

Bit-Reduzierung:

Verdeckungseffekte

Mithörschwellen-angepaßte Teilbandcodierung hochwertiger Tonsignale

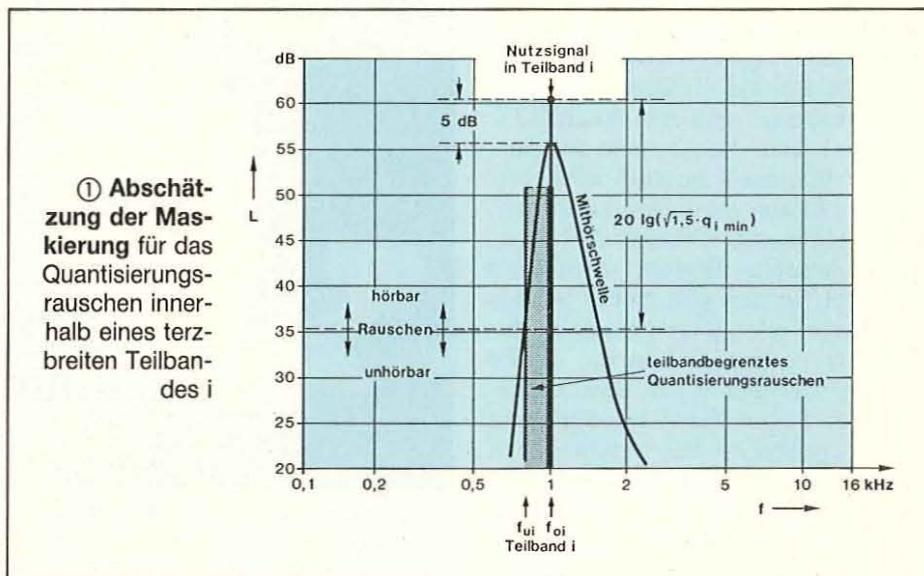
Die digitale Signalverarbeitung hat in den letzten Jahren einen Qualitätsstandard für die Aufzeichnung, Bearbeitung und Übertragung von hochwertigen Tonsignalen geschaffen. Ein wesentlicher Vorteil gegenüber den analogen Verfahren ist die bei ausreichendem Fehlerschutz hervorragende Reproduzierbarkeit des gespeicherten Signals.

klärt die Tatsache, warum mit breitbandig arbeitenden digitalen Kompondierungstechniken nur eine verhältnismäßig geringe Reduktion der Datenrate möglich ist, ohne daß die Codierungsfehler vom Gehör wahrgenommen werden können. Vor kurzem wurden drei Quellencodier-

Bei der digitalen Tonsignalübertragung ergeben sich aufgrund der relativ hohen Datenraten entsprechend hohe Kosten bzw. Probleme mit der begrenzten Kapazität der zur Verfügung stehenden Kanäle. Im Falle des digitalen Hörfunks über Rundfunksatelliten wurde die Nettodatenrate auf 448 kBit/s pro Monokanal reduziert. Dies konnte mit einer Abtastfrequenz von 32 kHz und einer Codierung der Abtastwerte mit 16/14-Bit-Gleitkomma-technik realisiert werden. Für einen digitalen terrestrischen Hörrundfunk wird jedoch noch eine weit niedrigere Datenrate gefordert.

Musik und Sprache enthalten neben der relevanten Information sowohl Redundanz als auch Irrelevanz. Der redundante Signalanteil ist für die vollständige Rekonstruktion des Signals im Empfangsgerät nicht notwendig, während der irrelevante Signalanteil zwar für die vollständige Rekonstruktion erforderlich ist, jedoch für den eigentlichen Empfänger, das menschliche Gehör, für eine subjektiv ungestörte Empfangsqualität nicht übertragen werden muß. Es ergeben sich Unterschiede zwischen dem Ein- und Ausgangssignal; diese werden jedoch nicht wahrgenommen.

Neuere Untersuchungen zeigen, daß insbesondere die Irrelevanzreduktion ein effizientes Mittel darstellt, ein bitsparendes Codierverfahren für eine hochqualitative Tonsignalübertragung zu entwickeln. Um die Irrelevanz eines Tonsignals zu reduzieren, müssen



bei der Codierung der Signale die Systemeigenschaften und insbesondere die Analyseigenschaften des menschlichen Gehörs berücksichtigt werden. Hierbei zeigt sich, daß die vom Nachrichtenempfänger „Gehör“ geforderte Genauigkeit der Signaldarstellung meist überraschend gering ist, dabei aber sehr vom spektralen und zeitlichen Verlauf abhängt. So erscheint es als sinnvoll, die Codierungsfehler digitaler Tonsignale dem spektralen und zeitlichen Auflösungsvermögen des Gehörs anzupassen.

Übliche PCM-Systeme erzeugen ein breitbandiges, näherungsweise weißes Quantisierungsrauschen, das dieser Forderung nicht entspricht. Dies er-

verfahren vorgestellt, die genau diese Möglichkeit nutzen, das Quantisierungsrauschen der spektralen Selektivität des Gehörs anzupassen. Die Datenraten, die mit diesen Codierverfahren erreicht werden, liegen in einem Bereich von 80...130 kBit/s für einen Monokanal. Die Verfahren beruhen auf einer spektralen Zerlegung und einer frequenzabhängigen Zuweisung der Quantisierung des Tonsignals. Zwei der Verfahren benützen zur Spektralanalyse eine diskrete Fourier-Transformation (DFT), das dritte Verfahren verwendet hierzu digitale Bandpaßfilter, deren Bandbreiten dem spektralen Auflösungsvermögen des Gehörs angepaßt werden. Das

Bandpaß-Verfahren, „MASCAM“⁽¹⁾ (Masking-pattern Adapted Subband Coding And Multiplexing) genannt, wird im folgenden ausführlich beschrieben.

Wie Mithörschwellen des Gehörs genutzt werden

Um beliebige Tonsignale für eine Speicherung oder Übertragung mit geringer Kanalkapazität im Sinne einer Irrelevanzminderung zu codieren, müssen die Analyseigenschaften des Gehörs berücksichtigt werden. Aufgrund des begrenzten Auflösungsvermögens werden vom Gehör nur ganz

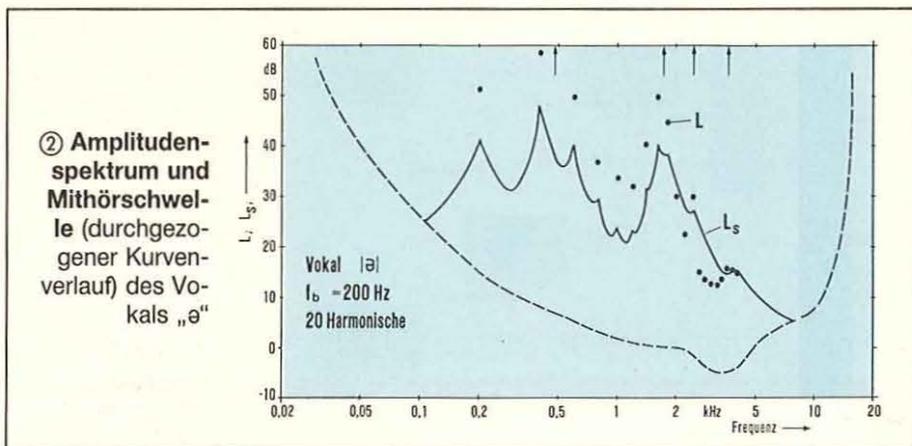
stimmte wichtige spektrale Komponenten der Tonsignale wahrgenommen. Man spricht hier von der spektralen Selektivität des Gehörs. Sie ist in erster Linie gekennzeichnet durch die Mithörschwellen, d. h. durch die Tatsache, daß laute Töne benachbarte leise Töne verdecken. Der Effekt ist zweifach nutzbar:

Schall unterhalb der Mithörschwelle eines anderen, maskierenden Schalles wird nicht wahrgenommen und ist für das Gehör irrelevant. Werden diese Maskierungsverhältnisse bei der Codierung eines Tonsignals berücksichtigt, so folgt daraus, daß der unterhalb der Mithörschwelle liegende Signalanteil, d. h. die verdeckten Komponenten des Tonsignals, nicht codiert und folglich keine Informationskapazität für die Übertragung dieses Anteils verwendet werden muß.

Störanteile, die bei einer Codierung des Tonsignales zum Einsparen von Informationsfluß von diesem überlagert werden, können vom Gehör nicht wahrgenommen werden, solange sie die Mithörschwelle des Tonsignales nicht überschreiten. Dies bedeutet, daß bei einer spektralen Zerlegung des breitbandigen Tonsignals in einzelne Teilbänder die Codierung gerade so gewählt werden muß, daß die innerhalb eines Teilbandes entstehenden Verzerrungen unter die Mithörschwelle des Nutzsignales fallen.

Simultanverdeckung

Ein Beispiel für die notwendige Quantisierung eines abgetasteten Tones in-



② Amplitudenspektrum und Mithörschwelle (durchgezogener Kurvenverlauf) des Vokals „ə“

nerhalb eines terzbreiten Teilbandes ist in *Bild 1* dargestellt. Die Differenz aus Nutzsignalpegel und Schnittpunkt des bandpaßbegrenzten Quantisierungsrauschens mit der unteren Flanke der Mithörschwelle kennzeichnet den notwendigen Rauschabstand, damit das Rauschen nicht wahrgenommen wird, und legt somit die minimale Quantisierung des Tonsignals innerhalb des Teilbandes *i* fest. Der kritische Fall bezüglich der Wahrnehmung des Quantisierungsrauschens ist durch eine Lage des Nutztones an der oberen Durchlaßgrenze des Filters gegeben, da die untere Flanke der Mithörschwelle steiler ist als die obere Flanke. Die notwendige Auflösung des Tonsignals innerhalb eines Teilbandes *i* hängt allein von der Bandbreite des verwendeten Filters ab. Die für eine Verdeckung des Quantisierungsrauschens erforderliche Quantisierung eines Abtastwertes beträgt:

$$\lg q_{i \min} = 32 \lg(f_{oi}/f_{ui}) + 1 \quad (1)$$

(Bit/Abtastwert)

wobei f_{oi} , f_{ui} die obere bzw. untere Grenzfrequenz eines Teilbandes *i* darstellen. Um eine weitergehende Reduktion der Irrelevanz des Tonsignals zu erreichen, wird anstelle einer festen Quantisierung innerhalb eines Teilbandes eine dynamisch gesteuerte Quantisierung vorgesehen, die den Effekt der „gegenseitigen Verdeckung“ benachbarter Teilbandsignale berücksichtigt. Damit kann die zur Verfügung stehende Informationskapazität optimal auf die einzelnen Teilbänder aufgeteilt werden.

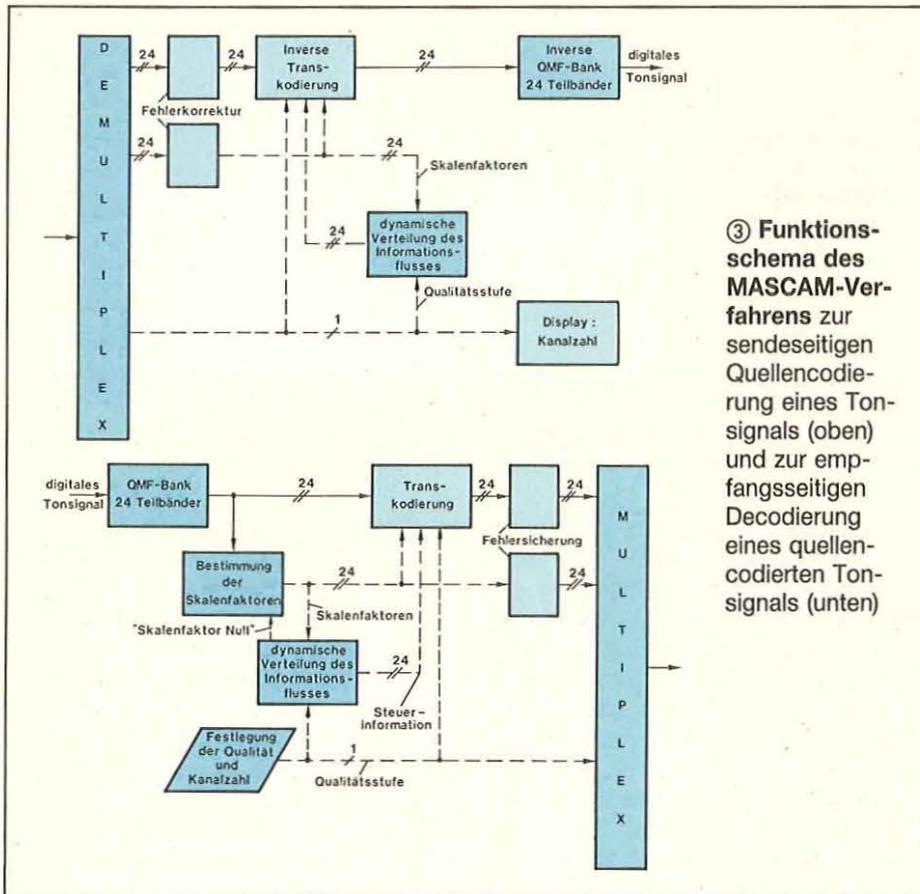
In *Bild 2* ist für einen stationären Vokal „ə“, bestehend aus 20 Harmo-

nischen mit einer Grundfrequenz von 200 Hz, die Mithörschwelle, die sich aufgrund der gegenseitigen Verdeckung aller Teiltöne ergibt, durch eine ausgezogene Kurve gekennzeichnet. In diesem Diagramm sind die Teiltonpegel der Harmonischen durch Punkte und die Formantfrequenzen des Vokals durch Pfeile dargestellt. Die gegenseitige Verdeckung führt zu zwei Konsequenzen, die bei der Quellencodierung genutzt werden können. Zum einen müssen sämtliche Teiltöne des Vokals, die unterhalb der gemeinsamen Mithörschwelle liegen, nicht übertragen werden. Dies trifft in dem vorliegenden Fall auf die 11. und 13. bis 17. Harmonischen zu, die durch den 2. bzw. 3. Formanten des Vokals eine Maskierung erfahren. Deshalb können Teilbänder, die verdeckte, d. h. unhörbare Signalanteile enthalten, zu Null gesetzt werden und beanspruchen keine Informationskapazität. Zum anderen kann das durch eine grobe Auflösung vorhandene Quantisierungsrauschen bezüglich der spektralen Verteilung so gesteuert werden, daß es sich immer unterhalb der in *Bild 2* dargestellten Mithörschwelle befindet. Somit erhält man ein dem Mithörschwellenverlauf des Tonsignals optimal angepaßtes Quantisierungsrauschen.

Zeitliche Verdeckung

Neben der Simultanverdeckung kann zusätzlich die sogenannte Vor- und Nachverdeckung ausgenutzt werden. Die Vorverdeckung findet in einem Zeitabschnitt statt, der vor dem Auftreten des Maskierers liegt. Die Dauer

¹⁾ Ursprüngliche Bezeichnung „MASC“



③ Funktions-schemata des MASCAM-Verfahrens zur sendeseitigen Quellencodierung eines Tonsignals (oben) und zur empfangsseitigen Decodierung eines quellen-codierten Tonsignals (unten)

bank, der sogenannten Quadrature-Mirror-Filterbank (QMF). Diese Filter ermöglichen bei äußerst geringer Welligkeit im Durchlaßbereich die vollständige Nutzung der Abtastfrequenz: Die für jedes Teilfilter erforderliche Abtastfrequenz ist genau doppelt so groß wie die Bandbreite des Filters. Die durch die endliche Flankensteilheit entstehenden Aliasing-Verzerrungen werden bei der inversen Filterung unter der Bedingung einer gleichförmigen Quantisierung in allen Teilbändern vollständig eliminiert. Um die spektralen Verdeckungseigenschaften des Gehörs möglichst optimal auszunutzen, wurde eine Aufteilung des Tonsignals in Teilbänder entsprechend vorgenommen. Diese Aufteilung stellt eine Annäherung der Bandbreiten der Teilbänder an die Frequenzgruppen des Gehörs dar.

Das Verfahren arbeitet mit einer Abtastfrequenz von 32 kHz und einer Audio-Bandbreite von 16 kHz. Eine größere Bandbreite ist nicht sinnvoll, da eine Tiefpaßbegrenzung des Tonsignals bei 15 kHz aufgrund der Mithörschwellen nicht wahrnehmbar ist. Eine Erhöhung der Abtastfrequenz auf z. B. 48 kHz würde zwar nur eine unerhebliche Zunahme der Datenrate bedeuten, jedoch eine unnötige Erhöhung des Hardware-Aufwandes zur Folge haben.

Die für die Reduktion der Datenrate notwendigen, wichtigsten Operationen sind in Bild 3 dargestellt. Nach der Aufteilung des Tonsignals in die 24 Teilbänder wird für jedes Teilband ein sogenannter Skalenfaktor gebildet. Ein Skalenfaktor kennzeichnet die Pegelklasse, in der die höchste Teilbandsignalamplitude innerhalb des Zeitraumes auftritt, für den dieser Skalenfaktor gilt (Blocklänge). Der Skalenfaktor beschreibt den zu quantisierenden Amplitudenbereich und ermöglicht daher einen näherungsweise signalunabhängigen Signal/Rauschabstand. Die Blocklängen in den einzelnen Teilbändern liegen frequenzabhängig im Bereich 1...20 ms und entsprechen der Dauer der Vorverdeckung. Die Skalenfaktoren müssen nicht für alle Abtastwerte, sondern nur blockweise übertragen werden. Sie enthalten alle Informationen, die für die Berechnung der Mithörschwelle erforderlich sind. Nach Maßgabe

der Vorverdeckung liegt zwischen 1 und 20 ms. Sie hängt vom Spektrum und Pegel des Maskierers ab. Die Nachverdeckung findet in einem Zeitabschnitt nach dem Abklingen des Maskierers statt; ihre Dauer liegt zwischen 100 und 200 ms, ebenfalls abhängig von den Eigenschaften des Maskierers.

Die Vorverdeckung ermöglicht eine Datenreduktion, weil schnelle Pegelanstiege nicht vollständig rekonstruiert werden müssen und deshalb die Abtastwerte blockweise zusammengefaßt werden können. Die Blocklängen müssen so gewählt werden, daß bei schnellen Pegelanstiegen die Dauer der Vorverdeckung nicht überschritten wird.

Die Nachverdeckung ermöglicht eine zusätzliche Datenreduktion, weil die Auflösung eines schwachen Signalanteiles, der kurz nach dem stärkeren, maskierenden Signalanteil folgt, wesentlich geringer sein muß als ohne vorausgehenden, maskierenden Signalanteil. In vielen Fällen wird der nachfolgende Signalanteil sogar vollständig

verdeckt, so daß während der Dauer der Nachverdeckung dessen Übertragung nicht notwendig ist.

Funktionsschema des MASCAM-Verfahrens

Die Funktionsweise des MASCAM-Verfahrens ist für die sendeseitige Codierung und empfangsseitige Decodierung in Bild 3 dargestellt. Der erste Schritt der Codierung besteht aus der Filterung des Tonsignals. Die Änderung der Quantisierung, die eigentliche Datenreduktion, findet im Sender (Bild 3, oben) in der 2. Stufe statt, der sogenannten Transcodierungsstufe. Danach erfolgt dann in einer weiteren Stufe die Übertragung der weitgehend von Irrelevanz befreiten Teilbandsignale. Im Empfänger (Bild 3, unten) wird die inverse Transcodierung und Filterung zu dem wiederum 16-Bit-linear-codierten Tonsignal vorgenommen.

Die Aufspaltung des Tonsignals in die 24 Teilbänder erfolgt mittels einer speziellen digital realisierten Filter-

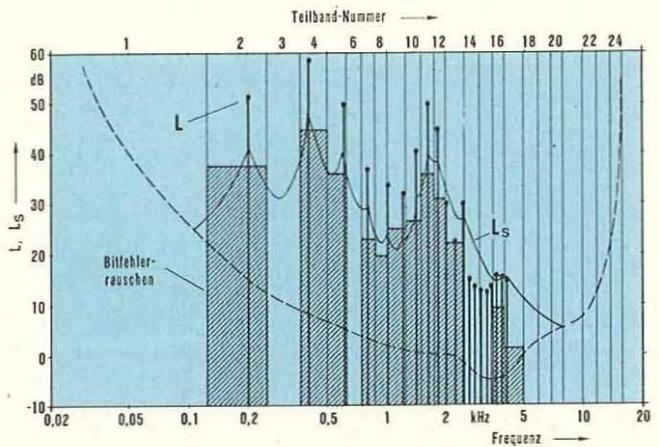
der Skalenfaktoren erfolgt die Ermittlung der Steuerinformation für die Transcodierung („dynamische Verteilung“ des Gesamtinformationsflusses auf die Teilbänder). Die dynamische Verteilung bewirkt eine über Gleichung 1 hinausgehende Reduktion der Datenrate. Die maximale Auflösung pro Teilband, die wie Gleichung 1 zeigt, nur von der Bandbreite des jeweiligen Teilbandes abhängig ist, kann je nach spektraler und zeitlicher Verteilung des Tonsignals erheblich unterschritten werden. Abschätzungen haben ergeben, daß mittels dieser dynamischen Verteilung des Informationsflusses die resultierende maximale Nettodatenrate nur noch etwa 60 kBit/s beträgt, zuzüglich einer Datenrate von ca. 20 kBit/s für die Übertragung der Skalenfaktoren. Die gesamte maximal erforderliche Datenrate beträgt somit etwa 80 kBit/s.

Übertragung und Decodierung

Vor Bildung des Multiplexsignals muß eine Fehlersicherung für jedes der 24 Teilbandsignale durchgeführt werden. Es ist nur für 24 Skalenfaktoren, welche die gesamte Nebeninformation für den Empfänger enthalten, eine besonders wirkungsvolle Fehlersicherung notwendig (s. u.).

Empfangsseitig wird durch inverse Transcodierung und Filterung der 24 Teilbandsignale ein 16-Bit-linear-codiertes Tonsignal zurückgewonnen (Bild 3). Nach der Demultiplex-Bildung werden für die noch in codierter Form vorliegenden Teilbandsignale und insbesondere für die Skalenfaktoren Fehlerkorrekturen durchgeführt. Die Skalenfaktoren stellen wie auf der Sendeseite die maßgeblichen Größen zur Berechnung der Mithörschwellen dar. Die Ermittlung der Steuerinformation für die inverse Transcodierung erfolgt identisch zur Sendeseite. So wird erreicht, daß keine zusätzlichen Steuerinformationen übertragen werden müssen. Abhängig von der im Empfänger berechneten Steuerinformation werden die reduzierten Abtastwerte der einzelnen Teilbänder wieder in die 16-Bit-Ebene umcodiert und anschließend mittels inverser Filterung zu einem Tonsignal voller Bandbreite zusammengesetzt.

④ **Amplitudenspektrum L des Vokals „ə“**, Mithörschwelle L_s und Quantisierungsrauschen, das bei der Quantisierung der Abtastwerte des Vokals bei einer Bitrate von



16 kBit/s für die einzelnen Abtastwerte und 20 kBit/s für die Skalenfaktoren auftritt

Qualitative Eigenschaften

Computer-Simulationen haben gezeigt, daß die QM-Filterung keinerlei Qualitätsbeeinträchtigungen verursacht, wenn 64-Koeffizienten-Filter benutzt werden. Sie haben ferner gezeigt, daß die Datenreduktion in der Transcodierungsstufe keine subjektiv wahrnehmbaren Qualitätsbeeinträchtigungen verursacht, wenn die bisher bekannten Daten der Vor-, Simultan- und Nachverdeckung wie beschrieben zugrunde gelegt werden.

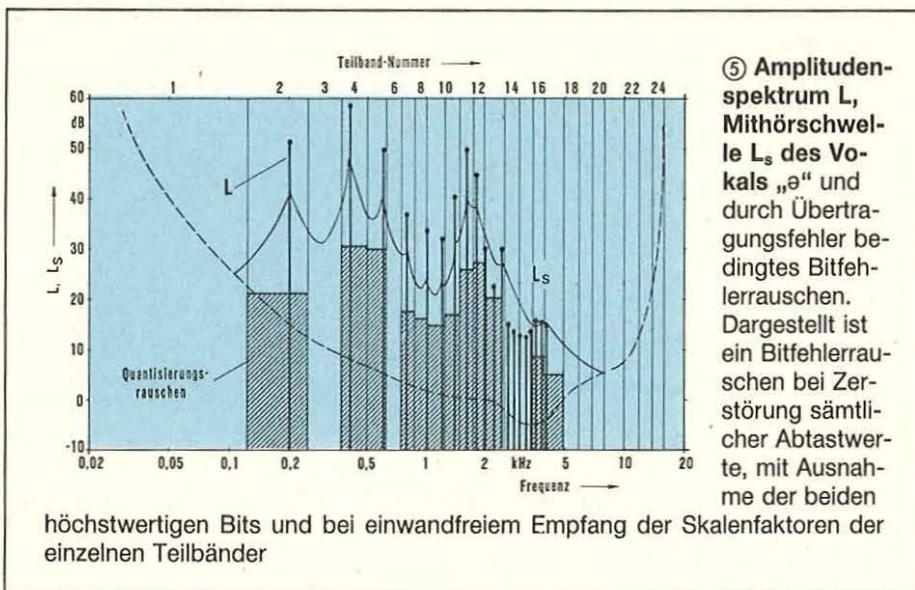
Ein wichtiger Schritt der weiteren Untersuchungen ist die Bestimmung des sogenannten „Worst-case“-Signals, welches das Maximum an relevanter Information enthält. Es läßt sich abschätzen, daß dieses (synthetische) Tonsignal mit etwa 80 kBit/s ohne merkbliche Qualitätseinbuße übertragbar ist. Alle anderen, insbesondere natürliche Mikrofonsignale, enthalten weniger relevante Information und erfordern prinzipiell eine geringere Datenrate. Dies ist in *Bild 4* am Beispiel des Vokales „ə“ (vgl. *Bild 2*) dargestellt. Das als schraffierte Fläche gekennzeichnete Quantisierungsrauschen in den Teilbändern bleibt im gesamten Spektrum unter der resultierenden Mithörschwelle (durchgezogene Kurve), wenn nur 16 kBit/s für die Übertragung der Abtastwerte (sowie 20 kBit/s für die Übertragung der Skalenfaktoren) zur Verfügung stehen würden. Diese geringe Datenrate ergibt sich, weil hier insgesamt 12 Teilbänder zu Null gesetzt werden können.

Die relevante Information des Tonsignals und damit die erforderliche Datenrate schwanken stark signalabhängig. Falls nicht gerade das Worst-case-Signal übertragen wird, ist der konstante Informationsfluß von 80 kBit/s größer als notwendig. Normalerweise steht also eine „Informationsfluß-Reserve“ zur Verfügung, die nach bestimmten Gesichtspunkten auf die Teilbänder verteilt wird. Vorrangig erfolgt die dynamische Verteilung so, daß die Abstände der Quantisierungsrauschpegel zur resultierenden Mithörschwelle in allen Teilbändern etwa gleich sind. Hierdurch ergibt sich normalerweise (bei nahezu allen natürlichen Tonsignalen) eine sogenannte „Noise-to-Mask“-Reserve, deren Größe signalabhängig schwankt. Die Schwankungen lassen sich jedoch durch den Einsatz von Pufferspeichern verringern.

Die dynamische Verteilung des Informationsflusses auf die relevanten Teilbänder und die resultierende „Noise-to-Mask“-Reserve führen zu praktischen Vorteilen:

○ **Pegelton-Meßfestigkeit:**

Sinustöne oder Schmalbandrauschen werden mit der vollen Auflösung (16 Bit) übertragen. Auch bei gleichzeitiger Übertragung mehrerer Sinustöne ergibt sich durch die Konzentration des Datenstromes auf die relevanten Teilbänder noch die höchstmögliche Quantisierung. Der analoge Teil eines Empfängers kann deshalb in althergebrachter Weise mit Testsendungen überprüft werden.



MASCAM-Verfahren wie ein 24-Kanal-Vocoder, wobei ein breitbandiges Rauschen in 24 Teilbänder zerlegt und in jedem Teilband mit dem Skalenfaktor amplitudenmoduliert wird. Damit ist auch bei einer sehr stark gestörten Übertragungsstrecke eine ausreichende Sprachverständlichkeit möglich.

Um im Falle sehr starker Störungen die Tonsignalqualität zu verbessern, und die wichtigsten tonalen Eigenschaften des Signals zu übertragen, werden die höherwertigen Bits der Abtastwerte in den Teilbändern im Frequenzbereich bis ca. 2 kHz vorrangig geschützt. In Bild 5 ist für den Vokal „a“ das Störspektrum durch Bitfehler dargestellt, wenn beispielsweise nicht nur die Skalenfaktoren, sondern auch jeweils die zwei höchstwertigen Bits gegenüber Übertragungsfehlern geschützt werden. Mit diesem geringen Aufwand an Fehlersicherung ist es möglich, das Störrauschen spektral so zu gewichten, daß es größtenteils maskiert wird.

MASCAM-codierte Tonsignale weisen eine größere Unempfindlichkeit gegenüber Bitfehlern auf als die bisher bekannten PCM-Signale.

Günther Theile, Gerhard Stoll,
Martin Link
(Institut für Rundfunktechnik,
München)

○ Kaskadierung / Tonsignalnachbearbeitung:

Sowohl bei mehrfacher Codierung/Decodierung des Tonsignals (beispielsweise einmal für die Übertragung und ein zweites Mal für die Speicherung) als auch bei empfangsseitiger Frequenzgangänderung könnte das Quantisierungsrauschen die Mithörschwelle überschreiten und wahrnehmbare Qualitätsbeeinträchtigungen verursachen, wenn keine ausreichende Noise-to-Mask-Reserve vorhanden wäre.

Die durchschnittliche Noise-to-Mask-Reserve hängt natürlich primär davon ab, welche Datenrate für die Übertragung zur Verfügung steht. Bei 80 kBit/s ist die Reserve für vielleicht 90 % aller Tonsignale völlig ausreichend; nur das Worst-case-Signal enthält keine Reserve. Stehen dagegen vielleicht 160 kBit/s für jeden Monokanal zur Verfügung, so wären diese MASCAM-codierten Signale voll „studiotauglich“, d. h. alle Tonsignalbearbeitungen wären problemlos möglich.

Aber auch reduzierte Datenraten sind sinnvoll: Bei 40 kBit/s beispielsweise treten nur bei wenigen kritischen Signalen wenig störende Qualitätsbeeinträchtigungen auf. Sprache würde mit 40 kBit/s ohne jede Qualitätsbeeinträchtigung übertragen werden, mit 26 kBit/s wäre die Qualität immer noch viel besser als die heutige Telefonqualität.

Unempfindlichkeit gegenüber Bitfehlern

MASCAM-codierte Tonsignale haben den wesentlichen Vorteil, daß Störungen des Tonsignals infolge von Bitfehlern sich nicht breitbandig, sondern nur innerhalb der Frequenzlage und -breite desjenigen Teilbandes auswirken, das die gestörten Abtastwerte enthält. Somit werden die typischen Knackstörungen vermieden, die je nach Wertigkeit des zerstörten Bits mehr oder weniger laut empfunden werden. Ein zweiter entscheidender Vorteil ergibt sich aus der Tatsache, daß die Skalenfaktoren der 24 Teilbänder getrennt von den eigentlichen Tondaten im Multiplexsignal übertragen werden. Unter der Voraussetzung, daß die Skalenfaktoren mit einem hohen Fehlerschutz versehen werden, beschränken sich die durch Bitfehler der Abtastwerte erzeugten Störungen auf den durch den Skalenfaktor gekennzeichneten Signalpegel. Diese Vorteile bewirken, daß Bitfehler-Störungen durch das Tonsignal selbst stark verdeckt werden.

Sogar bei maximaler Bitfehler-Rate, d. h. bei Zerstörung sämtlicher Abtastwerte, wird (wenn der Skalenfaktor im Empfänger richtig vorliegt) das Bitfehler-Störrauschen durch die 24 Skalenfaktoren frequenzabhängig so gewichtet, daß die spektrale Hüllkurve derjenigen des Tonsignals entspricht. In diesem Fall arbeitet das

ZUSAMMENGEFASST

Die Erfahrungen, die mit verschiedenen Quellencodierverfahren, die im wesentlichen eine Irrelevanzreduktion bewirken, für hochwertige Tonsignale gewonnen wurden, zeigen, daß die Übertragung von 16-Bit-linear-codierten Abtastwerten mit etwa 700 kBit/s eine hohe Verschwendung an Übertragungskapazität darstellt, wenn es darum geht, eine Tonübertragung ohne subjektive Qualitätseinbuße durchzuführen. Die Fähigkeiten des spektralen und zeitlichen Auflösungsvermögens des Gehörs sind begrenzt und auch weitgehend bekannt. Dieser Vorteil kann im Sinne einer Datenreduktion für künftige Übertragungssysteme genutzt werden. Der entscheidende Schritt hierbei ist die Anpassung der Verarbeitungsprozesse bei der Tonsignalcodierung an die Analyseigenschaften des Gehörs.