Lineares Modell der peripheren Schallübertragung im Gehör

Ernst Terhardt

Fachgebiet Akustische Kommunikation, Technische Universität München

1. PET: Grundlagen. Das *Peripheral-Ear Transduction* (PET) System besteht aus (1) einem Filter zur Nachbildung des Außen-Mittelohr-Frequenzganges (*Ear-Canal Resonances* ECR) und (2) einer Filterbank, welche die cochleären Übertragungsfunktionen nachbildet (*Cochlear Transmission Functions* CTF), Abb. 1.

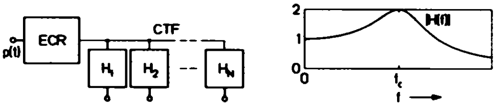

Abb. 1: PET-System (links) und Frequenzgang des Betrags der ÜF eines CTF-Einfachfilters für $\eta = 2$ (rechts)

Das ECR besteht aus zwei aufeinanderfolgenden Filtern des Typs

$$H_r(s) = 1 + \alpha \frac{s}{s^2 + 2as + a^2 + \omega_0^2}, \qquad (1)$$

wo $s = j\omega$, α eine reelle Konstante, a einen reellen Dämpfungsfaktor und ω_0 eine die Resonanz kennzeichnende Eigenfrequenz bedeuten. Die Resonanz des ersten Teilfilters liegt bei ungefähr 3,3 kHz, diejenige des zweiten bei 10 kHz.

Der Absolutbetrag von H_r nähert sich außerhalb der Resonanz von oben her dem Wert 1 und nimmt bei der Kreisfrequenz

$$\omega_r = \sqrt{a^2 + \omega_0^2} \qquad (2)$$

den Maximalwert $1 + \alpha/(2a)$ an. Die Resonanzbandbreite B_r ist näherungsweise durch $B_r = a/\pi$ gegeben. Damit kann α anhand der erforderlichen Bandbreite B_r und Resonanzüberhöhung bestimmt werden. Aus der Resonanzfrequenz f_r und der Bandbreite B_r ergibt sich

$$\omega_0 = \pi\sqrt{4f_r^2 - B_r^2}. \qquad (3)$$

Als geeigneter Typ der *cochleären* Übertragungsfunktion CTF erweist sich die Funktion

$$H_n(s) = \left(\frac{s_n s_n^*}{(s - s_n)(s - s_n^*)}\right)^k. \qquad (4)$$

Sie steht für eine Kette aus k Einfachfiltern des in (4) unter der Potenzfunktion stehenden Typs.

Mit der Zerlegung

$$s_n = -a_n + j\omega_n; \; s_n^* = -a_n - j\omega_n \qquad (5)$$

nimmt der Absolutbetrag der Übertragungsfunktion H_n bei der Kreisfrequenz

$$\omega_c = 2\pi f_c = \sqrt{\omega_n^2 - a_n^2} \qquad (6)$$

den Maximalwert (die Resonanzüberhöhung)

$$|H_n|_{max} = \eta^k = \left(\frac{a_n^2 + \omega_n^2}{2a_n\omega_n}\right)^k = \left(\frac{2a_n^2 + \omega_c^2}{2a_n\sqrt{a_n^2 + \omega_c^2}}\right)^k \qquad (7)$$

an. Um jedem *Einfach*filter des n-ten Kanals nach Festsetzung der charakteristischen Frequenz f_c eine bestimmte Resonanzüberhöhung η zu verleihen, muß gemäß (7)

$$a_n = f_c\pi\sqrt{2}\sqrt{\sqrt{1 + 1/(\eta^2 - 1)} - 1} \qquad (8)$$

gesetzt werden.

Die effektive Bandbreite (3 dB-Bandbreite) des Einfachfilters ergibt sich aus (4) zu $B_n' = a_n/\pi$. Bei gegebener Eigenfrequenz ω_n werden sowohl das Maximum der Resonanzüberhöhung als auch die Bandbreite des Einfachfilters durch die Dämpfungskonstante a_n bestimmt. Zwischen Resonanzüberhöhung und Bandbreite besteht daher ein fester Zusammenhang derart, daß ein hohes Resonanzmaximum eine geringe Bandbreite erfordert und umgekehrt.

Durch Kaskadierung von k Einfachfiltern pro Kanal kann die Bandbreite innerhalb gewisser Grenzen unabhängig von der Resonanzüberhöhung eingestellt werden. Die effektive Bandbreite des Mehrfachfilters beträgt

$$B_n = \frac{\sqrt{\omega_n^2 - a_n^2}}{\pi\sqrt{2}}\sqrt{1 - \sqrt{1 - \frac{4a_n^2\omega_n^2(\sqrt[k]{2} - 1)}{(\omega_n^2 - a_n^2)^2}}}. \qquad (9)$$

Das Einschwingverhalten der CTF-Filter ergibt sich aus der effektiven Dauer T der zu (4) gehörenden Impulsantwort. Eine Analyse derselben ergibt

$$Tf_c = \frac{(k-1)!\mathrm{e}^{k-1}}{\pi\sqrt{2}(k-1)^{k-1}} \cdot \frac{1}{\sqrt{\sqrt{1 + 1/(\eta^2 - 1)} - 1}}. \qquad (10)$$

Darin steht η für die geforderte Resonanzüberhöhung des Filters. Soll das Gesamtfilter die Überhöhung g aufweisen, so ist in (10) $\eta = \sqrt[k]{g}$ einzusetzen. In Abb. 2 ist das Produkt Tf_c für Resonanzüberhöhungen von 30, 50, 70 und 90 dB dargestellt.

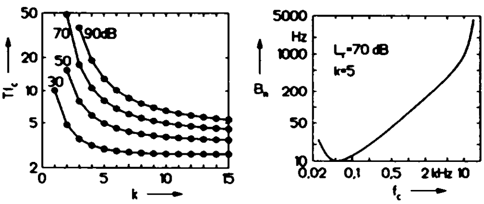

Abb. 2: Links: Normierte Einschwingzeit T der CTF-Filter als Funktion von k für die angegebenen Dynamikwerte. Rechts: Bandbreite der CTF-Filter als Funktion der char. Frequenz f_c

2. Dimensionierung. Für das erste Teilfilter des ECR erweisen sich die Parameterwerte $\alpha = 10000$, $\omega_0 = 20000$, $a = 5000$ als günstig; für das zweite die Werte $\alpha = 20000$, $\omega_0 = 65000$, $a = 20000$.

Zur Dimensionierung der CTF-Filter sind die charakteristischen Frequenzen f_c und die Dämpfungskoeffizienten a_n festzulegen. Die f_c werden so über den Hörbereich verteilt, daß sich eine gleichmäßige Überlappung der Durchlaßkurven ergibt. Die a_n werden so ausgelegt, daß sich aus den dazugehörigen Resonanzüberhöhungen unmittelbar die Abhängigkeit der Absoluthörschwelle von der Frequenz ergibt, wenn zuvor der Beitrag der Gehörgangresonanzen davon abgezogen wird. Das bedeutet, daß der Schwellenpegel in allen Kanälen *per definitionem* gleich ist.

Der Verlauf der Hörschwelle abzüglich der ECR, $L^{*}_{AS}(f_c)$, kann nach Terhardt (1979) durch

$$L^{*}_{AS}(f_c)/\mathrm{dB} = 3,64\left(\frac{f_c}{1000\mathrm{Hz}}\right)^{-0,8} + 10^{-3}\left(\frac{f_c}{1000\mathrm{Hz}}\right)^{4} \quad (11)$$

ausgedrückt werden. Aus der Festsetzung, daß die Resonanzüberhöhungen η^k der Filter als Funktion von f_c den *inversen* Verlauf von L^{*}_{AS} aufweisen sollen, ergibt sich die Bedingung

$$\begin{aligned} \lg\eta(f_c) &= \frac{L_r/\mathrm{dB}}{20k} + \frac{0,182}{k}\left[1-\left(\frac{f_c}{1000\mathrm{Hz}}\right)^{-0,8}\right] \\ &+ \frac{10^{-4}}{2k}\left[1-\left(\frac{f_c}{1000\mathrm{Hz}}\right)^{4}\right]. \end{aligned} \quad (12)$$

Darin ist für L_r die bei der Bezugsfrequenz 1000 Hz gewünschte Resonanzüberhöhung – die *Referenzüberhöhung* – einzusetzen. Bei gegebenen charakteristischen Frequenzen f_c erhält man aus (12) die Überhöhungen $\eta(f_c)$ und daraus mit (8) die a_n der Kanäle. Aus f_c und a_n ergeben sich sodann mit (6) die ω_n.

Die Referenzüberhöhung L_r ist den physiologischen und psychophysikalischen Daten des Gehörs anzupassen. Sie geht beispielsweise aus der Resonanzüberhöhung hervor, welche an Tuningkurven zu beobachten ist. Danach erscheint L_r = 70 dB angemessen. Die Zahl k der kaskadierten Einzelfilter pro Kanal wird aus der Forderung nach geringer Einschwingzeit abgeleitet, vgl. Abb. 2. Der Wert $k = 5$ erscheint angemessen. Die Bandbreiten B_n, welche sich damit ergeben, sind in Abhängigkeit von f_c in Abb. 2 rechts dargestellt.

3. Nachbildung der Absolutschwelle. Zur Verifizierung wird nach der Festlegung aller Parameter die Absolutschwelle folgendermaßen simuliert. Als Hörschwelle bei 1000 Hz wird $L_{AS} = 3$ dB zugrundegelegt. Unter Berücksichtigung des darin enthaltenen kleinen Beitrags von $|H_{ECR}|$ ergibt sich für den Kanal-Schwellenpegel

$$L_{nAS} = 3,3\mathrm{dB} + L_r. \quad (13)$$

Die Hörschwelle gilt als erreicht, wenn in irgend einem Kanal der bei $f = f_c$ erreichte Signalpegel gleich dem Kanal-Schwellenpegel nach (13) ist. Daraus ergibt sich für die Hörschwelle die Formel

$$\begin{aligned} L_{AS}(f_c)/\mathrm{dB} &= L_r/\mathrm{dB} + 3,3 - 20\lg|H_{ECR}(f_c)| \\ &+ 20\lg\eta^k(f_c). \end{aligned} \quad (14)$$

Abb. 3 (links) zeigt die nach (14) berechnete Absoluthörschwelle.

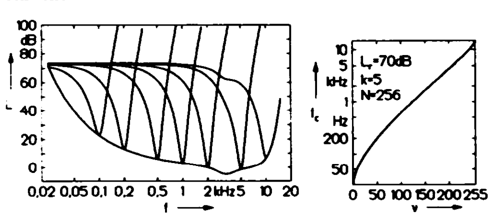

Abb. 3: Links: Absoluthörschwelle und Tuningkurven, simuliert mit dem PET-System (L_r = 70 dB, $k = 5$). Rechts: Charakteristische Frequenzen f_c der Kanäle für $N = 256$

4. Nachbildung der Tuningkurven. Ersetzt man in (14) η^k durch $|H_n(f)|$, so erhält man den invertierten Frequenzgang des Kanals mit dem Index n beziehungsweise der charakteristischen Frequenz f_c, und zwar derart, daß das bei der Frequenz f_c auftretende Minimum gerade die Höhe der Absoluthörschwelle hat. Der Frequenzgang wird durch

$$\begin{aligned} L(f)/\mathrm{dB} &= L_r/\mathrm{dB} + 3,3 - 20\lg|H_{ECR}(f)| \\ &+ 20\lg|H_n(f)| \end{aligned} \quad (15)$$

beschrieben. Er entspricht dem Frequenzgang der Tuningkurven des Gehörs. Die darin enthaltene Funktion $H_n(f)$ geht unter Zugrundelegung der wie oben zu ermittelnden Parameter a_n, ω_n aus (4) hervor. Eine Anzahl von Tuningkurven, welche mit (15) und den Parametern L_r = 70 dB, $k = 5$ berechnet wurden, ist in Abb. 3 dargestellt.

5. Frequenzskalierung. Die Zuordnung der f_c zu den Kanälen ergibt sich aus den Bandbreiten der CTF-Filter. Dem Kanal mit dem Index $\nu = 0$ wird die tiefste gerade noch mögliche oder sonstwie erwünschte charakteristische Frequenz zugeordnet. Bezeichnet man dieselbe mit $f_c(0)$, dann ergibt sich die zu $\nu = 1$ gehörende charakteristische Frequenz $f_c(1)$ aus

$$f_c(1) = f_c(0) + \varepsilon B[f_c(0)], \quad (16)$$

und so weiter. Darin ist $\varepsilon < 1$ ein fester Faktor. Beispielsweise ergeben sich die charakteristischen Frequenzen eines PET-Systems mit $N = 256$, L_r = 70 dB, $k = 5$, $f_c(0)$ = 25 Hz und $f_c(N-1)$ = 15000 Hz für $\varepsilon = 0,2562526$. Abb. 3 zeigt die dazugehörige Funktion $f_c(\nu)$.

6. Digitale Berechnung. Die Impulsantwort des ECR-Einzelfilters mit der Übertragungsfunktion (1) lautet

$$h_r(t) = \delta(t) + \alpha e^{-at}\left(\cos\omega_0 t - \frac{a}{\omega_0}\sin\omega_0 t\right), \quad (17)$$

und diejenige des CTF-Einfachfilters

$$h_n(t) = \frac{a_n^2 + \omega_n^2}{\omega_n} e^{-a_n t}\sin\omega_n t. \quad (18)$$

Die auf diesen Impulsantworten basierenden Faltungsintegrale kann man rekursiv berechnen, indem man die Rekursion

$$P^{\circ}_{m+1} = e^{-(a-j\omega)T_x} P^{\circ}_m + p_{m+1}\frac{1 - e^{-(a-j\omega)T_x}}{a - j\omega} \quad (19)$$

benutzt, denn im Real- bzw. Imaginärteil von P° kommen gerade diejenigen zeitabhängigen Terme vor, welche in (17,18) bzw. den Faltungsintegralen benötigt werden (p_{m+1} aktueller Abtastwert; T_x Abtastintervall). Die diskrete Spektralfunktion P° entspricht der mit $e^{j\omega t}$ multiplizierten FTT-Funktion mit Exponentialfenster (Terhardt, 1985).

Für die Abtastwerte des Ausgangssignals $q(t)$ eines ECR-Einzelfilters mit dem Eingangssignal $p(t)$ ergibt sich so die Formel

$$q_{m+1} = p_{m+1} + \alpha X_{m+1} - \frac{\alpha a}{\omega_n} Y_{m+1}, \quad (20)$$

worin X_{m+1}, Y_{m+1} Real- und Imaginärteil von P°_{m+1} bedeuten.

Für das Ausgangssignal des CTF-Einfachfilters ergibt sich

$$q_{m+1} = \frac{a_n^2 + \omega_n^2}{\omega_n} Y_{m+1}. \quad (21)$$

Der k–fachen Kaskadierung mehrerer CTF-Einfachfilter entspricht die Kaskadierung der Formel (21).

Literaturangaben

Terhardt, E. (1979). Calculating virtual pitch. Hear. Res. 1, 155-182.

Terhardt, E. (1985). Fourier transformation of time signals: Conceptual revision. Acustica 57, 242-256.