klaviere mit dem Verfahren bearbeitet. In drei nach Tonlagen getrennten Hörversuchen sollten die Versuchspersonen angeben, ob sie einen Unterschied zwischen den in Paaren dargebotenen originalen und bearbeiteten Tönen wahrnehmen konnten. An den Hörversuchen nahmen zehn normalhörende Versuchspersonen im Alter von 24 bis 40 Jahren teil, die über eine gute bis sehr umfassende musikalische Vorbildung verfügen. Die Versuchsergebnisse zeigen, daß in keiner der drei Tonlagen eine Unterscheidung zwischen originalen und bearbeiteten Tönen möglich ist. Somit werden in allen drei Tonlagen nur gehörirrelevante Partialtonzeitstrukturen durch das Verfahren vernachlässigt. Außerdem zeigt sich, daß die Hüllkurve eines Partialtons, dessen modulationsartige Feinstruktur gehörirrelevant ist, durch einen knicklinienförmigen Pegelverlauf ersetzt werden kann, ohne dabei hörbare Unterschiede zu verursachen.

Mit dem Bewertungsverfahren wurden in der tiefen Tonlage im Mittel etwa 88% der Partialtöne durch einen knicklinienförmigen Pegelverlauf ersetzt, in der mittleren Tonlage rund 68% und in der hohen Tonlage etwa 55%. Da für die Beschreibung des Ausklingverhaltens eines Partialtons durch eine Knicklinie nur 4 Kenngrößen erforderlich sind (diese können beispielsweise der Spitzenpegel, der Pegel an der Knickstelle und die beiden Pegelabfälle je Zeiteinheit sein, die die Steilheiten des Sofort- und Nachklangs kennzeichnen), kann mit diesem Verfahren der Speicherbedarf bei der Imitation von Klaviertönen mittels additiver Synthese deutlich reduziert werden. Bei einer Wortlänge von 8 bit für die Darstellung von Pegelwerten benötigt man nur 32 bit für die Codierung einer Knicklinie und damit eines beliebig lang ausklingenden Partialtons. Speichert man hingegen bei einer Zeitrasterung von 2,5 ms jeden Pegelwert mit 8 bit, so benötigt man für die Darstellung des Pegelverlaufs eines Partialtons, der 1 s lang ausklingt, 3,4 kbit Speicherplatz.

Für die kritische Durchsicht des Manuskripts danke ich Herrn Prof. Dr.-Ing. E. Terhardt.

Literatur

- [1] Abgarjan, T., Linsmeier, K.-D. (1997). Spektrum der Wissenschaft, November 1997: 74-84.
- [2] Martin, D.W. (1947). J. Acoust. Soc. Am., 19: 535-541.
- [3] Meyer, J., Melka, A. (1983). Das Musikinstrument, 32: 1049-1064.
- [4] Risset, J.C. (1965). J. Acoust. Soc. Am., 38: 912.
- [5] Terhardt, E. (1968). Acustica, 20: 210-214.
- [6] Terhardt, E. (1979). Hearing Research, 1: 155-182.
- [7] Valenzuela, M.N. (1997). Fortschritte der Akustik, DAGA'97: 321-322.
- [8] Zwicker, E., Feldtkeller, R. (1967). Hirzel, Stuttgart, 2. Edition.
- [9] Zwicker, E., Graf, L. (1987). Acustica, 64: 148-154.