


 INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation⁴ : H04B 1/66	A1	(11) Internationale Veröffentlichungsnummer: WO 88/ 04117 (43) Internationales Veröffentlichungsdatum: 2. Juni 1988 (02.06.88)
(21) Internationales Aktenzeichen: PCT/EP87/00723 (22) Internationales Anmeldedatum: 20. November 1987 (20.11.87) (31) Prioritätsaktenzeichen: P 36 39 753.9 (32) Prioritätsdatum: 21. November 1986 (21.11.86) (33) Prioritätsland: DE (71) Anmelder (für alle Bestimmungsstaaten ausser US): BAYERISCHE RUNDFUNKWERBUNG [DE/DE]; Arnulfstraße 42, D-8000 München 2 (DE). (72) Erfinder;und (75) Erfinder/Anmelder (nur für US) : THEILE, Günther [DE/DE]; Hauptstraße 26 A, D-8195 Thanning (DE). STOLL, Gerhard [DE/DE]; Ahornweg 21, D-8051 Zolling (DE). LINK, Martin [DE/DE]; Mistralstraße 22, D-8044 Unterschleißheim (DE).	(74) Anwalt: KONLE, Tilmar; Benderstraße 23a, D-8000 München 60 (DE). (81) Bestimmungsstaaten: AT (europäisches Patent), BE (europäisches Patent), CH (europäisches Patent), DE (europäisches Patent), DK, FI, FR (europäisches Patent), GB (europäisches Patent), IT (europäisches Patent), JP, KR, LU (europäisches Patent), NL (europäisches Patent), SE (europäisches Patent), US. Veröffentlicht <i>Mit internationalem Recherchenbericht. Vor Ablauf der für Änderungen der Ansprüche zugelassenen Frist. Veröffentlichung wird wiederholt falls Änderungen eintreffen.</i>	
(54) Title: PROCESS FOR TRANSMITTING DIGITAL AUDIO-SIGNALS		
(54) Bezeichnung: VERFAHREN ZUM ÜBERTRAGEN DIGITALISIERTER TONSIGNALE		
(57) Abstract <p>For purposes of transmission audio-signals are numerically represented by a plurality of spectral partial band signals by means of quadrature mirror-filtering. The quantification of the scanning values in partial bands, for example 24 partial bands, is regulated in such a manner that the quantifying noise levels of the individual partial band signals have approximately the same intervals as the hearing threshold of the human ear, which results from individual partial band signals. The intervals between the quantifying noise levels of the partial band signals with respect to the resulting hearing threshold are determined by the difference between the total information flow required for the coding and the total information flow available for coding. The available total information flow is determined and can vary depending on the signals.</p> (57) Zusammenfassung <p>Zum Übertragen von Tonsignalen wird mittels Quadrature-Mirror-Filterung das Tonsignal durch eine Vielzahl spektraler Teilbandsignale digital dargestellt. Die Quantisierung der Abtastwerte in den Teilbändern, z.B. 24 Teilbänder, wird dahingehend gesteuert, daß die Quantisierungsrauschpegel der einzelnen Teilbandsignale näherungsweise gleiche Abstände zu der aus einzelnen Teilbandsignalen resultierenden Mithörschwelle des menschlichen Gehörs aufweisen. Dabei werden die Abstände der Quantisierungsrauschpegel der Teilbandsignale bezüglich der resultierenden Mithörschwelle durch die Differenz zwischen dem für die Codierung erforderlichen Gesamtinformationsfluß und dem für die Codierung verfügbaren Gesamtinformationsfluß festgelegt. Der verfügbare Gesamtinformationsfluß wird festgelegt und kann signalabhängig schwanken.</p>		

LEDIGLICH ZUR INFORMATION

Code, die zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AT	Österreich	FR	Frankreich	MR	Mauritanien
AU	Australien	GA	Gabun	MW	Malawi
BB	Barbados	GB	Vereinigtes Königreich	NL	Niederlande
BE	Belgien	HU	Ungarn	NO	Norwegen
BG	Bulgarien	IT	Italien	RO	Rumänien
BJ	Benin	JP	Japan	SD	Sudan
BR	Brasilien	KP	Demokratische Volksrepublik Korea	SE	Schweden
CF	Zentrale Afrikanische Republik	KR	Republik Korea	SN	Senegal
CG	Kongo	LI	Liechtenstein	SU	Soviet Union
CH	Schweiz	LK	Sri Lanka	TD	Tschad
CM	Kamerun	LU	Luxemburg	TG	Togo
DE	Deutschland, Bundesrepublik	MC	Monaco	US	Vereinigte Staaten von Amerika
DK	Dänemark	MG	Madagaskar		
FI	Finnland	ML	Mali		

1

5

BESCHREIBUNG

10

Die Erfindung bezieht sich auf ein Verfahren gemäß dem Oberbegriff des Anspruchs 1.

15

20

25

Bei der Übertragung digitaler Tonsignale ist es bekannt ("Journal of the Audio Engineering Society", Nov. 1979, Bd. 27, Nr. 11, S. 855 - 865; "The Bell System Technical Journal", Sept. 1981, S. 1633 - 1653; "IEEE International Conference on ASSP, 1982, Proceedings, S. 1684 - 1687), das breitbandige digitale Tonsignal mittels einer QMF- (Quadrature-Mirror-Filter) Filterbank in eine Anzahl von Teilbandsignalen entsprechend einer linearen Quantisierung zu unterteilen und die resultierenden Teilbandsignale beispielsweise mittels adaptiver PCM- oder DPCM-Codierung einer Datenreduktion zu unterziehen.

30

35

Es ist ferner aus der DE-PS 34 40 613 bekannt, innerhalb jedes Teilbandsignales die Quantisierung der Nutzinformation so zu wählen, daß das Quantisierungsrauschen durch die Nutzinformation in demselben Teilband gerade verdeckt ist, wodurch sich ebenfalls eine Datenreduktion erzielen läßt. Der mit dem bekannten Verfahren erreichbare Datenreduktionsfaktor hat etwa den Wert vier, d.h., der Informationsfluß eines hochwertigen digitalen Tonsignals reduziert sich von etwa 500 Kbit/s auf etwa

1

125 Kbit/s, ohne subjektive Beeinträchtigung der Quali-
5 tät.

Zur Erzielung einer noch weitergehenden Datenreduktion
ist es ferner bekannt (EP-OS 193 143, DE-OS 35 06 912
sowie "Rundfunktechnische Mitteilungen", Bd. 30 (1986),
10 Heft 3, S. 117 - 123), eine Spektralanalyse des breit-
bandigen Tonsignals mit Hilfe einer diskreten Fourier-
Transformation (beispielsweise durch Fast-Fourier-Trans-
formation) durchzuführen und bestimmte, innerhalb ver-
schiebener Frequenzgruppen relevante Spektralwerte nach
15 Betrag und Phase so zu codieren, daß unter Berücksichti-
gung der durch Mithörschwellen beschriebenen Verdeckungs-
eigenschaften des menschlichen Gehörs und gemäß unter-
schiedlicher Qualitätskriterien eine höhere Datenreduk-
tion erreicht wird.

20

Das für die Fourier-Transformation erforderliche Analyse-
Zeitfenster beträgt indessen etwa 25 ms. Dieser Wert.
stellt einen Kompromiß dar, um einerseits dem spektralen
Auflösungsvermögen und andererseits dem zeitlichen Auf-
25 lösungsvermögen des menschlichen Gehörs gerecht zu wer-
den. Die mit diesem Analyse-Zeitfenster erreichbare
spektrale Auflösung beträgt lediglich 40 Hz, so daß
im Bereich tiefer Frequenzen, wo die Frequenzgruppen-
breite des Gehörs etwa 100 Hz umfaßt, nur zwei Spektral-
30 werte übertragen werden können. Die dabei entstehenden
Seitenbänder liegen deshalb in den benachbarten Frequenz-
gruppen, so daß wahrnehmbare Qualitätsbeeinträchtigungen
nicht auszuschließen sind. Andererseits ist das kompro-
mißweise gewählte 25 ms-Analyse-Zeitfenster für das
35 zeitliche Auflösungsvermögen des Gehörs zu lang. Da

- 3 -

1 bei impulshaltigen Nutzsignalen diese Ungenauigkeit im
Zeitbereich zu wahrnehmbaren Verzerrungen führt, muß
zur Verzerrungsverminderung eine Anhebung der Amplitu-
denwerte der zeitlich vorausgehenden Spektralanteile
5 durchgeführt werden, was jedoch nicht in allen Fällen
zum gewünschten Erfolg führt. Ferner dürfen in der di-
gitalen Tonstudioteknik Blocklängen von etwa 5 ms
nicht überschritten werden, um beim Schneiden von digi-
talisierten Tonsignalen unhörbare Schnitte zu ermöglichen.
10 Hinzu kommt ferner, daß der Prozessoraufwand insbesonde-
re im Empfänger zur Rücktransformation der sendeseitig
transformierten Signale im Bereich hoher Frequenzen un-
nötig hoch ist, da die Berücksichtigung psychoakusti-
scher Kriterien nur frequenzgruppenweise erfolgt.

15

Desweiteren ist bei dem zuletzt betrachteten, bekannten
Verfahren für die empfangsseitige Rückgewinnung der
Spektralwerte nach Betrag und Phase sowie für die empfangs-
20 seitige inverse Fourier-Transformation die Übertragung
von Nebeninformationen erforderlich. Diese Nebeninformationen
stellen einen relativ hohen Anteil des gesamten Netto-Informa-
tionsflusses dar und erfordern einen besonders wirksamen
Fehlerschutz, welcher den Informationsfluß des zu über-
25 tragenden, codierten Signals entsprechend erhöht.
Schließlich ist bei dem bekannten Verfahren das quellen-
codierte Signal empfindlich gegenüber Bitfehler-Störungen,
weil sowohl Betrag als auch Phase jedes Spektralwertes
blockweise, d.h., nur etwa alle 25 ms übertragen werden,
30 so daß ein Bitfehler ein Störspektrum innerhalb dieses
Zeitintervalls erzeugt. Die Störwirkung eines 25 ms-
Impulses ist wesentlich höher als beispielsweise die-
jenige eines 1 ms-Impulses, welche sich bei den eingangs

35

1 erwählten, bekannten Teilband-Verfahren für die
fehlerhafte Übertragung eines Teilband-Abtastwertes
ergibt.

5 Die Aufgabe der Erfindung besteht demgegenüber darin,
unter Vermeidung einer Fourier-Transformation mit
einem Verfahren der eingangs erwähnten Art die Ver-
deckungseigenschaften des menschlichen Gehörs voll-
ständig zu nutzen, dahingehend, daß ein begrenzter
10 Gesamtinformationsfluß qualitätsoptimal auf die
Spektralanteile des Nutzsignals dynamisch verteilt
wird, ohne einen hohen Aufwand hinsichtlich der Über-
tragung von Nebeninformationen, des Fehlerschutzes
oder der empfangsseitigen Signalverarbeitung in Kauf
15 nehmen zu müssen.

Diese Aufgabe wird erfindungsgemäß durch die kenn-
zeichnenden Merkmale des Anspruchs 1 gelöst.

20 Vorteilhafte Ausgestaltungen und Weiterbildungen des
erfindungsgemäßen Verfahrens ergeben sich aus den
Unteransprüchen.

25 Die Erfindung beruht auf der Überlegung, daß sowohl
die gegenseitige Verdeckung der Spektralanteile des
Nutzsignals als auch die Verdeckung des Quantisie-
rungsrauschens nicht nur innerhalb der jeweiligen
Teilbänder, sondern über mehrere, benachbarte Teilbän-
der hinweg erfolgt. Um die Verdeckung vollständig zu
30 nutzen, müssen die Quantisierungen von Teilbandsigna-
len nach Maßgabe der aus verschiedenen Spektralantei-
len des Nutzsignals resultierenden Mithörschwelle ge-
steuert werden. Die Berechnung der dazu erforderlichen
Steuerinformationen geschieht signalabhängig unter
35 Berücksichtigung der Vor-, Simultan- und Nachverdeckung

1

des menschlichen Gehörs. Da einerseits der erforderliche
Gesamtinformationsfluß eines derartig codierten Tonsigna-
les signalabhängig schwankt und wenn andererseits für den
Anwendungsfall der Übertragung der Informationsfluß des
übertragenen codierten Tonsignals konstant gehalten werden
soll, kann die daraus resultierende, signalabhängige In-
formationsfluß-Reserve sendeseitig unter Berücksichtigung
zusätzlicher Kriterien verwendet werden. Insbesondere kann
die signalabhängige Informationsfluß-Reserve ganz oder
teilweise für die Quantisierung der Teilbandsignale zur
Verfügung gestellt werden, so daß sich die Abstände der
Quantisierungsrauschpegel der Teilbandsignale zu der resul-
tierenden Mithörschwelle vergrößern. Sie kann ferner ganz
oder teilweise für den Fehlerschutz der im Multiplexbetrieb
übertragenen, codierten Teilbandsignale sowie für den
Fehlerschutz eines Multiplexrahmens zur Verfügung gestellt
werden, so daß sich der Fehlerschutzgrad des Multiplex-
signals erhöht.

Die Erfindung wird nachstehend anhand der Zeichnungen
näher erläutert. Es zeigt:

25

Fig. 1a Blockschaltbilder der sendeseitigen
und 2a und empfangsseitigen Schaltungsteile
zur Durchführung des erfindungsgemäßen
Verfahrens gemäß einer ersten Aus-
führungsform;

30

Fig. 1b Blockschaltbilder wie in Fig. 1a und
und 2b 2a gemäß einer zweiten Ausführungsform;

35

1

Fig. 1c und 2c Blockschaltbilder wie in Fig. 1b und 2b gemäß einer dritten Ausführungsform;

5

Fig. 1d und 2d Blockschaltbilder wie in Fig. 1c und 2c gemäß einer vierten Ausführungsform, mit zusätzlicher Spektral-Analyse (FFT) sendeseitig;

10

Fig. 1e Blockschaltbild zur Durchführung einer stufenweisen Datenreduktion gemäß der ersten Ausführungsform des erfindungsgemäßen Verfahrens

15

Fig. 3 ein Blockschaltbild für die in den Fig. 1a, 1b und 2a vorgesehene Stufe zur dynamischen Verteilung des Informationsflusses;

20

Fig. 4 Beispiele von drei verschiedenen Mithörschwellenkurven, welche sich in der Frequenzlage des maskierenden Tonsignals unterscheiden;

25

Fig. 5 die Abhängigkeit der mittleren Mithörschwellenkurve gemäß Fig. 4 von fünf unterschiedlichen Pegeln des maskierenden Tonsignals;

30

Fig. 6 ein Frequenzdiagramm, in welchem für den Vokal /ə/ die Harmonischen als Punkte sowie die daraus resultierende Mithörschwelle als durchgezogene Kurve aufgetragen sind;

35

- 1
- 5
- 10
- 15
- 20
- 25
- 30
- 35
- Fig. 7 ein Zeitdiagramm für den zeitlichen Verlauf der Vor-, Simultan- und Nachverdeckung des menschlichen Gehörs;
- Fig. 8 ein Teilband-Frequenzschema eines Ausführungsbeispiels mit insgesamt 24 Teilbändern, wobei die Mithörschwellenkurven gemäß Fig. 4 eingetragen sind;
- Fig. 9 Kurven gleicher Lautstärke von Schmalbandrauschen in Abhängigkeit von der Bandbreite;
- Fig. 10 den strukturellen Aufbau der sende- und empfangsseitigen QMF-Filterbänke gemäß Fign. 1a und 2a bzw. 1b und 2b;
- Fig. 11 ein Blockschaltbild der Transcodierungsstufe gemäß Fign. 1a und 1b;
- Fig. 12 eine schematische Darstellung für die Bestimmung des Skalenfaktors für je acht Abtastwertes eines digitalisierten Tonsignals;
- Fig. 13 eine schematische Darstellung ähnlich wie in Fig. 12, wobei jedoch nur jeder zweite Skalenfaktor für Übertragungszwecke verwendet wird und die nichtübertragenen Skalenfaktoren anhand von sog. Zuordnungsbits rekonstruiert werden, welche angeben, ob der betreffende,

1

nicht-übertragene Skalenfaktor dem vorangehenden oder nachfolgenden, übertragenen Skalenfaktor zugeordnet ist.

5

Fig. 14 Beispiel für den tonsignalabhängigen Zeitverlauf des für die Codierung erforderlichen Gesamtinformationsflusses;

10

Fig. 15 Beispiel für den tonsignalabhängigen Zeitverlauf des für die Codierung erforderlichen Gesamtinformationsflusses sowie der Informationsfluß-Reserve bei konstant gehaltenem Informationsfluß des Multiplexsignals;

15

Fig. 16 Beispiel für den tonsignalabhängigen Zeitverlauf des für die Codierung erforderlichen Gesamtinformationsflusses gemäß Fig. 15, wobei die Informationsfluß-Reserve im wesentlichen für einen dynamischen Fehlerschutz genutzt wird.

20

25

GESAMTSHEMA

Wie Fig. 1a zeigt, wird ein digitalisiertes Tonsignal, z.B. ein hochwertiges Rundfunksignal, in eine Anzahl von Teilbandsignalen, beispielsweise 24 Teilbandsignale, unterteilt. Die Aufspaltung des digitalen Tonsignals in Teilbänder erfolgt vorzugsweise mit Hilfe einer aus Quadrature-Mirror-Filtern (QMF) bestehenden Filterbank 1, deren Aufbau und Funktionsweise in Fig. 10 näher darge-

30

35

1

stellt ist und später erläutert werden soll. Das Eingangssignal der Filterbank 1 weist beispielsweise eine Bandbreite von 16 kHz auf und ist mit einer Auflösung von 16 Bit linear quantisiert. Auch größere Bandbreiten und Auflösungen sind möglich, beispielsweise 20 kHz und 18 Bit.

5

10

Die hohen Informationsflüsse der am Ausgang der Filterband 1 vorliegenden 24 Teilbandsignale werden in einer nachgeschalteten Transcodierungsstufe 2 reduziert. Die transcodierten Teilbandsignale werden einer speziellen, später noch näher beschriebenen Fehlersicherung 7 unterzogen und einem Multiplexer 3 zugeführt, welcher die 24 Teilbandsignale des Ausführungsbeispiels im Zeitmultiplex auf eine Übertragungsstrecke, beispielsweise eine Rundfunkübertragungsstrecke, gibt.

15

20

Für die Informationsflußreduktion können sowohl redundanzreduzierende als auch irrelevanzreduzierende Verfahren eingesetzt werden, wobei auch Kombinationen beider Verfahren denkbar sind. Redundanzreduzierende Verfahren unterdrücken solche Informationen, die für die Rekonstruktion des ursprünglichen Signals nicht benötigt werden. Demgegenüber unterdrücken irrelevanzreduzierende Verfahren diejenigen Informationen, welche zur Unterscheidung des rekonstruierten Signals vom ursprünglichen Signal von dem menschlichen Gehör nicht benötigt werden. Beispielsweise wird bei einem irrelevanzreduzierenden Verfahren die Quantisierung des Nutzsignals innerhalb jedes Teilbandes so gewählt, daß das Quantisierungsrauschen durch das Nutzsignal gerade verdeckt ist. Bei einem irrelevanz- und redundanzreduzierenden Verfahren werden redundanzreduzie-

25

30

35

1

rende Verfahren, beispielsweise adaptive PCM oder DPCM-Prozessoren, in den Teilbändern eingesetzt und so ausgelegt, daß die Rekonstruktion des ursprünglichen Signals nicht vollständig möglich ist, aber das resultierende Fehler-signal durch das Nutzsignal gerade verdeckt ist.

5

10

Ein zweckmäßiges Ausführungsbeispiel einer kombinierten Irrelevanz- und Redundanzreduktion in den Teilbändern wird im folgenden beschrieben. Die Quantisierung der Teilbandsignale erfolgt nach Maßgabe der Mithörschwelle und nach Maßgabe von Skalenfaktoren. Die Skalenfaktoren klassieren die Spitzenwerte der Teilbandsignalpegel innerhalb eines Zeitintervalls, welches dem zeitlichen Auflösungsvermögen des Gehörs entspricht. Sie werden einer zusätzlichen Irrelevanz- und Redundanzreduktion unterworfen. Die Übertragung erfolgt im Multiplexsignal, wobei die Fehlersicherung 8 spezielle, später noch zu beschreibende Eigenschaften aufweist.

15

20

Dynamische Informationsfluß-Verteilung

25

Eine invariante Verteilung des Informationsflusses auf die Teilbänder ist im Sinne einer optimalen Irrelevanzreduktion ungünstig, weil der Verlauf der Mithörschwelle von der spektralen und zeitlichen Struktur des Nutzsignals abhängt. Deshalb werden erfindungsgemäß die Quantisierungen der Teilbandsignale in der Stufe 2 gesteuert ("dynamische Verteilung" des zur Verfügung stehenden Informationsflusses auf die Teilbänder).

30

35

Die Steuersignale für die Quantisierungen der Teilbandsignale in der Transcodierungsstufe 2 werden in einer

1

Stufe 5 nach Maßgabe der spektralen und zeitlichen Mit-
hörschwellen des menschlichen Gehörs gewonnen. Unter
5 der Steuerung dieser Steuersignale in der Stufe 2 er-
folgt die dynamische Verteilung des Informationsflusses
auf die einzelnen Teilbänder unter Berücksichtigung
der Vor-, Simultan- und Nachverdeckung des menschlichen
10 Gehörs, wie im einzelnen anhand der Figuren 5 bis 8
noch näher erläutert wird. Dabei werden

- a) Teilbänder des Nutzsignals, die durch benachbarte
Teilbänder desselben Nutzsignals verdeckt werden,
15 nicht oder nur teilweise übertragen, und
- b) Teilbänder des Nutzsignals, die durch benachbarte
Teilbänder desselben Nutzsignals nicht (vollständig)
verdeckt werden, nur so fein quantisiert, daß das
20 resultierende Quantisierungsrauschen durch den
höheren Nutzsignalpegel der benachbarten Teilband-
signale verdeckt ist.

Damit läßt sich eine gegenüber dem Stand der Technik
25 höhere Irrelevanzreduktion erzielen.

25

Bekanntere Verfahren erreichen eine Reduktion des
Informationsflusses dadurch, daß die Quantisierung
der Teilbandsignale nach Maßgabe der innerhalb eines
Teilbandes wirksamen Verdeckung erfolgt. Beispielsweise
30 gilt für ein reines irrelevanzreduzierendes Verfahren
näherungsweise die in der DE-PS 34 40 613 angegebene
Beziehung

35

1

$$\text{ld } q_{i \text{ min}} = 32 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 1 \left[\frac{\text{bit}}{\text{Abtastwert}} \right] \quad (1)$$

5

wobei

10

$q_{i \text{ min}}$ die minimale Zahl der Quantisierungsstufen im Teilband i ,

f_{oi} , f_{ui} die obere bzw. untere Grenzfrequenz des Teilbandes i

15

bedeuten. Der resultierende erforderliche Informationsfluß für alle Teilbänder ergibt sich dann zu:

$$r_{\text{min}} = 2 \cdot \sum_{i=1}^n (f_{oi} - f_{ui}) \cdot \text{ld } q_{i \text{ min}} \quad [\text{bit/s}] \quad (2)$$

20

Der resultierende erforderliche Informationsfluß ist konstant, er beträgt für eine Auflösung in 24 Teilbänder etwa 100 Kbit/s.

25

Dieser Wert reduziert sich unter der Berücksichtigung der gegenseitigen Verdeckung der Teilbandsignale. Beispielsweise kann für ein etwa gleichmäßig verdeckendes, breitbandiges Nutzsignal die Auflösung aller Teilbänder, die auf der Frequenzachse oberhalb des ersten Teilbandes liegen, so reduziert werden, daß sich der erforderliche Informationsfluß um etwa 30% vermindert. Für ein schmalbandiges Nutzsignal ist prinzipiell eine wesentlich größere Datenreduktion möglich, weil der Informationsfluß in vielen Teilbändern zu Null gesetzt werden kann.

35

1

Die signalabhängig gesteuerten Quantisierungen der Teilbandsignale führen im Prinzip zu einem erforderlichen Informationsfluß, der weitgehend von der spektralen und zeitlichen Struktur des Nutzsignals abhängt und daher signalabhängig schwankt, etwa im Bereich 20 ... 70 Kbit/s. Auf die vorteilhafte Nutzung eines signalabhängig schwankenden Informationsflusses wird später näher eingegangen.

10

Skalenfaktor

Die Transcodierung der Teilbandsignale erfolgt bei dem Ausführungsbeispiel nach Fig. 1a nicht nur nach Maßgabe des Mithörschwellenkriteriums im Sinne einer Irrelevanzreduktion unter der Steuerung der Stufe 5, sondern ferner nach Maßgabe von Skalenfaktoren, welche die Spitzenwerte der Teilbandsignalpegel innerhalb eines bestimmten Zeitintervalls klassieren und welche für die Transcodierung jedes Teilbandsignals dessen Auflösung innerhalb des Zeitintervalls festlegen. Die Einzelheiten für die Bestimmung der Skalenfaktoren in der Stufe 4 und deren Auswertung bei der Transcodierung in der Stufe 2 sowie deren Übertragung werden im folgenden noch näher erläutert. Skalenfaktoren für die Teilbandsignale erweisen sich aus drei Gründen vorteilhaft:

25

1. Sie Skalenfaktoren für die Teilbandsignale enthalten alle Informationen, welche für die Bestimmung der Steuergröße nach dem Mithörschwellenkriterium in Stufe 5 erforderlich sind. Die Übertragung der Skalenfaktoren als Nebeninformation reicht daher aus, um empfangsseitig eine inverse Transcodierung vorzunehmen (vgl. Fig. 2a). Und zwar werden aus den empfangenen

30

35

1

Skalenfaktoren die Informationen über die Verteilung des
Gesamtinformationsflusses auf die Teilbänder gewonnen.

5

2. Da der durch den Skalenfaktor klassifizierte Spitzenwert
des Teilbandsignalpegels durch die Abtastwerte des be-
treffenden Teilbandsignals nicht überschritten werden
kann, kann folglich der Störpegel innerhalb jedes Teil-
bandes, welcher durch Bitfehler verursacht wird, den
Spitzenwert des Teilbandsignalpegels nicht über die
Klassifizierungsgenauigkeit hinaus überschreiten. Dadurch
ergibt sich durch Bitfehler grundsätzlich ein Störspek-
trum, welches durch das Nutzsinal größtenteils verdeckt
ist. Selbst bei Auftreten maximaler Bitfehler, d.h. voll-
ständiger Zerstörung aller Abtastwerte, ist durch den in
diesem Falle nur noch vorhandenen Skalenfaktor gewähr-
leistet, daß die spektrale Hüllkurve des Störspektrums der
Hüllkurve des Nutzsignals näherungsweise entspricht,
mithin eine Sprachverständlichkeit noch erhalten bleibt.

10

15

20

25

30

3. Da die Skalenfaktoren der Teilbandsignale alle wesent-
lichen Informationen enthalten, welche für die Bestimmung
der Steuergröße nach dem Mithörschwellenkriterium erforder-
lich sind, ist es notwendig, diese Information bei der
Übertragung gegen Bitfehler stark zu schützen. Dies kann
z.B. durch Einfügung von Redundanz geschehen. Aufgrund
dieser Einfügung der Redundanz ist es besonders wichtig,
eine Minimierung der Übertragung der Skalenfaktoren vorzu-
nehmen. Dem kann durch eine minimal notwendige Wortlänge
(sh. Abschnitt Skalenfaktorbildung) und eine von der
Statistik des Signals und den Erfordernissen des Gehörs
abhängige Übertragungshäufigkeit entsprochen werden.

35

1

Den Erfordernissen des Gehörs kann Rechnung getragen werden, indem die zeitliche Maskierung, vor allem der Effekt der Nachverdeckung genutzt wird (Irrelevanzreduktion). Dies bedeutet, daß bei schnellem Ausklingverhalten des Tonsignals die Skalenfaktoren nicht genau bestimmt werden müssen, sondern z.B. aufgrund von Interpolationen von zu früheren und späteren Zeitpunkten ermittelten Skalenfaktoren angenähert werden können. Da die zeitliche Vorverdeckung im allgemeinen sehr kurz und stark signalabhängig ist (1 bis 20 ms), müssen in Fällen schneller Signalanstiege die Skalenfaktoren häufiger übertragen werden.

15

Einer Redundanzreduktion bei der Übertragung der Skalenfaktoren wird entsprochen, indem bei solchen Teilbandsignalen, deren Pegel sich über einen bestimmten Zeitraum nicht, oder nur geringfügig ändern, die Skalenfaktoren nur selten übertragen werden.

20

Dadurch, daß nicht notwendigerweise für alle Teilbänder pro zeitlichen Block Skalenfaktoren übertragen werden müssen, sondern abhängig von den Erfordernissen des Gehörs bzw. des Tonsignals auch näherungsweise durch Interpolationsmechanismen bestimmt werden können, ergibt sich eine Übertragungsrate aller Skalenfaktoren von ca. 10 bis 20 kbit/s.

25

30

Da empfangsseitig die Transcodierung allein durch die detektierten Skalenfaktoren gesteuert wird, sind die teilbandcodierten Multiplexsignale besonders unempfindlich gegenüber Bitfehlerstörungen, wenn die Skalenfaktoren sendeseitig in der Stufe 8 einer wirkungsvollen Fehlersicherung unterzogen werden. Der Vorteil der Verwendung von Skalenfaktoren gegenüber anderen datenreduzierenden Methoden, beispielsweise adaptive PCM- oder DPCM, liegt also darin, daß eine hohe Unempfindlichkeit gegenüber Bitfehlern erreicht

35

1

wird, wenn nur der durch die Skalenfaktoren verursachte
5 Nebeninformationsfluß effektiv geschützt wird. Die Höhe die-
ses Nebeninformationsflusses schwankt signalabhängig etwa im
Bereich 10 ... 20 Kbit/s, weil zum Zwecke einer Irrelevanz-
und Redundanzreduktion die für die Bestimmung der Skalenfak-
10 toren zugrunde gelegten Zeitintervalle (Blocklängen) der
zeitlichen Verdeckung des menschlichen Gehörs und der zeit-
lichen Struktur des jeweiligen Teilbandsignals entsprechen.

Für die Übertragung der transcodierten Teilbandsignale sowie
der Skalenfaktoren ist also ein Gesamtinformationsfluß er-
15 forderlich, der signalabhängig im Bereich 30 ... 90 Kbit/s
schwankt.

Nutzung des schwankenden Informationsflusses

20 Durch die signalabhängige Schwankung des für die Codierung
erforderlichen Gesamtinformationsflusses kann das erfindungs-
gemäße Verfahren vorteilhaft ausgestaltet werden.

Fig. 14 zeigt beispielhaft den zeitlichen Verlauf des erfor-
25 derlichen Gesamtinformationsflusses. Die gestrichelte Linie
stellt den durchschnittlich erforderlichen Gesamtinforma-
tionsfluß (etwa 60 Kbit/s) dar. Dieser Wert könnte zugrunde
gelegt werden, wenn sämtliche Schwankungen innerhalb eines
großen Zeitabschnittes mit Hilfe eines entsprechend großen
30 Pufferspeichers ausgeglichen werden können. Dies ist für
den Fall der Tonsignalübertragung wegen der entsprechend
langen Verzögerungszeit nicht möglich, jedoch ergibt sich
für die Tonsignalspeicherung eine erste vorteilhafte Ausge-
staltung des erfindungsgemäßen Verfahrens:

35

1

1) Der Informationsfluß des Multiplexsignales schwankt.

5

10

Wird das Multiplexsignal so ausgelegt, daß am Ausgang des Multiplexers 3 in Fig. 1A der Informationsfluß in ähnlicher Weise schwankt wie der für das codierte Tonsignal aufgewendete Gesamtinformationsfluß an den Eingängen der Stufen 7 und 8, so resultiert daraus eine besonders hohe Datenreduktion für den Fall der Speicherung. Für spezielle Speichertechniken, beispielsweise Speicherung auf Rechner-Magnetplatten, können sogar die Stufen 3, 7 und 8 entfallen, so daß der zu speichernde Gesamtinformationsfluß nicht höher sein muß als der erforderliche Gesamtinformationsfluß.

15

20

Der Langzeit-Durchschnittswert des zu speichernden Gesamtinformationsflusses kann sogar unter 60 Kbit/s liegen, weil bei Speicherung eines vollständigen Programmes alle darin enthaltenen kurzzeitigen Pausen (beispielsweise in der Sprache) nur einen sehr geringen Gesamtinformationsfluß erfordern (etwa 10 bis 15 Kbit/s, hauptsächlich für die Skalenfaktoren). Für die Speicherung von Sprache können daher weniger als 40 Kbit/s zugrunde gelegt werden.

25

2) Der Informationsfluß des Multiplexsignals ist konstant.

30

Im Fall der Übertragung des Multiplexsignals ist dessen konstanter Informationsfluß vorteilhaft. Da die durch Pufferspeicher verursachte Verzögerungszeit nur gering sein darf, können die Schwankungen des erforderlichen Gesamtinformationsflusses nur geringfügig ausgeglichen werden. Die verbleibenden Schwankungen sind beispielhaft in Fig. 15 dargestellt. Der konstante Informationsfluß des Multiplexsignales ist durch die gestrichelte Gerade angegeben (90 Kbit/s). Der

35

1

5 obere Bereich stellt somit eine Informationsfluß-Reserve dar, die signalabhängig schwankt und auf verschiedene Weise genutzt werden kann:

10

a) Für einen erhöhten Abstand der Quantisierungsrauschpegel zu der resultierenden Mithörschwelle ("Erhöhter Rauschabstand").

15

b) Für den Fehlerschutz 7, 8 und die Bildung des Multiplexsignals in 3 ("Dynamischer Fehlerschutz").

20

c) Für die Übertragung beliebiger Zusatzinformationen im Multiplexsignal, die nicht zeitkritisch und unabhängig vom Tonsignal sind, z.B. Programminformationen oder Radiotext-Information ("Zusatzsignal-Übertragung").

25

Die Informationsfluß-Reserve kann natürlich in beliebiger Kombination der drei Möglichkeiten genutzt werden. Im folgenden werden der erhöhte Rauschabstand und der dynamische Fehlerschutz noch näher beschrieben.

25

a) Erhöhter Rauschabstand

30

Quellencodierungsverfahren, welche die Irrelevanz der digitalen Tonsignale vollständig beseitigen (d.h., die den Effekt der spektralen Vor-, Simultan- und Nachverdeckung vollständig nutzen), können in bestimmten Anwendungsfällen Probleme verursachen:

35

1

5 - Im Falle einer Kaskadierung derartiger Quellen-
codierungsverfahren kann das Quantisierungsrauschen
die Mithörschwellen überschreiten. Würde beispiels-
weise sowohl die Speicherung als auch die Übertra-
gung eines Rundfunk-Programmsignals mit Hilfe eines
derartigen Quellencodierungsverfahrens erfolgen,
10 können für kritische Tonsignale Qualitätsbeein-
trächtigungen wahrnehmbar werden.

15

- Im Falle einer nachträglichen empfängerseitigen
Anhebung oder Absenkung bestimmter Frequenzanteile
des Nutzsignals kann sich die spektrale Verdeckung
des Nutzsignals so verändern, daß Qualitätsbeein-
trächtigungen wahrnehmbar werden. Diese Gefahr be-
steht, wenn der Pegel eines Teilbandsignals, welches
benachbarte Teilbandsignale verdeckt, empfängersei-
tig abgesenkt wird, oder wenn der Pegel eines Teil-
bandsignales, welches durch ein benachbartes Teil-
bandsignal ganz oder teilweise verdeckt wird,
20 empfängerseitig angehoben wird.

20

25

Um Qualitätsbeeinträchtigungen für diese Anwendungs-
fälle auszuschließen, erfolgt die erfindungsgemäße
dynamische Verteilung des Informationsflusses nicht
allein nach Maßgabe einer maximalen Datenreduktion,
sondern auch nach Maßgabe einer Reserve an Rauschab-
stand (sog. "mask-to-noise Reserve"). Die "mask-to-noise
30 Reserve" schwankt signalabhängig, und zwar etwa propor-
tional der Informationsfluß-Reserve, wie beispielhaft
in Fig. 15 dargestellt. Ist das Nutzsignal z.B. so
schmalbandig, daß der Informationsfluß in vielen Teil-
bändern zu Null gesetzt wird, erfolgt eine entsprechende
35 Erhöhung des Informationsflusses (und damit der Auf-

35

1

lösung) für diejenigen Teilbandsignale, welche eine Verdeckung hervorrufen. Eine derartige Erhöhung erfolgt soweit, wie es die Informationsfluß-Reserve zuläßt. Daher ist die Auflösung der verdeckenden Teilbandsignale unter Umständen wesentlich höher als das Mithörschwellenkriterium es erfordert.

5

10

Der Vorteil dieser Form der dynamischen Verteilung des Informationsflusses auf die Teilbänder liegt darin, daß beispielsweise Pegeltöne mit sehr hoher Auflösung (z.B. 16 bis 18 bit linear quantisiert) übertragen werden. Die Übertragung einer einzelnen Spektrallinie mit 16 bit Auflösung in einem 500 Hz breiten Teilband erfordert theoretisch 16 Kbit/s, allerdings wegen der zu berücksichtigenden Aliasing Verzerrungen (auf welche später eingegangen wird, vgl. Fig. 3) einen etwa doppelt so hohen Bit-Fluß. Damit werden je nach Frequenzlage zwei oder mehr Spektrallinien gleichzeitig ohne meßtechnisch erfaßbare Qualitätseinbuße übertragen, wenn etwa 500 Hz breite Teilbänder und ein Bit-Fluß des Multiplexsignals von etwa 90 Kbit/s zugrundegelegt werden. Wesentliche Funktionen der in Fign. 1a und 2a dargestellten Teilband-Übertragungstrecken lassen sich daher meßtechnisch einfach überprüfen, indem beispielsweise Sinustöne beliebiger Frequenz und Amplitude übertragen werden.

15

20

25

30

35

1

Für den dynamischen Fehlerschutz gilt folgendes:

5

Für die Kanalcodierung, d.h. für die Fehlersicherung der transcodierten Abtastwerte in Stufe 7, der Skalenfaktoren in Stufe 8, sowie für die Bildung des Multiplexsignals in Stufe 3 ist ein zusätzlicher Informationsfluß erforderlich. Entsprechend setzt sich der Informationsfluß des Multiplexsignals aus den für die Quellen- und Kanalcodierung eingesetzten Informationsflüssen zusammen.

10

15

20

25

Die bei konstantem Informationsfluß des Multiplexsignals vorhandene dynamische Informationsfluß-Reserve ist für die Kanalcodierung nutzbar, dergestalt, daß der Fehlerschutzgrad des Multiplexsignals abhängig von der momentan vorhandenen Informationsfluß-Reserve gesteuert wird (dynamischer Fehlerschutz). Die Steuerung des Fehlerschutzgrades erfolgt vorteilhaft stufenweise. Fig. 1b zeigt beispielsweise einen stufenweise, signalabhängig schwankenden Informationsfluß für den dynamischen Fehlerschutz (Bereich zwischen gestrichelter Gerade und stufenförmiger Kurve). Der durch die stufenförmige Kurve beschriebene Informationsfluß stellt gleichzeitig den für die Quellencodierung verfügbaren Gesamtinformationsfluß dar; er ist etwas größer als der erforderliche Gesamtinformationsfluß.

30

35

Der dynamische Fehlerschutz führt zu einer Erhöhung des durchschnittlichen Fehlerschutzgrades, entsprechend der durchschnittlich vorhandenen Informationsfluß-Reserve. Damit wird die Wahrscheinlichkeit für Störungen durch Bitfehler reduziert. Der dynamische Fehler-

1

5

10

15

20

25

30

35

schutz bewirkt ferner, daß diejenigen Tonsignale, welche einen geringen Gesamtinformationsfluß erfordern, mit hohem Fehlerschutzgrad übertragen werden und Tonsignale, welche einen hohen Gesamtinformationsfluß erfordern, mit geringem Fehlerschutzgrad übertragen werden. Diese Fehlerschutzeigenschaft wirkt sich besonders vorteilhaft aus, weil Tonsignale, die einen geringen Gesamtinformationsfluß benötigen, die durch Bitfehler verursachten Störsignale nur schwach verdecken. Gerade diese empfindlichen Tonsignale werden stark geschützt. Beispielsweise ist für ein Sprachsignal, insbesondere in den Sprachpausen, oder während bestimmter Musikpassagen ("Stille im Konzertsaal") der erforderliche Gesamtinformationsfluß besonders gering und damit der Fehlerschutzgrad besonders hoch.

Die Auslegung des dynamischen Fehlerschutzgrades erfolgt zweckmäßigerweise derart, daß bei gegebener Bitfehlerrate die subjektive Störwirkung näherungsweise unabhängig vom Tonsignal ist.

Qualitätsabstufung

Ein weiteres Merkmal des erfindungsgemäßen Quellencodierungsverfahrens ist die Möglichkeit, sendeseitig in der Stufe 6 (Fig. 1b) die Qualität der Quellencodierung festzulegen. Die in der Stufe 5 festgelegten und später anhand der Fign. 4 bis 8 näher erläuterten Kriterien für die Quantisierungen der Teilbandsignale werden mit Hilfe der Qualitätsfestlegung (Stufe 6) bewertet. Dies geschieht auf folgende Weise:

1

a) Der für die Codierung des Tonsignales verfügbare Gesamtinformationsfluß wird durch die Qualitätsfestlegung bestimmt.

5

b) Der für die dynamische Verteilung bereits vorbeschriebene Gesichtspunkt "Erhöhter Rauschabstand" wird bewertet. Die "mask-to-noise-Reserve" wird abhängig von der Qualitätsfestlegung dimensioniert.

10

c) Das Mithörschwellenkriterium wird abhängig von der Qualitätsfestlegung so ausgelegt, daß bestimmte kritische Nutzsignale, die einen hohen Gesamtinformationsfluß erfordern und selten vorkommen, wahrnehmbare, aber nicht störende Qualitätsbeeinträchtigungen aufweisen. Beispielsweise werden Qualitätsabstufungen durch die Wahrscheinlichkeit festgelegt, mit welcher diese Qualitätsbeeinträchtigungen auftreten.

15

20

d) Abhängig von der Qualitätsfestlegung wird für kritische Nutzsignale eine Anzahl der Teilbandsignale zu Null gesetzt. Dies geschieht vorrangig für Teilbandsignale mit größerer Bandbreite bzw. größerem Informationsfluß und zusätzlich nach Maßgabe minimaler Störwirkung: Die verhältnismäßig hohe Störwirkung bei nicht ausreichender Auflösung bestimmter Teilbandsignale wird vermieden, indem zusätzlicher Informationsfluß zugunsten dieser Teilbandsignale durch Nullsetzen unwesentlicher Teilbandsignale gewonnen wird. Unwesentliche Teilbandsignale sind solche, die im Verhältnis zu anderen Teilbandsignalen einen geringen Pegel aufweisen und nur einen unbe-

25

30

35

1

deutenden Beitrag zur Wahrnehmung der Klangfarbe
liefern.

5

10

Wesentlich ist, daß eine Reduktion des Gesamtinfor-
mationsflusses mit einer minimalen Reduktion der Qualität
verbunden ist, weil die dynamische Verteilung des
Informationsflusses auf die Teilbänder durch Stufe 5
nicht nur nach Maßgabe des zur Verfügung stehenden
Gesamtinformationsflusses, sondern auch nach Maßgabe
von qualitätsstufenspezifischen Kriterien erfolgt.

15

20

25

30

Durch eine stufenweise Reduktion der Qualität ist es
möglich, anstelle eines Nutzsignals zwei oder mehrere
Nutzsignale gleichzeitig mit demselben Informations-
fluß des Multiplexsignals zu übertragen. Da durch die
Qualitätsfestlegung in der Stufe 6 der für jedes Nutz-
signal zur Verfügung stehende Gesamtinformationsfluß
festgelegt ist, wird die Qualitätsabstufung so gewählt,
daß mit Wahl der Qualität die Zahl der übertragbaren
Kanäle festliegt und umgekehrt. Dazu werden die Kanal-
sperre 9 (Fig. 1b) sowie die Multiplexstufe 3 ent-
sprechend geschaltet. Die Schaltinformation wird im
Multiplexsignal mitübertragen, um empfangsseitig eine
qualitätsstufenadaptive Decodierung zu ermöglichen.
Eine Fehlersicherung der Schaltinformation erfolgt
dadurch, daß diese Schaltinformation häufiger als be-
nötigt übertragen wird, beispielsweise im 100 ms-Inter-
vall.

35

1

Folgende Korrelation zwischen Qualitätsabstufungen und Kanalzahl ist denkbar:

5

10

Anwendung	Qualitätsstufe	Kanalzahl	Bitfluß pro Kanal
Speicherung, Übertragung/Studio	1	1	200 kbit/s
Speicherung, Übertragung/Standard	2	2	100 kbit/s
Übertragung/Kommentar	3	3	65 kbit/s
Übertragung/Telefon	4	6	33 kbit/s

15

20

25

30

35

Eine Reduktion des Informationsflusses pro Kanal führt in charakteristischer Weise nur für solche Nutzsignale zu einer Qualitätsbeeinträchtigung, welche auch nach erfolgter Irrelevanzunterdrückung und Redundanzreduktion einen höheren Informationsfluß benötigen als beispielsweise gemäß vorstehender Tabelle zugelassen ist. Die vorliegend vorgesehene stufenweise Reduktion des Informationsflusses zeichnet sich dadurch aus, daß die Wahrscheinlichkeit für das Auftreten qualitätsbeeinträchtigter Signale geringer ist als bei bekannten Verfahren zur stufenweisen Reduktion des Informationsflusses. Die Wahrscheinlichkeit für das Auftreten qualitätsbeeinträchtigter Sprachsignale beispielsweise ist für die Qualitäts-

1

5

stufen 1 bis 3 der Tabelle gleich null und für die Qualitätsstufe 4 kleiner als 100%, wobei der Grad der Qualitätsbeeinträchtigung wesentlich niedriger liegt als bei der für Telefonsignale üblichen Bandbreitenreduktion.

Empfänger

10

15

20

25

30

35

Empfangsseitig werden gemäß Fig. 2a in der Demultiplex-Stufe 13 die fehlergesicherten transkodierte Teilbandsignale, die dazugehörigen fehlergesicherten Skalenfaktoren sowie die Information zur sendeseitig festgelegten Qualitätsstufe zurückgewonnen. Mit Hilfe der Qualitätsstufen-Information kann die Anzahl der alternativen Kanäle im Display 16 (Fig. 2b) angezeigt werden, so daß am Kanal-Wahlschalter 19 die von der Demultiplex-Stufe 13 auszugebenden Daten vom Hörer festgelegt werden können. In den Stufen 17 und 18 werden die Fehlersicherungsdaten der transkodierte Teilbandsignale bzw. der Skalenfaktoren entfernt sowie Fehlerkorrekturmaßnahmen durchgeführt. Die Skalenfaktoren dienen wie auf der Sendeseite als Eingangsinformation der Stufe 15 zur Steuerung der Verteilung des Informationsflusses auf die Teilbänder. Die Stufe 15 ist daher identisch mit der sendeseitigen Stufe 5 (vgl. Fig. 3). Nach Maßgabe der in Stufe 15 gewonnenen Steuerinformationen, der Skalenfaktoren sowie der Qualitätsstufen-Information erfolgt in der Stufe 12 die zur sendeseitigen Stufe 2 inverse Transkodierung, so daß 16 bis 18 bit-linear-quantisierte Teilbandsignale an der inversen QMF-Filterbank 11 anliegen und hier die Rückgewinnung des breitbandigen digitalen Tonsignals erfolgen kann. Einzelheiten zum Aufbau und zur Arbeitsweise der QMF-Filterbank 11 werden später anhand der Fig. 10 noch näher erläutert.

1

Zweites Ausführungsbeispiel

5

10

15

20

Ein weiteres Ausführungsbeispiel des erfindungsgemäßen Verfahrens ist in den Fig. 1c (Sendeseite) und Fig. 2c (Empfängerseite) veranschaulicht. Die sendeseitige Kodierung gemäß Fig. 1c ist identisch mit dem vorstehend bereits erläuterten Ausführungsbeispiel gemäß Fig. 1b. Abweichend hiervon ist in Fig. 1c vorgesehen, die Steuerinformationen für die Verteilung des Informationsflusses auf die Teilbänder im Multiplexsignal zu übertragen, wobei in Stufe 8b eine Fehlersicherung der Steuerinformation erfolgt, deren Wirksamkeit ähnlich bemessen wird wie diejenige der Skalenfaktor-Fehlersicherung in Stufe 8a. Bei der empfangsseitigen Dekodierung gemäß Fig. 2c werden in Abweichung von Fig. 2b die Steuerinformationen nicht nach Maßgabe der Skalenfaktoren neu ermittelt (wie dies in Stufe 15 gemäß Fig. 2b erfolgt), sondern direkt dem Multiplexsignal entnommen. In den Stufen 18a und 18b werden die Fehlersicherungsdaten der Skalenfaktoren bzw. der Steuerinformationen entfernt sowie Fehlerkorrekturmaßnahmen durchgeführt.

25

30

Das betrachtete zweite Ausführungsbeispiel erfordert gegenüber dem Ausführungsbeispiel gemäß Fig. 1b und 2b empfangsseitig einen geringeren technischen Aufwand, da die für die inverse Transkodierung in Stufe 12 erforderlichen Steuerinformationen nicht neu ermittelt zu werden brauchen. Der für die Übertragung und Fehlersicherung der Steuerinformationen erforderliche zusätzliche Informationsfluß ist etwa so groß wie für den Skalenfaktor.

35

1

Weitere Ausführungsformen

5

Weitere Ausführungsformen sowie vorteilhafte Ausgestaltungen und Weiterbildungen des erfindungsgemäßen Verfahrens, insbesondere gemäß Fig. 1d, 2d und 1e, werden im Verlauf der folgenden Beschreibung erläutert.

10

EINZELHEITEN DES VERFAHRENSSteuerung des Informationsflusses

15

Die anhand von Fig. 3 näher erläuterte Stufe 5 zur Ableitung des Mithörschwellenkriteriums umfaßt die getrennte Ermittlung der spektralen Mithörschwellen im Block 5.1 und der zeitlichen Mithörschwellen in Block 5.2. Dies geschieht unter Berücksichtigung des verfügbaren Gesamtinformationsflusses durch die Stufe 5.5 sowie in Abhängigkeit qualitätsstufenspezifischer Kriterien in Stufe 5.6. Dabei sind die Stufen 5.1, 5.2 und 5.5 in Serie geschaltet, wobei die Stufe 5.1 mit dem Ausgang der Stufe 4 und ein Steuereingang der Stufe 5.5 mit der Stufe 6 verbunden sind. Desweiteren ist ein erster Steuereingang der Stufe 5.1 ebenfalls mit der Stufe 6 verbunden, während ein zweiter Steuereingang der Stufe 5.1 von der Stufe 5.6 gespeist wird.

20

25

30

35

1

Da in der QMF-Filterbank 1 bei stark unterschiedlicher
5 Quantisierung benachbarter Teilbandsignale Aliasing-
Verzerrungen auftreten, erfolgt durch eine Stufe 5.4
eine Sollvorgabe bezüglich der maximal zulässigen
Quantisierungsunterschiede der benachbarten Teilbandsigna-
le dahingehend, daß die Aliasing-Verzerrungen unhörbar
10 bleiben. Die Stufe 5.4 steuert hierzu die Ausgangsstufe
5.3, die mit dem Ausgang der Stufe 5.5 verbunden ist.
Auf diese Weise wird in der Ausgangsstufe 5.3 die Ver-
teilung der Quantisierung unter Berücksichtigung der
Mithörschwellenvorgabe durch die Blöcke 5.1 und 5.2, der
15 Vorgabe des verfügbaren Gesamtinformationsflusses durch die
Stufe 5.5 sowie der Aliasing-Verzerrungsvorgabe durch die
Stufe 5.4 festgelegt.

Mithörschwellenkriterium

20

Eine statische Verteilung der Quantisierung der Ausgangs-
signale innerhalb der einzelnen Teilbänder nach dem
Stand der Technik berücksichtigt nur die Maskierung des
innerhalb dieser Teilbänder begrenzten Quantisierungs-
rauschens durch das Nutzsignal in demselben Band. Allein
25 durch die bekannte Berücksichtigung der Maskierung des
Quantisierungsrauschens innerhalb der Teilbänder läßt
gemäß Gleichung (1) bei der gewählten Aufteilung des
breitbandigen Signals 24 Teilbänder eine Reduktion der
Datenrate gegenüber einer 16-bit-linear-PCM-Codierung
30 um etwa 400 kbit/s erreichen. Demgegenüber
ermöglicht die vorliegend vorgesehene dynamische (signal-
abhängige) Steuerung der Quantisierung der Teilbandsignale
einerseits bei gleicher subjektiver Qualität eine weiter-
gehende Reduzierung der Datenrate sowie eine zusätzliche
35

ERSATZBLATT

1

signalspezifische Qualitätsreserve, andererseits bei nicht verringerter Datenrate eine weitergehende Qualitätsreserve.

5

Spektrale Mithörschwellen

10

Die vorliegend vorgesehene Steuerung der Quantisierung erfolgt gemäß der maskierenden Wirkung eines Teilbandsignals mit hohem Signalpegel auf die benachbarten Teilbänder. Diese maskierende Wirkung kann bei Anwendung des Teilbandschemas nach Fig. 8 im wesentlichen im Frequenzbereich oberhalb von 2 kHz genutzt werden, wobei im Bereich bis 8 kHz Teilbänder gleicher absoluter Breite von 500 Hz zugrundegelegt sind. In Abweichung hiervon bezieht sich die Frequenzselektivität des menschlichen Gehörs und damit auch die spektrale Maskierung auf konstante relative Bandbreiten, den sogenannten Frequenzgruppen. Die in dem vorliegenden Teilbandschema nach Fig. 8 gewählten absolut konstanten Bandbreiten entsprechen im Bereich von 2 bis 4 kHz in etwa der Frequenzgruppenbreite und sind im Frequenzbereich oberhalb von 4 kHz wesentlich schmaler als die Frequenzgruppen des Gehörs. In diesem Bereich ist folglich eine stärkere Maskierung benachbarter Teilbandsignale zu erwarten, wie Fig. 8 schematisch zeigt.

15

20

25

30

35

Fig. 4 zeigt für frequenzgruppenbreites Rauschen mit den Mittenfrequenzen 250 Hz, 1 kHz und 4 kHz als maskierender Schall die Mithörschwellen. Der Pegel des maskierenden Rauschens beträgt in allen drei dargestellten Fällen $L = 60$ dB. Die Mithörschwellen, aufgetragen über eine logarithmische Frequenzachse weisen bei den Mittenfrequenzen von 1 kHz und 4 kHz näherungsweise dieselbe Form auf; dagegen ist der Verlauf bei der Mittenfrequenz

1

von 250 Hz deutlich verbreitert. Obwohl der Pegel des
5 Störrauschens konstant gewählt wurde, ist die Differenz
des Maximums der Mithörschwelle nur gestrichelt einge-
tragen (60 dB-Linie bei 250 Hz nur etwa 2 dB, bei 4 kHz
wächst die Differenz jedoch auf 5 dB an). Desweiteren
steigen die Mithörschwellen an der unteren Flanke mit
10 etwa 100 dB/Oktave sehr steil an und fallen nach hohen
Frequenzen hin wesentlich flacher ab. Dies bedeutet, daß
tiefe laute Töne hauptsächlich höhere leise Töne ver-
decken.

15 Die Steilheit der oberen Flanke ist vom Pegel des maskie-
renden Schalles abhängig. Diese Abhängigkeit ist in
Fig. 5 dargestellt. Bei kleinen Pegeln fallen die Mithör-
schwellen zu höheren Frequenzen hin steil ab, während bei
mittleren Pegeln und noch ausgeprägter bei hohen Pegeln
20 der Abfall flacher wird. Dieser Abfall beträgt bei einem
Pegel von 70 dB etwa 40 dB/Oktave. Die Frequenzabhängig-
keit der Mithörschwelle ist also zusätzlich vom Pegel des
Störschalles abhängig.

25 Typischerweise bestehen die zu übertragenden Signale
nicht nur aus einem einzigen Ton, sondern aus einer Viel-
zahl von Harmonischen (z.B. Musikinstrument, stimmhafte
Sprachlaute), oder breitbandigem Rauschen (z.B. Zisch-
laute). Je nach Zusammensetzung der Amplituden
30 dieser Harmonischen sind die von solchen Signalen hervor-
gerufenen Mithörschwellen sehr verschieden. So verur-
sacht eine Trompete mit vielen Harmonischen eine viel
breitbandigere Verdeckung als eine Flöte, deren Ton ein
Spektrum besitzt, das fast nur aus einer einzigen Linie
35 besteht. Fig. 6 zeigt z.B. die Mithörschwellen des

1 Vokals /ə/. Der Pegel der einzelnen Harmonischen ist
 durch die schwarzen Punkte markiert, die resultierende
 Mithörschwelle durch die ausgezogene Linie angedeutet.
 5 Trotz gegenseitiger teilweiser Maskierung sind die ersten
 neun Harmonischen wahrnehmbar, während die 10. und 11.
 Harmonische vor allem durch die 8. Harmonische verdeckt
 sind. Die Harmonischen Nr. 13 bis Nr. 17 sind aufgrund
 ihrer schwachen Pegel durch die relativ starke 12. Har-
 monische maskiert.

10

Die spektrale Verdeckung eines beliebigen Schallsignals
 kann in dem Block 5.1 (Fig. 3) folgendermaßen berechnet
 werden (Einzelheiten zur Berechnung der Mithörschwellen
 sind in der Literaturstelle "Algorithm for Extraction
 15 of Pitch and Pitch Saliency from Complex Tonal Signals";
 Terhardt, E., Stoll, G., Seewann, M.; J. Acoust. Soc.
 Am 71, 1982, pp 679-688, beschrieben).

$$\begin{aligned}
 20 \quad L_{RHS_i}(f_{ui}) &= \\
 &= \left\{ 3,64 (f_{ui}/\text{kHz})^{-0,8} - 6,5 \exp \left[-0,6 (f_{ui}/\text{kHz} \cdot 3,3)^2 \right] \right. \\
 &\quad \left. + 10^{-3} (f_{ui}/\text{kHz})^4 \right\} \quad [\text{dB}] \quad (3)
 \end{aligned}$$

25

Dabei bedeutet L_{RHS_i} die Ruhehörschwelle des i-ten
 30 Teilbandes, welche tabellarisch an den Grenzfrequenzen
 f_{ui} des Teilbandes i vorliegt.

Für die Tonheit z als Maß für die Frequenzgruppe des
 Gehörs gilt:

35

1

$$z = \left\{ 13 \arctan \left[0,76 \left(f/\text{kHz} \right) \right] + 3,5 \arctan \left(f/7,5 \text{ kHz} \right)^2 \right\} \quad [\text{Bark}] \quad (4)$$

10

wobei z ebenfalls tabellarisch vorliegt.

Für die Steilheit S der unteren Flanke der Mithörschwelle gilt:

$$S = 27 \text{ dB/Bark} \quad (5)$$

15

Für die Steilheit S der oberen Flanke der Mithörschwelle gilt:

$$s = \left[-24 - (0,23 \text{ kHz}/f_{oi}) + (0,2 \cdot L_i/\text{dB}) \right] [\text{dB/Bark}] \quad (6)$$

20

wobei f_{oi} die obere Grenzfrequenz und L_i der Signalpegel des betreffenden Teilbandes i bedeuten.

25

Für die Erregungsverteilung innerhalb des Teilbandes i , welche die Grundlage für die spektrale Verdeckung darstellt, gilt:

$$L_{Ek}(f_{ui}) = L_k - S(z_k - z_i) \quad (7)$$

30

Die Gleichung (7) beschreibt die Verdeckung eines Teilbandes k auf ein Teilband i . Die gesamte gegenseitige Maskierung aus den verschiedenen Teilbandsignalen ergibt sich aus der Summation der Amplituden der Erregungsverteilung innerhalb der einzelnen Teilbänder. Die Maskierung des Signalpegels L_i innerhalb des

35

1

Teilbandes i aufgrund der Signale in allen anderen
23 Teilbändern errechnet sich zu:

5

$$LX_i = L_i - 10 \lg \left[\left(\sum_{\substack{k=1 \\ i \neq k}}^{24} 10^{L_{Ek}(f_{ui})/20\text{dB}} \right)^2 + 10^{L_{TH}(f_{ui})/10\text{dB}} \right] + 6 \text{ dB} \quad (8)$$

10

Ein Teilbandsignal mit dem Signalpegel L_i ist dann vollständig verdeckt, wenn $LX_i < 0\text{dB}$ ist. Eine teilweise Maskierung findet statt, wenn

15

$$L_i > LX_i > 0 \text{ dB} \quad (9)$$

ist.

Zeitliche Mithörschwellen

20

Die Verteilung der Quantisierung auf die verschiedenen Teilbänder erfolgt nicht nur in Abhängigkeit der spektralen, sondern auch der zeitlichen Verdeckung. Drei Zeitbereiche der Verdeckung, die in Fig. 7 dargestellt sind, können unterschieden werden. Die Vorverdeckung findet in einem Zeitbereich statt, bevor der Maskierer eingeschaltet ist. Wenn Maskierer und Testschall zur gleichen Zeit eingeschaltet sind, spricht man von Simultanverdeckung. Nach Abschalten des Maskierers findet eine Nachverdeckung statt.

30

Die typische Dauer der Vorverdeckung liegt im Bereich von 10 ms bis 20 ms. Pegelanstiege müssen in dieser relativ kurzen Zeit voll rekonstruiert werden.

35

1

Die Simultanverdeckung ist bei kurzen Testschallimpulsen von der Impulsdauer T abhängig; bei längeren Impulsdauern ist die Mithörschwelle unabhängig von der Dauer. Bei Verkürzung der Impulsdauer ($T < 200$ ms) steigt die Mithörschwelle mit einer Steigung von $- 10$ dB/Dekade an.

10

Die Nachverdeckung erstreckt sich bis etwa 200 ms nach Ausschalten des Maskierers. Die Nachhörschwellen halten innerhalb der ersten 5 ms nach Abschalten den Wert der Simultanhörschwelle und erreichen nach 200 ms den Wert der Ruhehörschwelle. Der Effekt der Nachverdeckung spielt aufgrund seiner Dauer eine viel wesentlichere Rolle als der Effekt der Vorverdeckung. Die Nachverdeckung ist außerdem von der Einschaltdauer des Maskierers abhängig. Bei sehr kurzer Dauer des Maskierers ($T < 5$ ms) fällt die Nachhörschwelle bereits nach 20 ms auf den Wert der Ruhehörschwelle ab.

15

20

Anwendung der Mithörschwellen

25

Die spektralen und zeitlichen Verdeckungseigenschaften werden in dem senderseitigen Transcoder 2 gemäß Fig. 1a genutzt, um die Signale derjenigen Teilbänder, welche durch benachbarte Teilbandsignale nicht oder nur wenig verdeckt werden, feiner zu quantisieren, als solche Signale, welche sehr stark verdeckt sind und somit auch kaum wahrnehmbar sind. Signale in Teilbändern, die vollkommen verdeckt sind, d.h. unterhalb der resultierenden Mithörschwelle liegen, brauchen nicht übertragen zu werden. Das Signal des betreffenden Teilbandes kann zu Null gesetzt werden. Um das transcodierte Signal empfangs-

35

1

5

10

15

seitig in der Stufe 12 decodieren zu können (Fig. 2a) muß die in der Stufe 5 erzeugte Steuerinformation an der Stufe 12 vorliegen. Um eine Mitübertragung dieser Steuerinformation zu vermeiden, werden als Eingangsinformation der Stufe 5 die Skalenfaktoren die einzelnen Teilbänder benutzt, welche ohnehin als Nebeninformation im Multiplexsignal übertragen werden. Damit ist für den empfangsseitigen inversen Transcoder 12 allein aufgrund der Kenntnisse über die Skalenfaktoren der einzelnen Teilbänder und unter Einbeziehung der gleichen Kriterien für die resultierende Mithörschwelle wie in der senderseitigen Stufe 5 eine genaue Zuweisung der Rückquantisierung in die Ebene mit linearer Quantisierung (z.B. 16 bis 18 bit Auflösung pro Abtastwert) möglich.

Berücksichtigung der Aliasing-Verzerrungen

20

25

30

35

Unabhängig von der gegenseitigen Verdeckung benachbarter Teilbänder muß zur Verteilung der Quantisierung innerhalb der Teilbänder auch die Wahrnehmbarkeit von Aliasing-Verzerrungen berücksichtigt werden. Diese Verzerrungen entstehen durch die nicht ideale Bandpaßfilterung der QMF-Filterbank 1, deren Ausgangssignal gerade mit der für ein ideales Filter minimal möglichen Abtastfrequenz der zweifachen Bandbreite abgetastet wird. Die in den Durchlaßbereich eines Filters der Filterbank 1 gefalteten Aliasing-Komponenten werden jedoch aufgrund der QMF-Filterstruktur der Bank 1 nur bei gleicher Auflösung des Signals in benachbarten Bandpässen praktisch vollständig eliminiert. In dem gewählten Teilbandschema gemäß Fig. 8 sind Aliasing-Verzerrungen bezüglich ihrer Wahrnehmbarkeit besonders im Bereich der

1

unteren sechs Teilbänder kritisch, da die Bandbreiten
der verwendeten Quadrature-Mirror-Filter größer als
die Breite einer Frequenzgruppe sind (vgl. Fig. 8).
Im Bereich hoher Frequenzen sind die Aliasing-Ver-
zerrungen weit unkritischer, da sie aufgrund der wesent-
lich geringeren Breite der betreffenden Teilband-
Quadrature-Mirror-Filter gegenüber den Frequenzgruppen
eine stärkere Maskierung durch das Nutzsignal erfahren.

10

Um die Hörbarkeit von Aliasing-Verzerrungen zu vermei-
den, müssen die Höhen der Quantisierungsstufen in den
unteren Teilbändern zwischen benachbarten Bändern mög-
lichst ähnlich gewählt werden, während bei den oberen
Teilbändern größere Unterschiede in den Höhen der
Quantisierungsstufen zwischen benachbarten Teilbändern
auftreten dürfen.

20

25

30

35

1

5

Pufferspeicher

10

Für die zusätzliche Einsparung am Informationsfluß werden sende- und empfangsseitig in den Transkodierern 2 und 12 Pufferspeicher für die einzelnen Teilbandsignale benötigt, wie später anhand von Fig. 11 noch näher erläutert wird. Derartige Pufferspeicher ermöglichen es, die Signale in etwa der Zeitdauer, wie sie den Nachhörschwellen des menschlichen Gehörs entspricht, zu verzögern.

15

20

Eine Verzögerung des Tonsignals um einen Wert zwischen 200 ms und 500 ms ist einerseits den zeitlichen Mithörschwellen des Gehörs optimal angepaßt und andererseits für die Anwendung in der Praxis gut zu vertreten. Diese Verzögerung des Signals ist notwendig, um eine effiziente Verteilung des Informationsflusses nicht nur auf die verschiedenen Teilbandsignale, sondern auch auf die zeitlichen Blöcke, für welche jeweils ein Skalenfaktor bestimmt wurde, zu gewährleisten. Die innerhalb des Zeitfensters variable Aufteilung des Informationsflusses gewinnt vor allem bei der Berücksichtigung der zeitlichen Mithörschwellen an Bedeutung. So können schnelle Pegelanstiege der Teilbandsignale sehr genau, schnelle Pegelabsenkungen, welche durch die relativ langsam abfallenden Nachhörschwellen eine starke Maskierung erfahren, hinreichend genau transkodiert werden. Die

25

30

35

1

zeitliche Verteilung der Quantisierung erfolgt hierbei
ebenso wie die spektrale Verteilung auf die einzelnen
Teilbänder nach Maßgabe des Mithörschwellenkriteriums,
wobei in Stufe 5.5 gemäß Fig. 3 berücksichtigt wird,
daß ein minimaler "mask-to-noise" Abstand gehalten wird.
Dies kann beispielsweise dadurch erreicht werden, indem
vom Pufferspeicher ein Steuersignal mit Information über
den Ladezustand des Pufferspeichers an Stufe 5.5 über-
geben wird.

15

20

25

30

35

1

5

Fehlersicherung der transkodierten Teilbandsignale

Die in der Stufe 7 (Fig. 1a) erfolgende Sicherung der Abtastwerte gegenüber Übertragungsfehlern kann aus folgenden Gründen stark vereinfacht werden:

10

Die Störwirkungen von fehlerhaft empfangenen Abtastwerten erstrecken sich nicht auf die gesamte Audio-Bandbreite, sondern sind auf die Breite des zugehörigen Teilbandes beschränkt. Die maximale Amplitude der Störung in dem betreffenden Teilband ist durch die Übertragung eines Skalenfaktors in ihrer Höhe beschränkt. Dies bedeutet, daß die spektrale Verteilung fehlerhaft empfangener Abtastwerte sich der spektralen Grobstruktur des Nutzsignals annähert und somit eine größtmögliche Verdeckung bei Störungen durch Übertragungsfehler von Abtastwerten resultiert.

15

20

25

30

Da bei bandpassbegrenztem Rauschen mit konstantem Pegel eine Lautstärkeabhängigkeit von der Bandbreite besteht, ist es zweckmäßig, Störungen in Teilbändern mit Bandbreiten, die größer sind als die Frequenzgruppen des Gehörs, eher zu vermeiden als Störungen in schmäleren Teilbändern. Ein Rauschen mit einer Bandbreite über 2 Frequenzgruppen erzeugt bei mittlerem Schalldruckpegel die gleiche Lautstärkeempfindung wie ein im Pegel um 3 dB angehobenes Rauschen, jedoch nur

35

1

5 frequenzgruppenbreit (vgl. Fig. 9). Damit sich Störungen lautstärkemäßig minimal auswirken, sollte das Stör-
spektrum auf die Breite einer Frequenzgruppe reduziert
werden. Da die Bandbreite in dem vorliegenden Teilband-
10 schema (Fig. 8) in den untersten 5 Teilbändern (das
1. Teilband erstreckt sich über insgesamt 5 Frequenz-
gruppen) die Frequenzgruppenbreite z.T. wesentlich
überschreitet, müssen die Abtastwerte der untersten
2 Teilbänder stark und die der 3. bis 5. Teilbänder
hinreichend gesichert werden.

15

Quadrature-Mirror-Filterung in Stufen 1 und 11

Die in den Fig. 1a und 2a dargestellten
20 sende- und empfangsseitigen Verarbeitungsschritte sind
anhand des Funktionsschemas gemäß Fig. 10 veranschau-
licht.

Die mit dem digitalen Tonsignal gespeiste QMF-Filter-
bank 1 besteht aus einer Kaskade von Mirrorfiltern MF,
25 welche sukzessive das Tonsignalspektrum aufteilen, im
betrachteten Beispielsfalle in 24 Teilbandsignale.
Jedes Mirrorfilter MF stellt einen sog. "Finite-Impulse-
Response" (FIR)-Filter dar, welcher ein digitales
Eingangssignal in zwei/ spiegelsymmetrisch zur Grenz-
30 frequenz liegende Teilbänder zerlegt. Hierzu wird die

35

1

Ersatz-
Grenzfrequenz des/tiefpasses so gewählt, daß sie der
5 halben Bandbreite des Eingangssignals entspricht.

Die
Flankensteilheit der Tiefpaßcharakteristik ist propor-
tional der Anzahl der Koeffizienten des FIR-Tiefpasses.
10 Die Berechnung der Koeffizienten ist in der Literatur-
stelle "Multirate Digital Signal Processing" von
Chrochiere und Rabiner, Prentice-Hall-Verlag, Englewood
Cliffs, N.J., U.S.A., beschrieben. Jedes Mirrorfilter
MF teilt sein Eingangssignal in zwei gleichgroße Teil-
15 bänder und halbiert dabei dessen Abtastrate, so daß die
das Filter durchlaufende Informationsmenge prinzipiell
unverbändert bleibt. Lediglich die Wortlänge der zwi-
schen dem kaskadierten Filtern MF übergebenen Abtast-
werte ist mit 20 bis 23 bit etwas größer als die vom
20 Eingangssignal mit üblicherweise 16 bis 18 bit, um
Rundungsfehler infolge der Kaskadierung zu eliminieren.
Eine Frequenzaufteilung in 24 Bänder erfolgt bei einer
Abtastrate von 32 kHz durch Teilung des Bereiches
0 bis 8 kHz in 16 Bänder zu je 500 Hz und von 8 bis
25 16 kHz in 8 Bänder zu je 1 kHz (vgl. Fig. 8). Hierfür
sind 5 bzw. 4 kaskadierte Filter MF nötig. Die unver-
meidbaren Alias-Verzerrungen als Folge der nicht unend-
lich steilen Flanken der Ersatz-Tiefpaß-Filtercharak-
teristik der Mirrorfilter MF kompensieren sich bei der
30 Rückfilterung, falls die Quantisierung in den benach-
barten Teilbändern gleich ist.

Durch die Kaskadierung wird die Qualität der Filterbank
stets optimal ausgenutzt, da bei gleicher Koeffizienten-
35 zahl die zeitliche Länge des FIR-Tiefpasses und somit

1

die Steilheit pro Hertz in späteren Kaskadenstufen
zunimmt. Für die Filterung sind 64 Koeffizienten je
Filter ausreichend, denn es hat sich gezeigt, daß die
Steilheiten der Filterflanken nicht größer sein müssen
als die maximale Steilheit der Mithörschwellenkurve.

10

Wird in jeder Filterstufe der gleiche Koeffizientensatz
verwendet, so wird die Filtersteilheit umso größer, je
geringer die Abtastrate durch vorangegangene Spektral-
teilungen geworden ist. In den letzten Filterkaskaden ist
diese Steilheit nicht unbedingt nötig, da sie nicht an
allen Bandgrenzen erreicht wird. Verwendet man daher in
den letzten Filterkaskaden Filter mit weniger Koeffizien-
ten, so verringert sich in erster Linie die Signallaufzeit,
was bei einer Echtzeitanwendung, wie z.B. bei einer Rund-
funkübertragung von Bedeutung ist. Um diese für das Gehör
erforderliche Flankensteilheit zu erhalten, sind 16 Filter-
koeffizienten in der letzten Filterkaskade ausreichend.

20

25

Bei realen Filtern ist eine Restwelligkeit vorhanden, die
sich zwar nicht spürbar im Frequenzgang auswirkt (Rest-
welligkeit $\leq 0.002\text{dB}$), jedoch bei zeitlich sehr kurzen
Impulsen zu einem noch hörbaren "Filterklingeln" führen
kann. Dies sind Nebenimpulse, die 10 ms bis 100 ms vor und
nach dem Hauptimpuls mit einem Pegel von ca. -80 dB
gegenüber dem Nutzsinal entstehen. Um diese u.U. hörbaren
Signalfehler zu vermeiden, wird das Signal vor der
eigentlichen Codierung einer Hin- und Rückfilterung, nicht
jedoch einer Datenreduktion unterzogen. Das durch diese
Hin- und Rückfilterung entstandene Fehlersignal wird
extrahiert und dem zu codierenden Originalsignal invers
hinzugefügt. Die störenden Eigenschaften der Filter im
Coder und Decoder sind somit ausreichend unterdrückt.

35

1

5 Transkodierung und Bestimmung der Skalenfaktoren in
Stufen 2 und 4

10 In der Transkodierungsstufe 2 erfolgt eine Reduktion
der Auflösung der ankommenden Teilbandsignale von
16 (18) bit bis zu 1,5 bit pro Abtastwert. Dies be-
deutet eine Herabsetzung der Stufenzahl der Abtast-
werte in jedem Teilband. Hierzu wird pro Teilband zu-
nächst aus einer bestimmten Anzahl zeitlich aufeinander-
folgender Abtastwerte (= Block) der Betrag des
15 maximalen Abtastwertes gesucht, welcher dann einem
Klassierungsschema zugeordnet wird. Dieses Schema be-
steht aus 64 Klassen entsprechend 6 bit, welche den
Dynamikumfang einer 16 bit-PCM (96 dB) in $96/64$ dB =
1,5 dB pro Stufe unterteilen. Die dem maximalen Abtast-
wert entsprechende Stufennummer stellt den Skalenfaktor
20 der betrachteten Folge von Abtastwerten dar. Der in
Stufe 4 ermittelte Skalenfaktor wird der Transkodierungs-
stufe 2 zugeführt, welcher ferner von der Stufe 5 eine
Information über die Anzahl der je Block erforderlichen
25 Quantisierungsstufen (aufgrund des dort angewendeten
Mithörschwellenkriteriums) sowie von der Stufe 6 eine
weitere Information über die gewünschte Qualität der
Quellkodierung zugeführt werden.

30 Wie aus Fig. 11 im einzelnen hervorgeht, weist die
Stufe 2 einen regelbaren Verstärker 2.1 für die linear
quantisierten Teilbandsignale auf, welcher vom Skalen-
faktor aus Stufe 4 gesteuert wird. Die verstärkten

35

1

Teilbandsignale durchlaufen ein steuerbares Verzögerungs-
glied 2.2, das für die Berücksichtigung der zeitlichen
Mithörschwellen notwendig ist (vgl. Abschnitt
"Pufferspeicher"). Die verstärkten und verzögerten Teil-
bandsignale werden in dem Quantisierer 2.3 unter der
Steuerung der Stufen 5 und 6 unquantisiert.

10

Der durch den Skalenfaktor repräsentierte Bereich in
positiver und negativer Richtung wird durch die jeweils
vorgegebene Anzahl von Quantisierungsstufen unterteilt.
Dies bewirkt, daß die Stufenhöhe bei kleinen Skalenfak-
toren entsprechend kleinen Wertebereichen geringer und
damit die Auflösung höher ist als umgekehrt bei großen
Skalenfaktoren entsprechend großen Wertebereichen.
Durch die Unterteilung des breitbandigen Nutzsignals in
Teilbandsignale können deshalb Spektralanteile mit
kleinem Pegel individuell höher aufgelöst werden, als
bei einer Ingesamt-Quantisierung des breitbandigen
Nutzsignals, bei der solche kleinpegeligen Spektralan-
teile mit begrenzter gleicher Genauigkeit aufgelöst wer-
den wie Spektralanteile mit hohem Pegel.

25

Um einerseits Pegelanstiege wegen der relativ kurzen
Vorverdeckung des menschlichen Gehörs folgen zu können,
andererseits den Skalenfaktor nicht zu häufig übertra-
gen zu müssen, wird der Skalenfaktor z.B. zwar für je
4 Abtastwerte gebildet, aber nur jeder zweite Skalen-
faktor übertragen. Um die nicht-übertragenen Skalen-
faktoren empfängerseitig rekonstruieren zu können,
wird für die den weggelassenen Skalenfaktoren zugeord-

35

1

neten 4er-Blöcke ein Informationsbit übertragen, welches die Gültigkeit des vorangegangenen oder des folgenden Skalenfaktors für die Skalierung des betreffenden 4er-Blocks angibt. Diese Verhältnisse sind schematisiert anhand der Fig. 12 und 13 erläutert.

10 In Fig. 12 ist die Skalenfaktorbildung für einen Pegelverlauf mit Blöcken zu je 8 Abtastwerten dargestellt; Fig. 13 zeigt den gleichen Pegelverlauf mit Blöcken zu je 4 Abtastwerten, wobei nur jeder zweite Skalenfaktor zu 6 bit übertragen und ein zusätzliches Informationsbit aufgewendet wird, so daß statt 6 bit/8 Abtastwerte nunmehr 7 bit/8 Abtastwerte übertragen werden. Wie Fig. 13 zeigt, wird der zwischen den Skalenfaktoren S_n und S_{n+1} befindliche Pegelanstieg zeitlich exakter erfaßt als im Falle von Fig. 12. Das in Fig. 12 zwangsläufig stärkere Quantisierungsrauschen (großer Wert S_n) für die ersten 4 zum Skalenfaktor S_n zugehörigen Abtastwerte fällt wegen des in Fig. 13 niedrigeren Skalenfaktors mit entsprechend feineren Quantisierungsschritten geringer aus. Auch bei zeitlich sehr kurzen Pegelspitzen (innerhalb eines Zeitraums von weniger als 4 Abtastwerten) wird der Pegelanstieg mit Hilfe des zusätzlichen Informationsbits (Fig. 13) zeitlich besser angenähert, wobei es auf den zugehörigen Pegelabfall wegen der wesentlich längeren Nachdeckung des menschlichen Gehörs weniger ankommt und damit auch längere "Rauschfahnen" toleriert werden können.

Skalenfaktor und zusätzliche Informationsbits werden in jedem Teilband wie folgt erzeugt, wobei zum besseren Verständnis angenommen sei, daß hier nur die

35

1

Skalenfaktoren mit geradzahligem Index S_{2n} , S_{2n+2} ... übertragen werden.

5

$S_{2n} < S_{2n+1} < S_{2n+2}$ nächstliegender
Pegelwert

$S_{2n} < S_{2n+1} > S_{2n+2}$ S_{2n+2} aus S_{2n+1}

10

$S_{2n} > S_{2n+1} > S_{2n+2}$ nächstliegender
Pegelwert

$S_{2n} > S_{2n+1} < S_{2n+2}$ S_{2n+1} aus S_{2n}

15

Um das starre Klassierungsschema für den Skalenfaktor in 1,5 dB-Stufen zu verfeinern, ist die Übertragung einer Zusatzinformation vorgesehen, die den durch diese Klassierung entstehenden Fehler minimiert. Vorgesehen ist für die Frequenzbereiche von 1,5 bis 3,5 kHz sowie von 3,5 bis 8 kHz alle 8 Abtastwerte je eine 3 bit-Zusatzinformation, welche die gesamte Abweichung vom 1,5 dB-Raster in allen auf diese beiden Frequenzbereiche entfallenden Teilbänder gemäß Fig.8 in $(1/8) \cdot 1,5$ dB-Schritten angibt. Dies kommt z.B. Einzeltönen zugute, deren Pegeländerungen mit etwa 0,2 dB Genauigkeit übertragen werden können.

20

25

30

35

Die Wahl dieser Feinraasterung soll nach Maßgabe eines minimalen quadratischen Fehlers erfolgen, so daß nur die stark modulierten Teilbänder die Feinraasterung bestimmen. Die Zusatzinformation gilt also nicht für alle Teilbänder, was folgenden Grund hat: in dem gewählten Teilbandschema gemäß Fig. 8 ist die Bandbreite der unteren drei Teilbänder bis 1,5 kHz wesentlich größer als die der Frequenzgruppen des menschlichen

1

5 Gehörs, so daß die Verdeckung innerhalb dieser drei Teilbänder relativ gering ist. Es ist daher für diese drei Teilbänder eine relativ hohe Auflösung erforderlich, beispielsweise von 10,6 bit/Abtastwert für das Teilband von 0,5 bis 1 kHz. Unterhalb von 1 kHz werden die Nachteile der Klassierung für den Skalenfaktor wegen dieser
10 relativ hohen Auflösung für diese 3 Teilbänder bedeutungslos. Oberhalb von 8 kHz wird die 1,5 dB-Klassierung als ausreichend empfunden.

15 Eine weitere Eigenschaft der Transkodierungsstufe 2 besteht darin, daß ein Skalenfaktor der Größe Null so verarbeitet wird, daß für alle reduzierten Abtastwert-Angaben der betreffende Block des betreffenden Teilbandsignals nicht übertragen wird. Das gleiche gilt, wenn die Stufe 5 in einem Teilband feststellt,
20 daß der Skalenfaktor unterhalb der betreffenden Mithörschwelle liegt (gegenseitige Maskierung), worauf die Stufe 5 an die Stufe 4 den Steuerbefehl "setze den Skalenfaktor auf Null" gibt.

25 Ferner kann in der Stufe 4 für jedes Teilband eine zuschaltbare Rauschsperrung vorgesehen werden, um Leerkanalrauschen, Ungenauigkeiten bei der A/D-Wandlung und dgl. zu unterdrücken. Hierbei ist es denkbar, einstellbare Schwellwerte vorzusehen.

30

Bei der Quantisierung der Stufe 2 muß eine ungerade Anzahl von Quantisierungsstufen vorgesehen werden, um zu verhindern, daß Werte nahe Null in einen ständigen Wechsel zwischen zwei Quantisierungsstufen überführt werden. Ansonsten könnten Anteile entstehen, die weit
35 über dem Signalpegel des Teilbandes liegen. Nicht alle

1

stufenzahlen können direkt in digitale Werte umgewandelt werden, ohne daß man darstellbare Werte und somit Übertragungs-Kapazität verschenkt. Z.B. müßte man für die Darstellung von 3 möglichen Stufen 2 bit an Information aufwenden; die 4. Kombinationsmöglichkeit dieser 2 bit bliebe jedoch unbenutzt und würde hier 33% mehr Aufwand an Informationsmenge bedeuten.

10

Derartige Verluste an Übertragungskapazität lassen sich minimieren, indem man mehrere Abtastwerte zusammen in ein Datenwort kodiert. Beispielsweise ergeben 5 Abtastwerte mit je 3 Stufen $243 = 3^5$ Kombinationsmöglichkeiten, die mit 8 bit, also 256 Zuständen bei geringer Irrelevanz und wenig Kodierungsaufwand übertragen werden können.

15

20

Zusätzliche Spektralanalyse zur Gewinnung der Steuerinformation

Ein weiteres Ausführungsbeispiel des erfindungsgemäßen Verfahrens ist in den Fig. 1d (Sendeseite) und Fig. 2d (Empfängerseite) dargestellt. Ein wesentlicher Teil der Steuerinformation für die Quantisierung der Abtastwerte wird aus dem Spektralverlauf des Tonsignals gewonnen. Um diese Information zu gewinnen, wird sendeseitig eine Fourierzerlegung des Tonsignals durchgeführt, die z.B. durch eine Fast-Fourier-Transformation realisiert werden kann. Durch diese Fourier-Transformation ergeben sich im wesentlichen die folgenden Vorteile:

25

30

35

1

- genauere Spektraldarstellung gegenüber einer Bandpaß-
aufteilung des Tonsignals.

5

- Die tonalen Anteile können von den mehr rauschhaften
Bestandteilen eines Tonsignals unterschieden werden.

10

- Die genauere Spektralanalyse und Unterscheidung in
tonale und rauschhafte Bestandteile ermöglicht eine
effektivere Bestimmung der Mithörschwellen gegenüber
dem Fall, daß die Steuergrößen nur aufgrund einer be-
grenzten Anzahl von Teilbandsignalen gewonnen werden.

15

- Reduzierung des Hardwareaufwandes auf der Empfangsseite
bei gleichzeitig kürzerer Verzögerungszeit der Ton-
signalübertragung und gleicher Übertragungsrate des
Multiplexsignals wie ohne diese zusätzliche Maßnahme.

20

Die Reduzierung des Hardwareaufwandes besteht darin, daß
weniger Bandpässe verwendet werden können. Die geringere
Anzahl an Bandpässen bedeutet in erster Linie einen
geringeren Prozessoraufwand, was sich besonders bei der
Entwicklung eines kostengünstigen Empfängers positiv
auswirkt. Breitere Teilbänder ermöglichen eine kürzere
Gesamtverzögerungszeit des Systems. In den Fign. 1d und
2d ist der Reduzierung des Hardwareaufwandes insofern
Rechnung getragen, daß statt der 24 Teilbänder in der
vorigen Ausführungsform nun nur noch 16 Teilbänder gegen-
überstehen. Der durch die breiteren Teilbänder grundsätz-
lich höhere Datenfluß der transkodierten Teilbandsignale
kann durch die genauere Bestimmung der Mithörschwellen in
etwa wieder kompensiert werden.

25

30

35

ERSATZBLATT

1
5
10
15
20
25
30
35

Eine ausreichende Spektralauflösung im Bereich tiefer Frequenzen des Tonsignals wäre durch eine spektrale Darstellung benachbarter Stützwerte von etwa $\Delta f = 10$ Hz gegeben. Da der entscheidende Vorteil der genauen Spektralbestimmung für die Ermittlung der Mithörschwellen im unteren Teil des Frequenzbereiches gegeben ist (22 der insgesamt 24 Frequenzgruppen des Gehörs liegen im Frequenzbereich bis 10 kHz), genügt eine Spektralanalyse bis etwa 10 kHz. Bei einer Spektraldarstellung des Amplitudenspektrums mit 512 reellen Stützwerten ergibt sich damit ein Abstand benachbarter Stützwerte von $\Delta f = 20$ Hz. Mittels geeigneter Interpolationsalgorithmen ist für diskrete tonale Komponenten eine genauere Frequenzbestimmung aufgrund der Auswertung benachbarter Frequenzstützpunkte möglich.

Die Eingangsgrößen des Blocks dynamische Verteilung des Informationsflusses (sh. Fig. 1d) bestehen zusätzlich zu den Steuergrößen aus der FFT-Analyse auch aus Steuergrößen für die Festlegung der Qualität und Kanalzahl, sowie aus den Skalenfaktoren der einzelnen Teilbänder. Durch einen Vergleich der Steuergrößen aufgrund der FFT-Analyse und den Skalenfaktoren können die Aliasing-Verzerrungen, die sich durch die unterschiedliche Quantisierung in den Teilbandsignalen empfangsseitig nicht mehr vollständig kompensieren, bei der dynamischen Verteilung des Informationsflusses ausreichend genau berücksichtigt werden.

ERSATZBLATT

1

Stufenweise Datenreduktion

5

10

15

In verschiedenen Anwendungsfällen des erfindungsgemäßen Verfahrens ist eine stufenweise Datenreduktion der Tonsignale vorteilhaft. Für die Übertragung hochwertiger Tonsignale zwischen zwei Studios beispielsweise muß die Datenreduktion so erfolgen, daß eine ausreichende "mask-to-noise"-Reserve gewährleistet ist, um Nachbearbeitungen ohne Qualitätseinbuße zu ermöglichen (Studioqualität, vgl. Abschnitte "Nutzung des schwankenden Informationsflusses: Erhöhter Rauschabstand" sowie "Qualitätsabstufung"). Für die weitere Verteilung und/oder Speicherung ist ein erhöhter Rauschabstand nicht erforderlich, so daß die gemäß Studioqualität codierten Tonsignale einer weitergehenden Datenreduktion unterworfen werden dürfen.

20

25

Eine vorteilhafte Ausbildung des erfindungsgemäßen Verfahrens liegt darin, daß die verschiedenen Qualitätsabstufungen der erfindungsgemäß codierten Tonsignale "abwärtskompatibel" sind, d.h., daß beispielsweise ein 192 kbit/s-Multiplexsignal mit Hilfe eines speziellen Transcoders in ein 128 kbit/s-Multiplexsignal umgesetzt werden kann, wobei die 192 kbit/s-Codierung einen erhöhten Rauschabstand (Nachbearbeitungsfähigkeit) gewährleistet und die 128 kbit/s-Codierung einen kleineren Rauschabstand und einen höheren Fehlerschutzgrad vorsieht.

30

35

Fig. 1e zeigt ein Ausführungsbeispiel eines derartigen Transcoders. Er besteht aus den Stufen 12,13,15,17,18 des Decoders gemäß Fig. 2a sowie aus den Stufen 2,3,5,7,8 des Coders gemäß Fig. 1a. Er ist dadurch gekennzeichnet, daß das 192 kbit/s-codierte Tonsignal nicht vollständig

1

zurückgewonnen und neu codiert wird, sondern daß die Nebeninformationen (Skalenfaktoren), welche aus dem Originalsignal gewonnen sind, für die neue Codierung im rechten Teil des Transcoders (Stufen 2,3,5,7,8) zugrunde gelegt werden. Dadurch wird eine Verminderung des Rauschabstandes durch Kaskadierung vermieden. Da die Rückfilterung und Hinfilterung (Stufen 11 und 1 in Fig. 2a bzw. 1a) entfallen, ist die durch den Transcoder gemäß Fig. 1e verursachte Verzögerungszeit klein (etwa 4 ms).

15

20

25

30

35

54

1

5

VERFAHREN ZUM ÜBERTRAGEN DIGITALISIERTER
TONSIGNALE

10

PATENTANSPRÜCHE

15

1. Verfahren zum Übertragen oder Speichern digitalisierter Tonsignale, bei dem sende- bzw. produktionsseitig

20

a) das Tonsignal durch eine Vielzahl spektraler Teilbandsignale digital dargestellt wird, wobei für jedes Teilbandsignal zeitdiskrete, quantisierte Abtastwerte vorliegen;

25

b) die Quantisierung der Abtastwerte in den einzelnen Teilbändern nach Maßgabe der jeweiligen Mithörschwellen des menschlichen Gehörs im Sinne einer Informationsreduktion geändert (codiert) wird, wobei die Höhe des für die Übertragung oder Speicherung aller codierter Teilbandsignale erforderlichen Gesamtinformationsflusses in Abhängigkeit der spektralen und zeitlichen Struktur des Tonsignals schwankt, und

35

1

c) die codierten Teilbandsignale übertragen oder gespeichert werden, und bei dem wiedergabeseitig

5

d) die codierten Teilbandsignale decodiert werden, und

10

e) die decodierten Teilbandsignale zu einem breitbandigen digitalen Tonsignal zusammengefügt werden,

15

dadurch gekennzeichnet, daß nach Maßgabe der Pegelwerte jedes Teilbandsignals oder davon abgeleiteter Pegelinformationen die Quantisierung der Abtastwerte in den Teilbändern dahingehend gesteuert wird, daß die Quantisierungsrauschpegel der einzelnen Teilbandsignale näherungsweise gleiche Abstände zu der aus einzelnen Teilbandsignalen resultierenden Mithörschwelle aufweisen, und daß die Abstände der Quantisierungsrauschpegel der Teilbandsignale zu der resultierenden Mithörschwelle durch die Differenz zwischen dem erforderlichen Gesamtinformationsfluß und dem für die Codierung verfügbaren Gesamtinformationsfluß festgelegt sind.

20

25

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Höhe des verfügbaren Gesamtinformationsflusses variabel ist, ggf. auch etwa gleich dem erforderlichen Gesamtinformationsfluß ist, und wiedergabeseitig eine entsprechende Anpassung der Decodierung erfolgt.

30

35

1

3. Verfahren nach Anspruch 1 und 2, dadurch gekennzeichnet, daß die Quantisierungsrauschpegel der Teilbandsignale unterhalb oder oberhalb der resultierenden Mithörschwelle liegen.

5

10

4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß Teilbandsignale, welche die Klangfarbe des Tonsignals nur geringfügig beeinflussen und deren Pegelwerte oberhalb der resultierenden Mithörschwelle liegen, zu Null gesetzt werden, falls die Quantisierungsrauschpegel der Teilbandsignale oberhalb der resultierenden Mithörschwelle liegen.

15

20

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß nur soviele Teilbänder zu Null gesetzt werden, wie erforderlich ist, um den Quantisierungsrauschpegel der Teilbandsignale soweit zu reduzieren, daß er gerade unterhalb der resultierenden Mithörschwelle liegt.

25

6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß diejenigen Teilbänder zu Null gesetzt werden, welche die größere absolute Bandbreite aufweisen.

30

7. Verfahren nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß die Höhe des Informationsflusses eines codierten Nutzsignals stufenweise etwa um den ganzzahligen Faktor n reduziert wird und stattdessen $n-1$ zusätzliche Nutzsignale übertragen werden.

35

1

8. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die resultierende Mithörschwelle nach Maßgabe der Pegelwerte in jedem Teilband unter Berücksichtigung der Gesetzmäßigkeiten für die Vor-, Simultan- und Nachverdeckung gewonnen wird.

5

10

9. Verfahren nach Anspruch 8, dadurch gekennzeichnet, daß die Berücksichtigung der zeitlichen Mithörschwellen nach Maßgabe der innerhalb eines Zeitfensters von etwa 500 ms vorliegenden Teilband-Pegelwerte erfolgt.

15

10. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß Teilbandsignale, deren Pegelwerte unterhalb der resultierenden Mithörschwelle liegen, nicht oder nur mit geringer Auflösung codiert werden.

20

11. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Höhe der Quantisierungsstufen der Teilbandsignale so bemessen ist, daß die bei der Aufteilung des digitalisierten Tonsignals in Teilbandsignale entstehenden Aliasing-Verzerrungen unterhalb der Mithörschwelle des menschlichen Gehörs liegen.

25

12. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß

30

- sendeseitig für jedes digitalisierte Teilbandsignal ein Skalenfaktor bestimmt wird, welcher den Spitzenwert des Teilbandsignalpegels innerhalb eines bestimmten Zeitintervalls klassiert;

35

1

- bei der Codierung jedes Teilbandsignals dessen Auflösung nach Maßgabe des ermittelten Skalenfaktors festgelegt wird;

5

- die ermittelten Skalenfaktoren zusammen mit den codierten Teilbandsignalen übertragen werden, und

10

- die übertragenen Skalenfaktoren bei der Decodierung der empfangenen codierten Teilbandsignale zur Rekonstruktion des ursprünglichen Teilbandsignals verwendet werden.

15

13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, daß die Anzahl der Klassierungsstufen bei der Skalenfaktorbestimmung so gewählt ist, daß die Wahrnehmbarkeitsschwelle für Pegelsprünge innerhalb des zugeordneten Teilbandes unterschritten ist.

20

14. Verfahren nach Anspruch 12 oder 13, dadurch gekennzeichnet, daß das Zeitintervall zur Klassierung des Spitzenwertes des jeweiligen Teilbandsignales entsprechend der zeitlichen Verdeckung des menschlichen Gehörs und entsprechend der zeitlichen Struktur des Teilbandsignals innerhalb des zugeordneten Teilbandes bestimmt ist.

25

15. Verfahren nach einem der Ansprüche 12 bis 14, dadurch gekennzeichnet, daß für jeden aus einer Folge von Abtastwerten bestehenden Block jedes Teilbandsignales nur für die erste Blockhälfte der Skalenfaktor bestimmt wird, und daß für die zweite Blockhälfte der Skalenfaktor der ersten Hälfte desselben

30

35

1

oder des nachfolgenden Blockes verwendet wird.

5

16. Verfahren nach einem der Ansprüche 1 bis 15, dadurch gekennzeichnet, daß sende- und empfangsseitig aus den Skalenfaktoren der Teilbandsignale die Steuerinformationen für die Quantisierung der Abtastwerte in den Teilbändern gewonnen werden.

10

17. Verfahren nach einem der Ansprüche 1, 8, 9 und 16, dadurch gekennzeichnet, daß die Steuerinformationen für die Quantisierung der Abtastwerte nur sendeseitig gewonnen und für die empfangsseitige Rückquantisierung zusätzlich übertragen werden.

15

18. Verfahren nach Anspruch 17, dadurch gekennzeichnet, daß sendeseitig die Steuerinformationen für die Quantisierung der Abtastwerte nach Maßgabe einer ausgewählten Spektralanalyse des digitalen Tonsignals gewonnen werden.

20

19. Verfahren nach den Ansprüchen 1 und 12, dadurch gekennzeichnet, daß vor ihrer Übertragung die Skalenfaktoren einer höheren Fehlersicherung unterzogen werden als die transcodierte Abtastwerte der Teilbandsignale.

25

20. Verfahren nach Anspruch 1, 17 und 18 dadurch gekennzeichnet, daß vor ihrer Übertragung die Steuerinformationen einer höheren Fehlersicherung unterzogen werden als die transcodierte Abtastwerte der Teilbandsignale.

30

35

1

21. Verfahren nach Anspruch 19 und 20, dadurch gekennzeichnet, daß die Fehlersicherung der Skalenfaktoren und/oder der Steuerinformationen in Teilbändern größerer relativer Bandbreite höher ist als in Teilbändern kleinerer relativer Bandbreite.

5

10

22. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die transcodierten Abtastwerte in den Teilbändern einer Fehlersicherung unterworfen werden, derart, daß in Teilbändern größerer relativer Bandbreite ein höherer Fehlerschutz besteht als in Teilbändern kleinerer relativer Bandbreite.

15

23. Verfahren nach einem der Ansprüche 1 und 19 bis 22, dadurch gekennzeichnet, daß die Informationsmenge für den Fehlerschutz (Fehlerschutzgrad) nach Maßgabe der spektralen und zeitlichen Struktur des digitalen Tonsignals bemessen wird, derart, daß

20

- für Tonsignale, welche für die Übertragung einen geringen Gesamtinformationsfluß benötigen, ein hoher Fehlerschutzgrad vorgesehen wird, und

25

- für Tonsignale, welche für die Übertragung einen hohen Gesamtinformationsfluß benötigen, ein geringer Fehlerschutzgrad vorgesehen wird.

30

24. Verfahren nach Anspruch 23, dadurch gekennzeichnet, daß der signalabhängige Fehlerschutzgrad so bemessen wird, daß die durch Bitfehler verursachte subjektive Störung für wenig verdeckende Tonsignale nicht größer ist als für stark verdeckende Tonsignale.

35

1

25. Verfahren nach Anspruch 1, dadurch gekennzeichnet,
daß bei Unterschreitung eines Pegelschwellwertes
eines Teilbandsignales dessen Abtastwerte nicht
transcodiert werden.

5

10

26. Verfahren nach einem der Ansprüche 4 bis 6 und
12 bis 16, dadurch gekennzeichnet, daß der
Skalenfaktor eines Teilbandes die Information
darüber enthält, daß die Abtastwerte dieses Teil-
bandes innerhalb eines Blockes Null betragen und
nicht übertragen werden.

15

20

25

30

35

1/21

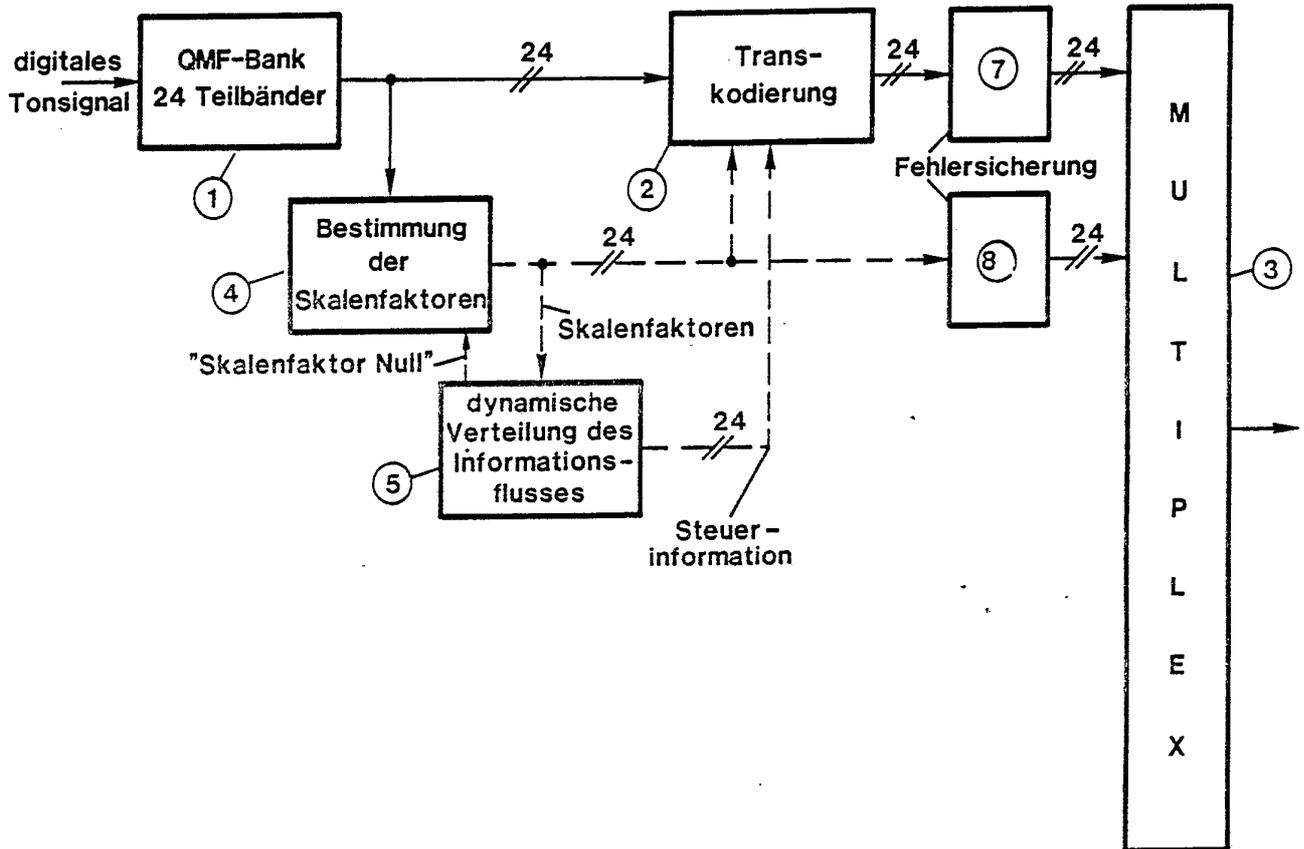


FIG. 1A

2/21

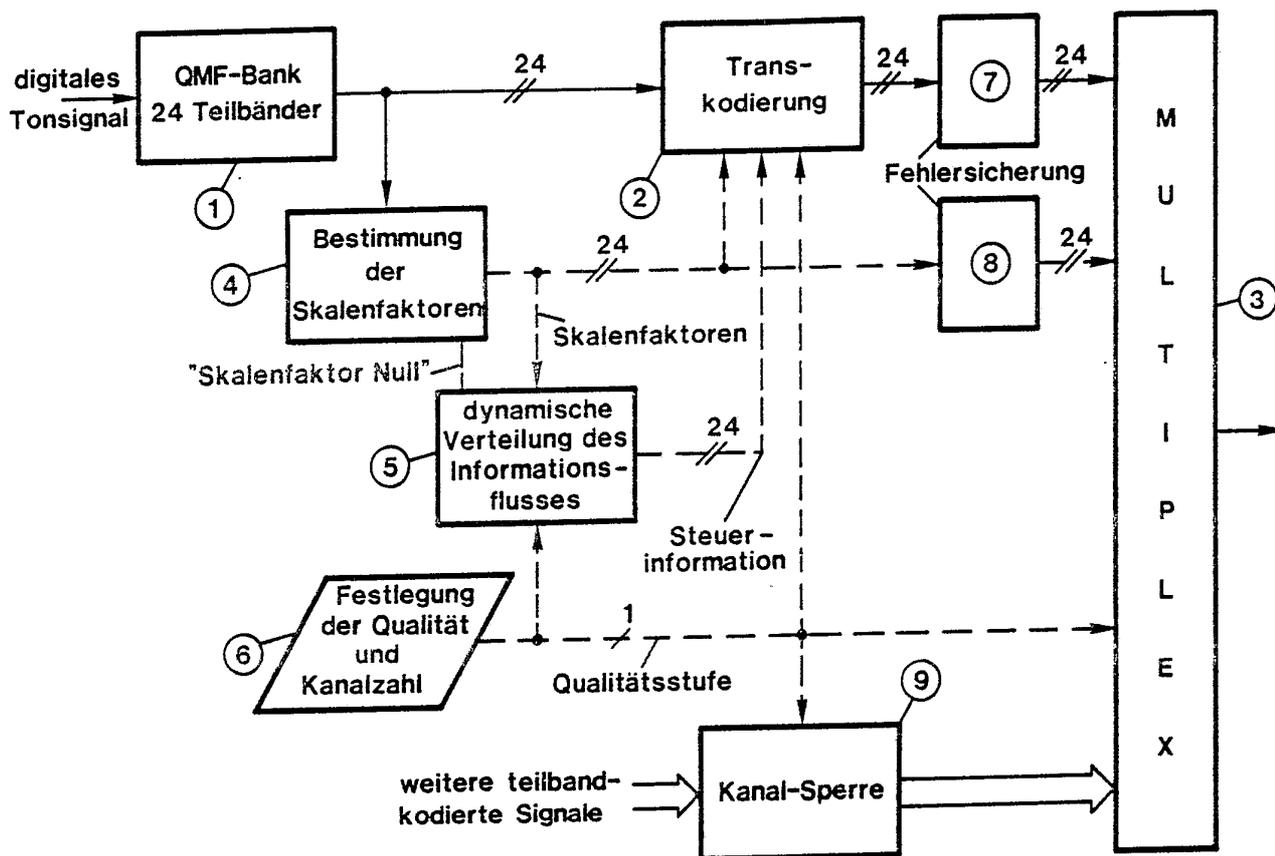


FIG. 1B

3/21

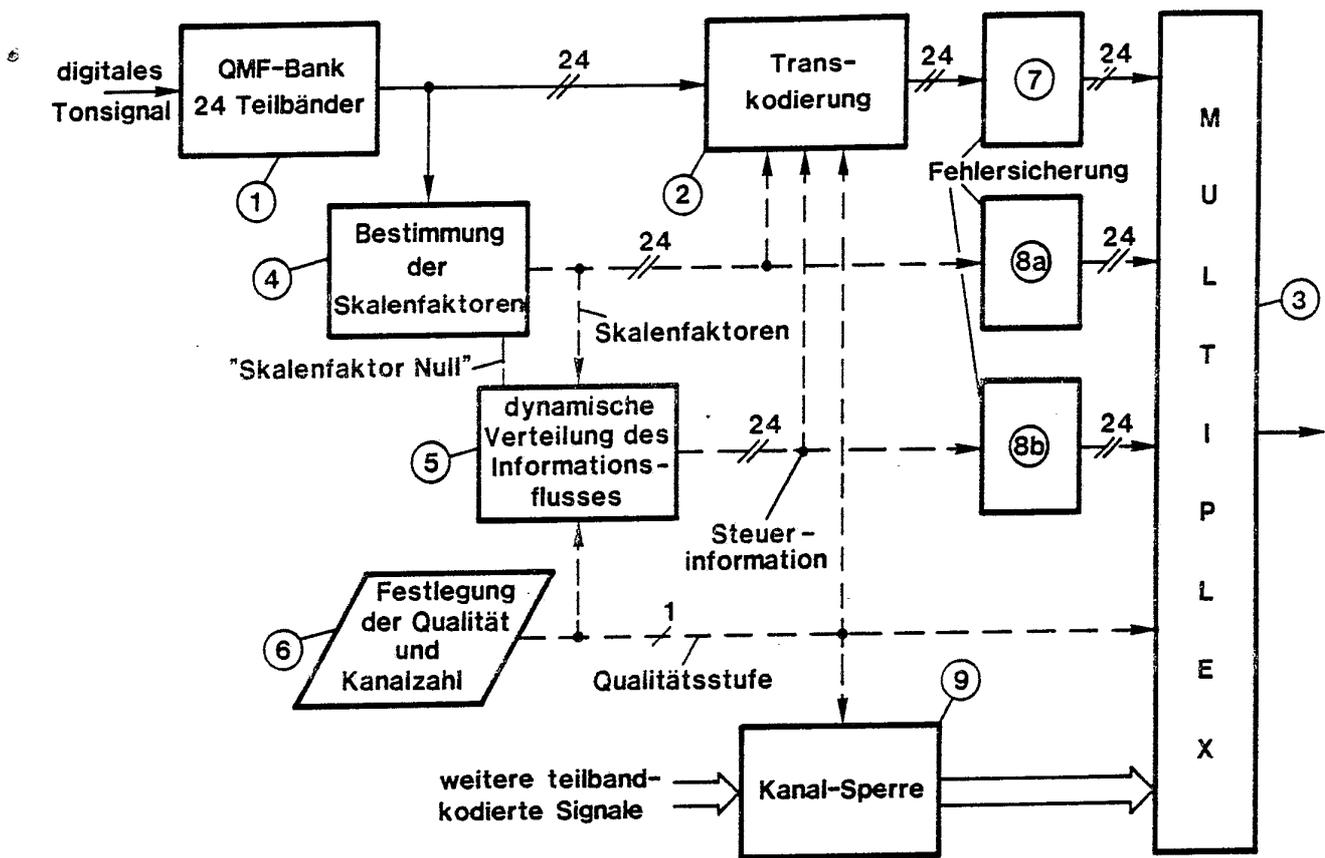
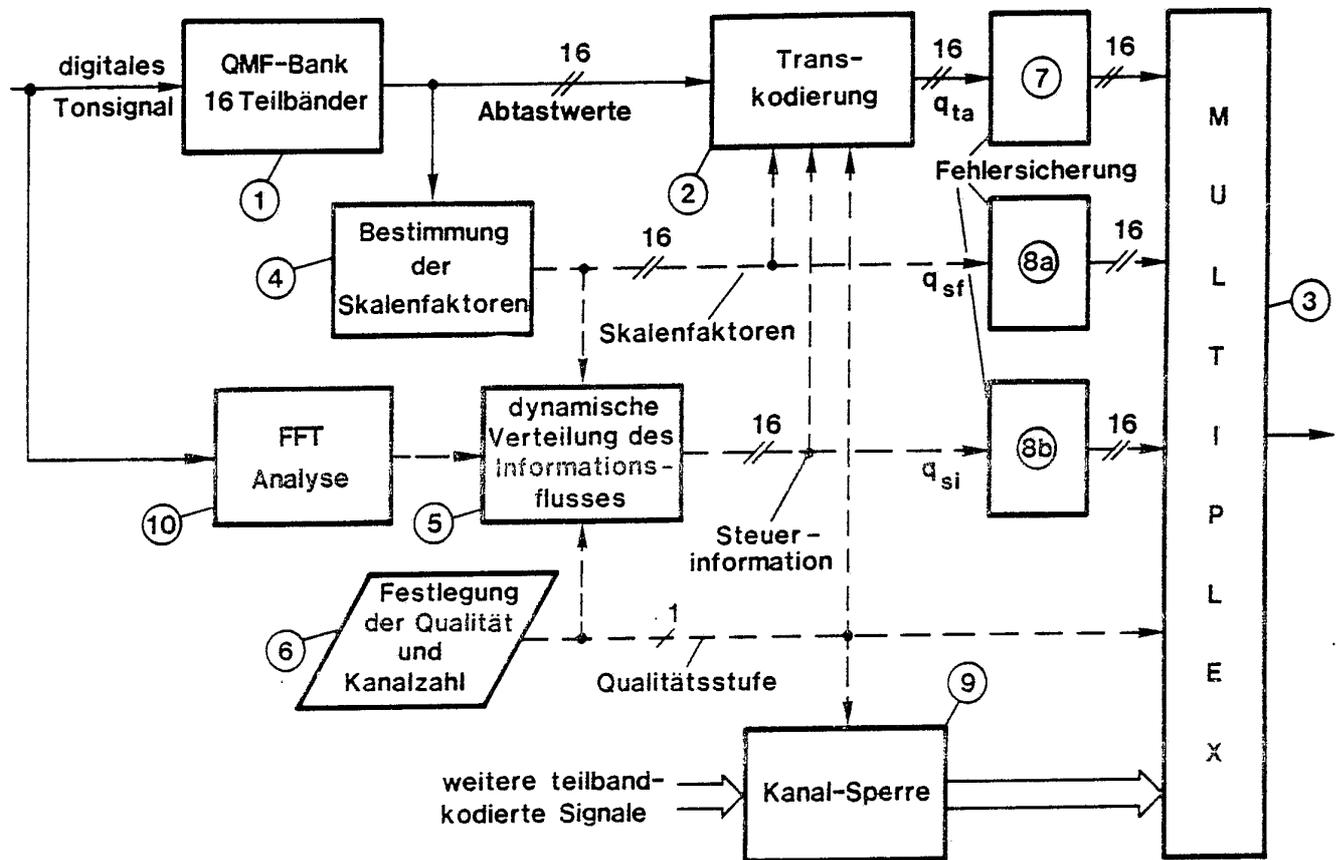


FIG. 1C

4/21



q_{ta} : Datenfluß der transkodierten Abtastwerte der Teilbänder

$q_{ta} + q_{sf} + q_{si}$: Datenfluß der kodierten Teilbandsignale

FIG. 1D

5/27

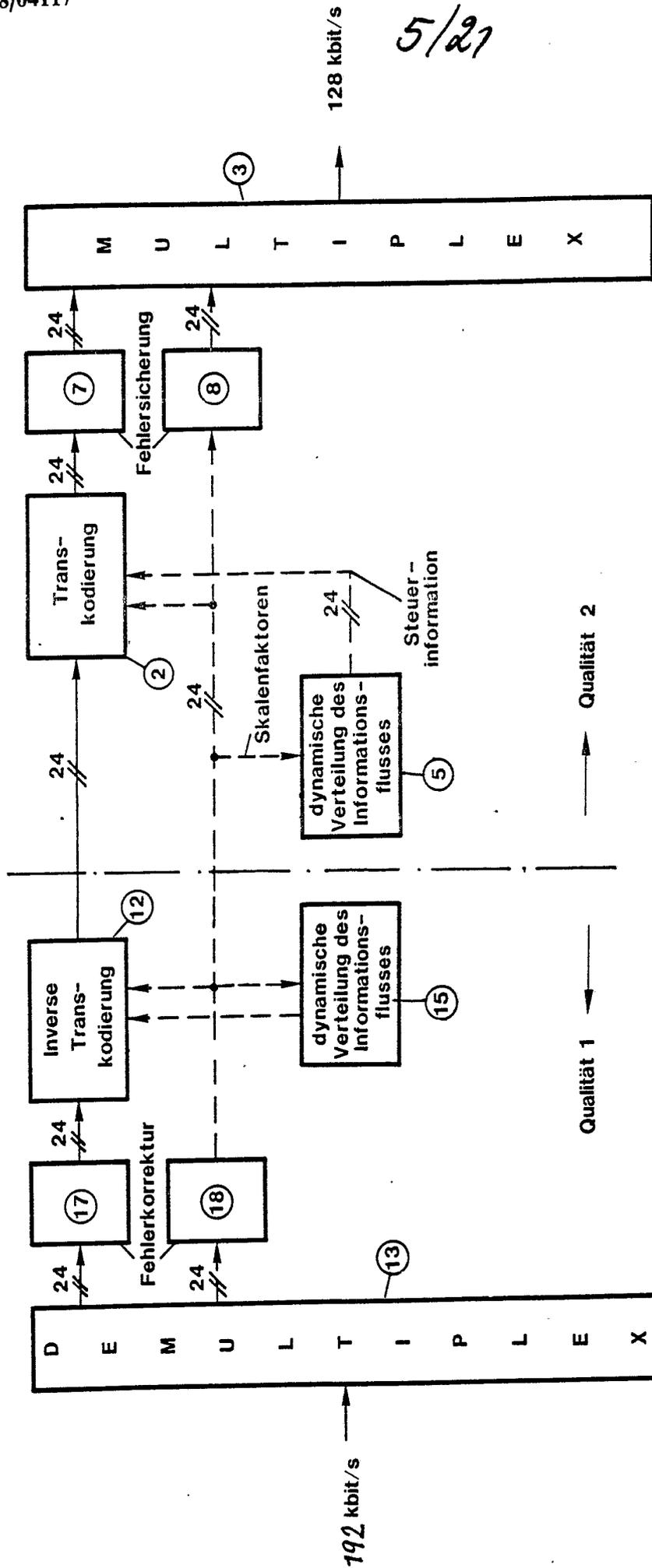


FIG. 1E

6/21

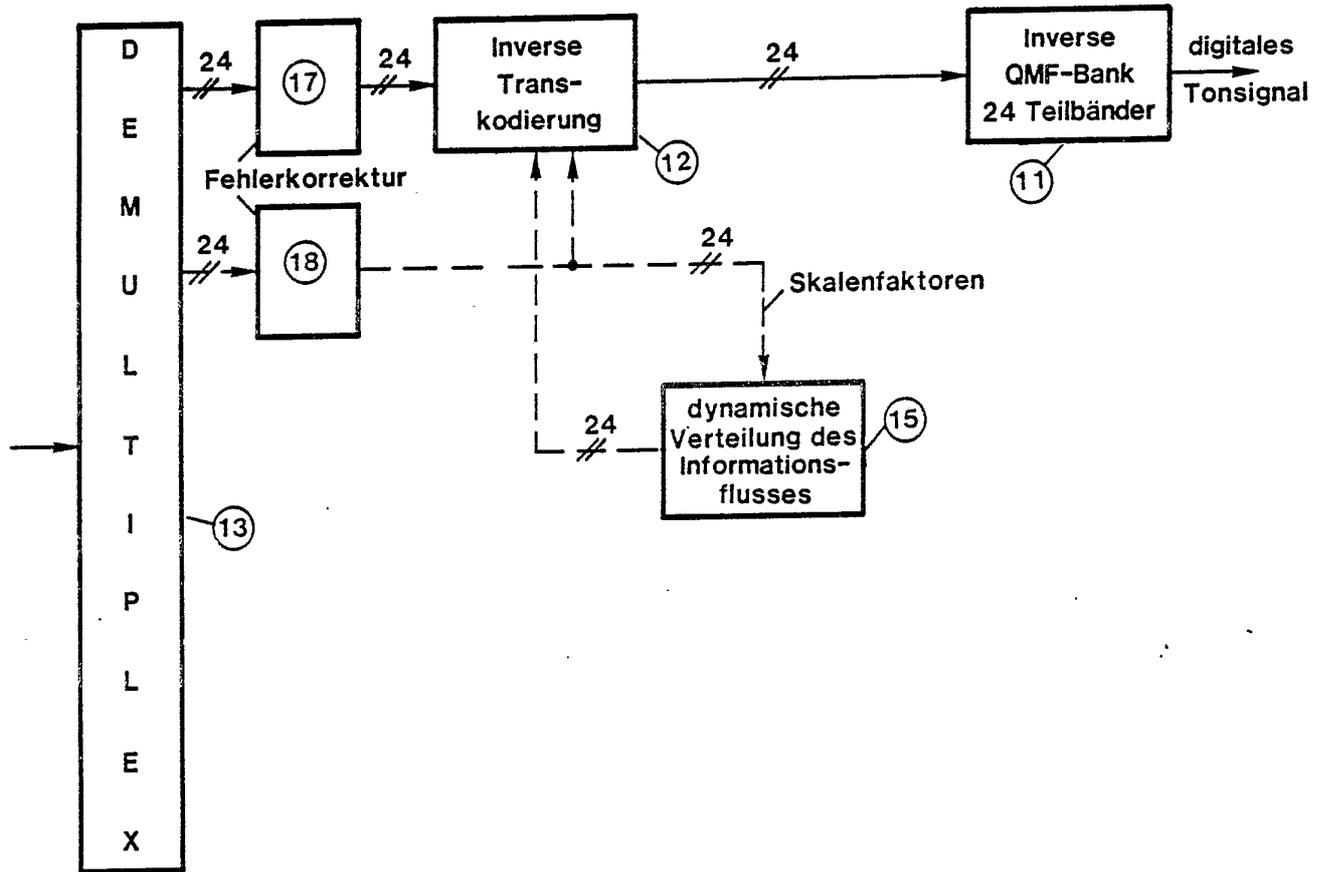


FIG. 2A

7/21

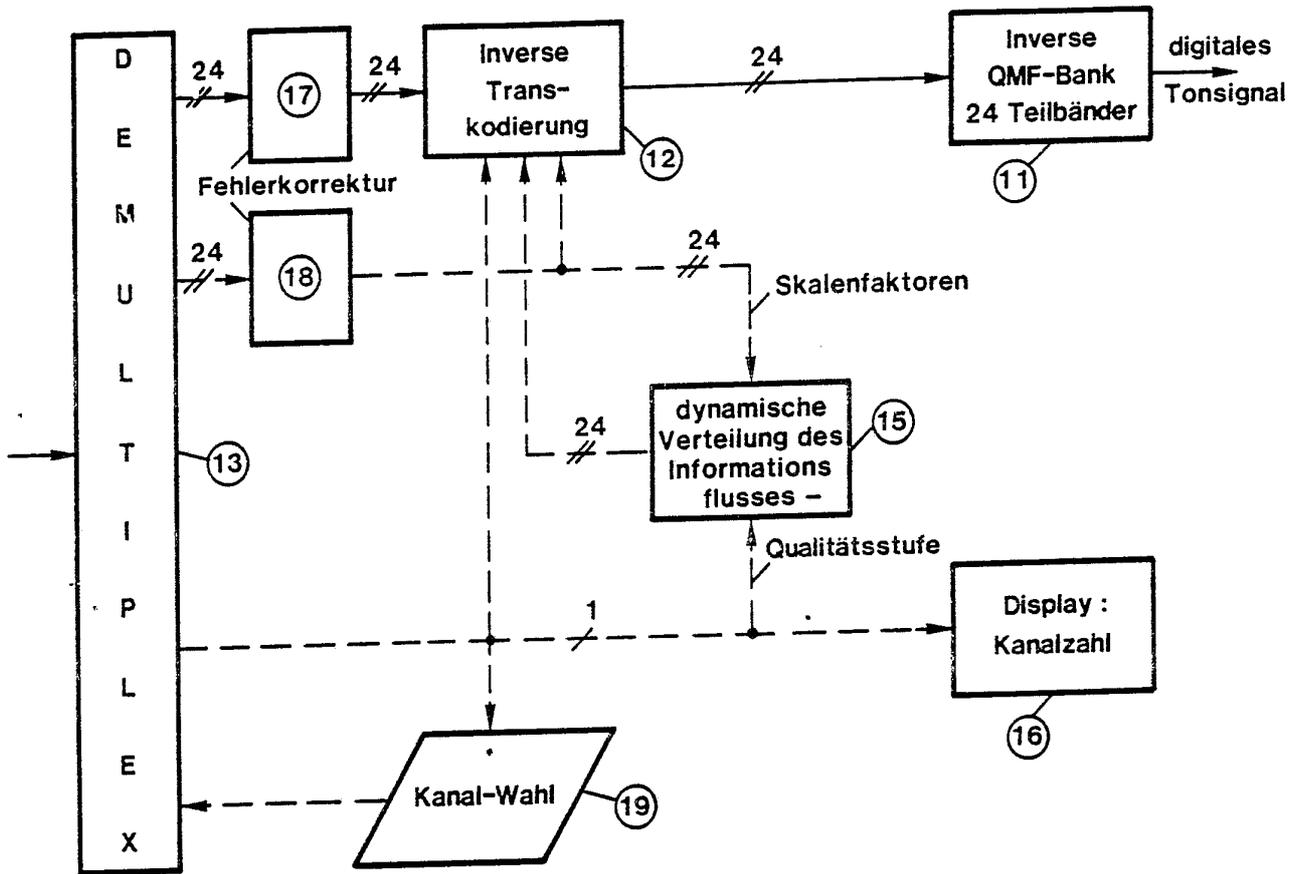


FIG. 2B

9/21

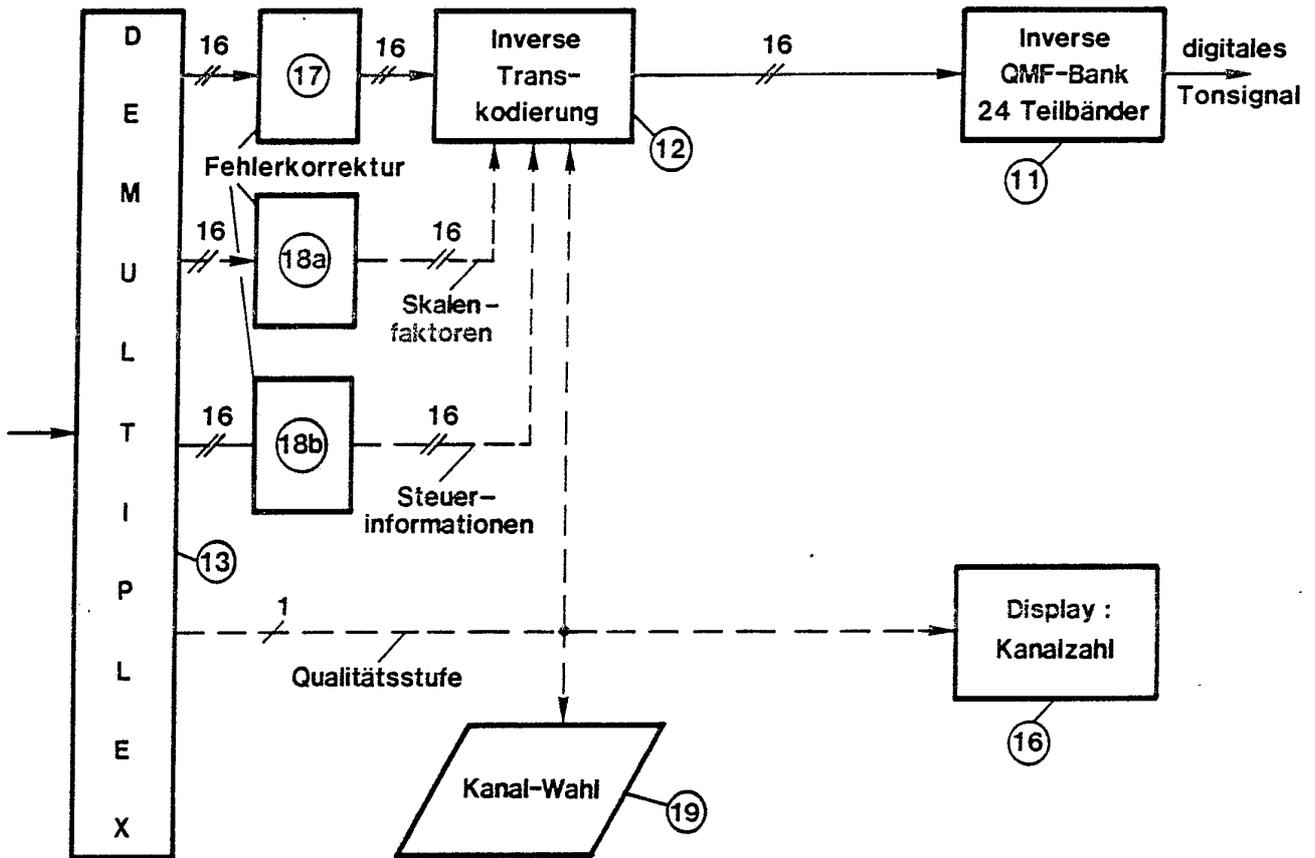


FIG. 2D

10/21

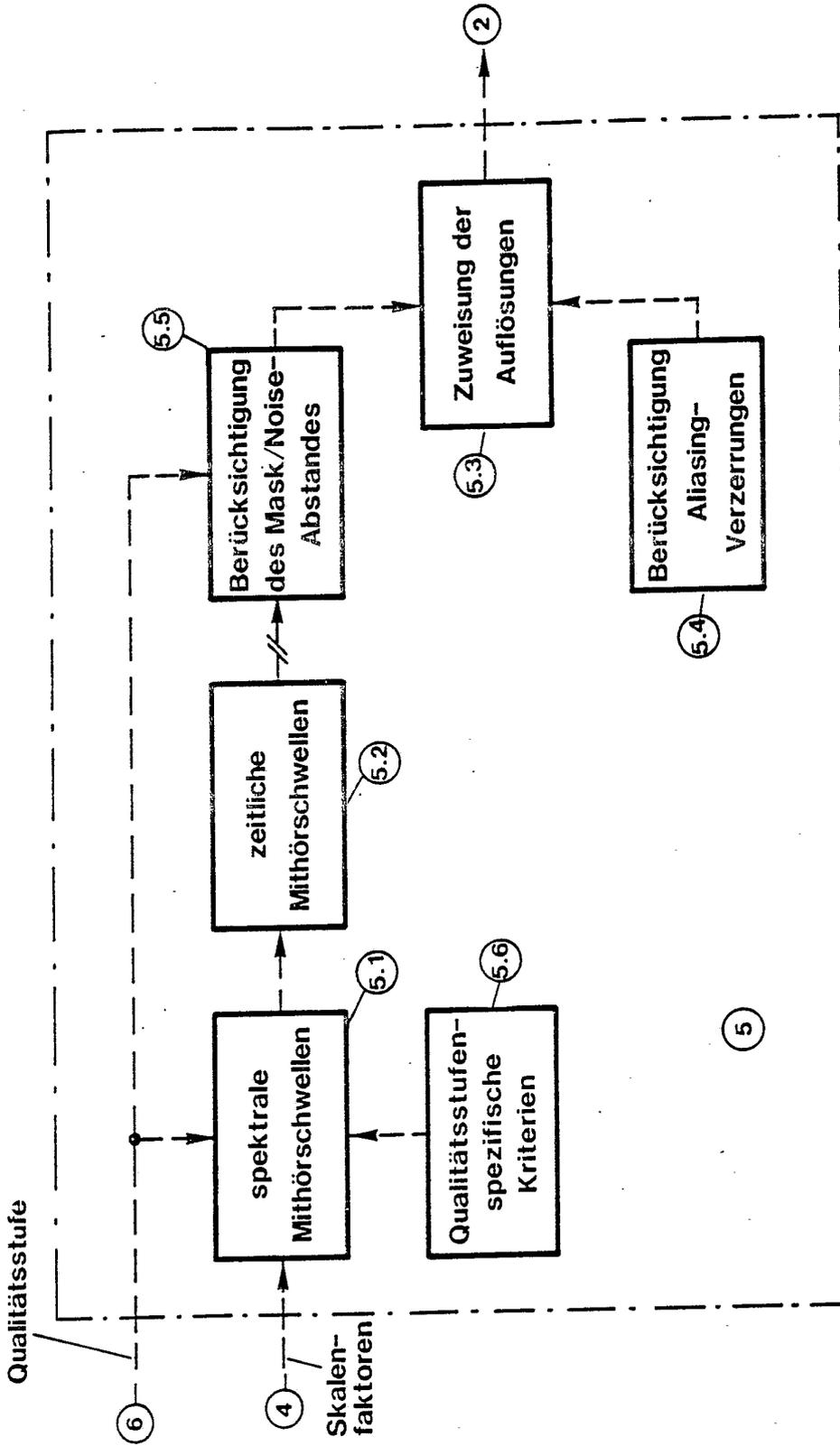


FIG. 3

11/21

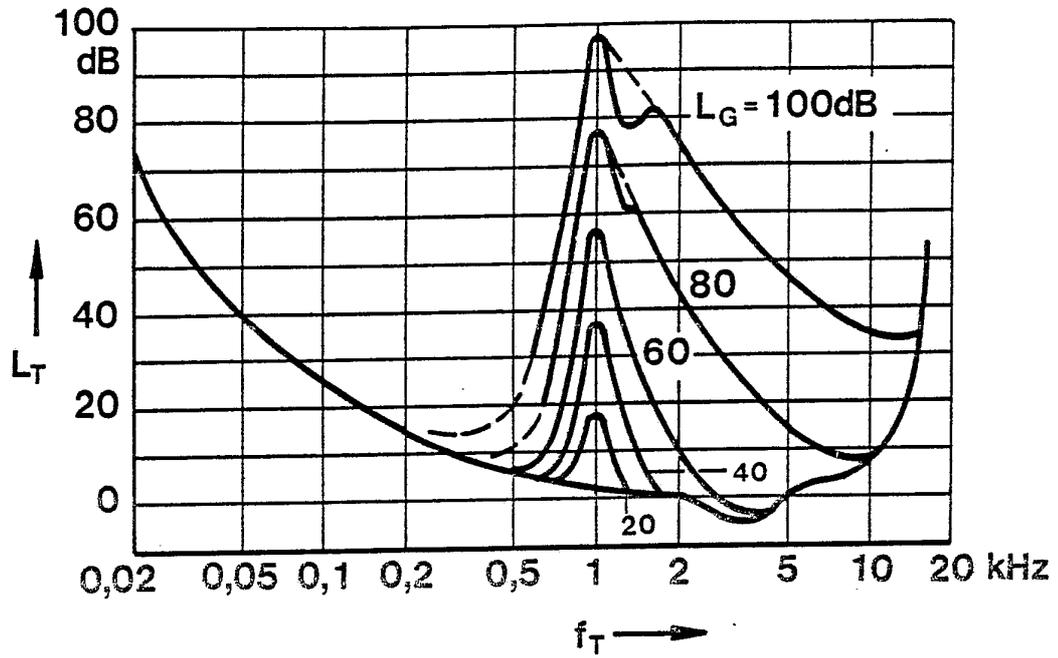


FIG. 5

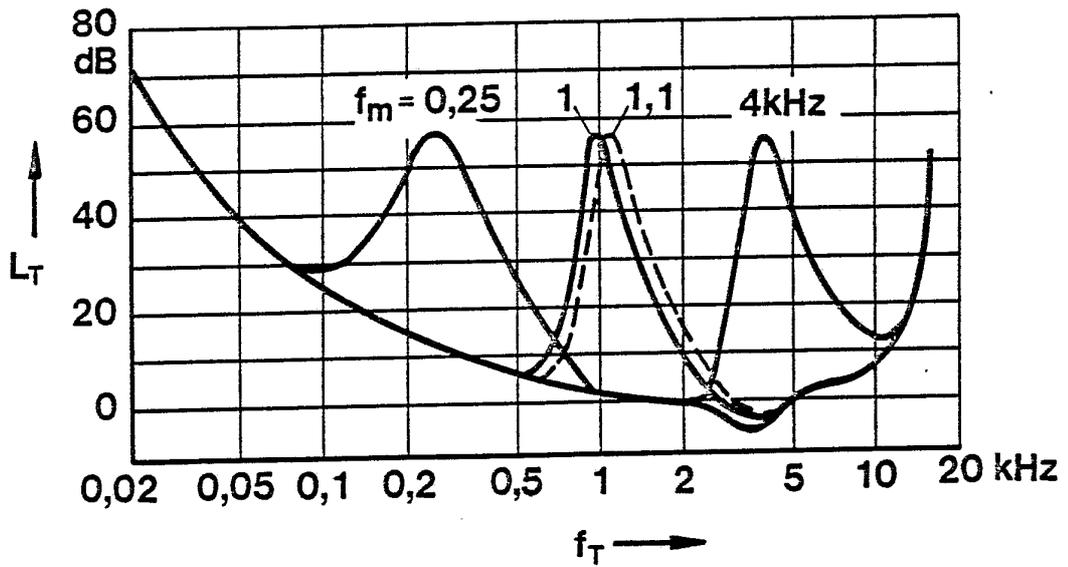


FIG. 4

12/21

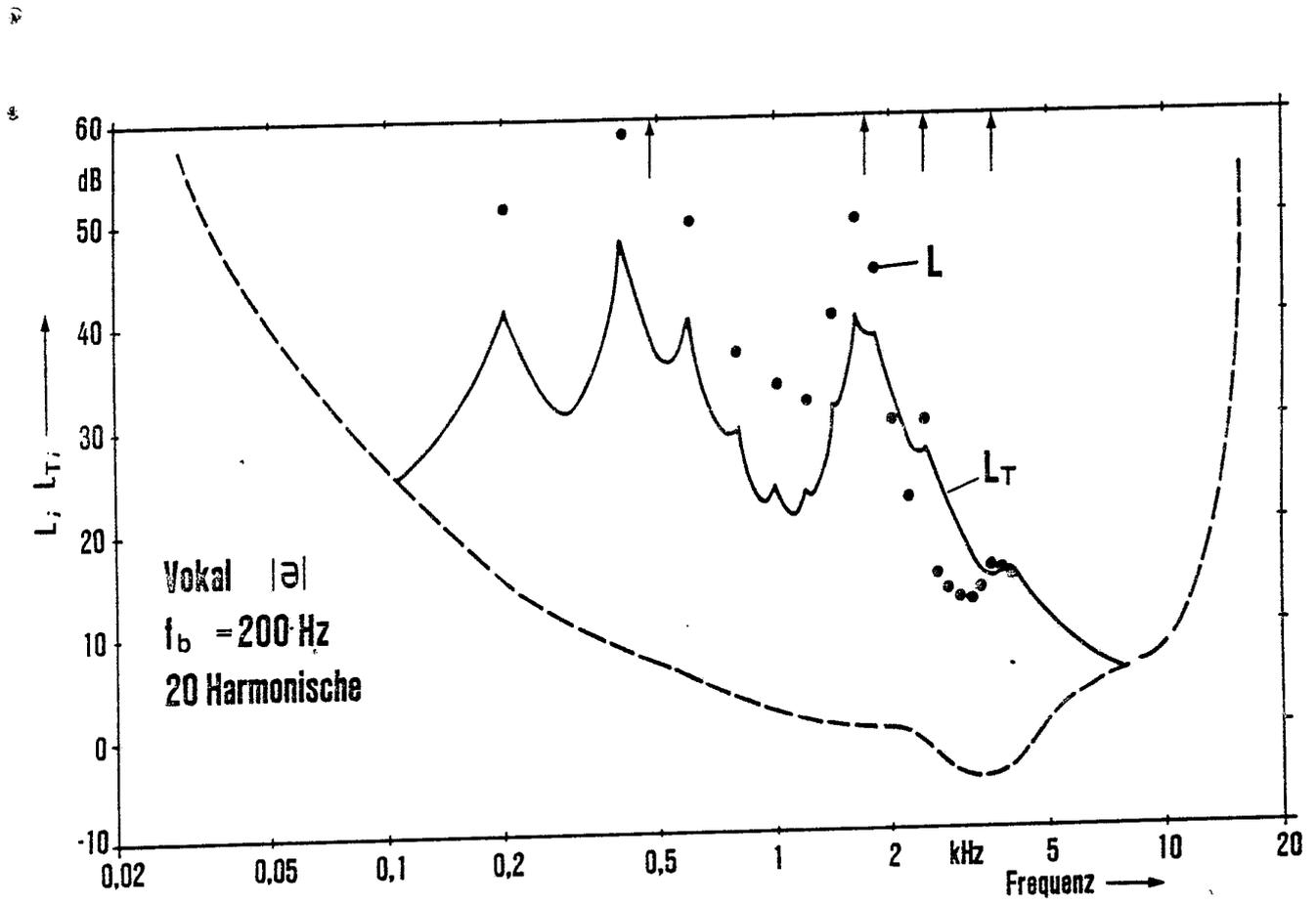


FIG. 6

13/21

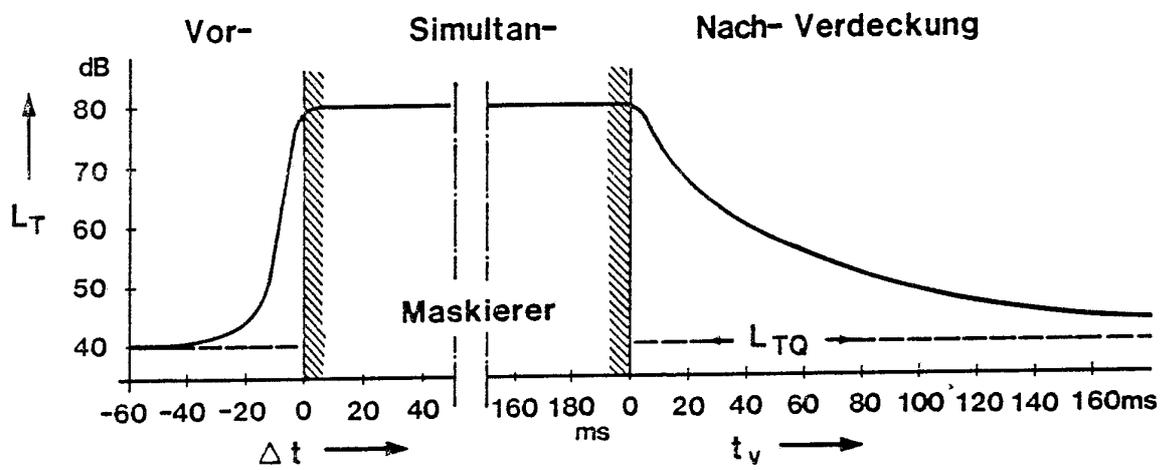


FIG. 7

14/21

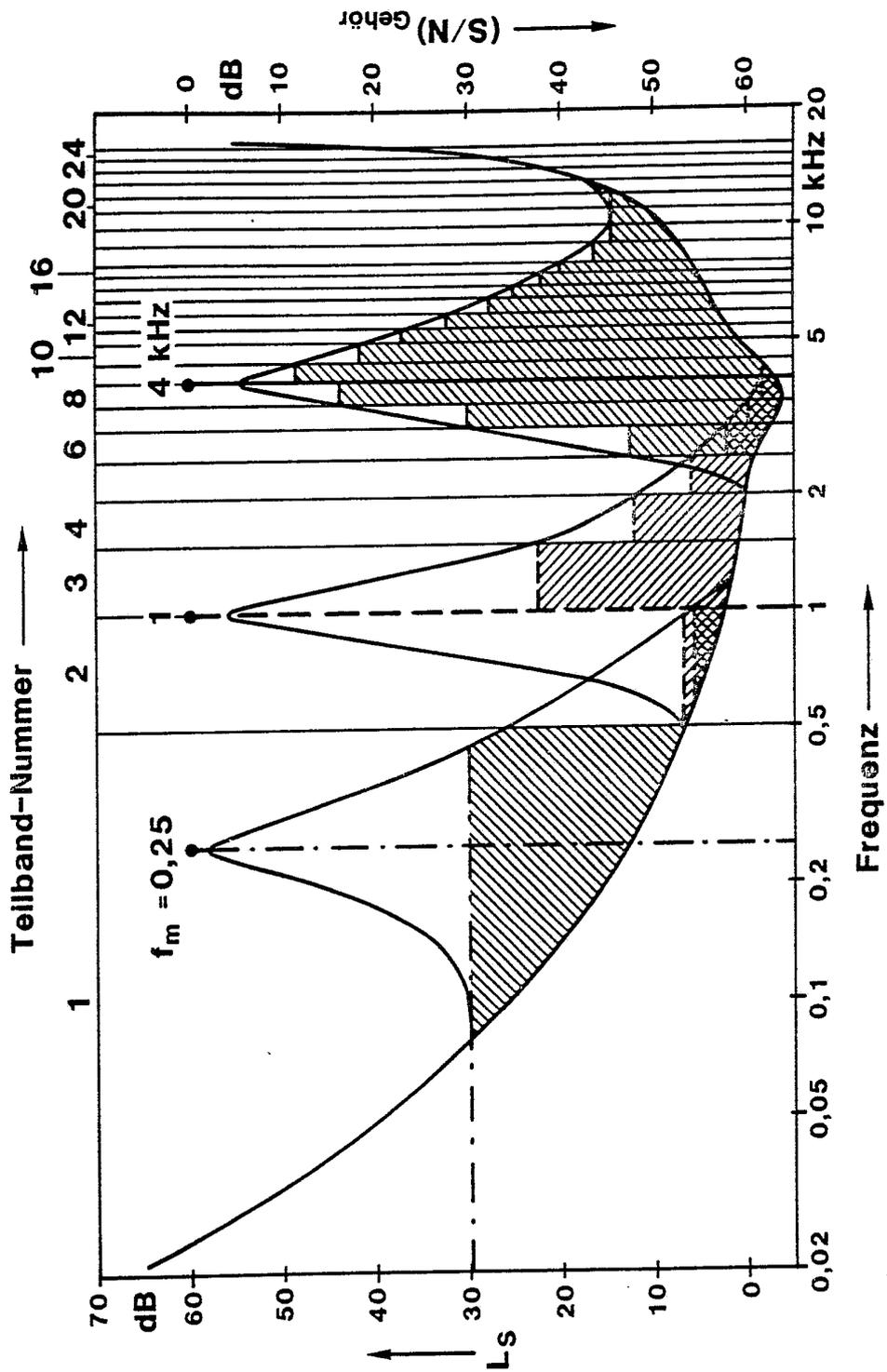


FIG. 8

15/21

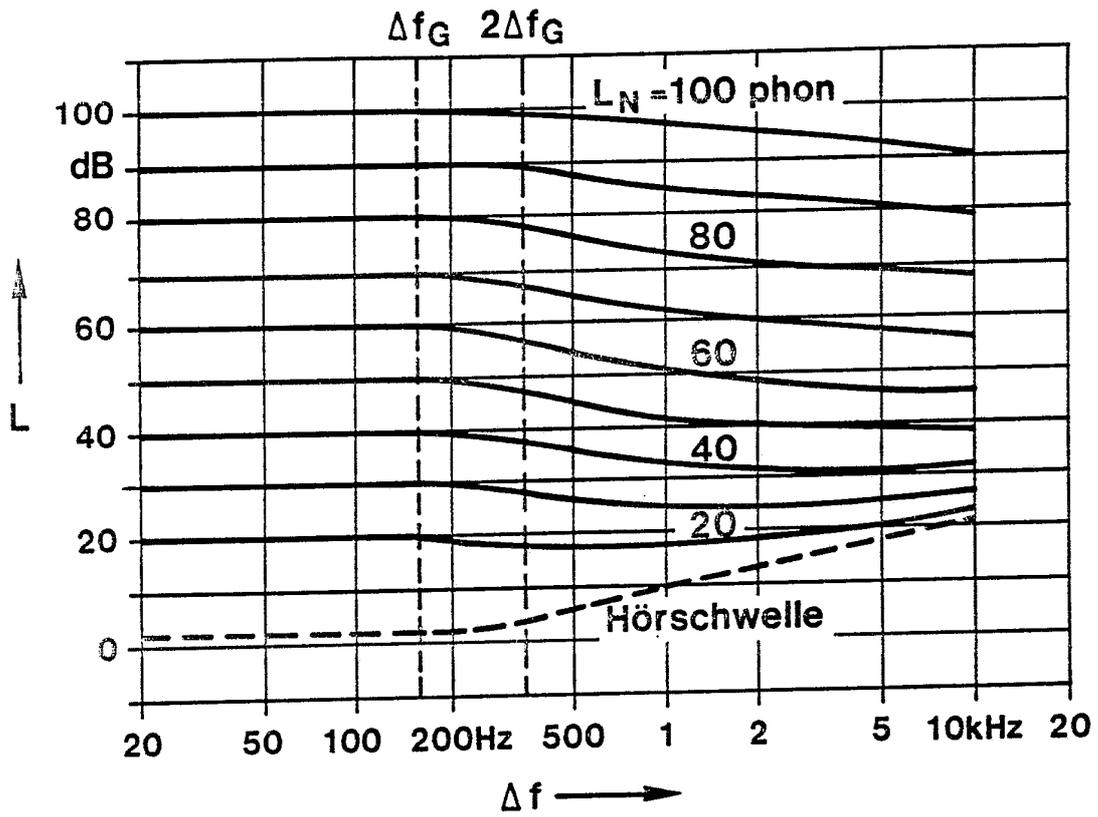


FIG. 9

16/21

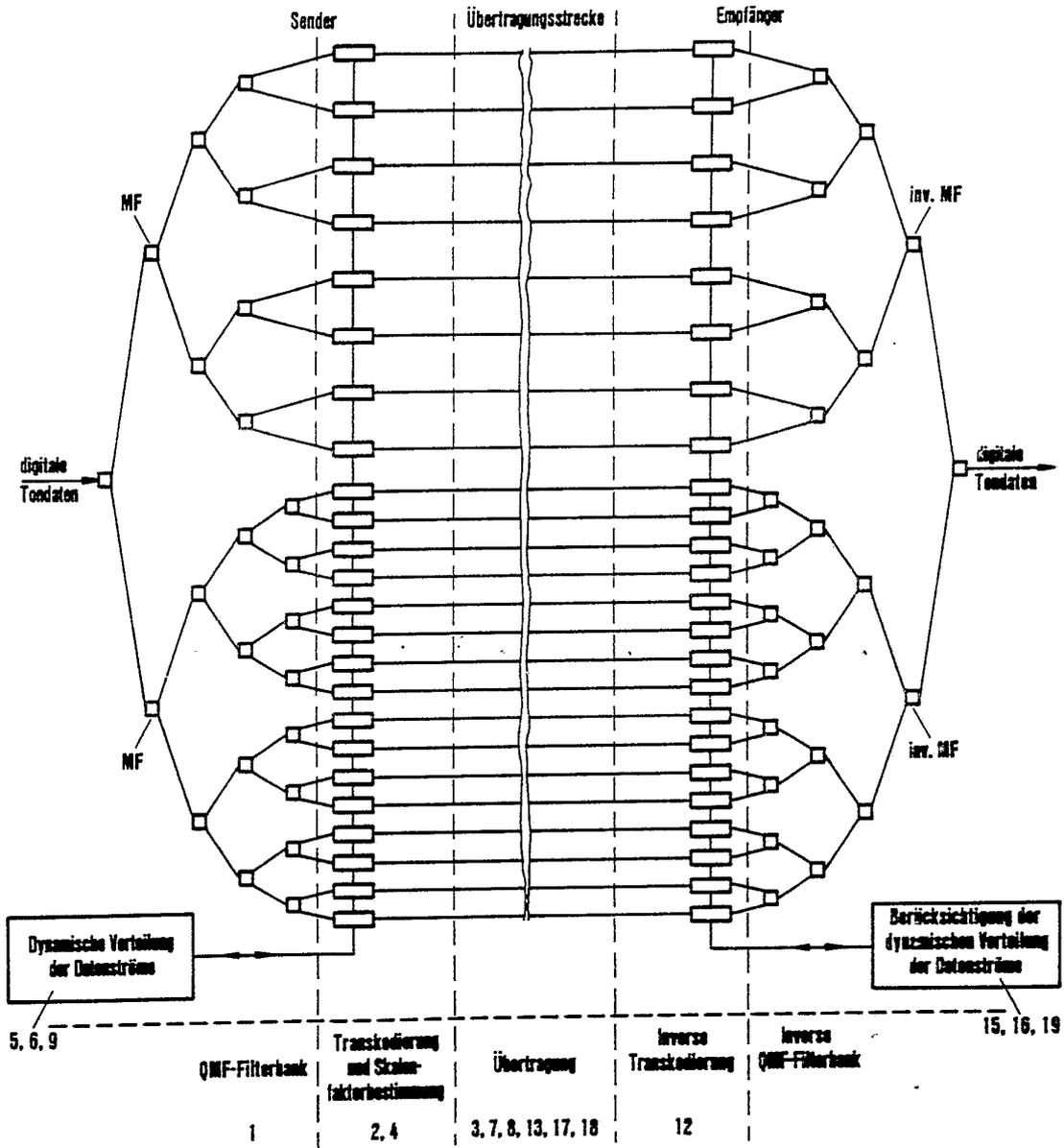


FIG. 10

17/21

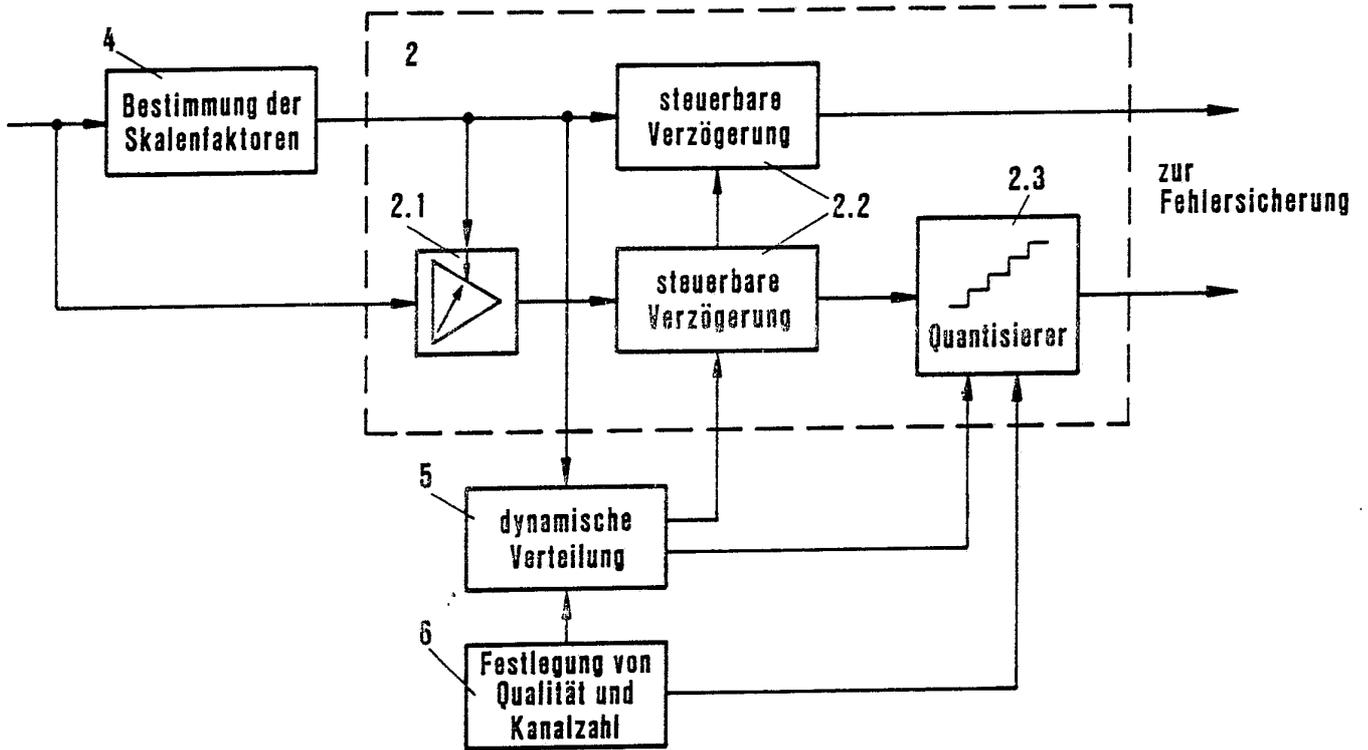


FIG. 11

78/21

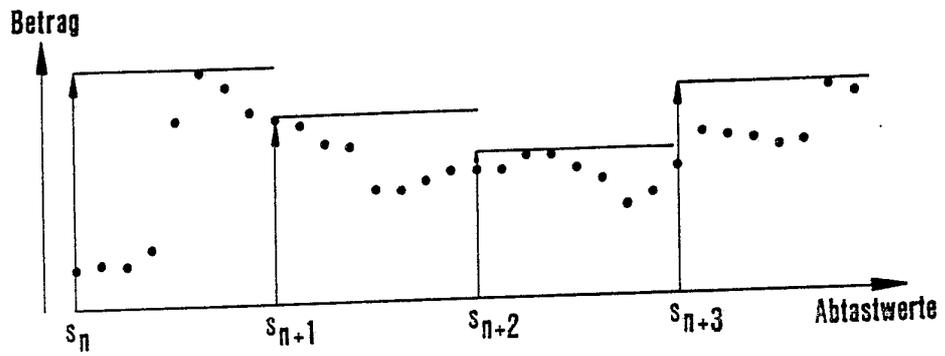


FIG. 12

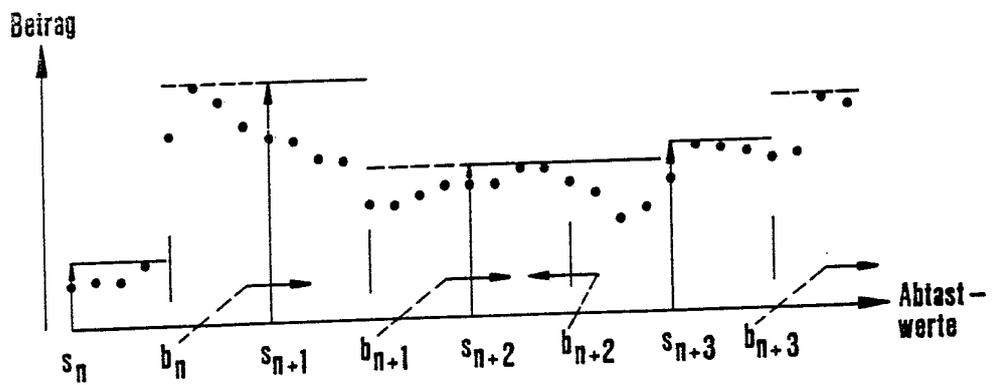
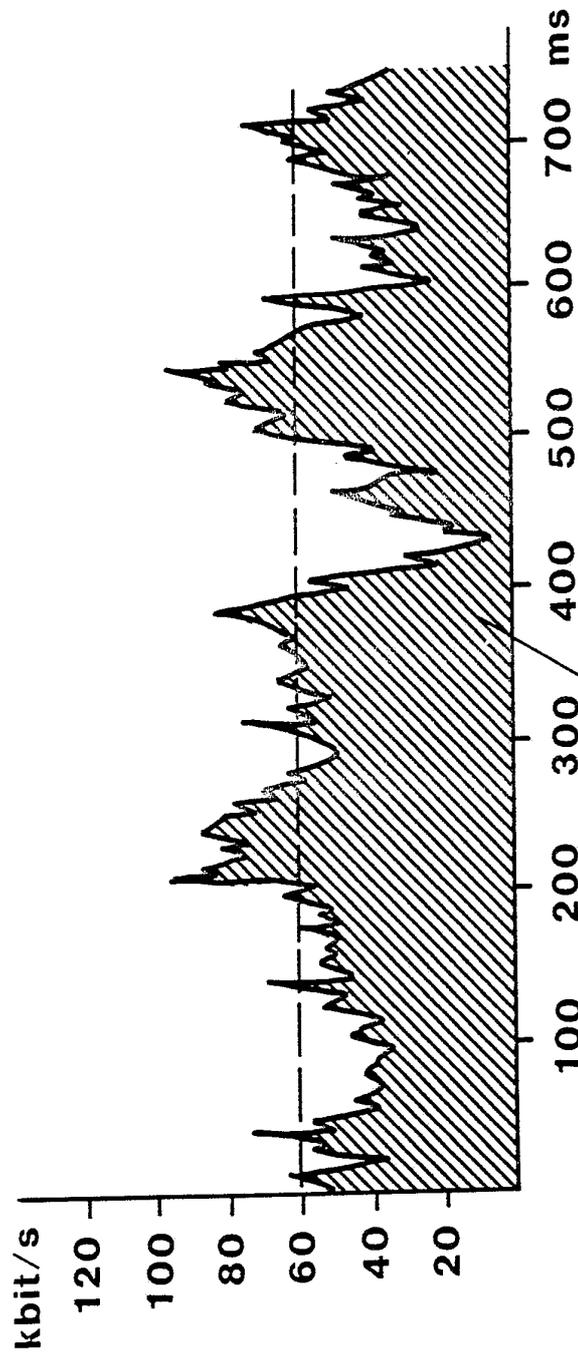


FIG. 13

19/21



erforderlicher Gesamtinformationsfluß

FIG. 14

20/21

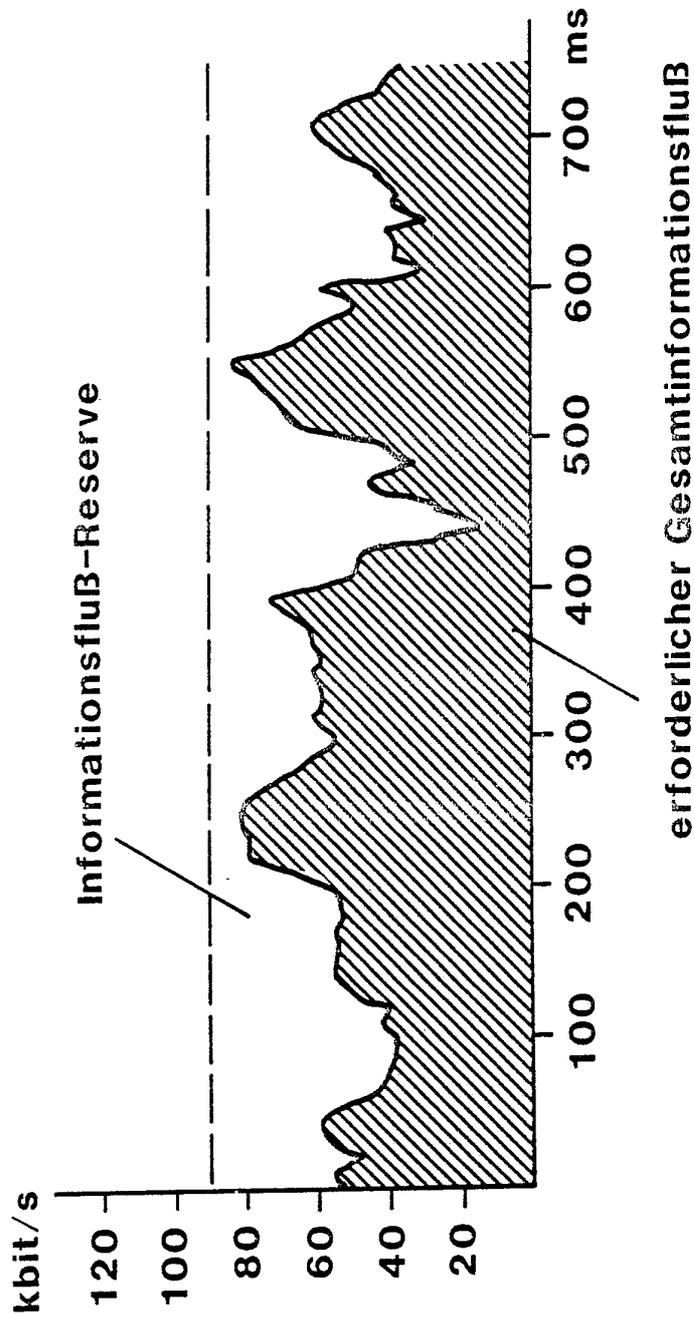
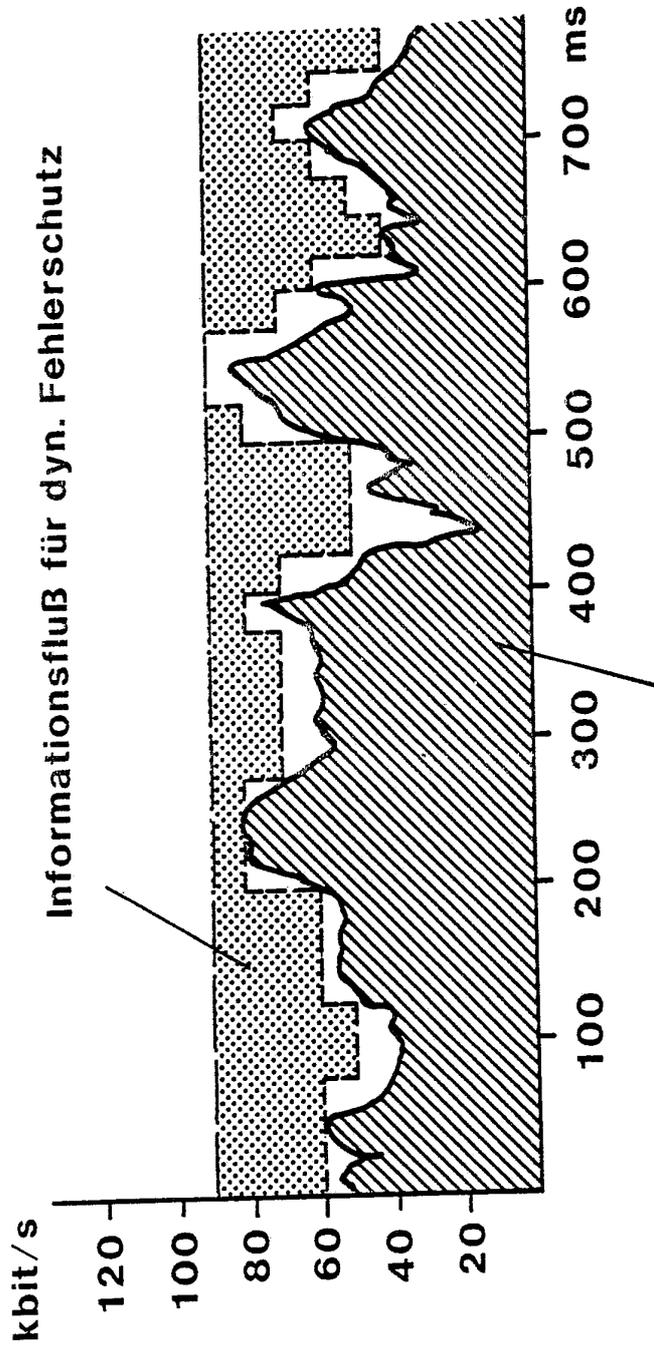


FIG. 15

21/21



erforderlicher Gesamtinformationsfluß

FIG. 16

INTERNATIONAL SEARCH REPORT

International Application No PCT/EP 87/00723

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) ⁶		
According to International Patent Classification (IPC) or to both National Classification and IPC		
Int.Cl. ⁴ : H 04 B 1/66		
II. FIELDS SEARCHED		
Minimum Documentation Searched ⁷		
Classification System	Classification Symbols	
Int.Cl. ⁴ :	H 04 B	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched ⁸		
III. DOCUMENTS CONSIDERED TO BE RELEVANT ⁹		
Category [*]	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No. ¹³
A	DE, Cl, 3440613 (INSTITUT FÜR RUNDFUNK-TECHNIK) 10 April 1986, see page 3, lines 52-63; page 5, lines 23-44 cited in the application --	1
A	Rundfunktechnische Mitteilungen, vol. 30, No: 3, May-June 1986, (Hamburg, DE), D.Krahe: "Ein Verfahren zur Datenreduktion bei digitalen Audiosignalen unter Ausnutzung psychoakustischer Phänomene", pages 117-123. see page 118, right column, paragraph 3; page 122, left column, paragraph 1, 2 and right column, lines 1-7 cited in the application --	1,8,12,18
A	IEEE Transactions on Communications, vol. COM-33, No: 10, October 1985, IEEE, (New York, US) K.Y. Kou et al.: "Digital speech interpolation for variable rate coders with application to subband	1,2,12 ./.
<p>[*] Special categories of cited documents: ¹⁰</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"Δ" document member of the same patent family</p>		
IV. CERTIFICATION		
Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report	
3 March 1988 (03.03.88)	15 April 1988 (15.04.88)	
International Searching Authority	Signature of Authorized Officer	
European Patent Office		

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
	coding", pages 1100-1108 see page 1101, left column, lines 14-47 and right column, before last line - page 1102, left column, line 12 -.-	
A	EP, A2, 0064119 (IBM) 10 November 1982 see page 4, lines 17-29; page 7, lines 17-20; page 9, line 20 - page 10, line 5 -----	1,12,15,16, 19

ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO.

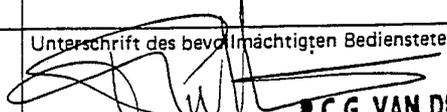
EP 8700723
SA 19760

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 07/04/88. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
DE-C- 3440613	10-04-86	Keine	
EP-A- 0064119	10-11-82	JP-A- 57183143	11-11-82
		US-A- 4464783	07-08-84
		CA-A- 1188423	04-06-85

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen PCT/EP 87/00723

I. KLASSIFIKATION DES ANMELDUNGSGEGENSTANDS (bei mehreren Klassifikationssymbolen sind alle anzugeben) ⁶		
Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC		
Int. Cl. 4 H 04 B 1/66		
II. RECHERCHIERTE SACHGEBIETE		
Recherchierter Mindestprüfstoff ⁷		
Klassifikationssystem	Klassifikationssymbole	
Int. Cl. 4	H 04 B	
Recherchierte nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Sachgebiete fallen ⁸		
III. EINSCHLÄGIGE VERÖFFENTLICHUNGEN⁹		
Art*	Kennzeichnung der Veröffentlichung ¹¹ , soweit erforderlich unter Angabe der maßgeblichen Teile ¹²	Betr. Anspruch Nr. ¹³
A	DE, C1, 3440613 (INSTITUT FÜR RUNDFUNKTECHNIK) 10. April 1986 siehe Seite 3, Zeilen 52-63; Seite 5, Zeilen 23-44 in der Anmeldung erwähnt --	1
A	Rundfunktechnische Mitteilungen, Band 30, Nr. 3, Mai-Juni 1986, (Hamburg, DE), D. Krahe: "Ein Verfahren zur Datenreduktion bei digitalen Audiosignalen unter Aus- nutzung psychoakustischer Phänomene", Seiten 117-123 siehe Seite 118, rechte Spalte, Abschnitt 3; Seite 122, linke Spalte, Abschnitte 1, 2 und rechte Spalte, Zeilen 1-7 in der Anmeldung erwähnt --	1, 8, 12, 18
A	IEEE Transactions on Communications, Band COM-33, Nr. 10, Oktober 1985, IEEE, (New York, US), ./. .	1, 2, 12
<p>* Besondere Kategorien von angegebenen Veröffentlichungen¹⁰:</p> <p>"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist</p> <p>"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist</p> <p>"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)</p> <p>"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht</p> <p>"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist</p> <p>"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist</p> <p>"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden</p> <p>"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist</p> <p>"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist</p>		
IV. BESCHEINIGUNG		
Datum des Abschlusses der internationalen Recherche	Absenddatum des internationalen Recherchenberichts	
3. März 1988	15 APR 1988	
Internationale Recherchenbehörde	Unterschrift des bevollmächtigten Bediensteten	
Europäisches Patentamt	 P.C.G. VAN DER PUTTEN	

III. EINSCHLÄGIGE VERÖFFENTLICHUNGEN (Fortsetzung von Blatt 2)		
Art *	Kennzeichnung der Veröffentlichung, soweit erforderlich unter Angabe der maßgeblichen Teile	Betr. Anspruch Nr.
A	<p>K.Y. Kou et al.: "Digital speech interpolation for variable rate coders with application to subband coding", Seiten 1100-1108 siehe Seite 1101, linke Spalte, Zeilen 14-47 und rechte Spalte, vorletzte Zeile - Seite 1102, linke Spalte, Zeile 12</p> <p>EP, A2, 0064119 (IBM) 10. November 1982 siehe Seite 4, Zeilen 17-29; Seite 7, Zeilen 17-20; Seite 9, Zeile 20 - Seite 10, Zeile 5</p> <p>-----</p>	<p>1,12,15,16, 19</p>

ANHANG ZUM INTERNATIONALEN RECHERCHENBERICHT
 ÜBER DIE INTERNATIONALE PATENTANMELDUNG NR.

EP 8700723
 SA 19760

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengenannten internationalen Recherchenbericht angeführten Patentdokumente angegeben.
 Die Angaben über die Familienmitglieder entsprechen dem Stand der Datei des Europäischen Patentamts am 07/04/88
 Diese Angaben dienen nur zur Unterrichtung und erfolgen ohne Gewähr.

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
DE-C- 3440613	10-04-86	Keine	
EP-A- 0064119	10-11-82	JP-A- 57183143	11-11-82
		US-A- 4464783	07-08-84
		CA-A- 1188423	04-06-85

EPO FORM P6873

Für nähere Einzelheiten zu diesem Anhang : siehe Amtsblatt des Europäischen Patentamts, Nr.12/82