

# **Gehörrichtige Qualitätsbewertung von Audiosignalen - Übersicht und Einschätzung der gegenwärtigen Verfahren**

## **1. Einleitung**

Bei der digitalen Übertragung und Speicherung von Audiosignalen werden in zunehmendem Maße Datenreduktionsverfahren eingesetzt, die an die Eigenschaften des menschlichen Gehörs angepaßt sind und insbesondere Verdeckungseffekte ausnutzen. Dabei wird in erster Linie nicht versucht, die absolute Größe der bei einer vorgegebenen Datenrate auftretenden Fehler zu minimieren, sondern diese Fehler so zu formen, daß sie möglichst wenig hörbar sind. Daher würde ein konventionelles Meßverfahren einem derart verarbeiteten Audiosignal stets eine bedeutend geringere Qualität zuordnen als tatsächlich wahrgenommen wird.

Aus diesem Grunde sind zur Beurteilung derartiger Datenreduktionsverfahren, die oft als "perceptual coder" (gehörrichtige Codierverfahren) bezeichnet werden, heutzutage noch subjektive Hörtests üblich. Da diese Versuche sehr umfangreich sein müssen, um eine hinreichende Genauigkeit und Reproduzierbarkeit der Ergebnisse zu erreichen, ist der damit verbundene Zeit- und Arbeitsaufwand sehr hoch.

Die Entwicklung geeigneter Meßverfahren, die auch die Beurteilung von solchen gehörriichtigen Codierverfahren ermöglichen sollen, erfolgte weitgehend parallel zur Coderentwicklung. Die ersten gehörangepaßten Meßverfahren wurden Ende der 70er Jahre als Werkzeug für die Entwicklung und Optimierung von Sprach-Codierungsverfahren vorgestellt. Eine Anwendung auf Musiksignale erfolgte erst deutlich später in den 80er Jahren. Die Genauigkeitsanforderungen an diese Meßverfahren waren zu diesem Zeitpunkt relativ gering, da sie ausschließlich als Hilfsmittel in der Coderentwicklung, nicht aber zur abschließenden Qualitätsbeurteilung von Audio-Codecs verwendet wurden.

Durch die Standardisierungsbemühungen für Audiocodierverfahren (MPEG Layer I-III) und die absehbare Verbreitung psychoakustischer Datenreduktionsverfahren im Rundfunkbereich ergab sich in den letzten Jahren ein gesteigertes Interesse an entsprechenden Meßverfahren, so daß nochmals eine Vielzahl weiterer Verfahren entwickelt wurde, die nun auch eine für den Vergleich unterschiedlicher Codecs ausreichende Genauigkeit zum Ziel haben. Im folgenden

wird die prinzipielle Arbeitsweise solcher Meßverfahren erläutert und eine Übersicht über die zur Zeit vorhandenen Ansätze auf diesem Gebiet gegeben.

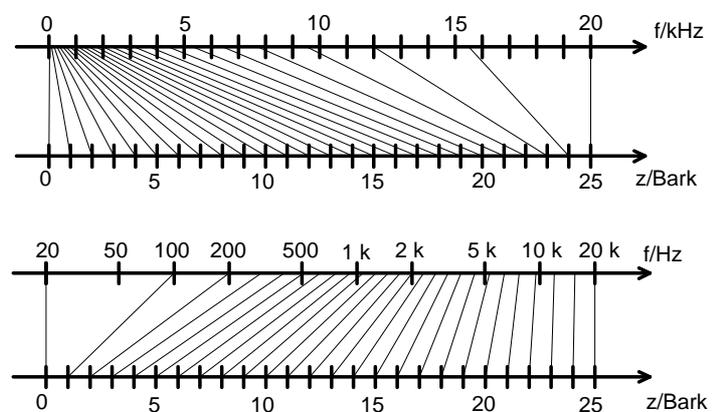
## 2. Arbeitsweise von gehörrichtigen Meßverfahren

Gemeinsames Kennzeichen aller gehörangepaßten Meßverfahren ist die Nachbildung von Verdeckungseffekten. Dabei wird eine Verdeckung zwischen gleichzeitig vorhandenen Signalanteilen (Simultanverdeckung) in allen Verfahren berücksichtigt, während zeitliche Verdeckungseffekte in vielen Verfahren vernachlässigt werden. Auch die Tonhöhenempfindung des menschlichen Gehörs wird immer nachgebildet, was allerdings eher wegen der sich daraus ergebenden Verringerung des Rechenaufwandes<sup>1</sup> geschieht als wegen des Einflusses der Tonhöhenkalen auf die Wahrnehmung von Störungen.

Als Eingangssignale dienen ein ungestörtes Referenzsignal sowie das durch das Testobjekt veränderte Signal (meist ein codiertes/decodiertes Signal). Wie diese Signale auszusehen haben, ist nicht von vornherein festgelegt. Üblicherweise werden die gleichen Signale verwendet wie in subjektiven Hörtests, es gibt aber Versuche, geeignete synthetische Meßsignale zu definieren.

Ebenso wie fast alle gehörrichtigen Codierverfahren arbeiten auch die bekannten gehörrichtigen Meßverfahren hauptsächlich im Frequenzbereich. Die Eingangssignale werden aus Gründen des Rechenaufwands meist mittels einer FFT aber z. T. auch durch (FIR-) Filterbänke in Kurzzeitspektren zerlegt. Oft folgt dann eine Gewichtung der Spektralkoeffizienten mit der Übertragungsfunktion zwischen Außen- und Innenohr, die dem Verlauf der Ruheshwelle bei

Frequenzen oberhalb von etwa 1 kHz entspricht. Anschließend wird die Frequenzkala auf eine Tonhöhenkala umgerechnet (Abb. 1). Da der Verlauf der Intensitätsdichte durch diese

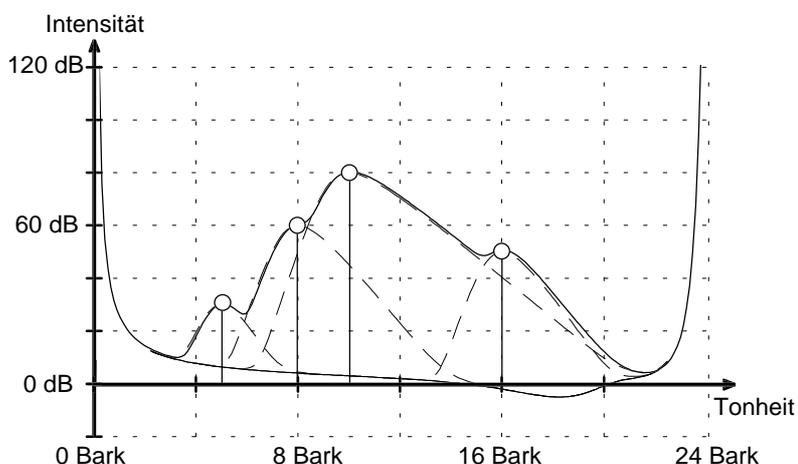


**Abb. 1: Skalentransformation von Frequenz auf Tonheit**  
(oben: lineare Frequenzkala, unten: logarithmische Frequenzkala)

<sup>1</sup> Zum einen lassen sich die Verläufe der Verdeckungsmuster einfacher beschreiben und zum anderen verringert sich die Anzahl der zum Erreichen einer bestimmten Genauigkeit erforderliche Anzahl von Stützstellen bei der Verdeckungsberechnung.

Skalentransformation verzerrt wird, wird die entsprechende Rechengröße hier in Anlehnung an Zwicker [19] als Anregungsdichte bezeichnet.

Zur Nachbildung der Frequenzselektivität des menschlichen Gehörs wird jeder Stützstelle dieser Funktion ein Mithörschwellenverlauf zugeordnet, der meist als Erregungsverlauf interpretiert wird und den man allgemein als "spreading function" (Verschmierungsfunktion) bezeichnet (Abb. 2). Für den Fall, daß diese Verläufe

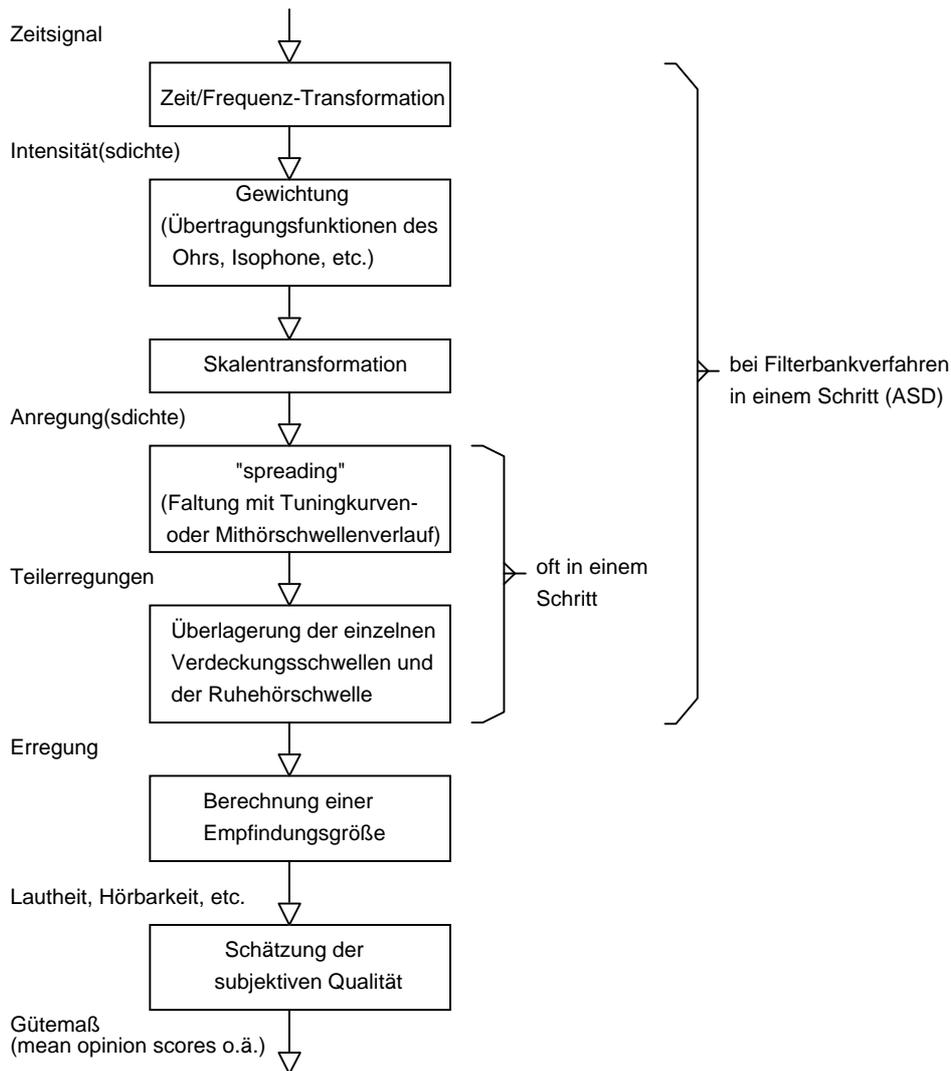


**Abb. 2: Verschmierung im Frequenzbereich durch Faltung eines Linienspektrums mit einer "spreading function".**  
*Beispiel: 4 reine Töne + Ruhehörschwelle, Überlagerung durch lineare Addition.*

linear addiert werden, entspricht diese Operation einer Faltung der Anregungsdichte mit der Verschmierungsfunktion, es werden aber oft auch andere Arten der Überlagerung gewählt, die weiter unten betrachtet werden. Die resultierende Funktion wird als Verdeckungs- oder Erregungsmuster bezeichnet. In einem Teil der Verfahren wird hieraus über eine weitere Umformung eine Empfindungsgröße berechnet, meist die spezifische Lautheit des Signals. Der bis hierher beschriebene Ablauf wird in Abb. 3 gezeigt.

Zeitliche Verdeckungseffekte werden meist durch eine exponentielle Mittelung über aufeinanderfolgende Signalblöcke nachgebildet.

Die vom Testobjekt verursachten Störungen werden aus der Differenz von Test- und Referenzsignal ermittelt. An welcher Stelle dies im Laufe der beschriebenen Umformungen geschieht und auf welche Weise hieraus ein Maß für den Einfluß dieser Störungen auf den Höreindruck berechnet wird, unterscheidet sich zwischen den einzelnen Meßverfahren allerdings erheblich.



**Abb 3: Modellierung des Hörvorganges in einem gehörrihtigen Beurteilungsverfahren**

### 3. Gegenüberstellung typischer Vorgehensweisen

Die zur Zeit bekanntesten gehörrihtigen Meßverfahren sind das am Institut für Integrierte Schaltungen der Fraunhofergesellschaft FhG-IIS entwickelte *NMR-Verfahren* ("Noise to Mask Ratio"), das etwa seit 1987 angewendet wird, und das beim Forschungslaboratorium der niederländischen PTT entwickelte *PAQM-Verfahren* ("Perceptual Audio Quality Measure"), das in seiner heutigen Ausführung erstmals 1992 vorgestellt wurde. Daneben fand kurzzeitig auch das an der Universität von Sherbrooke in Kanada entwickelte *PERCEVAL-Verfahren* ("PERCeptual EVALuation") Beachtung.

Diese drei Verfahren repräsentieren in den meisten Ablaufschritten typische Ansätze, die in unterschiedlicher Kombination auch bei den übrigen Meßverfahren auftauchen.

### 3.1 Auflösung im Frequenzbereich

Der zwischen den verschiedenen Verfahren am auffälligsten variierende Parameter ist die durch die Anzahl der Stützstellen der Anregungsfunktion vorgegebene *Auflösung im Frequenzbereich*.

Der zu der niedrigsten Auflösung von einem Bark führende Ansatz ist die im *NMR-Verfahren* gewählte Orientierung an der Bandbreite, innerhalb derer das Gehör bei der Bildung der Hörschwelle und der

Lautheitsempfindung spektrale Signalanteile zusammenfaßt

(Frequenzgruppen; engl.: "critical bands"). Selbst unter der Voraussetzung, daß eine Aufteilung in

Frequenzgruppen bei der Verdeckungsrechnung zulässig wäre, was durchaus zweifelhaft ist, führt diese Betrachtungsweise zu erheblichen Fehlern, da sich Frequenzgruppen um beliebige Mittenfrequenzen bilden können, während sie in den entsprechenden Verfahren durch feststehende Frequenzbänder nachgebildet werden.

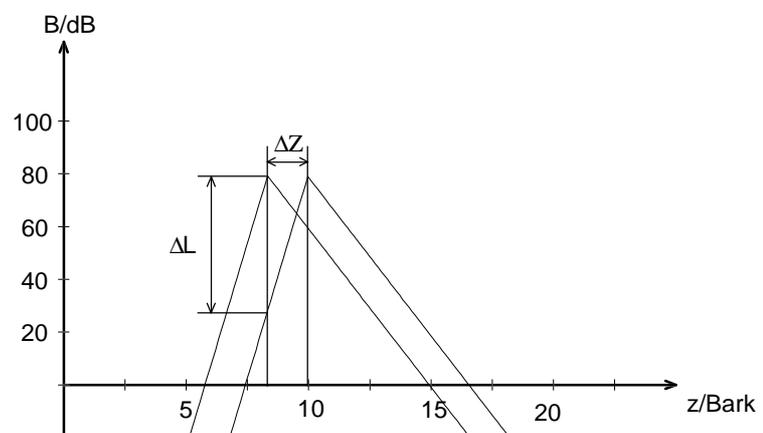
Da die hierdurch verursachten Fehler entsprechend der Flankensteilheit der Mithörschwellenverläufe in ungünstigen Fällen in einer Größenordnung von über 20 dB liegen können (siehe *Abb. 4, Gl. 1*), ist dieser Ansatz nicht zu empfehlen.

Die im *PERCEVAL-Verfahren* vorgesehene Orientierung an der Anzahl von Sinneszellen

entlang der Basilarmembran führt dagegen mit einem hundertstel Bark zu der mit Abstand größten spektralen Auflösung. Dieser Wert entspricht dem durch den Aufbau des Innenohres vorgegebenen maximalen Tonhöhenunterscheidungsvermögen. Traditionell wird zwar ein

**Größenordnung der als Folge der begrenzten Auflösung im Frequenzbereich möglichen Fehler**

$$\Delta L = 27 \frac{\text{dB}}{\text{Bark}} \cdot \Delta z \quad \left| \begin{array}{l} \Delta z : \text{Stützstellenabstand in Bark} \\ \Delta L : \text{mögliche Abweichungen (als Pegel)} \end{array} \right. \quad (1)$$



**Abb. 4: Einfluß eines Frequenzversatzes auf die Höhe der Verdeckungsschwelle an einem Punkt des Spektrums**

Die Mithörschwellen werden hier zur Vereinfachung als Dreiecksfunktionen angenommen.

deutlich größerer Wert von 0,04 Bark nach [19] angenommen, neuere Messungen von Wier et al. [18] und Terhardt [17] führten jedoch zu Werten, die tatsächlich in der oben genannten Größenordnung liegen.

Das Problem bei einer derart hohen Auflösung liegt einerseits im Rechenaufwand und andererseits darin, daß die Stützstellenanzahl der eingangsseitigen FFT zur Wahrung einer ausreichenden zeitlichen Auflösung auf eher kleine Werte (hier: 1024 Samples/Block) begrenzt ist. Daher ist die hohe Stützstellenanzahl der Anregungsfunktion nicht durch eine entsprechende Zahl direkt berechneter Werte abgedeckt<sup>2</sup>.

Im *PAQM-Verfahren* wird der eher pragmatische Ansatz gewählt, die Stützstellendichte der Anregungsfunktion auf einen Wert zu begrenzen, der gerade noch durch die Anzahl der bei der FFT berechneten Spektralkoeffizienten abgedeckt wird. Bei der gewählten Blocklänge von 2048 Samples führt dies zu einer Auflösung von 0,2 Bark, was einen Kompromiß zwischen den beiden zuvor genannten Werten darstellt.

### **3.2 Pegelabhängigkeit der Mithörschwellen**

Der Verlauf der Mithörschwellen ist von der Frequenz und dem Pegel des verdeckenden Signals abhängig<sup>3</sup>. In der Darstellung über einer gehörrichtigen Tonhöenskala verschwindet die Frequenzabhängigkeit fast völlig. Der verbleibende Verlauf wird meist nach einem Vorschlag von Terhardt [16] durch ein exponentielles Abklingen der Verdeckungswirkung bei zunehmendem Abstand zwischen verdeckendem und verdecktem Signal nachgebildet, was bei der in der Akustik üblichen Pegeldarstellung zu einer Dreiecksfunktion führt. Die ansteigende Flanke wird dabei als pegelunabhängig angenommen, während sich die abfallende Flanke bei zunehmendem Pegel des Verdeckers abflacht (*Abb. 5*). Diese *pegelabhängige Modellierung* wird u.a. im *PAQM-Verfahren* verwendet. Sofern der Abhörpegel des Testsignals im voraus bekannt ist, liefert diese Vorgehensweise zweifellos die genauesten Ergebnisse; diese Bedingung ist jedoch praktisch meist nicht erfüllt.

---

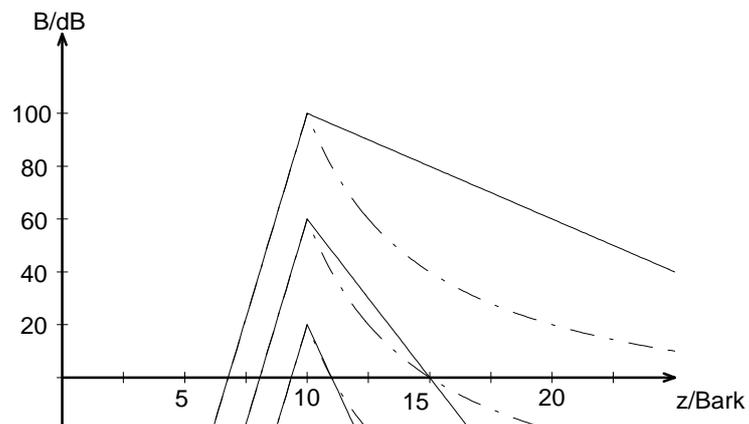
<sup>2</sup> Wegen des nichtlinearen Zusammenhanges zwischen Frequenz und Tonhöhe wäre hierzu eine FFT mit 65536 Samples/Block erforderlich (*vgl. Abb. 1*), was zu einer Blocklänge von mehr als einer Sekunde und zu einem gewaltigen Rechenaufwand führen würde.

<sup>3</sup> Es besteht auch eine erhebliche Abhängigkeit von der Art des verdeckenden Signals, auf die weiter unten eingegangen wird.

Im *PERCEVAL-Verfahren* wird dieses Problem durch Wahl eines festen Mithörschwellenverlaufes, der einem *durchschnittlichen Signalpegel* entspricht, vermieden. Die dabei auftretenden Fehler können jedoch erheblich sein, wie sich durch Vergleich der in *Abb. 5* gezeigten Verläufe zeigen läßt. Selbst wenn tatsächlich der für die verwendete Verdeckungskurve zugrunde gelegte Signalpegel eingestellt wird, können die für die Verdeckung von Störsignalen entscheidenden Anteile deutlich von diesem mittleren Pegel abweichen, so daß auch dann noch Fehler auftreten.

Hinzu kommt, daß eine Überschätzung von Verdeckungseffekten zu einer Einstufung von hörbaren Störungen als nicht wahrnehmbar führen kann und daher unbedingt vermieden werden sollte. Dementsprechend müßten für eine pegelunabhängige Beschreibung sicherheits- halber die steilsten vorkommenden Mithörschwellenflanken verwendet werden, was jedoch bei mittleren und hohen Signalpegeln zu noch größeren Fehleinschätzungen führen würde.

Ein guter Kompromiß zwischen Genauigkeit und Unabhängigkeit vom Abhörpegel wird im *NMR-Verfahren* durch die Verwendung einer *"worst case"-Kurve* erreicht, die für alle Signalpegel zu einer Verdeckung knapp unterhalb der pegelabhängigen Kurven führt (*Abb. 5*). Wie man in der Abbildung erkennt, weicht



**Abb. 5: Vergleich zwischen pegelabhängigen Mithörschwellenverläufen und einer "worst-case"-Kurve entsprechend [3].**

diese Kurve insgesamt deutlich weniger von den pegelabhängigen Verläufen ab, als es bei Verwendung einer Verdeckungskurve für einen niedrigen Signalpegel der Fall wäre.

### **3.3 Addition von Mithörschwellen / Abhängigkeit des Verdeckungsmaßes von der Signalform**

Um aus den zu den einzelnen Stützstellen berechneten Verdeckungskurven das vollständige Verdeckungs- bzw. Erregungsmuster zu bestimmen, muß bekannt sein, wie sich die einzelnen Teilerregungen addieren.

Aus Versuchen zum Zusammenwirken von Mithörschwellen ([7], [20]) ist bekannt, daß diese in Anwesenheit mehrerer Verdeckter überproportional ansteigen. Diesem Umstand wird im *PAQM-Verfahren* durch eine Potenzierung der Teilerregungen vor ihrer Addition Rechnung getragen. Im *PERCEVAL-Verfahren* und im *NMR-Verfahren* wird dieser Effekt dagegen nicht berücksichtigt. Hier werden die Erregungsmuster einfach durch eine lineare Addition der Teilerregungen (*PERCEVAL*) bzw. durch Verwendung ihres Maximalwertes (*NMR*) berechnet.

Ein eng hiermit zusammenhängender Effekt ist die Abhängigkeit des *Verdeckungsmaßes*<sup>4</sup> von der Signalform des maskierenden und des maskierten Signals: das Verdeckungsmaß von frequenzgruppenbreitem Rauschen liegt im Mittel um etwa 20 dB über dem von reinen Tönen.

Diese Abhängigkeit wird im *NMR-Verfahren* vernachlässigt. Im *PAQM-Verfahren* ergibt sich dieser Unterschied indirekt durch die nichtlineare Addition der Teilerregungen. Im *PERCEVAL-Verfahren* ergibt sich ein ähnlicher Zusammenhang ebenfalls indirekt, hier durch die Art der Berechnung der Wahrnehmbarkeitswahrscheinlichkeit aus den Erregungsmustern. Bei Verfahren, die wie das *NMR-Verfahren* mit frequenzgruppenbreiten Bändern arbeiten, ist auch eine direkte Berücksichtigung der unterschiedlichen Verdeckungswirkungen möglich. Dieser Ansatz wird in einem der weiter unten aufgeführten Verfahren (*Verfahren nach Kapust*) verwendet.

### 3.4 Zeitliche Verdeckungseffekte

Einige Modelle wie z. B. das *PERCEVAL-Verfahren* verzichten völlig auf die Berücksichtigung zeitlicher Verdeckungseffekte. Beim *NMR-* und *PAQM-Verfahren* wird eine sehr grobe Modellierung durch ein exponentielles Abklingen der berechneten Verdeckungs- bzw. Erregungsverläufe angewendet. Diese Näherung ist zwar bekanntermaßen sehr ungenau (siehe z. B. [19]), eine genauere Formulierung dieser Effekte ist bei Verfahren, die mit einer FFT arbeiten und dabei den Zeitbereich entsprechend der Blocklänge der FFT aufrastern, aber auch kaum sinnvoll.

Die bisher mit diesen beiden Verfahren gemachten Erfahrungen scheinen darauf hinzuweisen, daß die Modellierung einer zeitlichen Verdeckung keinen allzu großen Einfluß auf die

---

<sup>4</sup> Der Abstand zwischen dem Pegel eines Maskierers und dem Maximalwert seiner Verdeckungswirkung wird in Pegelschreibweise als *Verdeckungsmaß*, sonst auch als *Schwellenfaktor* (z. B. in [19]) bezeichnet.

Qualität der erzielten Vorhersagen hat. Insbesondere sind Signale, bei denen ein deutlicher Einfluß solcher Effekte zu erwarten ist, für ungeübte Testhörer offenbar ohnehin nur schwer zu beurteilen, so daß die Unterschiede zwischen geübten und ungeübten Versuchspersonen größer sind, als die selbst bei groben Fehlern der Verfahren auftretenden Ungenauigkeiten. Diese Frage ist insgesamt jedoch noch nicht ausreichend geklärt, um etwa einen grundsätzlichen Verzicht auf die Modellierung von Vor- und Nachverdeckung zu begründen.

### 3.5 Berechnete Ausgangsgrößen

Die genannten Verfahren repräsentieren drei grundsätzlich verschiedene Ansätze zur Ermittlung eines Gütemaßes:

- a) Das *NMR-Verfahren* berechnet aus dem Referenzsignal ein Verdeckungsmuster und bestimmt die *Pegeldifferenz zwischen dem Fehlersignal und dieser Verdeckungsschwelle*. Das Fehlersignal kann dabei optional entweder aus der Differenz der Zeitsignale oder aus der Differenz der Betragsspektren ermittelt werden.

Mit diesem Ansatz kann nicht ohne weiteres ein zur Abschätzung des Höreindrucks geeigneter Einzahlwert gewonnen werden, da die auftretenden Pegeldifferenzen nicht der Arbeitsweise des menschlichen Gehörs entsprechen. Als Vorzug ist dagegen zu bemerken, daß bei dieser Verfahrensweise auch Störungen unterhalb der Verdeckungsschwelle erfaßt werden und dadurch eine Abschätzung von Tendenzen möglich ist. Dies kann insbesondere bei der Weiterentwicklung von Codierverfahren und bei der Beurteilung der Kaskadierbarkeit von Codern<sup>5</sup> von Bedeutung sein.

- b) Das *PAQM-Verfahren* bestimmt die *Lautheit des Fehlersignals* aus der Differenz der für Test- und Referenzsignal getrennt ermittelten spezifischen Lautheiten. Eine Besonder-

**Spezifische Lautheit**, Einheit: *sone/Bark* nach [19]

$$N' = k \cdot \left( \frac{1}{s} \cdot \frac{E_{HS}}{E_0} \right)^\gamma \cdot \left[ \left( 1 - s + s \cdot \frac{E}{E_{HS}} \right)^\gamma - 1 \right] \quad (2)$$

k : Skalierungskonstante; in [19]: k = 0.068

$\gamma$  : Kompression; in [19]:  $\gamma = 0.23$

s : Schwellenfaktor (siehe 3.3)

$E_{HS}$  : Erregung an der Hörschwelle

$E_0$  : Bezugswert

<sup>5</sup> Es ist allerdings fraglich, ob für eine solche Anwendungen der Einsatz eines gehörrichtigen Qualitätsbewertungsverfahrens überhaupt sinnvoll ist.

heit des Verfahrens liegt darin, daß die Skalen der spezifischen Lautheiten durch eine Änderung des bei der Lautheitsberechnung nach Zwicker [19] auftauchenden Exponenten (vgl. Gl. 2) gestaucht werden, was zu einer indirekten Berücksichtigung von Verdeckungseffekten führt ("masking by smearing and compression").

Ein solcher Ansatz bietet die Möglichkeit, hörbare Störungen sinnvoll zu klassifizieren und hieraus einen aussagekräftigen Einzahlwert zu gewinnen. Die subjektiv empfundene Lästigkeit muß zwar nicht der Störlautheit proportional sein, ihre Größe ist jedoch mit Sicherheit eine entscheidend zur Lästigkeit beitragende Empfindungsgröße.

- c) Beim *PERCEVAL-Verfahren* wird aus der Differenz der logarithmierten Erregungsmuster von Test- und Referenzsignal für kleine, jeweils einer Haarzelle entsprechende Frequenzbereiche eine *Wahrscheinlichkeit für die Hörbarkeit einer Störung* bestimmt ("probability of detection"). Nach den Regeln der Wahrscheinlichkeitsrechnung wird hieraus eine Gesamtwahrscheinlichkeit für die Wahrnehmbarkeit einer Störung berechnet.

Der Vorzug dieses Ansatzes liegt in der potentiell sehr genauen Abschätzung von Störungen nahe der Verdeckungsschwelle. Dagegen ist zwischen starken, eindeutig hörbaren Störungen keinerlei Differenzierung mehr möglich.

Als ein weiteres mögliches Qualitätsmaß kann auch die *Auftrittshäufigkeit hörbarer Störungen* verwendet werden, wie es im *NMR-Verfahren* durch die Ermittlung der "masking flags" geschieht.

#### **4. Übersicht über die bekannten Meßverfahren**

In diesem Abschnitt werden die vorhandenen Meßverfahren in chronologischer Reihenfolge aufgeführt und ihre wichtigsten Merkmale zusammengefaßt (die anfangs beschriebenen Umformungsschritte, die allen Verfahren gemeinsam sind, werden dabei nicht mehr gesondert erläutert).

Die ältesten unter den hier aufgeführten Verfahren, *Speech Signal Degradation* und *ASD*, wurden ursprünglich zur Beurteilung von Sprach-Codecs entwickelt. Sie weisen jedoch bereits die typischen Merkmale der hier behandelten Meßverfahren auf und sind als Vorgänger der folgenden Verfahren zu sehen.

#### **4.1 Verfahren nach Schroeder, Atal, Hall ("Speech Signal Degradation"):**

Dieses Verfahren ist eines der frühesten gehörrichtigen Beurteilungsverfahren, die aus der Fachliteratur bekannt sind. Es diente wahrscheinlich als Vorbild für viele der neueren Verfahren. Das Verfahren berechnet die Lautheit des Störsignals, wobei die Drosselung der Störung durch das Nutzsignal explizit modelliert wird. Der Mithörschwellenverlauf wird durch eine pegelunabhängige Funktion beschrieben, die im Gegensatz zu der meist üblichen Näherung durch eine Dreiecksfunktion einen abgerundeten Verlauf aufweist und aus eigenen Hörversuchen für einen 1-kHz-Ton bei einem Schalldruckpegel von 80 dB(SPL) gewonnen wurde.

Literatur: [14]

#### **4.2 ASD ("Auditory Spectrum Distance"):**

Das Verfahren berechnet die Lautheitsdifferenz zwischen gestörtem Signal und Referenzsignal. Es ist das einzige Filterbankverfahren unter den bekannten Ansätzen. Die zeitliche Auflösung ist daher sehr gut, während die Auflösung im Frequenzbereich mit 0,5 Bark eher gering ist. Dabei muß angemerkt werden, daß die durch eine begrenzte Anzahl von Stützstellen verursachten Fehler bei einem Filterbankverfahren weit geringer sind als es bei anderen Verfahren der Fall wäre.

Durch Grundverstärkung, Mittenfrequenz und Amplitudenfrequenzgang der Teilfilter werden die bei den übrigen Verfahren nacheinander ausgeführten Schritte der Gewichtung, Skalentransformation und Mithörschwellenberechnung direkt realisiert. Die Filtercharakteristiken wurden aus dem von Schroeder et al. (siehe 4.1) vorgeschlagenem Mithörschwellenverlauf abgeleitet. Durch eine nachfolgende Filterung der quadrierten Ausgangssignale mit einem nichtlinearen Tiefpaß werden auch zeitliche Verdeckungseffekte berücksichtigt.

Literatur: [12]

Die drei bekanntesten Verfahren, *NMR*, *PAQM* und *PERCEVAL*, sind oben bereits weitgehend beschrieben. Sie lassen sich kurz folgendermaßen charakterisieren:

#### **4.3 NMR ("Noise to Mask Ratio"):**

Das Verfahren berechnet eine Pegeldifferenz zwischen Verdeckungsschwelle und Störsignal. Die Auflösung im Frequenzbereich ist sehr gering, wodurch die erforderliche Rechenleistung begrenzt bleibt. Daher konnte das Verfahren bereits frühzeitig in Echtzeit realisiert werden. Andererseits sind, wie oben gezeigt wird, die hierdurch möglichen Fehleinschätzungen der Verdeckungsschwellen erheblich. Das Verfahren wird bereits seit 1987 praktisch verwendet, wodurch seine grundsätzliche Brauchbarkeit trotz der aufgezeigten Schwächen in gewissen Grenzen als erwiesen gelten kann. Die Modellierung des Hörvorgangs ist insgesamt eher grob.

Literatur: [4]

#### **4.4 PAQM ("Perceptual Audio Quality Measure"):**

Das Verfahren berechnet eine komprimierte Lautheitsdifferenz zwischen gestörtem Signal und Referenzsignal. Die Auflösung im Frequenzbereich ist durchschnittlich. Das benutzte Gehörmodell ist eines der vollständigsten unter den bekannten Verfahren (siehe oben). Bei der Betrachtung der Korrelation zwischen den Ausgangswerten des Verfahrens und den Ergebnissen subjektiver Hörtestreihen liefert dieses Verfahren die bislang mit Abstand besten Ergebnisse (vgl.: *Tabelle 2*). Wegen der experimentellen Optimierung der Parameter des Gehörmodells stimmt dieses in einigen Punkten nicht mehr mit grundlegenden Erkenntnissen der Psychoakustik überein. Daher könnten bei unerwarteten Signaltypen durchaus wieder größere Fehler auftreten, was allerdings bei der bisherigen Anwendung auf Musik-Codecs offenbar nicht vorkam, wohl aber bei Sprach-Codecs, für die daraufhin eine spezielle Modifikation dieses Verfahrens ("Perceptual Speech Quality Measure", *PSQM*) entwickelt wurde.

Literatur: [1]

#### **4.5 PERCEVAL ("PERCeptual EVALuation"):**

Das Verfahren berechnet eine Wahrscheinlichkeit für die Hörbarkeit einer Störung. Die Auflösung im Frequenzbereich ist sehr hoch, wird aber nicht durch eine entsprechende Anzahl tatsächlich berechneter Spektralkoeffizienten abgedeckt. Das benutzte Gehörmodell ist ansonsten relativ einfach. Insbesondere werden weder zeitliche Verdeckungseffekte noch die Pegelabhängigkeit der Mithörschwellen berücksichtigt.

Literatur: [13], [9]

Es gibt noch eine Reihe weniger bekannter Verfahren, von denen abschließend die neuesten und wahrscheinlich auch wichtigsten genannt werden sollen:

#### **4.6 POM ("Perceptual Objective Model"):**

Auch dieses Verfahren berechnet eine Wahrscheinlichkeit für die Hörbarkeit einer Störung. Die Auflösung im Frequenzbereich ist etwas größer als beim *PAQM-Verfahren*. Ansonsten ist das benutzte Gehörmodell nahezu mit dem im *PAQM-Verfahren* verwendeten identisch, wobei der Verlauf der Mithörschwellen durch eine etwas genauere Näherung beschrieben wird, die sowohl die Pegelabhängigkeit nach Terhardt [16] als auch den gerundeten Verlauf nach Schroeder et al. [14] berücksichtigt. Dennoch konnten mit diesem Verfahren bisher keine befriedigenden Ergebnisse bei der Vorhersage von Hörtestergebnissen erzielt werden.

Literatur: [5], [6]

#### **4.7 Verfahren nach Kapust:**

Das Verfahren berechnet wie das *NMR-Verfahren* eine Pegeldifferenz zwischen Verdeckungsschwelle und Störsignal. Es kann als Weiterentwicklung dieses Verfahrens angesehen werden. Die Auflösung im Frequenzbereich ist nach wie vor sehr gering, so daß hier bezüglich der erreichbaren Genauigkeit dieselben Beschränkungen gelten wie beim *NMR-Verfahren*. Die zeitliche Auflösung ist dagegen sehr gut. Verglichen mit dem *NMR-Verfahren* ist das verwendete Gehörmodell erheblich verbessert worden. Insbesondere wird die Pegelabhängigkeit der Verdeckungsschwellen und die unterschiedliche Verdeckungswirkung zwischen tonalen und rauschartigen Signalen modelliert.

Literatur: [11]

#### **4.8 Verfahren nach Stuart:**

Das in diesem Verfahren verwendete Gehörmodell unterscheidet sich sowohl in den verwendeten Parametern als auch in einigen Teilen des Ablaufs erheblich von den zuvor genannten Verfahren. So wird eine Verzerrungsfunktion zur Modellierung von Nichtlinearitäten des Gehörs verwendet und bei der Berechnung der Verdeckungsmuster wird der zur Frequenzselektivität des Gehörs beitragende Effekt der lateralen Inhibition explizit modelliert. Da die Eingangssignale als Spektren vorliegen müssen (Terzfilterbank oder externer FFT-Analysator), können offenbar *nur stationäre Signale* verarbeitet werden.

Leider existieren noch keine Vergleichstests mit subjektiven Beurteilungen wie sie für die ersten drei Verfahren in *Tabelle 2* vorliegen.

Literatur: [15]

### **5. Möglichkeiten und Grenzen der Verfahren**

Für Eingangssignale, bei denen die auftretenden Störungen nahe der Verdeckungsschwelle liegen, sind prinzipiell sehr gute Vorhersagen möglich, da die individuellen Schwankungen speziell der Simultanverdeckungsschwellen nicht sehr groß sind. Auch bei etwas stärker gestörten Signalen sollten immer noch Vorhersagen innerhalb der bei Hörversuchen auftretenden (dann aber deutlich größeren) Standardabweichungen möglich sein. Hier gewinnt dann die genaue Abschätzung der Lautheit der Störung eine immer größere Bedeutung, die aber bei keinem der heute aktuellen Verfahren in befriedigender Weise vorgenommen wird (eher schon in den älteren Verfahren *ASD* und *Speech Signal Degradation*). Wenn die Störungen dagegen so stark werden, daß sie nicht nur eindeutig wahrnehmbar sind, sondern bereits die Art der Störung erkannt wird, spielen immer mehr auch subjektive Vorlieben der Testhörer eine Rolle. In diesem Bereich kann daher keine brauchbare Vorhersage eines Höreindrucks mehr erwartet werden, es sei denn, es werden nur Testobjekte miteinander verglichen, die dem übertragenen Signal gleichartige Störanteile hinzufügen.

Bei der Beurteilung von Mehrkanalübertragungen tritt eine Reihe zusätzlicher Probleme auf, die zwar in erster Linie die Codierverfahren betreffen, aber von einem Beurteilungsverfahren ebenfalls in angemessener Weise berücksichtigt werden müssen. So kann hier die Verdeckungsschwelle durch Auslöschungen im freien Schallfeld und durch eine für das Gehör mögliche Trennung von Nutz- und Störsignal anhand der Richtungsinformation bekannterweise merklich herabgesetzt werden.

Auch die durch Nachbearbeitung (z. B. Klangregler, Equalizer) und Raumeigenschaften entstehenden Frequenzgänge können sich auf die Verdeckungsschwellen auswirken, was die Allgemeingültigkeit der gewonnenen Beurteilungen ebenfalls einschränkt (dies gilt natürlich auch für subjektive Hörtests).

## 6. Forderungen an ein gehörrichtiges Beurteilungsverfahren

Für die Anwendung eines gehörrichtigen Beurteilungsverfahrens ist allein die Ermittlung eines aussagekräftigen Maßes für die Qualität einer Audioübertragung entscheidend. Die konkrete (physikalische oder psychophysische) Bedeutung dieser Größe ist dabei letztendlich unerheblich. Um aber zunächst die Gültigkeit der verwendeten Modellierungen des Hörvorgangs nachzuweisen und mit anderen Verfahren zu vergleichen, müssen sich mit einem solchen Verfahren auch direkt oder indirekt Meßgrößen ermittelt lassen, die anhand subjektiver Hörtests überprüfbar sind. Dies sind insbesondere:

- a) Die *Vorhersage von subjektiven Bewertungen* der Qualität von Audiosignalen (meist als "mean opinion scores", MOS).

Dies ist von besonderer Bedeutung, da in der Ergänzung (bis hin zum teilweisen Ersatz) der entsprechenden Hörtestreihen der zur Zeit interessanteste Anwendungsbereich solcher Verfahren liegt.

- b) Die *Nachempfindung bekannter Gehöreigenschaften* (speziell: Lautheitsempfindung mit und ohne Drosselung, Frequenzgruppenbildung, Verdeckungsschwellen).

Ein funktionierendes Gehörmodell muß selbstverständlich in der Lage sein, die aus der psychoakustischen Literatur bekannten Gehöreigenschaften hinreichend genau wiederzugeben. Dies ist wichtig, weil bei einer ausschließlichen Überprüfung eines Gehörmodells anhand der Vorhersage von subjektiven Bewertungstests (wobei praktisch wegen des zu einer statistisch abgesicherten Auswertung notwendigen Umfangs immer die Testreihen zu den im ISO-Standard festgelegten Codierverfahren herangezogen werden) unerwartete Beschränkungen der Gültigkeit aufgrund der begrenzten Signalauswahl vorkommen können.

Um bei der späteren Anwendung der Verfahren Fehlbewertungen bei unerwarteten Eingangssignalen zu vermeiden, sollten die Verfahren weiterhin

- c) für *beliebige Signaltypen* (also z. B. auch für Sprache und impulshaltige Signale) anwendbar sein

und

d) sich auf ein *sinnvolles, zum Verständnis des Hörvorgangs beitragendes Modell* stützen.

Die Bedeutung dieser Forderung läßt sich an zwei Extremfällen illustrieren. Der erste wäre ein Modell, das versucht, die Eingangssignale in ein Schema von bekannten Hörversuchen einzuordnen und als Ausgangswerte die hierzu bekannten Versuchsergebnisse wiedergibt. Ein solches Modell könnte die in *b)* genannte Forderung perfekt erfüllen und würde dennoch bei unbekanntem Signalkonstellationen mit hoher Wahrscheinlichkeit völlig versagen. Dieser Teil der Forderung, nicht die Summe von Höreigenschaften wiederzugeben, sondern das dazu führende Funktionsprinzip, wurde bereits in ähnlicher Weise in [15] formuliert.

Der andere Extremfall wäre der gelegentlich erwähnte Ansatz, Verfahren der künstlichen Intelligenz (KI-Verfahren) zu verwenden. In diesem Fall könnten zwar zweifellos schnell sehr gute Ergebnisse insbesondere bei der Vorhersage von Bewertungstests erzielt werden, das Verfahren wäre aber letztlich auf die in der Trainingsphase verwendeten Signaltypen begrenzt. Es ist zwar prinzipiell möglich, daß sich dabei ein gutes und allgemeingültiges Hörmodell ergibt, da man praktisch jedoch nicht weiß, was das Verfahren wirklich berechnet, kann man hiervon auch niemals mit Sicherheit ausgehen.

## **7. Einschätzung der heute bekannten Meßverfahren**

Beim Vergleich der von den Meßverfahren gelieferten Ausgangsgrößen mit den Ergebnissen subjektiver Bewertungstests (*Tabelle 2*) zeigt sich eine deutliche Überlegenheit des *PAQM-Verfahrens* gegenüber den anderen getesteten Verfahren. Problematisch ist die hier vorgenommene experimentelle Parameteroptimierung, die zwar einerseits die guten Ergebnisse des Verfahrens erst ermöglicht hat, andererseits aber zu deutlichen Abweichungen des Hörmodells von bekannten Höreigenschaften führt. So ergibt sich durch die Art der Modellierung einer nichtlinearen Addition von Verdeckungsschwellen nicht nur eine (korrekterweise) erhöhte Verdeckungswirkung von Schmalbandrauschen gegenüber einem reinen Ton, sondern auch eine (fälschlicherweise) erhöhte Lautheit. Auch die vorgenommene Kompression der spezifischen Lautheitsfunktion hat sehr viel weitreichendere Auswirkungen, als es zunächst scheint, so daß hier letztendlich bereits die oben auf KI-Verfahren bezogene Kritik gilt: man weiß nicht mehr, was das Verfahren eigentlich berechnet. Diese Frage kann und sollte noch geklärt werden, ehe man sich für oder gegen die Verwendung des *PAQM-Verfahrens* entscheidet. Ein weiterer Nachteil des Verfahrens liegt darin, daß der Abhörpegel des Testsignals im voraus bekannt sein muß.

Testreihe	Bedingungen	NMR		PAQM		PERCEVAL	
		$\rho$	$\sigma$	$\rho$	$\sigma$	$\rho$	$\sigma$
ISO/ MPEG 1990	vollständige Auswahl (50 Teststücke)	-	-	0,97	0,35	0,92 (0,93)	0,57 (0,53)
	Teilauswahl (18 Teststücke)	0,94	0,27	0,98	0,17	-	-
ITU- TG/10-2 1993	contribution, distribution, emission	0,436	0,461	0,829 (0,91)	0,287 (0,22)	0,554	0,473
	commentary mono	0,491	0,465	0,850	0,281	-	-
	commentary stereo	0,330	0,797	0,493	0,735	-	-

$\rho, \sigma$  : Korrelationskoeffizient, Standardabweichung

( ) : verbesserte Ergebnisse nach Einführung der "basilar degradation" anstelle der "average probability of detection" im PERCEVAL-Verfahren [9] bzw. der "cognitive correction" im PAQM-Verfahren [2]

**Tabelle 2: Übersicht über die erreichte Übereinstimmung mit den Ergebnissen subjektiver Hörtests (nach [8] und [10])**

Hier liegt ein klarer Vorzug der im *NMR-Verfahren* verwendeten "worst case"-Kurve. Auch die bei diesem Verfahren vorgesehene Ausgabe der Pegeldifferenzen zur Verdeckungsschwelle kann für einige Anwendungen in der Coderentwicklung sicherlich sinnvoll sein. Dies kann die Berechnung einer Lautheit oder einer in vergleichbarer Weise an der Arbeitsweise des Gehörs orientierten Größe aber nicht ersetzen, sondern nur ergänzen.

Beim *NMR-Verfahren* sowie auch beim *Verfahren nach Kapust* wird die erreichbare Genauigkeit durch die grobe Auflösung im Frequenzbereich so stark eingeschränkt (vgl. Gl. 1), daß von diesen Verfahren ungeachtet weiterer Vor- oder Nachteile immer deutlich schlechtere Ergebnisse zu erwarten sind als von Verfahren, die mit einer angemessenen Auflösung arbeiten. Die als möglicher Ansatz für eine Weiterentwicklung des *NMR-Verfahrens* angekündigte Modellierung verschiebbarer Frequenzgruppen und die bereits im *Verfahren nach Kapust* vorgenommene Tonalitätsbestimmung zur Berücksichtigung der unterschiedlichen Verdeckungswirkung von Tönen und Schmalbandrauschen kann diese Nachteile nur für bestimmte Elementarsignale beseitigen, wird aber aller Wahrscheinlichkeit nach bei der Beurteilung komplexerer Signale scheitern (vgl. Punkt *d*) im vorhergehenden Abschnitt).

Bei Verfahren, die wie das *PERCEVAL-* und das *POM-Verfahren* eine Wahrscheinlichkeit für die Wahrnehmung einer Störung berechnen, liegt das Problem in der fehlenden Möglichkeit, zwischen verschiedenen eindeutig hörbaren Störungen zu differenzieren. Da besonders bei der

Beurteilung von Mehrkanalcodern und bei Übertragungen mit sehr niedrigen Bitraten davon ausgegangen werden muß, daß immer ein Rest an eindeutig hörbaren Störungen verbleiben kann, sollte für diese Anwendung auf eine Abschätzung hörbarer Störungen nicht verzichtet werden. Wegen des ansonsten sehr präzisen und vollständigen Gehörmodells können vom *POM-Verfahren* dennoch gute Ergebnisse erwartet werden. Ein Grund, warum das Verfahren bei einer ersten Anwendung noch keine befriedigenden Resultate geliefert hat [6], liegt wahrscheinlich in der Übernahme des im *PAQM-Verfahren* verwendeten Potenzgesetzes zur Erregungsaddition, das aus einer Reihe von Gründen hier nicht anwendbar ist.

## 8. Ausblick

Bedarf an weiterführenden Untersuchungen besteht insbesondere bei der Modellierung von binauralen Effekten und bei der Beschreibung der Addition von Mithörschwellen. Wünschenswert ist weiterhin eine Kombination der bei vielen der neueren Verfahren erreichbaren genauen Schwellenbestimmung mit der bei den älteren Verfahren *ASD* und *Speech Signal Degradation* bestehenden Möglichkeit, die subjektive Störlautheit auch bei deutlich hörbaren Störungen abzuschätzen. Für die Verbesserung bestehender Verfahren oder die Entwicklung neuer Verfahren sollte auch untersucht werden, warum das *PAQM-Verfahren* mit Parametern, die in Folge der experimentellen Optimierung z. T. erheblich von den aus der psychoakustischen Literatur bekannten Werten abweichen, die bisher vergleichsweise genauesten Vorhersagen subjektiver Beurteilungen von Musik-Codecs liefern konnte.

## Quellen

- [1] Beerends, J. G.; Stererdink, J. A.:  
A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation.  
J. AES, Vol. 40(1992 ), No. 12, December, S. 963-978.
- [2] Beerends, J. G.; Stererdink, J. A.:  
Modelling a Cognitive Aspect in the Measurement of the Quality of Music Codecs.  
Beitrag zur 96th AES Convention, Amsterdam, February 1994, Preprint 3800.
- [3] Brandenburg, K. H.:  
Ein Beitrag zu den Verfahren und der Qualitätsbeurteilung für hochwertige Musikcodierung.  
Dissertation am Lehrstuhl für Technische Elektronik der Universität Erlangen-Nürnberg, 1989.
- [4] Brandenburg, K. H.; Sporer, Th.:  
NMR and Masking Flag: Evaluation of Quality Using Perceptual Criteria.  
Proceedings of the AES 11th International Conference, Portland, Oregon, USA, 1992, S. 169-179.
- [5] Colomes, C.; Lever, M.; Rault, J. B.; Dehery, Y. F.:  
A Perceptual Model applied to Audio Bit-Rate Reduction.  
Beitrag zur 95th AES Convention, New York, October 1993, Preprint 3742.

- [6] Colomes, C.; Lever, M.; Dehery, Y. F.:  
A Perceptual Objective Measurement System (POM) for the Quality Assessment of Perceptual Codecs.  
Beitrag zur 96th AES Convention, Amsterdam, February 1994, Preprint 3801.
- [7] Humes, L. E.; Jesteadt, W.:  
Models of the Additivity of Masking.  
J. Acoust. Soc. Am., Vol. 85, No. 3, March 1989, S. 1285-1294.
- [8] ITU Radiocommunication Study Groups, Task Group 10/2:  
Comparison of three Objective Perceptual Models.  
Doc. 10-2/TEMP/4, 5 October 1993.
- [9] ITU Radiocommunication Study Groups, Task Group 10/2:  
Further Evaluation of PERCEVAL as a Predictor of Human Mean Opinion Scores.  
Doc. 10-2/44, 19 October 1993.
- [10] ITU Radiocommunication Study Groups, Task Group 10/2:  
Overview of the Results on the Correlation of PERCEVAL, NMR and PAQM with the Subjective Results of the ITU-R Task Group 10/2 Codec Test.  
Doc. 10-2/TEMP/20(Rev.1), November 1993.
- [11] Kapust, R.:  
Qualitätsbeurteilung codierter Audiosignale mittels einer BARK-Transformation.  
Dissertation an der Technischen Fakultät der Universität Erlangen-Nürnberg, 1993.
- [12] Karjalainen, M.:  
A New Auditory Model for the Evaluation of Sound Quality of Audio Systems.  
Proceedings of the IEEE International Conference of Acoustics, Speech and Signal Processing, ICASSP '85, March 26-29, Tampa, Florida, USA, 1985, S. 608-611.
- [13] Paillard, B.; Mabillean, P.; Morissette, S.; Soumagne, J.:  
PERCEVAL: Perceptual Evaluation of the Quality of Audio Signals.  
J. AES., Vol. 40(1992), No. 1/2, January/February, S. 21-31.
- [14] Schroeder, M. R.; Atal, B. S.; Hall, J. L.:  
Optimizing digital speech coders by exploiting masking properties of the human ear.  
J. Acoust. Soc. Am., Vol. 66(1979), No. 6, December, S. 1647-1652.
- [15] Stuart, J. R.:  
Noise: Methods for Estimating Detectability and Threshold.  
Beitrag zur 94th AES Convention, Berlin, March 1993, Preprint 3477.
- [16] Terhardt, E.:  
Calculating Virtual Pitch.  
Hearing Research, Vol. 1(1979), S. 155-182.
- [17] Terhardt, E.:  
The SPINC Function for Scaling of Frequency in Auditory Models.  
Acustica, Vol. 77(1992), S. 40-42.
- [18] Wier, C. C.; Jesteadt, W.; Mc Green, D.:  
Frequency Discrimination as a Function of Frequency and Sensation Level.  
J. Acoust. Soc. Am., Vol. 61(1977), No. 1, January, S. 178-184.
- [19] Zwicker, E.; Feldkeller, R.:  
Das Ohr als Nachrichtenempfänger.  
Stuttgart: Hirzel Verlag, 1967.
- [20] Zwicker, E.; Herla, S.:  
Über die Addition von Verdeckungseffekten.  
Acustica, Vol. 34(1975), S. 89-97.

# **Gehörrichtige Qualitätsbewertung von Audiosignalen - Übersicht und Einschätzung der gegenwärtigen Verfahren**

<b>1. Einleitung</b> .....	<b>1</b>
<b>2. Arbeitsweise von gehörrichtigen Meßverfahren</b> .....	<b>2</b>
<b>3. Gegenüberstellung typischer Vorgehensweisen</b> .....	<b>5</b>
3.1 Auflösung im Frequenzbereich.....	5
3.2 Pegelabhängigkeit der Mithörschwellen.....	7
3.3 Addition von Mithörschwellen / Abhängigkeit des Verdeckungsmaßes von der Signalform .....	8
3.4 Zeitliche Verdeckungseffekte .....	9
3.5 Berechnete Ausgangsgrößen.....	9
<b>4. Übersicht über die bekannten Meßverfahren</b> .....	<b>11</b>
4.1 Verfahren nach Schroeder, Atal, Hall ("Speech Signal Degradation"):.....	11
4.2 ASD ("Auditory Spectrum Distance"):.....	11
4.3 NMR ("Noise to Mask Ratio"): .....	12
4.4 PAQM ("Perceptual Audio Quality Measure"): .....	13
4.5 PERCEVAL ("PERCEptual EVAuation"):.....	13
4.6 POM ("Perceptual Objective Model"):.....	14
4.7 Verfahren nach Kapust: .....	14
4.8 Verfahren nach Stuart: .....	14
<b>5. Möglichkeiten und Grenzen der Verfahren</b> .....	<b>15</b>
<b>6. Forderungen an ein gehörrichtiges Beurteilungsverfahren</b> .....	<b>15</b>
<b>7. Einschätzung der heute bekannten Meßverfahren</b> .....	<b>17</b>
<b>8. Ausblick</b> .....	<b>19</b>
<b>Quellen</b> .....	<b>19</b>

	Transformation		Anregungsfunktion		Simultanverdeckung			zeitliche Verdeckung	
	Blocklänge	Blockabstand	Stützstellen	Auflösung	spreading function	Zusammenwirken von Erregungen	Unterschiedliche Verdeckungswirkung von Tönen und Rauschen	Vorverdeckung	Nachverdeckung
NMR	1024 Samples	512 Samples	27	0,9 Bark	worst case	Maximalwert	keine	keine	exponentiell
PAQM	2048 Samples	1024 Samples	120	0,2 Bark	pegelabhängig	Potenzgesetz	indirekt durch nichtlineare Erregungsaddition	keine	exponentiell, Zeitkonstante frequenzabhängig
PERCEVAL	3 · 1024 Samples	1024 Samples	2500	1 mel	mittlere Pegel	lineare Addition	indirekt durch statistischen Ansatz	keine	keine
Verfahren nach Schroeder, Atal, Hall	1024 Samples	1024 Samples	Aufteilung auf Stützstellen wird nicht beschrieben		mittlere Pegel	lineare Addition	nicht beschrieben (ergibt sich wahrscheinlich grob durch die Form der "spreading function")	keine	keine
ASD	256(*) Samples	1 Sample	48	0,5 Bark	mittlere Pegel	lineare Addition	indirekt durch die Impulsantworten der Teilfilter und Tiefpässe	keine	keine
Verfahren nach Kapust	512(**) Samples	1 Sample	24	1 Bark	pegelabhängig	Maximalwert	direkt durch Tonalitätsbestimmung	keine	Arcustangensfunktion
POM	2048 Samples	1024 Samples	620	0,04 Bark	pegelabhängig	Potenzgesetz	indirekt durch nichtlineare Erregungsaddition und statistischen Ansatz	keine	keine
Verfahren nach Stuart	nicht festgelegt		380	0,1 ERB	pegelabhängig	indirekt durch Nachbildung der Funktionsweise des Gehörs		keine	keine

\* : FIR-Filterbank

\*\* : variable Fensterlänge auch über mehrere Blöcke ("Polyphasenfilterung")

**Tabelle 1: Übersicht über die Parameter der in den Beurteilungsverfahren verwendeten Gehörmodelle**