



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 698 21 089 T2** 2004.11.11

(12)

Übersetzung der europäischen Patentschrift

(97) **EP 0 940 015 B1**

(51) Int Cl.⁷: **H04B 1/66**

(21) Deutsches Aktenzeichen: **698 21 089.1**

(86) PCT-Aktenzeichen: **PCT/IB98/00893**

(96) Europäisches Aktenzeichen: **98 921 697.3**

(87) PCT-Veröffentlichungs-Nr.: **WO 98/57436**

(86) PCT-Anmeldetag: **09.06.1998**

(87) Veröffentlichungstag
der PCT-Anmeldung: **17.12.1998**

(97) Erstveröffentlichung durch das EPA: **08.09.1999**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **14.01.2004**

(47) Veröffentlichungstag im Patentblatt: **11.11.2004**

(30) Unionspriorität:

| | | |
|----------------|-------------------|-----------|
| 9702213 | 10.06.1997 | SE |
| 9704634 | 12.12.1997 | SE |
| 9800268 | 30.01.1998 | SE |

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE**

(73) Patentinhaber:

Coding Technologies AB, Stockholm, SE

(72) Erfinder:

**LILJERYD, Gustaf, Lars, S-171 34 Solna, SE;
EKSTRAND, Rune, Per, S-116 31 Stockholm, SE;
HENN, Fredrik, Lars, S-168 31 Bromma, SE;
KJÖRLING, Magnus, Hans, S-752 27 Uppsala, SE**

(74) Vertreter:

**Schoppe, Zimmermann, Stöckeler & Zinkler, 82049
Pullach**

(54) Bezeichnung: **VERBESSERUNG VON QUELLENKODIERUNG UNTER VERWENDUNG VON SPEKTRAL-
BANDREPLIKATION**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

[0001] Bei Quellcodierungssystemen werden digitale Daten vor einer Übertragung oder Speicherung komprimiert, um die erforderliche Bitrate oder Speicherkapazität zu reduzieren. Die vorliegende Erfindung bezieht sich auf ein neues Verfahren und eine Vorrichtung zu der Verbesserung von Quellcodierungssystemen mittels einer Spektralbandreplikation (SBR). Es wird eine wesentliche Bitratenreduzierung erreicht, während die gleiche Wahrnehmungsqualität beibehalten wird oder umgekehrt eine Verbesserung bei einer Wahrnehmungsqualität bei einer gegebenen Bitrate erreicht wird. Dies wird mittels einer Spektralbandbreitenreduzierung an der Codierseite und einer nachfolgenden Spektralbandreplikation bei dem Decodierer erzielt, wodurch die Erfindung neue Konzepte einer Signalredundanz in dem Spektralbereich verwertet.

[0002] Audioquellcodierungstechniken können in zwei Klassen eingeteilt werden: natürliche Audiocodierung und Sprachcodierung. Natürliche Audiocodierung wird allgemein für Musik oder beliebige Signale bei mittleren Bitraten verwendet und bietet allgemein eine breite Audiobandbreite. Sprachcodierer sind im Grunde auf eine Sprachreproduktion begrenzt, können aber jedoch bei sehr niedrigen Bitraten verwendet werden, wenn auch mit einer niedrigen Audiobandbreite. Eine Breitbandsprache bietet eine größere subjektive Qualitätsverbesserung gegenüber einer Schmalbandsprache. Ein Erhöhen der Bandbreite verbessert nicht nur eine Verständlichkeit und Natürlichkeit einer Sprache, sondern erleichtert auch eine Sprechererkennung. Eine Breitbandsprachcodierung ist somit ein wichtiger Punkt bei Telefonsystemen der nächsten Generation. Ferner ist aufgrund des enormen Wachstums des Multimediagebiets eine Übertragung von Musik und anderen Nicht-Sprachsignalen mit einer hohen Qualität über Telefonsysteme ein erwünschtes Merkmal.

[0003] Ein lineares PCM-Signal mit hoher Wiedergabetreue ist hinsichtlich einer Bitrate gegenüber der Wahrnehmungsentropie sehr ineffizient. Der CD-Standard diktiert eine Abtastfrequenz von 44,1 kHz, eine Auflösung von 16 Bits pro Abtastwert und Stereo. Dies kommt einer Bitrate von 1411 kbit/s gleich. Um die Bitrate drastisch zu reduzieren, kann eine Quellcodierung unter Verwendung von Splitband-Wahrnehmungsaudiocodern durchgeführt werden. Diese natürlichen Audiocodern verwerten eine Wahrnehmungsirrelevanz und statistische Redundanz bei dem Signal. Unter Verwendung der besten Codec-Technologie kann bei einem Standard-CD-Format-Signal eine Datenreduzierung von näherungsweise 90% praktisch ohne wahrnehmbare Verschlechterung erreicht werden. Es ist somit eine sehr hohe Klangqualität in Stereo bei etwa 96 kbit/s möglich, d. h. einem Komprimierungsfaktor von näherungsweise 15 : 1. Einige Wahrnehmungscodecs bieten sogar höhere Komprimierungsverhältnisse. Um dies zu erreichen, ist es allgemein üblich, die Abtastrate und somit die Audiobandbreite zu reduzieren. Es ist auch üblich, die Anzahl von Quantisierungspegeln zu verringern, was eine gelegentlich hörbare Quantisierungsverzerrung ermöglicht, und eine Verschlechterung des Stereofelds zu verwenden, durch ein Intensitätscodieren. Eine übermäßige Verwendung derartiger Verfahren resultiert in einer störenden Wahrnehmungsverschlechterung. Eine gegenwärtige Codetechnologie befindet sich nahe einer Sättigung und ein weiterer Fortschritt bei einem Codiergewinn wird nicht erwartet. Um das Codierverhalten weiter zu verbessern ist ein neuer Ansatz notwendig.

[0004] Die menschliche Stimme und die meisten Musikinstrumente erzeugen quasistationäre Signale, die aus oszillierenden Systemen hervorgehen. Gemäß der Fourier-Theorie kann ein jegliches periodisches Signal als eine Summe von Sinuswellen mit den Frequenzen f , $2f$, $3f$, $4f$, $5f$ etc. ausgedrückt werden, wobei f die Grundfrequenz ist. Die Frequenzen bilden eine harmonische Reihe. Eine Bandbreitenbegrenzung eines derartigen Signals ist äquivalent zu einer Abschneidung der harmonischen Reihe. Eine derartige Abschneidung ändert das wahrgenommene Timbre, die Klangfarbe, eines Musikinstruments oder einer Stimme und ergibt ein Audiosignal, das „gedämpft“ oder „dumpf“ klingt, und eine Verständlichkeit kann reduziert sein. Die hohen Frequenzen sind somit wichtig für den subjektiven Eindruck einer Klangqualität.

[0005] Bekannte Verfahren sind hauptsächlich zu einer Verbesserung einer Sprachcodecleistung bestimmt und insbesondere zu einer Hochfrequenzregenerierung (HFR) bestimmt, ein Punkt bei einem Sprachcodieren. Derartige Verfahren verwenden lineare Breitbandfrequenzverschiebungen, Nichtlinearitäten oder ein Aliasing [US-Patent 5,127,054], ein Erzeugen von Intermodulationsprodukten oder anderen nicht-harmonischen Frequenzkomponenten, die eine ernste Dissonanz bewirken, wenn dieselben auf Musiksignale angewendet werden. Eine derartige Dissonanz wird in der Sprachcodierungsliteratur als „scharf“- und „rauh“-klingend bezeichnet. Andere synthetische Sprach-HFR-Verfahren erzeugen sinusförmige Oberschwingungen, die auf einer Grundtonhöhenschätzung basieren und somit auf tonale stationäre Klänge begrenzt sind [US-Patent 4,771,465]. Obwohl dieselben für Sprachanwendungen mit niedriger Qualität nützlich sind, arbeiten derartige bekannte Verfahren nicht bei Sprach- oder Musiksignalen mit hoher Qualität. Einige Verfahren versuchen die Leistung von Audioquellcodecs mit hoher Qualität zu verbessern. Eines verwendet synthetische Rauschsignale, die bei dem Decodierer erzeugt werden, um rauschähnliche Signale bei einer Sprache oder einer Musik zu

substituieren, die vorhergehend durch den Codierer ausgesondert wurden [„Improving Audio Codecs by Noise Substitution“ D. Schultz, JAES, Bd. 44, Nr. 7/8, 1996]. Dies wird innerhalb eines anderweitig normal übertragenen Hochbands auf einer intermittierenden Basis durchgeführt, wenn Rauschsignale vorhanden sind. Ein anderes Verfahren erzeugt einige fehlende Hochbandoberschwingungen wieder, die bei dem Codierungsprozeß verloren wurden [„Audio Spectral Coder“ A. J. S. Ferreira, AES Preprint 4201, 100th Convention, 11.–14. Mai 1996, Kopenhagen], und ist wieder abhängig von tonalen Signalen und einer Tonhöhen Erfassung. Beide Verfahren sind auf einer Basis eines niedrigen Belastungszyklus wirksam, was einen vergleichsweise begrenzten Codierungs- oder Leistungsgewinn bietet.

[0006] Das US-Patent Nr. 5,127,054 offenbart eine Sprachqualitätsverbesserung für Stimmcodierer und -synthesizer. Ein Codierer erzeugt ein codiertes Signal, das wesentliche Spektralinformationen aufweist, die aus dezimierten Abtastwerten der erheblichsten Spektralsubbänder bestehen. Diese Abtastwerte sind auf eine geeignete Weise in eine Mehrzahl von Rekonstruktionsfiltern interpoliert. Das Ausgangssignal der Filter wird summiert und einerseits einem Endsummierer zugeführt und andererseits in ein nicht-lineares Gerät zu einem Erzeugen eines harmonischen Signals eingegeben, das eine Mehrzahl von harmonischen Signalen aufweist, die aus den wesentlichen Spektralinformationen erzeugt werden, die in den ausgewählten Sprachsubbändern enthalten sind. Das harmonische Signal wird zu einer Mehrzahl von Subbandfiltern geleitet, die das harmonische Signal mit den speziellen interessierenden Bändern bandbegrenzen. Die gefilterten harmonischen Signale werden durch eine optionale Gewinnstufe eingestellt.

[0007] Das US-Patent Nr. 4,667,340 offenbart ein Sprachbenachrichtigungssystem mit einer tonhöhenkongruenten Basisbandcodierung. Ein Codierer erzeugt LPC-Parameter plus ein Restsignal, die übertragen werden. Um eine Bandbreite einzusparen, wird das Restsignal gefiltert, so daß lediglich ein Bruchteil der gesamten Bandbreite desselben von dem Codierer zu dem Decodierer übertragen wird. Bei dem Decodierer wird das übertragene Restsignal auf kopiert oder anderweitig zu höheren Frequenzen vergrößert, um das Anregungssignal zu liefern, das gemäß den LPC-Parametern gefiltert wird, um den wiedergebildeten Sprachausgang zu liefern.

[0008] Es ist die Aufgabe der vorliegenden Erfindung, ein Konzept zu einer verbesserten Spektralbandreplikation zu schaffen. Diese Aufgabe wird durch ein Verfahren zum Decodieren eines codierten Signals gemäß Anspruch 1 oder eine Vorrichtung zum Decodieren eines codierten Signals gemäß Anspruch 21 gelöst.

[0009] Die vorliegende Erfindung stellt ein neues Verfahren und eine Vorrichtung zu wesentlichen Verbesserungen von digitalen Quellcodierungssystemen und insbesondere zu den Verbesserungen von Audiocodern bereit. Das Ziel umfaßt eine Bitratenreduzierung oder eine verbesserte Wahrnehmungsqualität oder eine Kombination derselben. Die Erfindung basiert auf neuen Verfahren, die eine harmonische Redundanz verwerten, was die Möglichkeit bietet, Durchlaßbänder eines Signals vor einer Übertragung oder Speicherung auszusondern. Es wird keine Wahrnehmungsverschlechterung wahrgenommen, falls der Decodierer eine Spektralreplikation mit hoher Qualität gemäß der Erfindung durchführt. Die ausgesonderten Bits stellen den Codierungsgewinn bei einer festen Wahrnehmungsqualität dar. Alternativ können mehr Bits zu einem Codieren der Tiefbandinformationen bei einer festen Bitrate zugeteilt werden, wodurch eine höhere Wahrnehmungsqualität erreicht ist.

[0010] Die vorliegende Erfindung postuliert, daß eine abgeschnittene harmonische Reihe basierend auf der direkten Beziehung zwischen Tiefband- und Hochband-Spektralkomponenten erweitert werden kann. Diese erweiterte Reihe gleicht dem Original in einem Wahrnehmungssinn, falls bestimmte Regeln befolgt werden: Erstens müssen die extrapolierten Spektralkomponenten harmonisch auf die abgeschnittene harmonische Reihe bezogen sein, um dissonanzbezogene Artefakte zu vermeiden. Die vorliegende Erfindung verwendet eine Transposition als eine Einrichtung für den Spektralreplikationsprozeß, was sicherstellt, daß dieses Kriterium eingehalten wird. Für eine erfolgreiche Operation ist es jedoch nicht notwendig, daß die Tiefbandspektralkomponenten eine harmonische Reihe bilden, da neue replizierte Komponenten, die harmonisch auf dieselben des Tiefbands bezogen sind, die rauschähnliche oder transiente Beschaffenheit des Signals nicht ändern. Eine Transposition ist als ein Transfer von Partialen von einer Position zu einer anderen auf der Musikskala definiert, während die Frequenzverhältnisse der Partialen beibehalten sind. Zweitens muß die Spektralhüllkurve, d. h. die grobe Spektralverteilung, des replizierten Hochbands derselben des ursprünglichen Signals ziemlich gut gleichen. Die vorliegende Erfindung bietet zwei Betriebsmodi, SBR-1 und SBR-2, die sich in der Weise unterscheiden, auf die die Spektralhüllkurve eingestellt ist.

[0011] SBR-1, für die Verbesserung von Codecanwendungen einer mittleren Qualität vorgesehen, ist ein unsymmetrischer Prozeß, der sich ausschließlich auf die Informationen stützt, die in einem bei dem Decodierer empfangenen Tiefband- oder Tiefpaßsignal enthalten sind. Die Spektralhüllkurve dieses Signals wird bestimmt

und extrapoliert, z. B. unter Verwendung von Polynomen zusammen mit einem Regelsatz oder einem Codebuch. Diese Informationen werden verwendet, um das replizierte Hochband kontinuierlich einzustellen und zu entzerren. Das vorliegende SBR-1-Verfahren bietet den Vorteil eines Nachverarbeitens, d. h. an der Codierseite sind keine Modifizierungen notwendig. Ein Sender wird in einer Kanalbelegung gewinnen oder in der Lage sein, eine verbesserte Wahrnehmungsqualität anzubieten, oder eine Kombination von beiden. Eine existierende Bitstromsyntax und Standards können ohne eine Modifizierung verwendet werden.

[0012] SBR-2, für die Verbesserung von Codecsanwendungen mit hoher Qualität vorgesehen, ist ein symmetrischer Prozeß, bei dem zusätzlich zu dem gesendeten Tiefbandsignal gemäß SBR-1 die Spektralhüllkurve des Hochbands codiert und gesendet wird. Da die Variationen der Spektralhüllkurve eine viel niedrigere Rate aufweisen als die Hochbandsignalkomponenten, muß lediglich eine begrenzte Größe von Informationen gesendet werden, um die Spektralhüllkurve erfolgreich darzustellen. SBR-2 kann verwendet werden, um die Leistung von gegenwärtigen Codetechnologien ohne oder mit kleineren Modifizierungen einer existierenden Syntax oder Protokollen zu verbessern, und als ein wertvolles Werkzeug für eine zukünftige Codecentwicklung.

[0013] Sowohl SBR-1 als auch SBR-2 können verwendet werden, um kleinere Durchlaßbänder des Tiefbands zu replizieren, wenn derartige Bänder durch den Codierer heruntergefahren werden, wie es durch das psychoakustische Modell unter Bitknappen Bedingungen festgesetzt ist. Dies resultiert in einer Verbesserung der Wahrnehmungsqualität durch eine Spektralreplikation innerhalb des Tiefbands zusätzlich zu einer Spektralreplikation außerhalb des Tiefbands. Ferner können SBR-1 und SBR-2 auch bei Codecs verwendet werden, die eine Bitratenskalierbarkeit verwenden, wobei die Wahrnehmungsqualität des Signals bei dem Empfänger abhängig von Sendekanalbedingungen variiert. Dies bringt gewöhnlich lästige Variationen der Audiobandbreite bei dem Empfänger mit sich. Unter derartigen Bedingungen können die SBR-Verfahren erfolgreich verwendet werden, um eine konstant hohe Bandbreite beizubehalten, wobei wiederum die Wahrnehmungsqualität verbessert wird.

[0014] Die vorliegende Erfindung ist auf einer kontinuierlichen Basis wirksam und repliziert einen jeglichen Typ von Signalinhalten, d. h. tonal oder nicht-tonal (rauschähnliche und transiente Signale). Zusätzlich erzeugt das vorliegende Spektralreplikationsverfahren eine wahrnehmungsmäßig genaue Nachbildung der ausgesonderten Bänder aus bei dem Decodierer verfügbaren Frequenzbändern. Somit bietet das SBR-Verfahren einen wesentlich höheren Pegel eines Codiergewinns oder einer Wahrnehmungsqualitätsverbesserung verglichen mit bekannten Verfahren. Die Erfindung kann mit derartigen bekannten Codecverbesserungsverfahren kombiniert werden; aufgrund derartiger Kombinationen wird jedoch kein Leistungsgewinn erwartet.

[0015] Das SBR-Verfahren weist die folgenden Schritte auf:

- Codieren eines Signals, das von einem ursprünglichen Signal abgeleitet ist, wobei Frequenzbänder des Signals ausgesondert werden und das Aussondern vor oder während einem Codieren durchgeführt wird, wobei ein erstes Signal gebildet wird,
- während oder nach einem Decodieren des ersten Signals, Transponieren von Frequenzbändern des ersten Signals, wobei ein zweites Signal gebildet wird.
- Durchführen einer Spektralhüllkurveneinstellung, und
- Kombinieren des decodierten Signals und des zweiten Signals, wobei ein Ausgangssignal gebildet wird.

[0016] Die Paßbänder des zweiten Signals können festgelegt sein, um die Paßbänder des ersten Signals nicht zu überlappen oder teilweise zu überlappen, und können in Abhängigkeit von den zeitlichen Charakteristika des ursprünglichen Signals und/oder des ersten Signals oder Sendekanalbedingungen festgelegt sein. Die Spektralhüllkurveneinstellung wird basierend auf einer Schätzung der ursprünglichen Spektralhüllkurve von dem ersten Signal oder auf gesendeten Hüllkurveninformationen des ursprünglichen Signals durchgeführt.

[0017] Die vorliegende Erfindung umfaßt grundlegende Typen von Transponierern: Mehrbandtransponierer und zeitvariante Mustersuchvoraussagetransponierer, die unterschiedliche Eigenschaften aufweisen. Eine grundlegende Mehrbandtransposition kann gemäß der vorliegenden Erfindung durch das Folgende durchgeführt werden:

- Filtern des zu transponierenden Signals durch einen Satz von $N \geq 2$ Bandpaßfilter mit Durchlaßbändern, die die jeweiligen Frequenzen $[f_1, \dots, f_N]$ aufweisen, wobei N Bandpaßsignale gebildet werden,
- Verschieben der Bandpaßsignale frequenzmäßig in Regionen, die die Frequenzen $M[f_1, \dots, f_N]$ aufweisen, wobei $M \neq 1$ der Transpositionsfaktor ist, und
- Kombinieren der verschobenen Bandpaßsignale, wobei das transponierte Signal gebildet wird.

[0018] Alternativ kann diese grundlegende Mehrbandtransposition gemäß der vorliegenden Erfindung durch

das Folgende durchgeführt werden:

- Bandpaßfiltern des Signals, um ein transponiertes Signal zu sein, unter Verwendung einer Analysefilterbank oder einer Transformation einer solchen Art, daß reell- oder komplexwertige Subbandsignale eines Tiefpaßtyps erzeugt werden,
- eine beliebige Anzahl von Kanälen k der Analysefilterbank oder der Transformation werden mit Kanälen M_k , $M \neq 1$ in einer Synthesefilterbank oder einer Transformation verbunden, und
- das transponierte Signal wird unter Verwendung der Synthesefilterbank oder der Transformation gebildet.

[0019] Eine verbesserte Mehrbandtransposition gemäß der Erfindung umfaßt Phaseneinstellungen, wobei die Leistung der grundlegenden Mehrbandtranspositionen verbessert wird.

[0020] Die zeitvariante Mustersuchvoraussagetransposition gemäß der vorliegenden Erfindung kann durch das Folgende durchgeführt werden:

- Durchführen einer Transientenerfassung bei dem ersten Signal,
- Bestimmen, welches Segment des ersten Signals verwendet werden soll, wenn Teile des ersten Signals dupliziert/ausgesondert werden, abhängig von dem Ergebnis der Transientenerfassung,
- Einstellen von Zustandsvektor- und Codebucheigenschaften abhängig von dem Ergebnis der Transientenerfassung, und
- Suchen nach Synchronisationspunkten bei einem gewählten Segment des ersten Signals basierend auf dem Synchronisationspunkt, der bei der vorhergehenden Synchronisationspunktsuche gefunden wurde.

[0021] Die SBR-Verfahren und -Vorrichtungen gemäß der vorliegenden Erfindung bieten die folgenden Merkmale:

1. Die Verfahren und Vorrichtungen werten neue Konzepte einer Signalredundanz in dem Spektralbereich aus.
2. Die Verfahren und Vorrichtungen sind auf beliebige Signale anwendbar.
3. Jeder harmonische Satz ist einzeln erzeugt und gesteuert.
4. Alle replizierten Oberschwingungen sind auf eine derartige Weise erzeugt, um eine Fortführung der existierenden harmonischen Reihe zu bilden.
5. Der Spektralreplikationsprozeß basiert auf einer Transposition und erzeugt keine oder nicht wahrnehmbare Artefakte.
6. Die Spektralreplikation kann mehrere kleinere Bänder und/oder einen breiten Frequenzbereich abdecken.
7. Bei dem SBR-1-Verfahren wird die Verarbeitung lediglich an der Decodierseite durchgeführt, d. h. alle Standards und Protokolle können ohne eine Modifizierung verwendet werden.
8. Das SBR-2-Verfahren kann gemäß den meisten Standards und Protokollen ohne oder mit kleineren Modifizierungen implementiert werden.
9. Das SBR-2-Verfahren bietet dem Codec-Entwickler ein neues leistungsstarkes Komprimierungswerkzeug.
10. Der Codiergewinn ist erheblich.

[0022] Die anziehendste Anmeldung bezieht sich auf die Verbesserung von verschiedenen Typen von Codecs mit niedriger Bitrate, wie beispielsweise MPEG 1/2 Layer I/II/III [US-Patent. 5,040,217], MPEG 2/4 AAC, Dolby AC-2/3, NTT Twin VQ [US-Patent 5,684,920], AT & T/Lucent PAC etc. Die Erfindung ist auch bei Sprachcodecs mit hoher Qualität nützlich, wie beispielsweise Breitband-CELP und SB-ADPCM G.722, um eine Wahrnehmungsqualität zu verbessern. Die obigen Codecs sind bei Multimedia, in der Telefonindustrie, im Internet sowie bei professionellen Anwendungen in breiter Verwendung. T-DAB- (Terrestrial Digital Audio Broadcasting) Systeme verwenden Protokolle mit niedriger Bitrate, die durch ein Verwenden des vorliegenden Verfahrens in einer Kanalbelegung gewinnen oder eine Qualität bei FM und AM DAB verbessern. Satelliten-S-DAB gewinnt aufgrund der betroffenen übermäßigen Systemkosten durch ein Verwenden des vorliegenden Verfahrens, um die Anzahl von Programmkanälen in dem DAB-Multiplex zu erhöhen, erheblich. Ferner ist zum ersten Mal ein Audio-Streamen mit voller Bandbreite in Echtzeit über das Internet unter Verwendung von Telefonmodems mit niedriger Bitrate erreichbar.

[0023] Die vorliegende Erfindung wird nun durch darstellende Beispiele mit Bezug auf die zugehörigen Zeichnungen beschrieben, in denen:

[0024] Fig. 1 eine SBR darstellt, die in einem Codiersystem gemäß der vorliegenden Erfindung enthalten ist;

[0025] Fig. 2 eine Spektralreplikation von oberen Oberschwingungen gemäß der vorliegenden Erfindung dar-

stellt;

[0026] Fig. 3 eine Spektralreplikation von Inbandoberschwingungen gemäß der vorliegenden Erfindung darstellt;

[0027] Fig. 4 ein Blockdiagramm für eine Zeit-Bereich-Implementierung eines Transponierers gemäß der vorliegenden Erfindung ist;

[0028] Fig. 5 ein Flußdiagramm ist, das einen Betriebszyklus für den Mustersuchvorhersagetransponierer gemäß der vorliegenden Erfindung darstellt;

[0029] Fig. 6 ein Flußdiagramm ist, das die Suche nach einem Synchronisationspunkt gemäß der vorliegenden Erfindung darstellt;

[0030] Fig. 7a–7b das Codebuch-Positionieren während Transienten gemäß der vorliegenden Erfindung darstellt;

[0031] Fig. 8 ein Blockdiagramm für eine Implementierung von mehreren Zeit-Bereich-Transponierern in Verbindung mit einer geeigneten Filterbank für eine SBR-Operation gemäß der vorliegenden Erfindung ist;

[0032] Fig. 9a–9c Blockdiagramme sind, die ein Gerät für eine STFT-Analyse und -Synthese darstellen, konfiguriert zu einer Erzeugung von Oberschwingungen zweiter Ordnung, gemäß der vorliegenden Erfindung;

[0033] Fig. 10a–10b Blockdiagramme eines Subbands mit einer linearen Frequenzverschiebung bei dem STFT-Gerät gemäß der vorliegenden Erfindung sind;

[0034] Fig. 11 ein Subband zeigt, das einen Phasenmultiplizierer gemäß der vorliegenden Erfindung verwendet;

[0035] Fig. 12 darstellt, wie Oberschwingungen dritter Ordnung gemäß der vorliegenden Erfindung erzeugt werden;

[0036] Fig. 13 darstellt, wie Oberschwingungen zweiter und dritter Ordnung simultan gemäß der vorliegenden Erfindung erzeugt werden;

[0037] Fig. 14 eine Erzeugung einer nicht-überlappenden Kombination von mehreren Oberschwingungsordnungen gemäß der vorliegenden Erfindung darstellt;

[0038] Fig. 15 eine Erzeugung einer verschachtelten Kombination von mehreren Oberschwingungsordnungen gemäß der vorliegenden Erfindung darstellt;

[0039] Fig. 16 eine Erzeugung von linearen Breitbandfrequenzverschiebungen darstellt;

[0040] Fig. 17 darstellt, wie Suboberschwingungen gemäß der vorliegenden Erfindung erzeugt werden;

[0041] Fig. 18a–18b Blockdiagramme eines Wahrnehmungscodecs sind;

[0042] Fig. 19 eine Grundstruktur einer maximal dezimierten Filterbank zeigt;

[0043] Fig. 20 eine Erzeugung von Oberschwingungen zweiter Ordnung in einer maximal dezimierten Filterbank gemäß der vorliegenden Erfindung darstellt;

[0044] Fig. 21 ein Blockdiagramm für die verbesserte Mehrbandtransposition bei einer maximal dezimierten Filterbank ist, die an Subbandsignalen wirksam ist, gemäß der vorliegenden Erfindung;

[0045] Fig. 22 ein Flußdiagramm ist, das die verbesserte Mehrbandtransposition bei einer maximal dezimierten Filterbank darstellt, die an Subbandsignalen wirksam ist, gemäß der vorliegenden Erfindung;

[0046] Fig. 23 Subbandabtastwerte und Skalierungsfaktoren eines typischen Codec darstellt;

[0047] Fig. 24 Subbandabtastrwerte und Hüllkurveninformationen für SBR-2 gemäß der vorliegenden Erfindung darstellt;

[0048] Fig. 25 eine versteckte Übertragung von Hüllkurveninformationen bei SBR-2 gemäß der vorliegenden Erfindung darstellt;

[0049] Fig. 26 ein Redundanzcodieren bei SBR-2 gemäß der vorliegenden Erfindung darstellt;

[0050] Fig. 27 eine Implementierung eines Codec unter Verwendung des SBR-1-Verfahrens gemäß der vorliegenden Erfindung darstellt;

[0051] Fig. 28 eine Implementierung eines Codec unter Verwendung des SBR-2-Verfahrens gemäß der vorliegenden Erfindung darstellt; und

[0052] Fig. 29 ein Blockdiagramm eines „Pseudo-Stereo“-Generators gemäß der vorliegenden Erfindung ist.

[0053] Überall in der Erläuterung der Ausführungsbeispiele werden hierin Natürliche-Audioquelle-Codieranwendungen betont. Es ist jedoch klar, daß die vorliegende Erfindung auf einen anderen Bereich von Quellcodierungsanwendungen als auf denselben eines Codierens und Decodierens von Audiosignalen anwendbar ist.

Transpositionsgrundlagen

[0054] Eine Transposition, wie dieselbe gemäß der vorliegenden Erfindung definiert ist, ist das ideale Verfahren zu einer Spektralreplikation und weist mehrere größere Vorteile gegenüber dem Stand der Technik auf, wie beispielsweise: es ist keine Tonhöhenfassung erforderlich, es wird eine gleich hohe Leistung für einstimmiges und polyphones Programmaterial erhalten und die Transposition arbeitet gleich gut bei tonalen und nicht-tonalen Signalen. Im Gegensatz zu anderen Verfahren kann die Transposition gemäß der Erfindung bei beliebigen Audioquellcodierungssystemen für beliebige Signaltypen verwendet werden.

[0055] Eine exakte Transposition eines Faktors M eines diskreten Zeitsignals $x(n)$ in der Form einer Summe von Kosinussen mit zeitvariablen Amplituden ist definiert durch die Beziehung

$$x(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi f_i n / f_s + \alpha_i) \rightarrow \quad (1)$$

$$y(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi M f_i n / f_s + \beta_i) \quad (2)$$

wobei N die Anzahl von Sinuskurven ist, hierin im folgenden als Partiale bezeichnet, f_i , $e_i(n)$, α_i die einzelnen Eingangsfrequenzen, Zeithüllkurven bzw. Phasenkonstanten sind, β_i die beliebigen Ausgangsphasenkonstanten sind und f_s die Abtastfrequenz ist, und $0 \leq M f_i \leq f_s/2$.

[0056] In Fig. 2 ist die Erzeugung von Oberschwingungen der M-ten Ordnung gezeigt, wobei M eine Ganzzahl ≥ 2 ist. Der Ausdruck Oberschwingungen M-ter Ordnung wird zu einer Einfachheit verwendet, obgleich der Prozeß Oberschwingungen M-ter Ordnung zu allen Signalen in einer bestimmten Frequenzregion erzeugt, die in den meisten Fällen selbst Oberschwingungen einer unbekannten Ordnung sind. Das Eingangssignal mit der Frequenzbereichsdarstellung $X(f)$ ist auf den Bereich 0 bis f_{\max} bandbegrenzt **201**. Der Signalgehalt in dem Bereich f_{\max}/M bis $Q f_{\max}/M$, wobei Q der erwünschte Bandbreitenexpansionsfaktor $1 < Q \leq M$ ist, wird mittels eines Bandpaßfilters extrahiert, wobei ein Bandpaßsignal mit einem Spektrum $X_{BP}(f)$ gebildet wird **203**. Das Bandpaßsignal wird einen Faktor M transponiert, wobei ein zweites Bandpaßsignal mit einem Spektrum $X_T(f)$, das den Bereich $Q f_{\max}$ bis $Q f_{\max}$ abdeckt, gebildet wird **205**. Die Spektralhüllkurve dieses Signals wird mittels eines programmgesteuerten Entzerrers eingestellt, wobei ein Signal mit einem Spektrum $X_E(f)$ gebildet wird **207**. Dieses Signal wird dann mit einer verzögerten Version des Eingangssignals kombiniert, um die Verzögerung zu kompensieren, die durch das Bandpaßfilter und den Transponierer auferlegt wurde, wodurch ein Ausgangssignal mit einem Spektrum $Y(f)$, das den Bereich 0 bis $Q f_{\max}$ abdeckt, gebildet wird **209**. Alternativ kann ein Bandpaßfiltern nach der Transposition M unter Verwendung von Grenzfrequenzen f_{\max} und $Q f_{\max}$ durchgeführt werden. Durch ein Verwenden von mehreren Transponierern ist natürlich eine simultane Erzeugung von unterschiedlichen Oberschwingungsordnungen möglich. Das obige Schema kann auch verwendet werden, um Sperrbänder innerhalb des Eingangssignals „auszufüllen“, wie es in Fig. 3 gezeigt ist, wo das Eingangssignal

ein Sperrband aufweist, das sich von f_0 bis Qf_0 erstreckt **301**. Ein Paßband $[f_0/M, Qf_0/M]$ wird dann extrahiert **303**, einen Faktor M zu $[f_0, Qf_0]$ transponiert **305**, hüllkurveneingestellt **307** und mit dem verzögerten Eingangssignal kombiniert, wobei das Ausgangssignal mit einem Spektrum $Y(f)$ gebildet wird **309**.

[0057] Eine Annäherung einer exakten Transposition kann verwendet werden. Gemäß der vorliegenden Erfindung ist die Qualität derartiger Annäherungen unter Verwendung einer Dissonanztheorie bestimmt. Ein Kriterium für eine Dissonanz ist durch Plomp präsentiert [„Tonal Consonance and Critical Bandwidth“ R. Plomp, W. J. M. Levelt JASA, Bd. 38, 1965] und legt dar, daß zwei Partiale als Dissonanz betrachtet werden, falls der Frequenzunterschied innerhalb näherungsweise 5 bis 50% der Bandbreite des kritischen Bands gelegen ist, in dem die Partiale positioniert sind. Zu einer Bezugnahme kann die kritische Bandbreite für eine gegebene Frequenz angenähert werden durch

$$cb(f) = 25 + 75 \left(1 + 1,4 \left(\frac{f}{1000} \right)^2 \right)^{0,69} \quad (3)$$

mit f und cb in Hz. Ferner legt Plomp dar, daß das menschliche Hörsystem zwei Partiale nicht unterscheiden kann, falls dieselben sich in einer Frequenz um näherungsweise weniger als 5% der kritischen Bandbreite unterscheiden, in der dieselben positioniert sind. Die exakte Transposition in Gl. 2 wird angenähert durch

$$y_{approx}(n) = \sum_{i=0}^{N-1} e_i(n) \cos(2\pi(Mf_i \pm \Delta f_i)n / f_s + \beta_i) \quad (4)$$

wobei Δf_i die Abweichung von der exakten Transposition ist. Falls die Eingangspartiale eine harmonische Reihe bilden, legt eine Hypothese der Erfindung dar, daß die Abweichungen von der harmonischen Reihe der transponierten Partialen 5% der kritischen Bandbreite nicht überschreiten dürfen, in der dieselben positioniert sind. Dies würde erklären, warum bekannte Verfahren unzufriedenstellende „scharfe“ und „rauhe“ Ergebnisse ergeben, da lineare Breitbandfrequenzverschiebungen eine viel größere Abweichung ergeben, als es akzeptabel ist. Wenn bekannte Verfahren mehr als eine Partiale für lediglich eine Eingangspartiale erzeugen, müssen die Partialen dennoch innerhalb der oben dargelegten Abweichungsbegrenzung liegen, um als eine Partiale wahrgenommen zu werden. Dies erklärt wiederum die dürrtigen Ergebnisse, die mit bekannten Verfahren unter Verwendung von Nichtlinearitäten etc. erhalten werden, da dieselben Intermodulationspartiale erzeugen, die nicht innerhalb der Abweichungsbegrenzung liegen.

[0058] Wenn das obige transpositionsbasierte Verfahren einer Spektralreplikation gemäß der vorliegenden Erfindung verwendet wird, werden die folgenden wichtigen Eigenschaften erreicht.

- Normalerweise tritt zwischen replizierten Oberschwingungen und existierenden Partialen keine Frequenzbereichsüberlappung auf.
- Die replizierten Partiale sind harmonisch mit den Partialen des Eingangssignals verwandt und rufen keine störende Dissonanz oder Artefakte hervor.
- Die Spektralhüllkurve der replizierten Oberschwingungen bildet eine glatte Fortführung der Eingangssignal-Spektralhüllkurve und stimmt wahrnehmungsmäßig mit der ursprünglichen Hüllkurve überein.

Transposition basierend auf einer zeitvarianten Mustersuchvoraussage

[0059] Es existieren verschiedene Wege, um die erforderlichen Transponierer zu entwerfen. Typische Zeit-Bereich-Implementierungen erweitern das Signal in einer Zeit durch ein Duplizieren von Signalsegmenten basierend auf der Tonhöhenperiode. Dieses Signal wird nachfolgend bei einer unterschiedlichen Rate ausgelesen. Leider sind derartige Verfahren streng von einer Tonhöhenenerfassung zu einem genauen Zeitspleißen der Signalsegmente abhängig. Ferner macht die Beschränkung, an Tonhöhenperiode-basierten Signalsegmenten zu arbeiten, dieselben empfindlich für Transienten. Da die erfaßte Tonhöhenperiode viel länger sein kann als die tatsächliche Transiente, ist das Risiko des Duplizierens der gesamten Transiente und nicht nur eines Expandierens derselben in einer Zeit offensichtlich. Ein anderer Typ von Zeit-Bereich-Algorithmen erhält eine Zeitausdehnung/-komprimierung von Sprachsignalen durch ein Verwenden einer Mustersuchvoraussage des Ausgangssignals [„Pattern Search Prediction of Speech“ R. Bogner, T. Li, Proc. ICASSP '89, Bd. 1, Mai 1989, „Time-Scale Modification of Speech based on a nonlinear Oscillator Model“ G. Kubin, W. B. Kleijn, IEEE, 1994]. Dies ist eine Form einer granularen Synthese, bei der das Eingangssignal in kleine Teile, Granalien, geteilt wird, die verwendet werden, um das Ausgangssignal zu generieren. Diese Synthese wird gewöhnlich durch ein Durchführen einer Korrelation von Signalsegmenten vorgenommen, um die besten Spleißpunkte zu bestimmen. Dies bedeutet, daß die Segmente, die verwendet werden, um das Ausgangssignal zu bilden, nicht von der Tonhöhenperiode abhängig sind, und somit die nicht-triviale Aufgabe einer Abstandserfassung nicht

erforderlich ist. Dennoch bleiben Probleme bei sich schnell ändernden Signalamplituden bei diesen Verfahren und eine Transposition mit hoher Qualität neigt dazu, einen hohen Berechnungsbedarf hervorzurufen. Es wird jedoch nun ein verbesserter Zeit-Bereich-Abstandsverschieber/-transponierer vorgelegt, bei dem die Verwendung einer Transientenerfassung und dynamischen Systemparametern eine genauere Transposition für hohe Transpositions faktoren während sowohl stationären (tonal oder nicht tonal) als auch transienten Klängen bei niedrigen Berechnungskosten erzeugt.

[0060] Unter Bezugnahme auf die Zeichnungen, bei denen gleiche Bezugszeichen gleiche Elemente angeben, sind in **Fig. 4** neun getrennte Module gezeigt: ein Transientendetektor **401**, ein Fensterpositionseinsteller **403**, ein Codebuchgenerator **405**, ein Synchronisationssignalselektor **407**, ein Synchronisationspositionsspeicher **409**, ein Minimaldifferenzestimator **411**, ein Ausgangssegmentspeicher **413**, eine Mischeinheit **415** und ein Abwärtsabtaster **417**. Das Eingangssignal wird sowohl dem Codebuchgenerator **405** als auch dem Transientendetektor **401** zugeführt. Falls eine Transiente erfaßt wird, wird die Position derselben zu dem Fensterpositionsmodul **403** gesendet. Dieses Modul setzt die Größe und Position des Fensters fest, das bei einem Erzeugen des Codebuchs mit dem Eingangssignal multipliziert wird. Der Codebuchgenerator **405** empfängt eine Synchronisationsposition von dem Synchronisationsauswahlmodul **407**, vorausgesetzt, dasselbe wurde mit einem anderen Transponierer verbunden. Falls diese Synchronisationsposition innerhalb des Codebuchs liegt, wird dieselbe verwendet und ein Ausgangssegment wird erzeugt. Andernfalls wird das Codebuch zu dem Minimaldifferenzestimator **411** gesendet, der eine neue Synchronisationsposition zurückgibt. Das neue Ausgangssegment wird mit dem vorhergehenden Ausgangssegment in dem Mischmodul **415** zusammengefenstert und nachfolgend in dem Modul **417** abwärtsabgetastet.

[0061] Um die Erklärung zu verdeutlichen, ist eine Zustand-Raum-Darstellung eingebracht. Hier stellen die Zustandsvektoren oder Granalien die Eingangs- und Ausgangssignale dar. Das Eingangssignal ist durch einen Zustandsvektor $x(n)$ dargestellt:

$$x(n) = [x(n), x(n - D), x(n - 2D), \dots, x(n - (N - 1)D)] \quad (5)$$

der aus N verzögerten Abtastwerten des Eingangssignals erhalten wird, wobei N die Abmessung des Zustandsvektors ist und D die Verzögerung zwischen den Eingangsabtastwerten ist, die verwendet werden, um den Vektor zu bauen. Die granulare Abbildung ergibt den Abtastwert $x(n)$, der jedem Zustandsvektor $x(n - 1)$ folgt. Dies ergibt Gl. 6, wobei $a(\bullet)$ die Abbildung ist:

$$x(n) = a(x(n - 1)) \quad (6)$$

[0062] Bei dem vorliegenden Verfahren wird die granulare Abbildung verwendet, um den nächsten Ausgang basierend auf dem vorigen Ausgang zu bestimmen, wobei ein Zustandsübergang-Codebuch verwendet wird. Das Codebuch einer Länge L wird kontinuierlich wieder aufgebaut und enthält die Zustandsvektoren und den nächsten Abtastwert, der jedem Zustandsvektor folgt. Jeder Zustandsvektor ist von dem Nachbar desselben durch K Abtastwerte getrennt; dies ermöglicht es, daß das System die Zeitauflösung abhängig von den Charakteristika des aktuell verarbeiteten Signals einstellt, wobei K gleich 1 die feinste Auflösung darstellt. Das Eingangssignalsegment, das verwendet wird, um das Codebuch zu bauen, ist basierend auf der Position einer möglichen Transiente und der Synchronisationsposition bei dem vorhergehenden Codebuch gewählt.

[0063] Dies bedeutet, daß die Abbildung $a(\bullet)$ theoretisch für alle in dem Codebuch enthaltenen Übergänge ausgewertet wird:

$$a \left(\begin{bmatrix} x(n - L) \\ x(n - L + K) \\ \vdots \\ x(n - 1) \end{bmatrix} \right) = \begin{bmatrix} x(n - L + 1) \\ x(n - L + K + 1) \\ \vdots \\ x(n) \end{bmatrix} \quad (7)$$

[0064] Bei diesem Übergangscodbuch wird der neue Eingang $y(n)$ durch ein Suchen nach dem Zustandsvektor in dem Codebuch berechnet, der dem aktuellen Zustandsvektor $y(n - 1)$ am ähnlichst ist. Diese Nächster-Nachbar-Suche wird durch ein Berechnen der minimalen Differenz vorgenommen und ergibt den neuen Ausgangsabtastwert:

$$y(n) = a(y(n - 1)) \quad (8)$$

[0065] Das System ist jedoch nicht begrenzt, um auf einer Basis Abtastwert für Abtastwert zu arbeiten, sondern ist vorzugsweise auf einer Basis Segment für Segment betrieben. Das neue Ausgangssegment wird gefenstert und addiert, mit dem vorhergehenden Ausgangssegment gemischt und nachfolgend nach abwärtsabgetastet. Der Tonhöhentranspositionsfaktor ist durch das Verhältnis der Eingangssegmentlänge, die durch das Codebuch dargestellt ist, und der Ausgangssegmentlänge bestimmt, die aus dem Codebuch gelesen wird.

[0066] Unter erneuter Bezugnahme auf die Zeichnungen sind in **Fig. 5** und **Fig. 6** Flußdiagramme vorgelegt, die den Betriebszyklus des Transponierers anzeigen. Bei **501** werden die Eingangsdaten dargestellt, eine Transientenerfassung **503** wird an einem Segment des Eingangssignals durchgeführt; die Suche nach Transienten wird an einer Signallänge durchgeführt, die gleich der Ausgangssignallänge ist. Falls eine Transiente gefunden wird **505**, wird die Position der Transiente gespeichert **507** und die Parameter L (der die Codebuchlänge darstellt), K (der den Abstand in Abtastwerten zwischen jedem Zustandsvektor darstellt) und D (der die Verzögerung zwischen Abtastwerten bei jedem Zustandsvektor darstellt) werden eingestellt **509**. Die Position der Transiente wird mit der Position des vorhergehenden Ausgangssegments verglichen **511**, um zu bestimmen, ob die Transiente verarbeitet wurde. Falls dem so ist **513**, werden die Position des Codebuchs (Fenster L) und die Parameter K, L und D eingestellt **515**. Nach den notwendigen Parametereinstellungen findet basierend auf dem Ergebnis der Transientenerfassung die Suche nach einem neuen Synchronisations- oder Spleißpunkt statt **517**. Diese Prozedur ist in **Fig. 6** angezeigt. Zuerst wird basierend auf dem vorhergehenden ein neuer Synchronisationspunkt berechnet **601**, gemäß:

$$\text{Sync_pos} = \text{Sync_pos_old} + S \cdot M - S, \quad (9)$$

wobei Sync_pos und Sync_pos_old die neue bzw. die alte Synchronisationsposition sind, S die Länge des Eingangssegments ist, das verarbeitet wird, und M der Transpositionsfaktor ist. Dieser Synchronisationspunkt wird verwendet, um die Genauigkeit des neuen Spleißpunkts mit der Genauigkeit des alten Spleißpunkts zu vergleichen **603**. Falls die Übereinstimmung so gut wie oder besser als die vorhergehende ist **605**, wird dieser neue Synchronisationspunkt zurückgegeben **607**, vorausgesetzt, derselbe befindet sich innerhalb des Codebuchs. Falls nicht, wird in der Schleife nach einem neuen Synchronisationspunkt gesucht **609**. Dies wird mit einem Ähnlichkeitsmaß, bei diesem Fall einer Minimaldifferenzfunktion, durchgeführt **611**, es ist jedoch auch möglich, eine Korrelation bei dem Zeit- oder Frequenzbereich zu verwenden. Falls die Position eine bessere Übereinstimmung ergibt, als dieselbe der vorhergehend gefundenen Position **613**, wird die Synchronisationsposition gespeichert **615**. Wenn alle Positionen ausprobiert sind **617**, kehrt das System zu dem Flußdiagramm in **Fig. 5** zurück **619**. Der erhaltene neue Synchronisationspunkt wird gespeichert **519** und ein neues Segment wird von dem Codebuch beginnend bei dem gegebenen Synchronisationspunkt ausgelesen **521**. Dieses Segment wird gefenstert und zu dem vorhergehenden addiert **523**, durch den Transpositionsfaktor abwärtsabgetastet **525** und in dem Ausgangspuffer gespeichert **527**.

[0067] In **Fig. 7** ist das Verhalten des Systems unter transienten Bedingungen mit Bezug auf die Position des Codebuchs dargestellt. Vor der Transiente ist das Codebuch 1, das das Eingangssegment 1 darstellt, „zu der Linken“ des Segments 1 positioniert. Das Korrelationssegment 1 stellt einen Teil des vorhergehenden Ausgangs dar und wird verwendet, um einen Synchronisationspunkt 1 in dem Codebuch 1 zu finden. Wenn die Transiente erfaßt ist und der Punkt der Transiente verarbeitet ist, wird das Codebuch gemäß **Fig. 7a** bewegt und ist stationär, bis das aktuell verarbeitete Eingangssegment sich wieder „zu der Rechten“ des Codebuchs befindet. Dies macht es unmöglich, die Transiente zu duplizieren, da das System vor der Transiente nicht nach Synchronisationspunkten suchen darf.

[0068] Die meisten Tonhöhentransponierer oder Zeitexpandierer ergeben basierend auf einer Mustersuchvorraussage zufriedenstellende Ergebnisse bei einer Sprache und bei Einzeltonhöhenmaterial. Die Leistung derselben verschlechtert sich jedoch bei Signalen hoher Komplexität, wie beispielsweise Musik insbesondere bei großen Transpositionsfaktoren schnell. Die vorliegende Erfindung bietet mehrere Lösungen zu einer verbesserten Leistung und erzeugt daher hervorragende Ergebnisse für einen jeglichen Typ eines Signals. Entgegen anderen Entwürfen ist das System zeitvariant und die Systemparameter basieren auf den Eigenschaften des Eingangssignals und den während des vorhergehenden Betriebszyklus verwendeten Parametern. Die Verwendung eines Transientendetektors, der nicht nur die Codebuchgröße und -position steuert, sondern auch die Eigenschaften der enthaltenen Zustandsvektoren, ist ein sehr robustes und berechnungsmäßig effizientes Verfahren, um eine hörbare Verschlechterung während sich schnell verändernden Signalsegmenten zu vermeiden. Ferner ist eine Änderung der Länge des Signalsegments, das verarbeitet wird, was einen höheren Berechnungsbedarf hervorrufen würde, nicht erforderlich. Ferner verwendet die vorliegende Erfindung eine verbesserte Codebuchsuche basierend auf den Ergebnissen auf der vorhergehenden Suche. Dies bedeutet, daß entgegen einer gewöhnlichen Korrelation von zwei Signalsegmenten, wie es gewöhnlich bei Zeit-Bereich-Sys-

temen vorgenommen wird, die auf einer Mustersuchvoraussage basieren, die wahrscheinlichsten Synchronisationspositionen zuerst ausprobiert werden, anstelle eines aufeinanderfolgenden Ausprobierens aller Positionen. Dieses neue Verfahren zu einem Reduzieren der Codebuchsuche reduziert die Berechnungskomplexität des Systems drastisch. Wenn mehrere Transponierer verwendet werden, können ferner Synchronisationspositionsinformationen unter den Transponierern zu einer weiteren Reduzierung der Rechenkomplexität gemeinschaftlich verwendet werden, wie es in der folgenden Implementierung gezeigt ist.

[0069] Die Zeit-Bereich-Transponierer, wie dieselben oben erläutert sind, werden verwendet, um die SBR-1- und SBR-2-Systeme gemäß dem folgenden darstellenden aber nicht begrenzenden Beispiel zu verwenden. In **Fig. 8** werden drei Ausdehnungsmodule verwendet, um Oberschwingungen zweiten, dritten und vierten Grades zu erzeugen. Da bei diesem Beispiel jeder Zeit-Bereich-Expander/-Transponierer auf einem Breitbandsignal arbeitet, ist es vorteilhaft, die Spektralhüllkurve des Quellfrequenzbereichs vor einer Transposition einzustellen, unter der Betrachtung, daß es nach den Transpositionen keine Einrichtung geben wird, um dies zu tun, ohne ein getrenntes Entzerrersystem hinzuzufügen. Die Spektralhüllkurveinsteller **801**, **803** und **805** arbeiten jeweils auf mehreren Filterbankkanälen. Der Gewinn jedes Kanals bei den Hüllkurveinstellern muß festgelegt sein, so daß die Summe **813**, **815**, **817** bei dem Ausgang nach einer Transposition die erwünschte Spektralhüllkurve ergibt. Die Transponierer **807**, **809** und **811** sind verbunden, um Synchronisationspositionsinformationen gemeinschaftlich zu verwenden. Dies basiert auf der Tatsache, daß unter bestimmten Bedingungen eine hohe Korrelation zwischen den Synchronisationspositionen auftritt, die während einer Korrelation in den getrennten Transponiereinheiten in dem Codebuch gefunden werden. Man nehme als ein Beispiel und wieder ohne ein Begrenzen des Schutzbereichs der Erfindung an, der Transponierer der Oberschwingung vierter Ordnung arbeitet auf einer Zeitrahmenbasis, die die Hälfte derselben des Transponierers der Oberschwingung der zweiten Ordnung ist, aber mit dem doppelten Belastungszyklus. Man nehme ferner an, daß die Codebücher, die für die zwei Expander verwendet werden, die gleichen sind und daß die Synchronisationspositionen der zwei Zeit-Bereich-Expander `sync_pos4` bzw. `sync_pos2` etikettiert sind. Dies ergibt die folgende Beziehung:

$$\text{sync_pos2} = \text{sync_pos4} - n \cdot 4 \cdot S - \text{sync_offset}, \text{ für } n = 1, 2, 3, 4 \quad (10)$$

wobei

$$\text{sync_offset} = \text{sync_pos4} - \text{sync_pos2}, \text{ für } n = 0 \quad (11)$$

und S die Länge des Eingangssegments ist, das durch das Codebuch dargestellt ist. Dies ist gültig, solange keiner der Synchronisationspositionszeiger das Ende des Codebuchs erreicht. Während eines normalen Betriebs wird n für jeden Zeitrahmen um 1 erhöht, der durch den Transponierer der Oberschwingung zweiter Ordnung verarbeitet wird, und wenn das Codebuchende durch einen der Zeiger unausweichlich erreicht ist, wird der Zähler n auf $n = 0$ gesetzt und `sync_pos2` und `sync_pos4` werden einzeln berechnet. Ähnliche Ergebnisse werden für den Transponierer der Oberschwingung dritter Ordnung erhalten, wenn derselbe mit dem Transponierer der Oberschwingung vierter Ordnung verbunden ist.

[0070] Die oben vorgelegte Verwendung von mehreren verbundenen Zeit-Bereich-Transponierern zu der Erzeugung von Oberschwingungen höherer Ordnung bringt eine wesentliche Berechnungsreduzierung ein. Ferner präsentiert die vorgeschlagene Verwendung von Zeit-Bereich-Transponierern in Verbindung mit einer geeigneten Filterbank die Möglichkeit, die Hüllkurve des erzeugten Spektrums einzustellen, während die Einfachheit und die geringen Berechnungskosten eines Zeit-Bereich-Transponierers beibehalten sind, da diese mehr oder weniger unter Verwendung einer Festkommaarithmetik und nur additiven/subtraktiven Operationen implementiert sein können.

[0071] Andere darstellende aber nicht begrenzende Beispiele der vorliegenden Erfindung sind:

- die Verwendung eines Zeit-Bereich-Transponierers innerhalb jedes Subbands in einer Subbandfilterbank, wobei so die Signalkomplexität für jeden Transponierer reduziert ist.
- die Verwendung eines Zeit-Bereich-Transponierers in Kombination mit einem Frequenz-Bereich-Transponierer, wobei es so ermöglicht ist, daß das System unterschiedliche Verfahren zu einer Transposition verwendet, abhängig von den Charakteristika des Eingangssignals, das verarbeitet wird.
- die Verwendung eines Zeit-Bereich-Transponierers bei einem Breitband-Sprachcodec, der z. B. an dem Restsignal wirksam ist, das nach einer linearen Voraussage erhalten wird.

[0072] Es ist zu erkennen, daß das oben umrissene Verfahren lediglich durch ein Weglassen der Abtastratenumwandlung vorteilhaft zu einer Zeitskalierungsmodifikation verwendet werden kann. Es ist ferner klar, daß, obwohl das umrissene Verfahren auf ein Tonhöhentransponieren zu einer höheren Tonhöhe fokussiert, d. h.

eine Zeitausdehnung, die gleichen Prinzipien gelten, wenn zu einer niedrigeren Tonhöhe transponiert wird, d. h. eine Zeitkomprimierung, wie es Fachleuten auf dem Gebiet offensichtlich ist.

Filterbankbasierte Transposition

[0073] Verschiedene neue und innovative filterbankbasierte Transpositionstechniken werden nun beschrieben. Das zu transponierende Signal ist in eine Reihe von BP- oder Subbandsignalen geteilt. Die Subbandsignale werden dann exakt oder näherungsweise transponiert, was vorteilhafterweise durch eine Wiederverbindung von Analyse- und Synthesesubbändern erzielt wird, was hierin im folgenden als ein „Patch“ bezeichnet wird. Das Verfahren ist zuerst unter Verwendung einer Kurzzeit-Fourier-Transformation (STFT = Short Time Fourier Transform) gezeigt.

[0074] Der N-Punkt STFT eines Diskretzeitsignals $x(n)$ ist definiert durch

$$X_k(n) = \sum_{p=-\infty}^{\infty} x(p)h(n-p)e^{-j\omega_k p} \quad (12)$$

wobei $k = 0, 1, \dots, N-1$ und $\omega_k = 2\pi k/N$ und $h(n)$ ein Fenster ist. Falls das Fenster die folgenden Bedingungen erfüllt

$$\begin{cases} h(0) = 1 \\ h(n) = 0 \text{ für } n = \pm N, \pm 2N, \pm 3N, \dots \end{cases} \quad (13)$$

existiert eine inverse Transformation und ist gegeben durch

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_k(n)e^{j\omega_k n} \quad (14)$$

[0075] Die direkte Transformation kann als ein Analysator interpretiert werden, siehe **Fig. 9a**, der eine Bank von N BP-Filtern mit Impulsantworten $h(n)\exp(j\omega_k n)$ **901**, gefolgt durch eine Bank von N Multiplizierern mit Trägern $\exp(-j\omega_k n)$ **903** umfaßt, die die BP-Signale nach unten zu Regionen um 0 Hz verschieben, wobei die N Analysesignale $X_k(n)$ gebildet werden. Das Fenster wirkt als ein Prototyp-LP-Filter. $X_k(n)$ weisen kleine Bandbreiten auf und sind normalerweise abwärtsabgetastet **905**. Gl. 12 muß somit nur bei $n = rR$ ausgewertet werden, wobei R der Dezimierungsfaktor ist und r die neue Zeitvariable ist. $X_k(n)$ kann aus $X_k(rR)$ durch ein Aufwärtsabtasten wiedergewonnen werden, siehe **Fig. 9b**, d. h. eine Einbringung von Nullen **907**, gefolgt durch ein LP-Filtern **909**. Die inverse Transformation kann als ein Synthesizer interpretiert werden, der eine Bank von N Multiplizierern mit Trägern $(1/N)\exp(j\omega_k n)$ **911**, die die Signale $X_k(n)$ zu den ursprünglichen Frequenzen derselben hinauf verschiebt, gefolgt durch Stufen **913**, **Fig. 9c**, umfaßt, die die Beiträge $y_k(n)$ von allen Kanälen addieren. Die STFT und ISTFT können umgeordnet werden, um die DFT und IDFT zu verwenden, was die Verwendung von FFT-Algorithmen möglich macht [„Implementation of the Phase Vocoder using the Fast Fourier Transform“ M. R. Portnoff, IEEE ASSP, Bd. 24, Nr. 3, 1976].

[0076] **Fig. 9c** zeigt ein Patch **915** zu einer Erzeugung von zweiten Oberschwingungen, $M = 2$, mit $N = 32$. Der Einfachheit halber sind lediglich Kanäle 0 bis 16 gezeigt. Die Mittenfrequenz von BP 16 ist gleich der Nyquist-Frequenz, Kanäle 17 bis 31 entsprechen negativen Frequenzen. Die P **917** bezeichneten Blöcke und die Gewinnblöcke **919** werden später beschrieben und sollten gegenwärtig als überbrückt betrachtet werden. Das Eingangssignal ist bei diesem Beispiel bandbegrenzt, so daß lediglich Kanäle 0 bis 7 Signale enthalten. Analysatorkanäle 8 bis 16 sind somit leer und müssen nicht zu dem Synthesizer abgebildet werden. Analysatorkanäle 0 bis 7 sind mit Synthesizerkanälen 0 bis 7 verbunden, was einem Eingangssignalverzögerungsweg entspricht. Analysekanäle k, wobei $4 \leq k \leq 7$, sind ebenfalls mit Synthesizerkanälen Mk , $M = 2$, verbunden, die die Signale zu Frequenzregionen bei zweimal den Mittenfrequenzen von BP-Filtern k verschieben. Somit sind die Signale zu den ursprünglichen Bereichen derselben nach oben verschoben sowie eine Oktave nach oben transponiert. Um die harmonische Erzeugung hinsichtlich reellwertiger Filterantworten und Modulatoren zu untersuchen, müssen die negativen Frequenzen ebenfalls betrachtet werden, siehe die untere Verzweigung von **Fig. 10a**. Daher muß der kombinierte Ausgang der Neuabbildung $1 \rightarrow Mk$ **1001** und $N - k \rightarrow N - Mk$ **1003** wobei $4 \leq k \leq 7$ ausgewertet werden.

[0077] Dies ergibt

$$y(n) = \frac{2}{N} [x(n) * h(n) \cos(\omega_k n)] \cos((M-1)\omega_k n) + \quad (15)$$

$$- \frac{2}{N} [x(n) * h(n) \sin(\omega_k n)] \sin((M-1)\omega_k n)$$

wobei $M = 2$. Gl. 15 kann als ein BP-Filtern des Eingangssignals, gefolgt durch eine lineare Frequenzverschiebung oder eine Oberes-Seitenband-(OSB-)Modulation interpretiert werden, d. h. eine Einseitenbandmodulation unter Verwendung des oberen Seitenbands, siehe **Fig. 10b**, wobei **1005** und **1007** einen Hilbert-Transformator bilden, **1009** und **1011** Multiplizierer mit Cosinus- und Sinusträgern sind und **1013** eine Differenzstufe ist, die das obere Seitenband auswählt. Natürlich kann ein derartiges BP- und SSB-Mehrbandverfahren explizit implementiert sein, d. h. ohne ein Filterbank-Patching in dem Zeit- oder Frequenz-Bereich, was eine beliebige Auswahl von einzelnen Durchlaßbändern und Oszillatorfrequenzen ermöglicht.

[0078] Gemäß Gleichung 15 ergibt eine Sinuskurve mit der Frequenz ω_i innerhalb des Durchlaßbands des Analysekanals k eine Oberschwingung bei der Frequenz $M\omega_k + (\omega_i - \omega_k)$. Daher erzeugt das Verfahren, auf das als eine grundlegende Mehrbandtransposition Bezug genommen wird, lediglich exakte Oberschwingungen für Eingangssignale mit Frequenzen $\omega_i = \omega_k$, wobei $4 \leq k \leq 7$. Falls jedoch die Anzahl von Filtern ausreichend groß ist, ist die Abweichung von einer exakten Transposition vernachlässigbar, siehe Gl. 4. Ferner ist die Transposition für quasistationäre Tonalsignale von beliebigen Frequenzen durch ein Einbringen der Blöcke exakt gemacht, die **P 917** bezeichnet sind (**Fig. 9c**), vorausgesetzt jeder Analysekanal enthält maximal eine Partiale. In diesem Fall sind $X_k(rR)$ komplexe Exponentiale mit Frequenzen gleich den Frequenzen zwischen den Partialfrequenzen ω_i und den Mittenfrequenzen ω_k der Analysefilter. Um die exakte Transposition M zu erhalten, müssen diese Frequenzen durch einen Faktor M erhöht werden, wobei die obige Frequenzbeziehung zu $\omega_i \rightarrow M\omega_k + M(\omega_i - \omega_k) = M\omega_i$ modifiziert ist. Die Frequenzen von $X_k(rR)$ sind gleich den Zeitableitungen der jeweiligen ausgewickelten Phasenwinkel derselben und können unter Verwendung von Differenzen erster Ordnung von aufeinanderfolgenden Phasenwinkeln geschätzt werden. Die Frequenzschätzwerte werden mit M multipliziert und Synthesephasenwinkel werden unter Verwendung dieser neuen Frequenzen berechnet. Abgesehen von einer Phasenkonstante wird das gleiche Ergebnis jedoch auf eine vereinfachte Weise durch ein Multiplizieren der Analyseargumente direkt mit M erhalten, was den Bedarf nach einer Frequenzschätzung eliminiert. Dies ist in **Fig. 11** beschrieben, die die Blöcke **917** darstellt. Somit werden $X_k(rR)$, wobei bei diesem Beispiel $4 \leq k \leq 7$, von Rechteck- zu Polarkoordinaten umgewandelt, was durch die Blöcke $R \rightarrow P$, **1101** dargestellt ist. Die Argumente werden mit $M = 2$ multipliziert **1103** und die Beträge sind unverändert. Die Signale werden dann zurück zu Rechteckkoordinaten umgewandelt ($P \rightarrow R$) **1105**, wobei die Signale $Y_{mk}(rR)$ gebildet werden, und werden gemäß **Fig. 9c** Synthesizerkanälen zugeführt. Dieses verbesserte Mehrbandtranspositionsverfahren weist somit zwei Stufen auf: das Patch liefert eine grobe Transposition, wie bei dem grundlegenden Verfahren, und die Phasenmultiplizierer liefern feine Frequenzkorrekturen. Die obigen Mehrbandtranspositionsverfahren unterscheiden sich von herkömmlichen Tonhöhenverschiebungstechniken, die STFT verwenden, bei der Nachschlagtabellenoszillatoren für die Synthese verwendet werden oder, wenn die ISTFT für die Synthese verwendet wird, das Signal zeitgestreckt und dezimiert wird, d. h. kein Patch verwendet wird.

[0079] Das harmonische Patch von **Fig. 9c** ist ohne weiteres für andere Transpositionsfaktoren als 2 modifiziert. **Fig. 12** zeigt ein Patch **1203** zu einer Erzeugung von Oberschwingungen dritter Ordnung, wobei **1201** die Analysekanäle sind und **1205** die Synthesekanäle sind. Unterschiedliche Oberschwingungsordnungen können simultan erzeugt werden, wie es in **Fig. 13** gezeigt ist, wobei Oberschwingungen zweiter und dritter Ordnung verwendet werden. **Fig. 14** stellt eine nicht-überlappende Kombination von Oberschwingungen zweiter, dritter und vierter Ordnung dar. Die niedrigstmögliche Oberschwingungsanzahl wird bei einer möglichst hohen Frequenz verwendet. Über der oberen Begrenzung des Bestimmungsbereichs der Oberschwingung M , wird eine Oberschwingung $M + 1$ verwendet. **Fig. 15** zeigt ein Verfahren zu einem Abbilden aller Synthesizerkanäle ($N = 64$, gezeigte Kanäle 0–32). Alle Hochbandkanäle mit Nicht-Primzahl-Indizes werden gemäß der folgenden Beziehung zwischen Quell- und Bestimmungskanalanzahl abgebildet: $k_{\text{dest}} = M k_{\text{source}}$, wobei M die kleinste Ganzzahl ≥ 2 ist, die die Bedingung erfüllt, daß k_{source} in dem Tiefband und k_{dest} in dem Hochband liegt. Daher empfängt kein Synthesizerkanal ein Signal von mehr als einem Analysekanal. Primzahl-Hochbandkanäle können zu $k_{\text{source}} = 1$ oder Tiefbandkanälen $k_{\text{source}} > 1$ abgebildet werden, die gute Annäherungen der obigen Beziehung ergeben (in **Fig. 15** sind lediglich Nicht-Primzahl-Verbindungen mit $M = 2, 3, 4, 5$ gezeigt).

[0080] Es ist auch möglich, Amplituden- und Phaseninformationen von unterschiedlichen Analysatorkanälen zu kombinieren. Die Amplitudensignale $|X_i(rR)|$ können gemäß **Fig. 16** verbunden werden, während die Phasensignale gemäß dem Prinzip von **Fig. 15** verbunden werden. Auf diese Weise werden die Tiefbandfrequenzen immer noch transponiert, wodurch eine periodische Wiederholung der Quellregionhüllkurve anstelle der

gestreckten Hüllkurve erzeugt wird, die aus einer Transposition gemäß Gleichung 2 resultiert. Ein Ausblenden oder eine andere Einrichtung kann eingegliedert sein, um eine Verstärkung von „leeren“ Quellkanälen zu vermeiden. **Fig. 17** stellt eine andere Anwendung dar, die Erzeugung von Suboberschwingungen zu einem hochpaßgefilterten oder baßbegrenzten Signal durch ein Verwenden von Verbindungen von höheren zu niedrigeren Subbändern. Wenn die obigen Transpositionen verwendet werden, kann es vorteilhaft sein, ein adaptives Schalten eines Patch basierend auf den Charakteristika des Signals zu verwenden.

[0081] Bei der obigen Beschreibung wurde angenommen, daß die höchste Frequenz, die in dem Eingangssignal enthalten ist, erheblich niedriger als die Nyquist-Frequenz war. Somit war es möglich, eine Bandbreitenerweiterung ohne eine Erhöhung bei einer Abtastrate durchzuführen. Dies ist jedoch nicht immer der Fall, warum ein vorhergehendes Aufwärtsabtasten notwendig sein kann. Wenn Filterbankverfahren zu einer Transposition verwendet werden, ist es möglich, ein Aufwärtsabtasten in den Prozeß zu integrieren.

[0082] Die meisten Wahrnehmungscodecs verwenden in der Zeit maximal dezimierte Filterbanken zu einer Frequenzabbildung [„Introduction to Perceptual Coding“ K. Brandenburg, AES, Collected Papers on Digital Audio Bitrate Reduction, 1996]. **Fig. 18a** zeigt die Grundstruktur eines Wahrnehmungscodecsystems. Die Analysefilterbank **1801** teilt das Eingangssignal in mehrere Subbandsignale. Die Subbandabtastwerte werden einzeln quantisiert **1803**, wobei eine reduzierte Anzahl von Bits verwendet wird, wobei die Anzahl von Quantisierungspegeln aus einem Wahrnehmungsmodell bestimmt sind **1807**, das die minimale Maskierungsschwelle schätzt. Die Subbandabtastwerte werden normalisiert, mit optionalen Redundanzcodierungsverfahren codiert und mit Seiteninformationen kombiniert, die aus den Normalisierungsfaktoren, Bitzuweisungsinformationen und anderen Codec-spezifischen Daten bestehen **1805**, um den Reihenbitstrom zu bilden. Der Bitstrom wird dann gespeichert oder gesendet. Bei dem Decodierer, **Fig. 18b**, wird der codierte Bitstrom gedemultiplext **1809**, decodiert und die Subbandabtastwerte werden wieder zu der gleichen Anzahl von Bits quantisiert **1811**. Eine Synthesefilterbank kombiniert die Subbandabtastwerte, um das ursprüngliche Signal wieder zu erzeugen **1813**. Implementierungen, die maximal dezimierte Filterbanken verwenden, reduzieren Berechnungskosten drastisch. Bei den folgenden Beschreibungen wird auf Cosinus-modulierte Filterbanken fokussiert. Es ist jedoch zu erkennen, daß die Erfindung unter Verwendung anderer Typen von Filterbanken oder Transformationen implementiert werden kann, einschließlich Filterbankinterpretationen der Wavelet-Transformation, anderer Filterbanken oder Transformationen mit nicht-gleicher Bandbreite und mehrdimensionale Filterbanken oder Transformationen.

[0083] Bei den folgenden darstellenden aber nicht begrenzenden Beschreibungen wird angenommen, daß eine Cosinus-modulierte L-Kanal-Filterbank das Eingangssignal $x(n)$ in L Subbandsignale teilt. Die allgemeine Struktur einer maximal dezimierten Filterbank ist in **Fig. 19** gezeigt. Die Analysefilter sind mit $H_k(z)$ **1901** bezeichnet, wobei $k = 0, 1, \dots, L - 1$. Die Subbandsignale $v_k(n)$ sind maximal dezimiert **1903**, jeweils von einer Abtastfrequenz f_s/L , wobei f_s die Abtastfrequenz von $x(n)$ ist. Der Syntheseabschnitt fügt die Subbandsignale nach einer Interpolation **1905** und einem Filtern **1907** wieder zusammen, um $\hat{x}(n)$ zu erzeugen. Die Synthesefilter sind mit $F_k(z)$ bezeichnet. Zusätzlich führt die vorliegende Erfindung eine Spektralreplikation an $\hat{x}(n)$ durch, was ein verbessertes Signal $y(n)$ ergibt.

[0084] Ein Generieren der Subbandsignale mit einer QL-Kanal-Filterbank, wobei lediglich die L-Tiefbandkanäle verwendet werden und der Bandbreitenausdehnungsfaktor Q gewählt ist, so daß QL ein ganzzahliger Wert ist, resultiert in einem Ausgangsbitstrom mit einer Abtastfrequenz Qf_s . Daher wirkt die erweiterte Filterbank, als ob sie eine L-Kanal-Filterbank gefolgt von einem Aufwärtsabtaster ist. Da in diesem Fall die $L(Q - 1)$ -Hochbandfilter ungenutzt (mit Nullen gespeist) sind, verändert sich die Audiobandbreite nicht – die Filterbank baut lediglich eine aufwärtsabgetastete Version von $\hat{x}(n)$ wieder auf. Falls jedoch die L-Subbandsignale zu den Hochbandfiltern gepatcht werden, wird die Bandbreite von $\hat{x}(n)$ um einen Faktor Q erhöht, was $y(n)$ erzeugt. Dies ist die maximal dezimierte Filterbankversion des grundlegenden Mehrbandtransponierers gemäß der Erfindung. Unter Verwendung dieses Schemas ist der Aufwärtsabtastprozeß in das Synthesefiltern integriert, wie es später erläutert ist. Es ist zu bemerken, daß eine jegliche Größe der Synthesefilterbank verwendet werden kann, was in unterschiedlichen Abtastraten des Ausgangssignals und somit unterschiedlichen Bandbreitenausdehnungsfaktoren resultiert. Ein Durchführen einer Spektralreplikation an $\hat{x}(n)$ gemäß der vorliegenden Erfindung des grundlegenden Mehrbandtranspositionsverfahrens mit einem ganzzahligen Transpositionsfaktor M ist erzielt durch ein Patchen der Subbandsignale als

$$v_{Mk}(n) = e_{Mk}(n)(-1)^{(M-1)kn}v_k(n) \quad (16)$$

wobei $k \in [0, L - 1]$ und gewählt ist, so daß $Mk \in [L, QL - 1]$, $e_{Mk}(n)$ die Hüllkurvenkorrektur ist und $(-1)^{(M-1)kn}$ ein Korrekturfaktor für spektralinvertierte Subbänder ist. Eine Spektralinversion resultiert aus einer Dezimierung

von Subbandsignalen und die invertierten Signale können durch ein Ändern eines Vorzeichens bei jedem zweiten Abtastwert in diesen Kanälen reinvertiert werden. Unter Bezugnahme auf **Fig. 20** betrachte man eine 16-Kanal-Synthesefilterbank, gepatcht **2009** für einen Transpositionsfaktor $M = 2$, mit $Q = 2$. Die Blöcke **2001** und **2003** bezeichnen die Analysefilter $H_k(z)$ bzw. die Dezimatoren von **Fig. 19**. Auf eine ähnliche Weise sind **2005** und **2007** die Interpolatoren und Synthesefilter $F_k(z)$. Gl. 16 vereinfacht sich dann zu einem Patchen der vier oberen Frequenzsubbandsignale der empfangenen Daten in jeden zweiten der acht obersten Kanäle in der Synthesefilterbank. Aufgrund einer Spektral inversion muß jedes zweite gepatchte Subbandsignal vor der Synthese frequenzinvertiert werden. Zusätzlich müssen die Beträge der gepatchten Signale gemäß den Prinzipien von SBR-1 oder SBR-2 eingestellt werden **2011**.

[0085] Unter Verwendung des grundlegenden Mehrbandtranspositionsverfahrens gemäß der vorliegenden Erfindung sind die erzeugten Oberschwingungen im allgemeinen keine exakten Mehrfachen der Fundamentalen. Alle Frequenzen außer der niedrigsten in jedem Subband unterscheiden sich zu einem gewissen Ausmaß von einer exakten Transposition. Ferner enthält das replizierte Spektrum Nullen, da das Zielintervall einen breiten Frequenzbereich als das Quellintervall abdeckt. Außerdem verschwinden die Aliasaufhebungseigenschaften der Cosinus-modulierten Filterbank, da die Subbandsignale in einer Frequenz in dem Zielintervall getrennt sind. Das heißt, benachbarte Subbandsignale überlappen sich in dem Hochbandbereich nicht. Es können jedoch Aliasingreduzierungsverfahren verwendet werden, die Fachleuten auf dem Gebiet bekannt sind, um diesen Typ von Artefakten zu reduzieren. Vorteile dieses Transpositionsverfahrens sind eine leichte Implementierung und die sehr niedrigen Berechnungskosten.

[0086] Um eine perfekte Transposition von Sinuskurven zu erreichen, wird nun eine wirksame maximal dezimierte Filterbanklösung des verbesserten Mehrbandtranspositionsverfahrens präsentiert. Das System verwendet eine zusätzliche modifizierte Analysefilterbank, während die Synthesefilterbank Cosinus-moduliert ist, wie es durch Vaidyanathan [„Multirate Systems and Filter Banks“ P. P. Vaidyanathan, Prentice Hall, Englewood Cliffs, New Jersey, 1993, ISBN 0-13-605718-7] beschrieben ist. Die Schritte zu einer Operation, wobei das verbesserte Mehrbandtranspositionsverfahren gemäß der vorliegenden Erfindung verwendet wird, basieren auf maximal dezimierten Filterbanken, sind schematisch in **Fig. 21** und in dem Flußdiagramm von **Fig. 22** gezeigt und sind wie folgt:

1. Die L empfangenen Subbandsignale werden mit einer QL -Kanal-Filterbank generiert **2101**, **2201**, **2203**, wobei die $L(Q - 1)$ oberen Kanäle mit Nullen gespeist werden, um ein Signal $x_1(n)$ zu bilden, das somit durch den Bandbreitenerweiterungsfaktor Q überabgetastet wird.
2. $x_1(n)$ wird durch einen Faktor Q abwärtsabgetastet, um ein Signal $x_2(n')$ zu bilden **2103**, **2205**, d. h. $x_2(n') = x_1(Qn')$.
3. Ein ganzzahlwertiges K wird als die Größe einer Synthesefilterbank gewählt, eingeschränkt, so daß $T = KM/Q$ eine Ganzzahl ist, wobei T die Größe der modifizierten Analysefilterbank ist und M der Transpositionsfaktor ist **2207**, **2209**, **2211**. K sollte vorzugsweise für stationäre (tonale) Signale groß und für dynamische (transiente) Signale kleiner gewählt sein.
4. $x_2(n')$ wird durch eine T -Kanal-modifizierte Analysefilterbank gefiltert **2107**, **2213**, wobei die T Analysefilter exponentiell moduliert sind, was einen Satz von komplexwertigen Subbandsignalen erzeugt. Die Subbandsignale werden durch einen Faktor T/M abwärtsabgetastet, was Subbandsignale $v_k^{(M)}(n'')$, $k = 0, 1, \dots, T - 1$ ergibt. Daher wird die Filterbank durch einen Faktor M überabgetastet.
5. Die Abtastwerte $v_k^{(M)}(n'')$ werden in eine polare Darstellung (Betrag und Phasenwinkel) umgewandelt. Die Phasenwinkel werden mit dem Faktor M multipliziert und die Abtastwerte werden gemäß dem Schema von **Fig. 11** zurück zu einer rechteckigen Darstellung umgewandelt. Die reellen Teile der komplexwertigen Abtastwerte werden genommen, was die Signale $s_k^{(M)}(n'')$ ergibt **2109**, **2215**. Nach dieser Operation werden die Signale $s_k^{(M)}(n'')$ kritisch abgetastet.
6. Die Gewinne der Signale $s_k^{(M)}(n'')$ werden gemäß den Prinzipien von SBR-1 oder SBR-2 eingestellt **2111**, **2217**.
7. Die Subbandsignale $s_k^{(M)}(n'')$, wobei $k \in [T/M, \min(K, T) - 1]$ werden mit einer gewöhnlichen Cosinusmodulierten K -Kanal-Filterbank generiert, wobei die Kanäle 0 bis $T/M - 1$ mit Nullen gespeist werden **2105**, **2221**. Dies erzeugt das Signal $x_3^{(M)}(n)$.
8. $x_3^{(M)}(n)$ wird schließlich zu $x_1(n)$ addiert, um $y(n)$ zu ergeben **2223**, was das erwünschte spektralreplizierte Signal ist.

[0087] Die Schritte 3 bis 6 können für unterschiedliche Werte des Transpositionsfaktors M wiederholt werden, wobei so mehrere Oberschwingungen zu $x_1(n)$ addiert werden. Dieser Betriebsmodus ist durch die gepunkteten Figuren von **Fig. 21** und in **Fig. 22** durch ein iteratives Wiederholen der Schleife über den Kästen **2211–2219** dargestellt. In diesem Fall ist k gewählt, um T für alle Auswahlmöglichkeiten von M ganzzahlwertig zu machen – für ein ganzzahlwertiges $M : s$; man wähle K vorzugsweise aus, um aus K/Q eine positive Ganz-

zahl zu machen. Alle Subbandsignale

$$s_k^{(M_i)}(n'')$$

wobei $i = 1, 2, \dots, m$, und m die Anzahl von Transpositionsfaktoren ist, werden gemäß

$$s_k(n'') = \sum_{i=1}^m s_k^{(M_i)}(n'') \quad (17)$$

für jedes anwendbare k addiert. Bei der ersten Iteration der Schleife von **Fig. 22** können die Signale $s_k(n'')$ als Subbandabtastwerte von lediglich Nullen betrachtet werden, wobei $k = 0, 1, \dots, K-1$. Bei jeder Schleife werden die neuen Abtastwerte zu $s_k(n'')$ addiert **2219** als

$$s_k(n'') = s_k(n'') + s_k^{(M_i)}(n'') \quad (18)$$

wobei $k = K/Q, K/Q + 1, \dots, \min(K, T_l) - 1$. Die Subbandsignale $s_k(n'')$ werden einmal mit einer K -Kanal-Filterbank gemäß Schritt 7 generiert.

[0088] Die modifizierte Analysefilterbank von Schritt 4 ist durch die Theorie von Cosinus-modulierten Filterbanken abgeleitet, wobei die modulierte überlappte Transformation (MLT = modulated lapped transform) [„Lapped Transforms for Efficient Transform/Subband Coding“ H. S. Malvar, IEEE Trans ASSP, Bd. 38, Nr. 6, 1990] ein spezieller Fall ist. Die Impulsantworten $h_k(n)$ der Filter in einer Cosinus-modulierten T -Kanal-Filterbank können geschrieben werden

$$h_k(n) = C p_0(n) \cos \left[\frac{\pi}{2T} (2k + 1) \left(n - \frac{N-1}{2} \right) + \Phi_k \right] \quad (19)$$

wobei $k = 0, 1, \dots, T-1$, N die Länge des Tiefpaßprototypfilters $p_0(n)$ ist, C eine Konstante ist und Φ_k ein Phasenwinkel ist, der eine Aliasaufhebung zwischen benachbarten Kanälen sicherstellt. Die Beschränkungen auf Φ_k sind

$$\Phi_0 = \pm \frac{\pi}{4}, \quad \Phi_{T-1} = \pm \frac{\pi}{4} \quad \text{und} \quad \Phi_k = \Phi_{k-1} \pm \frac{\pi}{2} \quad (20a-c)$$

die vereinfacht werden können zu dem geschlossenen Ausdruck

$$\Phi_k = \pm (-1)^k \frac{\pi}{4} \quad (21)$$

[0089] Bei dieser Auswahl von Φ_k können perfekte Rekonstruktionssysteme oder näherungsweise Rekonstruktionssysteme (Pseudo-QMF-Systeme) erhalten werden, die Synthesefilterbanken verwenden mit Impulsantworten wie

$$f_k(n) = C p_0(n) \cos \left[\frac{\pi}{2T} (2k + 1) \left(n - \frac{N-1}{2} \right) - \Phi_k \right] \quad (22)$$

[0090] Man betrachte die Filter

$$h'_k(n) = C p_0(n) \sin \left[\frac{\pi}{2T} (2k + 1) \left(n - \frac{N-1}{2} \right) + \Phi_k \right] \quad (23)$$

wobei $h'_k(n)$ die Sinus-modulierten Versionen des Prototypfilters $p_0(n)$ sind. Die Filter $H'_k(z)$ und $H_k(z)$ weisen identische Durchlaßbandträger auf, aber die Phasenantworten unterscheiden sich. Die Durchlaßbänder der Filter sind tatsächlich Hilbert-Transformationen voneinander (dies gilt nicht für Frequenzen nahe an $\omega = 0$ und $\omega = \pi$). Ein Kombinieren von Gleichung 19 und Gleichung 23 gemäß

$$h_k^\alpha(n) = h_k(n) + j h'_k(n) = C p_0(n) \exp \left[\frac{j\pi}{2T} (2k + 1) \left(n - \frac{N-1}{2} \right) + j\Phi_k \right] \quad (24)$$

ergibt Filter, die die gleiche Form der Betragsantworten wie $H_k(z)$ für positive Frequenzen aufweisen, aber für

negative Frequenzen 0 sind. Ein Verwenden einer Filterbank mit Impulsantworten wie in Gl. 24 ergibt einen Satz von Subbandsignalen, die als die analytischen (komplexen) Signale entsprechend den Subbandsignalen interpretiert werden können, die von einer Filterbank mit Impulsantworten wie in Gl. 19 erhalten werden. Analytische Signale sind zu einer Manipulation geeignet, da die komplexwertigen Abtastwerte in einer polaren Form geschrieben sein können, d. h. $z(n) = r(n) - ji(n) = |z(n)| \exp\{j \arg(z(n))\}$. Wenn jedoch die komplexe Filterbank zu einer Transposition verwendet wird, muß die Beschränkung auf Φ_k verallgemeinert werden, um die Aliashebungseigenschaft zu behalten. Die neue Beschränkung auf Φ_k , um eine Aliashebung in Kombination mit einer Synthesefilterbank mit Impulsantworten wie in Gl. 22 sicherzustellen, ist

$$\Phi_k = \pm(-1)^k \frac{\pi}{4M} \quad (25)$$

was sich zu Gl. 21 vereinfacht, wenn $M = 1$. Bei dieser Auswahl weisen transponierte Partiale die gleichen relativen Phasen auf, wie sie dieselben aufweisen würden, wenn $M = 1$ (keine Transposition).

[0091] Ein Kombinieren von Gl. 24 und Gl. 25 resultiert in

$$h_k^a(n) = C p_0(n) \exp\left\{j\pi \left[\frac{(2k+1)}{2T} \left(n - \frac{N-1}{2} \right) \pm \frac{(-1)^k}{4M} \right] \right\} \quad (26)$$

was die Filter sind, die bei der modifizierten Filterbank von Schritt 4 verwendet werden, gemäß der vorliegenden Erfindung.

[0092] Einige Klarstellungen bezüglich Schritt 5: ein Abwärtsabtasten der komplexwertigen Subbandsignale durch einen Faktor T/M macht dieselben durch M überabgetastet, was ein wesentliches Kriterium ist, wenn die Phasenwinkel nachfolgend mit dem Transpositionsfaktor M multipliziert werden. Das Überabtasten zwingt die Anzahl von Subbandabtastwerten pro Bandbreite nach einer Transposition zu dem Zielbereich, gleich der des Quellbereichs zu sein. Die einzelnen Bandbreiten der transponierten Subbandsignale sind aufgrund des Phasenmultiplizierers M -mal größer als dieselben in dem Quellbereich. Dies macht die Subbandsignale nach Schritt 5 kritisch abgetastet und zusätzlich gibt es keine Nullen in dem Spektrum bei einem Transponieren von tonalen Signalen.

[0093] Um trigonometrische Berechnungen zu vermeiden, d. h., die neuen Subbandsignale berechnen zu müssen als

$$\begin{aligned} s_k^{(M)}(n'') &= \text{real} \left\{ |v_k^{(M)}(n'')| \exp \left\{ jM \arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right\} \right\} \\ &= |v_k^{(M)}(n'')| \cos \left\{ M \arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right\} \end{aligned} \quad (27)$$

wobei $|v_k^{(M)}(n'')|$ der absolute Wert von $v_k^{(M)}(n'')$ ist, wird die folgende trigonometrische Beziehung verwendet:

$$\cos(M\alpha) = \cos^M(\alpha) - \binom{M}{2} \sin^2(\alpha) \cos^{M-2}(\alpha) + \binom{M}{4} \sin^4(\alpha) \cos^{M-4}(\alpha) - \dots \quad (28)$$

Angenommen

$$\cos(\alpha) = \cos \left(\arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right) = \frac{\text{real}\{v_k^{(M)}(n'')\}}{|v_k^{(M)}(n'')|} \quad (30)$$

und anmerkend, daß

$$\sin(\alpha) = \sin \left(\arctan \left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}} \right) \right) = \frac{\text{imag}\{v_k^{(M)}(n'')\}}{|v_k^{(M)}(n'')|} \quad (31)$$

und

$$\sin(\alpha) = \sin\left(\arctan\left(\frac{\text{imag}\{v_k^{(M)}(n'')\}}{\text{real}\{v_k^{(M)}(n'')\}}\right)\right) = \frac{\text{imag}\{v_k^{(M)}(n'')\}}{|v_k^{(M)}(n'')|} \quad (31)$$

können die Berechnungen von Schritt 5 ohne trigonometrische Berechnungen erzielt werden, was eine Berechnungskomplexität reduziert.

[0094] Wenn Transpositionen verwendet werden, bei denen M gerade ist, können Hindernisse bei dem Phasemultiplizierer entstehen, abhängig von den Charakteristika des Tiefpaßprototypfilters $p_0(n)$. Alle anwendbaren Prototypfilter weisen Nullen auf dem Einheitskreis in der z -Ebene auf. Eine 0 auf dem Einheitskreis erlegt eine 180° -Verschiebung bei der Phasenantwort des Filters auf. Für M gerade, translatiert der Phasemultiplizierer diese Verschiebungen zu 360° -Verschiebungen; d. h. die Phasenverschiebungen verschwinden. Die Partiale, die so in einer Frequenz positioniert sind, daß derartige Phasenverschiebungen verschwinden, rufen ein Aliasing bei dem generierten Signal hervor. Das Ungünstigste-Fall-Szenario ist, wenn eine Partiale an einem Punkt in einer Frequenz positioniert ist, der dem oberen Ende des ersten Seitenlappens eines Analysefilters entspricht. Abhängig von der Unterdrückung dieses Lappens bei der Betragsantwort, wird das Aliasing mehr oder weniger hörbar sein. Beispielsweise wird der erste Seitenlappen des Prototypfilters, das bei dem ISO/MPEG-Schicht-1 und -2-Standard verwendet wird, 96 dB unterdrückt, während die Unterdrückung für den ersten Seitenlappen des Sinusfensters, das bei dem MDCT-Schema des ISO/MPEG-Schicht-3-Standard verwendet wird, lediglich 23 dB beträgt. Es ist klar, daß dieser Typ eines Aliasing unter Verwendung des Sinusfensters hörbar sein wird. Eine Lösung zu diesem Problem wird präsentiert und wird als eine relative Phasenverriegelung bezeichnet.

[0095] Die Filter $h_k^a(n)$ weisen alle lineare Phasenantworten auf. Die Phasenwinkel Φ_k bringen relative Phasendifferenzen zwischen benachbarten Kanälen ein und die Nullen auf dem Einheitskreis bringen 180° -Phasenverschiebungen bei Positionen in einer Frequenz ein, die zwischen Kanälen unterschiedlich sein kann. Durch ein Überwachen der Phasendifferenz zwischen benachbarten Subbandsignalen bevor der Phasemultiplizierer aktiviert ist, ist es einfach, die Kanäle zu erfassen, die phaseninvertierte Informationen enthalten. Unter Betrachtung von tonalen Signalen beträgt die Phasendifferenz näherungsweise $\pi/2M$, gemäß Gl. 25, für nichtinvertierte Signale und folglich näherungsweise $\pi(1 - 1/2M)$ für Signale, bei denen eines der Signale invertiert ist. Die Erfassung von invertierten Signalen kann einfach durch ein Berechnen des Punktprodukts von Abtastwerten in benachbarten Subbändern erzielt werden als

$$v_k^{(M)}(n'') \cdot v_{k+1}^{(M)}(n'') = \text{real}\{v_k^{(M)}(n'')\}\text{real}\{v_{k+1}^{(M)}(n'')\} + \text{imag}\{v_k^{(M)}(n'')\}\text{imag}\{v_{k+1}^{(M)}(n'')\} \quad (32)$$

[0096] Falls das Produkt in Gleichung 32 negativ ist, ist die Phasendifferenz größer als 90° und eine Phaseninversionsbedingung liegt vor. Die Phasenwinkel der komplexwertigen Subbandsignale werden mit M multipliziert, gemäß dem Schema von Schritt 5, und schließlich werden die inversionsetikettierten Signale negiert. Das Relative-Phasenverriegelungsverfahren zwingt somit die 180° -verschobenen Subbandsignale, diese Verschiebung nach der Phasenmultiplikation zu behalten und somit die Aliasingaufhebungseigenschaften beizubehalten.

Spektralhüllkurveneinstellung

[0097] Die meisten Klänge, wie eine Sprache und eine Musik, sind als Produkte von langsam variierenden Hüllkurven und schnell variierenden Trägern mit einer konstanten Amplitude charakterisiert, wie es durch Stockham [„The Application of Generalized Linearity to Automatic Gain Control“ T. G. Stockham, Jr, IEEE Trans. on Audio and Electroacoustics, Bd. AU-16, Nr. 2, Juni 1968] und Gl. 1 beschrieben ist.

[0098] Bei Splitband-Wahrnehmungsaudiocodierern wird das Audiosignal in Rahmen segmentiert und in mehrere Frequenzbänder geteilt, wobei Subbandfilter oder eine Zeit-zu-Frequenz-Bereich-Transformation verwendet wird. Bei den meisten Codec-Typen wird das Signal nachfolgend in zwei größere Signalkomponenten zu einer Übertragung oder Speicherung getrennt, die Spektralhüllkurvendarstellung und die normalisierten Subbandabtastwerte oder Koeffizienten. Überall in der folgenden Beschreibung bezieht sich der Ausdruck „Subbandabtastwerte“ oder „Koeffizienten“ auf Abtastwerte, die von Subbandfiltern erhalten werden, sowie Koeffizienten, die von einer Zeit-zu-Frequenz-Transformation erhalten werden. Der Ausdruck „Spektralhüllkurve“ oder „Skalierungsfaktoren“ stellt Werte der Subbänder auf einer Zeit-Rahmen-Basis dar, wie beispielsweise den durchschnittlichen oder maximalen Betrag in jedem Subband, verwendet für eine Normalisierung der Subbandabtastwerte. Die Spektralhüllkurve kann jedoch auch unter Verwendung einer linearen Voraussage erhalten werden LPC, [US-Pat. 5,684,920]. Bei einem typischen Codec erfordern die normalisierten Subbandabtastwerte ein Codieren bei einer hohen Bitrate (ein Verwenden von näherungsweise 90% der verfü-

baren Bitrate) verglichen mit den langsam variierenden zeitlichen Hüllkurven und somit den Spektralhüllkurven, die bei einer sehr reduzierten Rate (ein Verwenden von näherungsweise 10% der verfügbaren Bitrate) codiert werden können.

[0099] Eine genaue Spektralhüllkurve der replizierten Bandbreite ist wichtig, falls die Klangfarbenqualitäten des ursprünglichen Signals bewahrt werden sollen. Die wahrgenommene Klangfarbe eines Musikinstruments oder einer Stimme ist hauptsächlich durch die Spektralverteilung unter einer Frequenz f_{lim} bestimmt, die in den höchsten Höroktaven positioniert ist. Die Spektraldetails über f_{lim} sind somit von geringerer Bedeutung und folglich erfordern die Hochband-Feinstrukturen, die durch die obigen Transpositionsverfahren erhalten werden, keine Einstellung, während die groben Strukturen dasselbe allgemein tun. Um eine derartige Einstellung zu ermöglichen, ist es nützlich, die Spektraldarstellung des Signals zu filtern, um die grobe Hüllkurvenstruktur von der feinen Struktur zu trennen.

[0100] Bei der SBR-1-Implementierung gemäß der vorliegenden Erfindung wird die grobe Hochbandspektralhüllkurve aus den bei dem Decodierer verfügbaren Tiefbandinformationen geschätzt. Diese Schätzung wird durch ein kontinuierliches Überwachen der Hüllkurve des Tiefbands und ein Einstellen der Hochbandspektralhüllkurve gemäß spezifischen Regeln durchgeführt. Ein neuartiges Verfahren, um die Hüllkurvenschätzung zu erzielen, verwendet Asymptoten in einem logarithmischen Frequenz-Betrag-Raum, was äquivalent zu einem Kurvenanpassen mit Polynomen einer variierenden Ordnung in dem linearen Raum ist. Der Pegel und die Steigung eines oberen Abschnitts des Tiefbandspektrums werden geschätzt und die Schätzwerte werden verwendet, um den Pegel und die Steigung eines oder mehrerer Segmente zu definieren, die die neue Hochbandhüllkurve darstellen. Die Asymptotenschnittpunkte sind in einer Frequenz fest und wirken als Schwenkpunkte. Obwohl es nicht immer notwendig ist, ist es vorteilhaft, Beschränkungen festzusetzen, um die Hochbandhüllkurvenauslenkungen innerhalb realistischer Grenzen zu halten. Ein alternativer Ansatz zu einer Schätzung der Spektralhüllkurve besteht darin, eine Vektorquantisierung, VQ, einer großen Anzahl von repräsentativen Spektralhüllkurven zu verwenden und dieselben in einer Nachschlagtabelle und einem Codebuch zu speichern. Eine Vektorquantisierung wird durch ein Trainieren der erwünschten Anzahl von Vektoren an einer großen Anzahl von Trainingsdaten durchgeführt, bei diesem Fall Audiospektralhüllkurven. Das Trainieren wird gewöhnlich mit dem verallgemeinerten Lloyd-Algorithmus (Generalised Lloyd Algorithm) [„Vector Quantization and Signal Compression“ A. Gersho, R. M. Gray, Kluwer Academic Publishers, USA 1992, ISBN 0-7923-9181-0] vorgenommen und ergibt Vektoren, die die Inhalte der Trainingsdaten optimal abdecken. Unter Betrachtung eines VQ-Codebuchs, das aus A Spektralhüllkurven besteht, die durch B Hüllkurven trainiert sind ($B \gg A$), dann stellen die A Hüllkurven die A wahrscheinlichsten Übergänge von der Tiefbandhüllkurve zu der Hochbandhüllkurve basierend auf B Beobachtungen einer breiten Vielfalt von Klängen dar. Das heißt theoretisch die A optimalen Regeln zu einem Voraussagen der Hüllkurve basierend auf den B Beobachtungen. Wenn eine neue Hochbandspektralhüllkurve geschätzt wird, wird die ursprüngliche Tiefbandhüllkurve verwendet, um das Codebuch zu durchsuchen, und der Hochbandteil des am besten übereinstimmenden Codebucheintrags wird angewendet, um das neue Hochbandspektrum zu erzeugen.

[0101] In Fig. 23 sind die normalisierten Subbandabstastwerte durch **2301** dargestellt und die Spektralhüllkurven sind durch die Skalierungsfaktoren **2305** dargestellt. Zu darstellenden Zwecken ist die Übertragung zu einem Decodierer **2303** in einer parallelen Form gezeigt. Bei dem SBR-2-Verfahren von Fig. 24 werden die Spektralhüllkurveninformationen gemäß Fig. 23 erzeugt und übertragen, wodurch lediglich die Tiefbandsbandabstastwerte übertragen werden. Übertragene Skalierungsfaktoren überspannen somit den vollen Frequenzbereich, während die Subbandabstastwerte lediglich einen beschränkten Frequenzbereich überspannen, ausschließlich des Hochbands. Bei dem Decodierer werden die Tiefbandsbandabstastwerte **2401** transponiert **2403** und mit den empfangenen Hochbandspektralhüllkurveninformationen **2405** kombiniert. Auf diese Weise ist die synthetische Hochbandspektralhüllkurve identisch mit derselben des Originals, während eine erhebliche Bitratenreduzierung beibehalten ist.

[0102] Bei einigen Codecs ist es möglich, die Skalierungsfaktoren für die gesamte Spektralhüllkurve zu übertragen, während die Hochbandsbandabstastwerte weggelassen werden, wie es in Fig. 24 gezeigt ist. Andere Codecstandards setzen fest, daß Skalierungsfaktoren und Subbandabstastwerte den gesamten Frequenzbereich abdecken müssen, d. h. Skalierungsfaktoren können nicht übertragen werden, falls die Subbandabstastwerte weggelassen sind. In derartigen Fällen gibt es mehrere Lösungen; die Hochbandspektralhüllkurveninformationen können in getrennten Rahmen übertragen werden, wobei die Rahmen eigene Kopfblöcke derselben und einem wahlweisen Fehlerschutz aufweisen, gefolgt durch die Daten. Reguläre Decodierer, die die vorliegende Erfindung nicht ausnutzen, werden die Kopfblöcke nicht erkennen und daher die zusätzlichen Rahmen aussondern. Bei einer zweiten Lösung werden die Hochbandspektralhüllkurveninformationen als Hilfsdaten innerhalb des codierten Bitstroms übertragen. Das verfügbare Hilfsdatenfeld muß jedoch groß genug sein,

um die Hüllkurveninformationen zu halten. In Fällen, bei denen keine der zwei ersten Lösungen anpaßbar ist, kann eine dritte Lösung angewendet werden, bei der die Hochbandspektralhüllkurveninformationen als Subbandabstastwerte versteckt werden. Subbandskalierungsfaktoren decken einen großen Dynamikbereich ab, typischerweise mehr als 100 dB. Es ist somit möglich, eine beliebige Anzahl von Subbandskalierungsfaktoren, **2505** in **Fig. 25**, zu sehr niedrigen Werten zu setzen und die Hochbandskalierungsfaktoren „getarnt“ als Subbandabstastwerte, **2501**, zu übertragen. Diese Weise eines Übertragens der Hochbandskalierungsfaktoren zu dem Decodierer **2503** stellt eine Kompatibilität zu der Bitstrom-Syntax sicher. Daher können beliebige Daten auf diese Weise übertragen werden. Es existiert ein verwandtes Verfahren, bei dem Informationen in den Subbandabstaststrom codiert werden [US-Pat. 5,687,191]. Eine vierte Lösung, **Fig. 26**, kann angewendet werden, wenn ein Codiersystem Huffman- oder eine andere Redundanzcodierung **2603** verwendet. Die Subbandabstastwerte für das Hochband sind dann auf null **2601** oder einen konstanten Wert gesetzt, um eine hohe Redundanz zu erreichen.

Transientenantwortverbesserungen

[0103] Transientenbezogene Artefakte sind allgemeine Probleme bei Audiocodern und ähnliche Artefakte treten bei der vorliegenden Erfindung auf. Im allgemeinen erzeugt ein Patchen spektrale „Nullen“ oder Einkerbungen, die voreilenden und nacheilenden Zeit-Bereich-Echos, d. h. störenden Transienten vor und nach „echten“ Transienten entsprechen. Obwohl die P-Blöcke bei langsam variierenden tonalen Signalen „die Nullen ausfüllen“, bleiben die voreilenden und nacheilenden Echos. Das verbesserte Mehrbandverfahren soll an diskreten Sinuskurven arbeiten, wobei die Anzahl von Sinuskurven auf eine pro Subband beschränkt ist. Transienten oder ein Rauschen in einem Subband kann als eine große Anzahl von diskreten Sinuskurven innerhalb dieses Subbands betrachtet werden. Dies erzeugt eine Intermodulationsverzerrung. Diese Artefakte werden als zusätzliche Quantisierungsrauschquellen betrachtet, die mit den replizierten Hochbandkanälen während Transientenintervallen verbunden sind. Herkömmliche Verfahren, um Artefakte eines voreilenden und nacheilenden Echos bei Wahrnehmungsaudiocodierern zu vermeiden, z. B. ein adaptives Fensterschalten, können somit verwendet werden, um die subjektive Qualität des verbesserten Mehrbandverfahrens zu verbessern. Durch ein Verwenden der Transientenerfassung, die durch den Codec oder einen getrennten Detektor bereitgestellt ist, und ein Reduzieren der Anzahl von Kanälen unter transienten Bedingungen ist das „Quantisierungsrauschen“ gezwungen, die zeitabhängige Maskierungsschwelle nicht zu überschreiten. Eine kleinere Anzahl von Kanälen wird während Transientendurchgängen verwendet, während eine größere während Tondurchgängen verwendet wird. Ein derartiges adaptives Fensterschalten wird allgemein bei Codern verwendet, um eine Frequenzauflösung mit einer Zeitauflösung zu tauschen. Unterschiedliche Verfahren können bei Anwendungen verwendet werden, bei denen die Filterbandgröße fest ist. Ein Ansatz besteht darin, das „Quantisierungsrauschen“ in einer Zeit über eine lineare Voraussage in dem Spektralbereich zu formen. Die Transposition wird dann an dem Restsignal durchgeführt, das der Ausgang des linearen Voraussagefilters ist. Nachfolgend wird simultan ein Inversvoraussagefilter auf die ursprünglichen und die spektral replizierten Kanäle angewendet. Ein anderer Ansatz verwendet ein Kompandersystem, d. h. eine dynamische Amplitudenkomprimierung des transienten Signals vor einer Transposition oder einem Codieren und eine komplementäre Ausdehnung nach einer Transposition. Es ist auch möglich, zwischen Transpositionsverfahren auf eine signalabhängige Weise zu schalten, z. B. wird ein hochauflösendes Filterbanktranspositionsverfahren für stationäre Signale verwendet und ein zeitvariantes Mustersuchvoraussageverfahren wird für transiente Signale verwendet.

Praktische Implementierungen

[0104] Unter Verwendung eines Standardsignalprozessors oder eines leistungsstarken PCs ist eine Echtzeitoperation eines SBR-verbesserten Codec möglich. Der SBR-verbesserte Codec kann auch auf einem kundenspezifischen Chip fest codiert sein. Derselbe kann auch in verschiedenen Arten von Systemen zu einer Speicherung oder Übertragung von Signalen, analog oder digital, unter Verwendung von beliebigen Codern, **Fig. 27** und **Fig. 28**, implementiert sein. Das SBR-1-Verfahren kann in einen Decodierer integriert sein oder als ein hinzufügbares Hardware- oder Softwarenachverarbeitungsmodul geliefert sein. Das SBR-2-Verfahren benötigt eine zusätzliche Modifizierung des Codierers. In **Fig. 27** wird das analoge Eingangssignal dem A/D-Wandler **2701** zugeführt, wobei ein digitales Signal gebildet wird, das einem beliebigen Codierer **2703** zugeführt wird, wo ein Quellcodieren durchgeführt wird. Das in das System eingespeiste Signal kann von einem derartigen Tiefpaßtyp sein, daß Spektralbänder innerhalb des Hörbereichs bereits ausgesondert wurden oder Spektralbänder in dem beliebigen Codierer ausgesondert werden. Die resultierenden Tiefbandsignale werden dem Multiplexer **2705** zugeführt, wobei ein Reihenbitstrom gebildet wird, der übertragen oder gesendet wird **2707**. Der Demultiplexer **2709** stellt die Signale wieder her und führt dieselben einem beliebigen Decodierer **2711** zu. Die Spektralhüllkurveninformationen **2715** werden bei dem Decodierer **2713** geschätzt und der SBR-1-Einheit **2713** zugeführt, die das Tiefbandsignal zu einem Hochbandsignal transponiert und ein Hüllkur-

veneingestelltes Breitbandssignal erzeugt. Schließlich wird das digitale Breitbandssignal in ein analoges Ausgangssignal umgewandelt **2717**.

[0105] Das SBR-2-Verfahren benötigt eine zusätzliche Modifizierung des Codierers. In **Fig. 28** wird das analoge Eingangssignal dem A/D-Wandler **2801** zugeführt, wobei ein digitales Signal gebildet wird, das einem beliebigen Codierer **2803** zugeführt wird, wo ein Quellcodieren durchgeführt wird. Die Spektralhüllkurveninformationen werden extrahiert **2805**. Die resultierenden Signale, Tiefbandsubbandabtastwerte oder Koeffizienten und Breitbandhüllkurveninformationen, werden dem Multiplexer **2807** zugeführt, wobei ein Reihenbitstrom gebildet wird, der gesendet oder gespeichert wird **2809**. Der Demultiplexer **2811** stellt die Signale wieder her, Tiefbandsubbandabtastwerte oder Koeffizienten und Breitbandhüllkurveninformationen, und führt dieselben einem beliebigen Decodierer **2815** zu. Die Spektralhüllkurveninformationen **2813** werden von dem Demultiplexer **2811** der SBR-2-Einheit **2817** zugeführt, die das Tiefbandssignal zu einem Hochbandssignal transponiert und ein hüllkurveneingestelltes Breitbandssignal erzeugt. Schließlich wird das digitale Breitbandssignal zu einem analogen Ausgangssignal umgewandelt **2819**.

[0106] Wenn lediglich sehr niedrige Bitraten verfügbar sind, (Internet und langsame Telefonmodems, AM-Rundsenden etc.) ist ein Mono-Codieren des Audioprogrammmaterials unvermeidbar. Um die wahrgenommene Qualität zu verbessern und das Programm angenehmer klingend zu machen, wird ein einfacher „Pseudo-Stereo“-Generator, **Fig. 29**, durch die Einbringung einer abgegriffenen Verzögerungsleitung **2901** erhalten. Dies kann um 10 ms und 15 ms verzögerte Signale bei näherungsweise -6 dB **2903** zusätzlich zu dem ursprünglichen Monosignal **2905** jedem Ausgangskanal zuführen. Der Pseudo-Stereo-Generator bietet eine wertvolle Wahrnehmungsverbesserung bei niedrigen Berechnungskosten.

[0107] Die oben beschriebenen Ausführungsbeispiele sind lediglich darstellend für die Prinzipien der vorliegenden Erfindung zu einer Audioquellcodierungsverbesserung. Es ist klar, daß Modifizierungen und Variationen der hierin beschriebenen Anordnungen und Details Fachleuten auf dem Gebiet offensichtlich sind. Es ist daher die Absicht, lediglich durch den Schutzbereich der folgenden Patentansprüche und nicht durch die spezifischen Details begrenzt zu sein, die durch eine Beschreibung und Erläuterung der Ausführungsbeispiele hierin präsentiert sind.

Patentansprüche

1. Ein Verfahren zum Decodieren eines codierten Signals, wobei das codierte Signal von einem ursprünglichen Signal abgeleitet ist und lediglich einen Abschnitt von Frequenzbändern darstellt, die in dem ursprünglichen Signal enthalten sind, das folgende Schritte aufweist:

Bereitstellen von Subbandabtastwerten für eine Mehrzahl von Subbändern oder einer Mehrzahl von Spektralkoeffizienten, wobei die Subbandabtastwerte oder Spektralkoeffizienten einen Frequenzgehalt des codierten Signals (**201, 301**) darstellen;

Extrahieren eines Bandpaßsignals (**203, 303**), wobei das Bandpaßsignal Subbandabtastwerte von einer vorbestimmten Anzahl von Analysesubbändern aufweist oder eine vorbestimmte Anzahl von Analysespektralkoeffizienten aufweist, wobei die vorbestimmte Anzahl kleiner ist als eine Gesamtanzahl von Subbändern oder Spektralkoeffizienten, die den Frequenzgehalt des codierten Signals (**201, 301**) darstellen;

Transponieren (**205, 305**) der Subbandabtastwerte der Analysesubbänder oder der Analysespektralkoeffizienten, die in dem Bandpaßsignal (**203, 303**) enthalten sind, in ein Rekonstruktionsfrequenzband (**205, 305**), das in dem ursprünglichen Signal enthalten ist und das in dem codierten Signal nicht enthalten ist, wobei das Rekonstruktionsfrequenzband Synthesesubbänder oder Synthesespektralkoeffizienten aufweist, wobei der Schritt des Transponierens ein Wiederverbinden von Analysesubbändern mit Synthesesubbändern oder ein Wiederverbinden von Analysespektralkoeffizienten mit Synthesespektralkoeffizienten umfaßt;

wobei die Subbandabtastwerte oder Spektralkoeffizienten, die in dem Bandpaßsignal (**203, 303**) enthalten sind, vor oder nach dem Schritt des Transponierens spektralhüllkurveneingestellt werden (**207, 307**), wobei Spektralhüllkurveninformationen verwendet werden, die von dem ursprünglichen Signal oder dem codierten Signal abgeleitet sind, um spektralhüllkurveneingestellte transponierte Subbandabtastwerte oder spektralhüllkurveneingestellte transponierte Spektralkoeffizienten zu erhalten, wobei die Spektralhüllkurveninformationen das Spektralhüllkurven einstellen steuern, so daß eine Spektralhüllkurve des Rekonstruktionsfrequenzbands durch die Spektralhüllkurveninformationen bestimmt ist; und

Kombinieren der Subbandabtastwerte und der eingestellten transponierten Subbandabtastwerte, oder der Spektralkoeffizienten und der eingestellten transponierten Spektralkoeffizienten, um ein decodiertes Ausgangssignal (**209, 309**) zu erhalten, wobei das decodierte Ausgangssignal einen Frequenzgehalt aufweist, der den Frequenzgehalt des codierten Signals und den Frequenzgehalt des Rekonstruktionsfrequenzbands umfaßt.

2. Ein Verfahren gemäß Anspruch 1, bei dem ein Durchlaßband (Durchlaßbänder) der eingestellten transponierten Subbandabtastwerte oder der eingestellten transponierten Spektralkoeffizienten festgelegt sind, um ein Durchlaßband (Durchlaßbänder) der Subbandabtastwerte oder Spektralkoeffizienten, die den Frequenzgehalt des codierten Signals darstellen, nicht zu überlappen oder lediglich teilweise zu überlappen.

3. Ein Verfahren gemäß Anspruch 1, das ferner den Schritt eines Schätzens der Spektralhüllkurveninformationen des Rekonstruktionsfrequenzbands unter Verwendung des codierten Signals aufweist.

4. Ein Verfahren gemäß Anspruch 1, bei dem das codierte Signal ferner übertragene Spektralhüllkurveninformationen auf dem Rekonstruktionsfrequenzband umfaßt, das bei einem Codieren des ursprünglichen Signals ausgesondert wurde, wobei das Verfahren ferner den Schritt eines Extrahierens der übertragenen Hüllkurveninformationen des ausgesonderten Frequenzbands (der ausgesonderten Frequenzbänder) des ursprünglichen Signals aufweist.

5. Ein Verfahren gemäß Anspruch 4, bei dem die Spektralhüllkurveninformationen als Skalierungsfaktoren ohne eine Übertragung der entsprechenden Subbandabtastwerte in dem codierten Signal übertragen werden.

6. Ein Verfahren gemäß Anspruch 4, bei dem die Spektralhüllkurveninformationen als Skalierungsfaktoren übertragen werden und die entsprechenden Subbandabtastwerte auf 0 oder einen konstanten Wert in dem codierten Signal festgelegt sind; wodurch die Entropie der Subbandabtastwerte reduziert wird.

7. Ein Verfahren gemäß Anspruch 1, bei dem das decodierte Ausgangssignal ein monophones Audiosignal ist, wobei das Verfahren ferner folgenden Schritt aufweist: Zerlegen des decodierten Ausgangssignals in zwei Signale, wobei jedes aus dem Ausgangssignal und verzögerten Versionen desselben zusammengesetzt ist, um ein Pseudostereosignal zu erhalten.

8. Ein Verfahren gemäß Anspruch 1, bei dem der Schritt des Bereitstellens einen Schritt eines Filterns eines Signals durch einen Satz von $N \geq 2$ Bandpaßfiltern mit Durchlaßbändern umfaßt, die jeweils die Frequenzen $[f_1, \dots, f_N]$ aufweisen, wobei N Bandpaßsignale gebildet werden; bei dem der Schritt des Transponierens den Schritt eines Verschiebens der Bandpaßsignale frequenzmäßig in Regionen umfaßt, die die Frequenzen $M[f_1, \dots, f_N]$ aufweisen, wobei $M \neq 1$ der Transpositionsfaktor ist.

9. Ein Verfahren gemäß Anspruch 8, bei dem der Schritt des Verschiebens durch eine USB-Modulation (USB = upper side band = oberes Seitenband) erhalten wird.

10. Ein Verfahren gemäß Anspruch 1, bei dem der Schritt des Bereitstellens den Schritt eines Bandpaßfilterns eines Signals unter Verwendung einer Analysefilterbank oder einer Transformation einer solchen Art, daß reell- oder komplexwertige Subbandsignale eines Tiefpaßtyps erzeugt werden, umfaßt; bei dem der Schritt des Transponierens den Schritt eines Einfügens einer willkürlichen Anzahl von Kanälen k der Analysefilterbank oder der Transformation zu Kanälen M_k , $M \neq 1$, in einer Synthesefilterbank oder einer Transformation umfaßt, wobei M ein Transpositionsfaktor ist; und bei dem die Synthesefilterbank oder die Transformation bei dem Schritt des Kombinierens verwendet wird.

11. Ein Verfahren gemäß Anspruch 10, bei dem die Filterbank maximal dezimiert ist und der Schritt des Einfügens gemäß der folgenden Beziehung durchgeführt wird

$$v_{Mk}(n) = (-1)^{(M-1)kn} v_k(n),$$

wobei $(-1)^{(M-1)kn}$ ein Korrekturfaktor ist, $v_k(n)$ das Subbandsignal des Kanals k ist und $v_{Mk}(n)$ das Subbandsignal des Kanals M_k ist, wodurch eine Kompensation von spektralinvertierten Subbandsignalen erhalten wird.

12. Ein Verfahren gemäß Anspruch 1, bei dem der Schritt des Transponierens folgende Schritte umfaßt: Einfügen von Phasen der Subbandabtastwerte der Analysesubbänder oder der Analysefrequenzkoeffizienten von Kanälen k einer Analysefilterbank oder Transformation als Phasen von Subbandabtastwerten, die Synthesekanälen M_k zugeordnet sind, wobei M ein Transpositionsfaktor und eine Ganzzahl $\neq 1$ ist und wobei k eine Kanalanzahl ist; und Einfügen von Beträgen der Subbandabtastwerte der Analysesubbänder oder der Analysespektralkoeffizienten von aufeinanderfolgenden Kanälen 1 einer Analysefilterbank oder Transformation als Beträge von Subbandab-

tastwerten oder Spektralkoeffizienten, die aufeinanderfolgenden Synthesekanälen $1 + S$ zugeordnet sind, wobei S eine Ganzzahl $\neq 1$ ist, und wobei 1 eine Kanalanzahl ist.

13. Ein Verfahren gemäß Anspruch 10 oder 12, bei dem bei dem Schritt des Transponierens Phasen der Subbandabtastwerte der Kanäle k durch den Faktor M multipliziert werden, bevor die Synthesefilterbank oder Transformation bei dem Schritt des Kombinierens verwendet wird.

14. Ein Verfahren gemäß Anspruch 12, bei dem $M = K^{\pm 1}$, wobei K eine Ganzzahl > 1 ist.

15. Ein Verfahren gemäß Anspruch 10, bei dem bei dem Schritt des Einfügens mehrere Werte des Transpositions faktors M verwendet werden.

16. Eine Vorrichtung zum Decodieren eines codierten Signals, wobei das codierte Signal von einem ursprünglichen Signal abgeleitet ist und lediglich einen Teil von Frequenzbändern darstellt, die in dem ursprünglichen Signal enthalten sind, die folgende Merkmale aufweist:

eine Einrichtung zum Bereitstellen von Subbandabtastwerten für eine Mehrzahl von Subbändern oder einer Mehrzahl von Spektralkoeffizienten, wobei die Subbandabtastwerte oder Spektralkoeffizienten einen Frequenzgehalt des codierten Signals (**201, 301**) darstellen;

eine Einrichtung zum Extrahieren eines Bandpaßsignals (**203, 303**), wobei das Bandpaßsignal Subbandabtastwerte von einer vorbestimmten Anzahl von Analysesubbändern aufweist oder eine vorbestimmte Anzahl von Analysespektralkoeffizienten aufweist, wobei die vorbestimmte Anzahl kleiner ist als eine Gesamtanzahl von Subbändern oder Spektralkoeffizienten, die den Frequenzgehalt des codierten Signals (**201, 301**) darstellen;

eine Einrichtung zum Transponieren (**205, 305**) der Subbandabtastwerte der Analysesubbänder oder der Analysespektralkoeffizienten, die in dem Bandpaßsignal (**203, 303**) enthalten sind, in ein Rekonstruktionsfrequenzband (**205, 305**), das in dem ursprünglichen Signal enthalten ist und das in dem codierten Signal nicht enthalten ist, wobei das Rekonstruktionsfrequenzband Synthesesubbänder oder Synthesespektralkoeffizienten aufweist, wobei der Schritt des Transponierens ein Wiederverbinden von Analysesubbändern mit Synthesesubbändern oder ein Wiederverbinden von Analysespektralkoeffizienten mit Synthesespektralkoeffizienten umfaßt;

wobei die Subbandabtastwerte oder Spektralkoeffizienten, die in dem Bandpaßsignal (**203, 303**) enthalten sind, bevor oder nachdem dieselben durch die Einrichtung zum Transponieren verarbeitet wurden, spektralhüllkurveeingestellt werden (**207, 307**), wobei Spektralhüllkurveninformationen verwendet werden, die von dem ursprünglichen Signal oder dem codierten Signal abgeleitet sind, um spektralhüllkurveeingestellte transponierte Subbandabtastwerte oder spektralhüllkurveeingestellte transponierte Spektralkoeffizienten zu erhalten, wobei die Spektralhüllkurveninformationen das Spektralhüllkurven einstellen steuern, so daß eine Spektralhüllkurve des Rekonstruktionsfrequenzbands durch die Spektralhüllkurveninformationen bestimmt ist; und eine Einrichtung zum Kombinieren der Subbandabtastwerte und der eingestellten transponierten Subbandabtastwerte, oder der Spektralkoeffizienten und der eingestellten transponierten Spektralkoeffizienten, um ein decodiertes Ausgangssignal (**209, 309**) zu erhalten, wobei das decodierte Ausgangssignal einen Frequenzgehalt aufweist, der den Frequenzgehalt des codierten Signals und den Frequenzgehalt des Rekonstruktionsfrequenzbands umfaßt.

17. Ein Decoder gemäß Anspruch 16, bei dem das codierte Ausgangssignal ein monophones Audiosignal ist, wobei der Decoder ferner folgende Merkmale aufweist:

eine erste Verzögerung und ein erstes Dämpfungsglied zum Bilden eines ersten verzögerten Signals aus dem decodierten Ausgangssignal;

eine zweite Verzögerung, die sich von der ersten Verzögerung unterscheidet, und ein zweites Dämpfungsglied zum Bilden eines zweiten verzögerten Signals aus dem decodierten Ausgangssignal;

einen ersten Addierer zum Addieren des decodierten Ausgangssignals und des ersten verzögerten Signals, wobei ein Links-Kanal-Ausgangssignal gebildet wird; und

einen zweiten Addierer zum Addieren des decodierten Ausgangssignals und des zweiten verzögerten Signals, wobei ein Rechts-Kanal-Ausgangssignal gebildet wird; wodurch ein pseudostereophones Signal erhalten wird.

Es folgen 13 Blatt Zeichnungen

Anhängende Zeichnungen

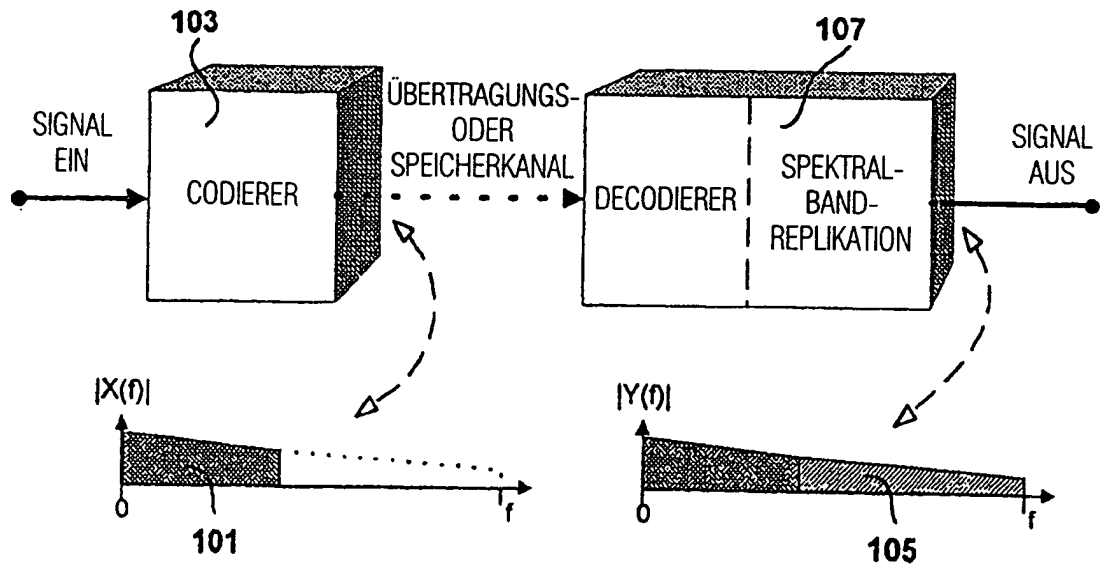


Fig. 1

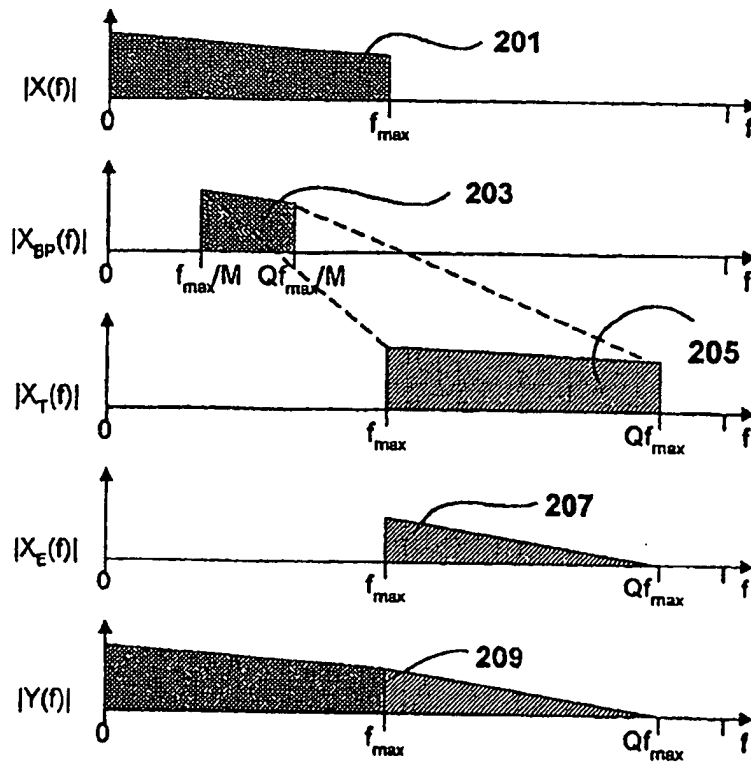
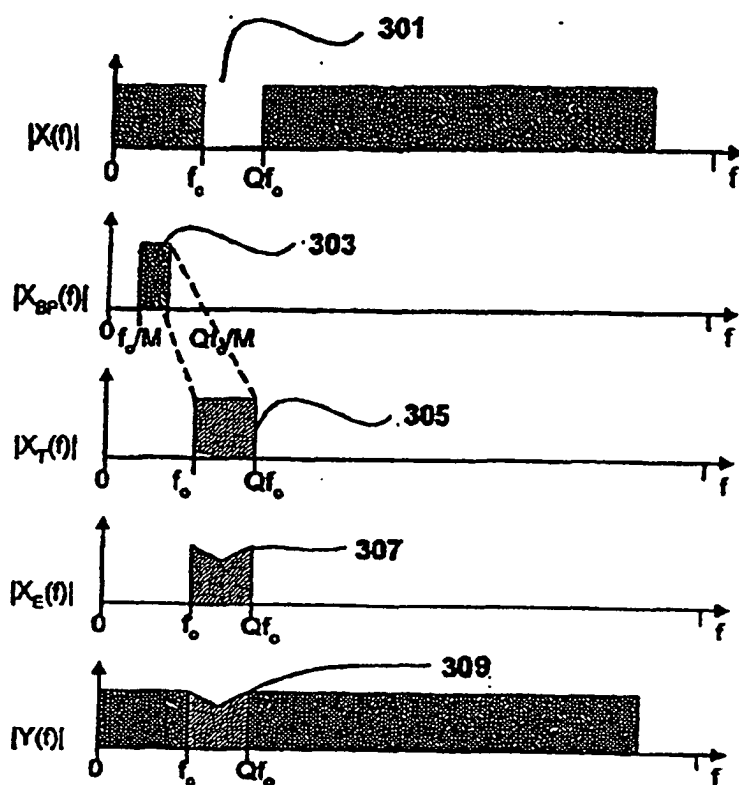
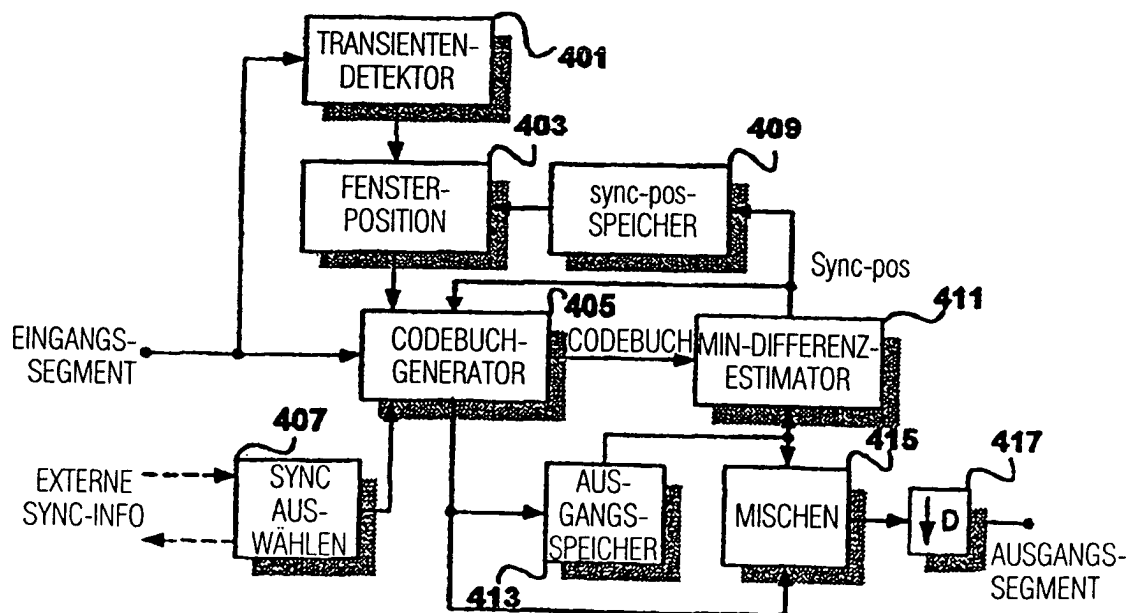
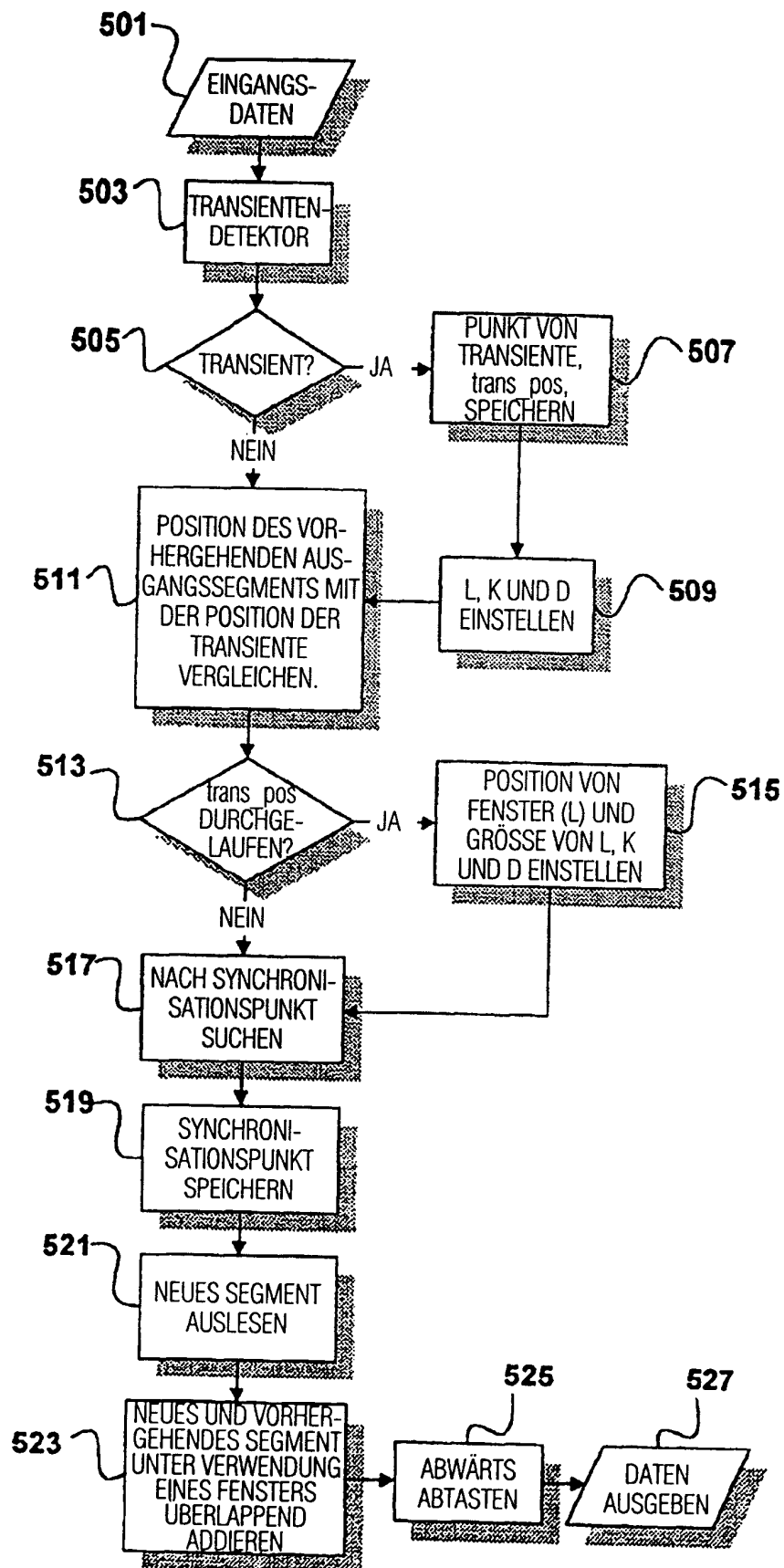
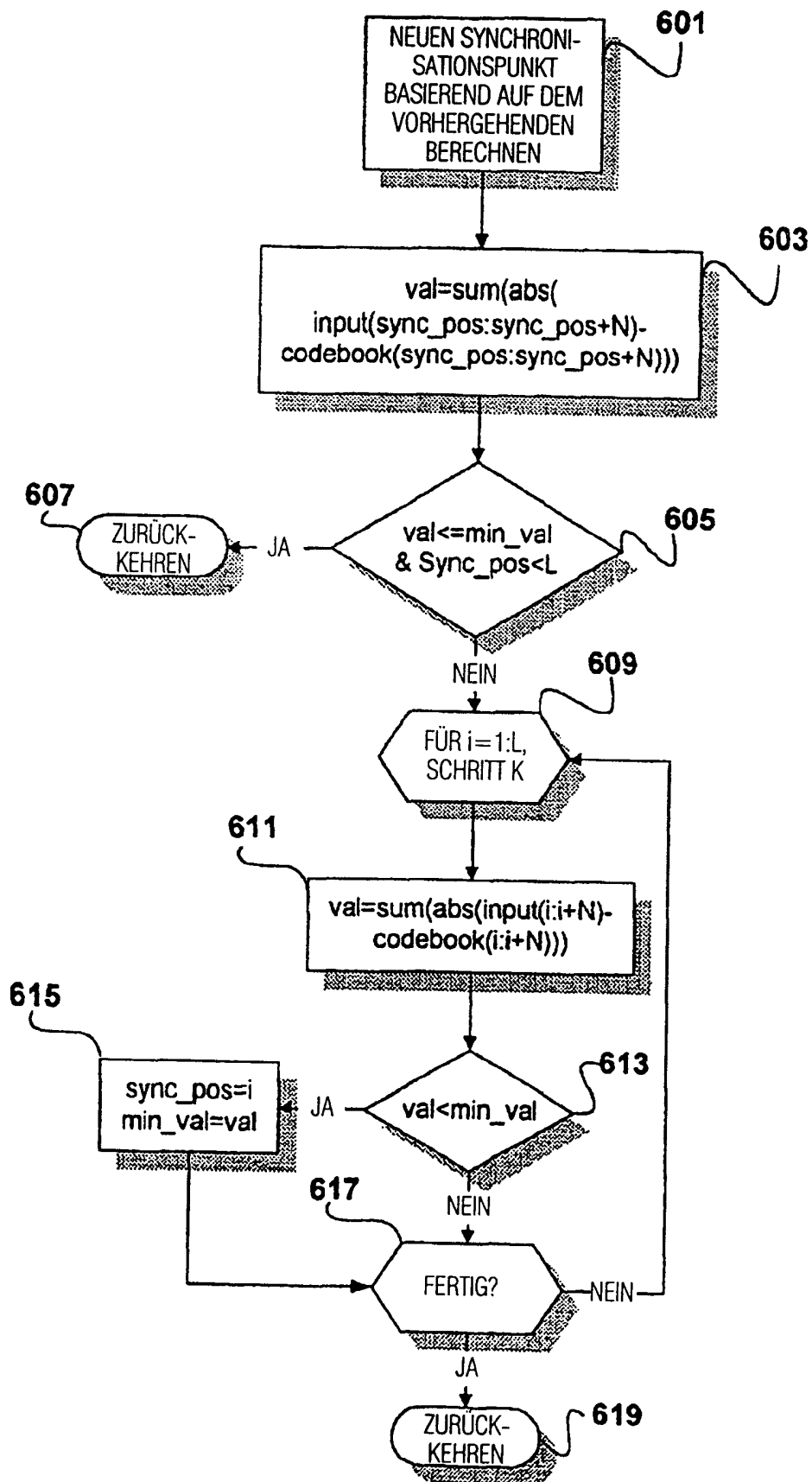


Fig. 2

**Fig. 3****Fig. 4**

*Fig. 5*

**Fig. 6**

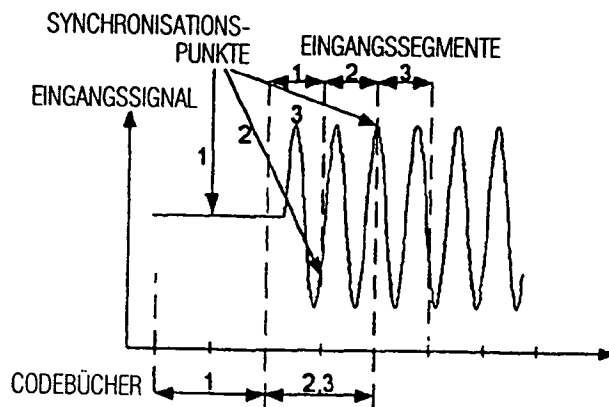


Fig. 7a

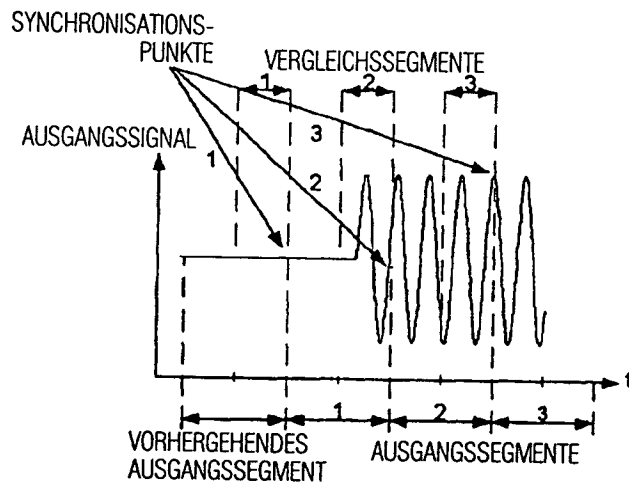


Fig. 7b

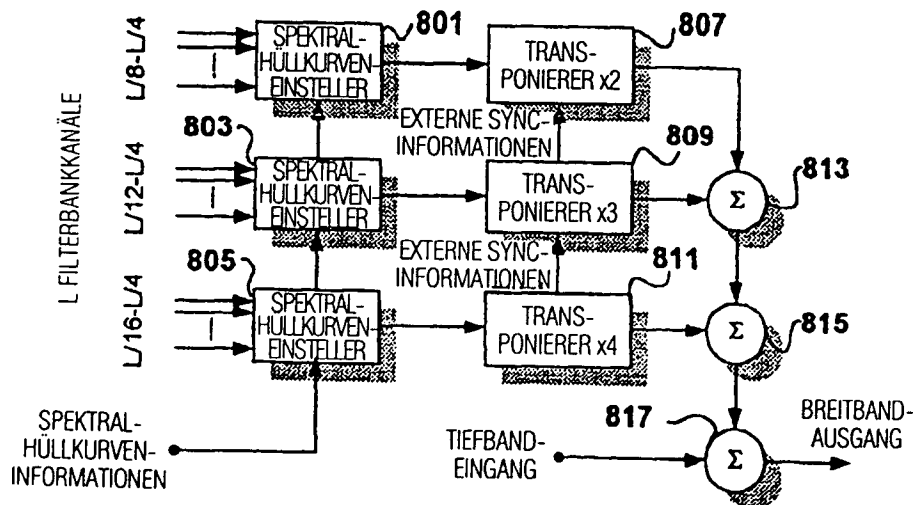


Fig. 8

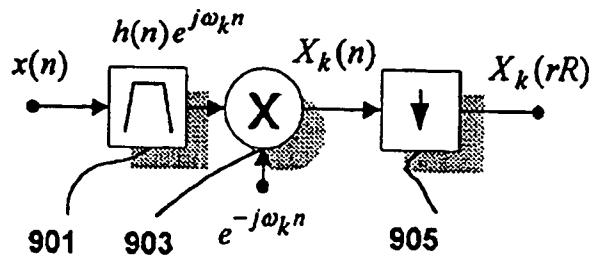


Fig. 9a

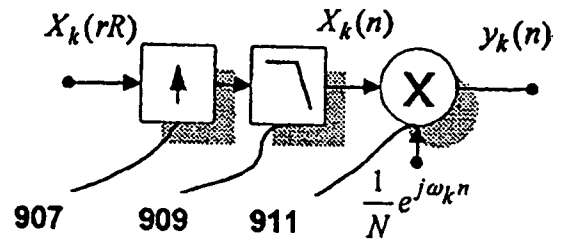


Fig. 9b

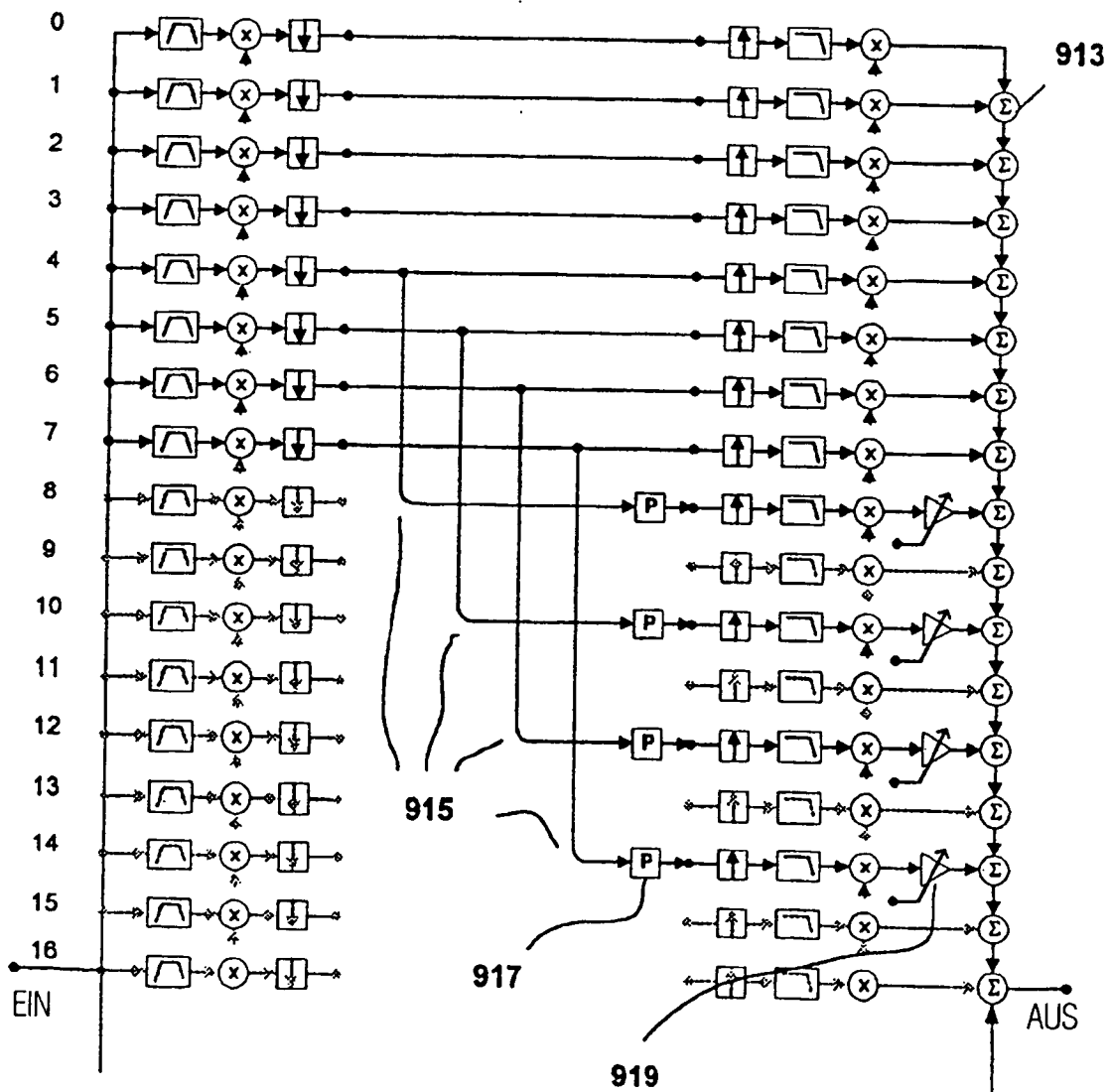


Fig. 9c

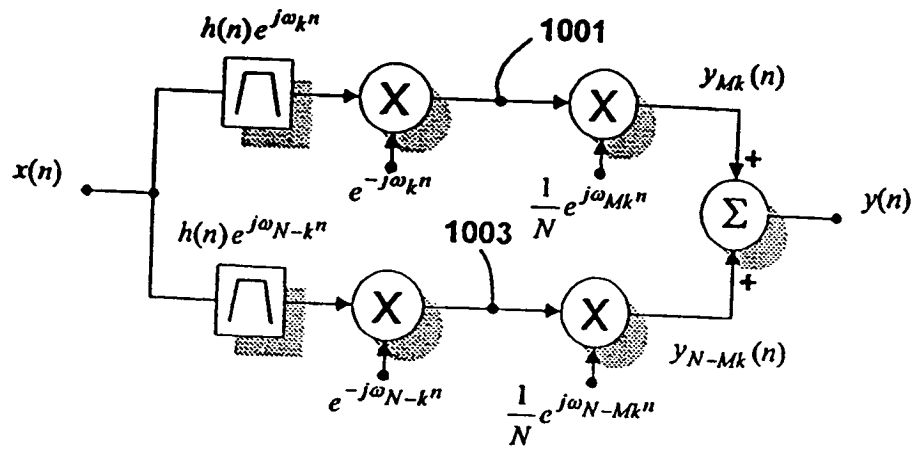


Fig. 10a

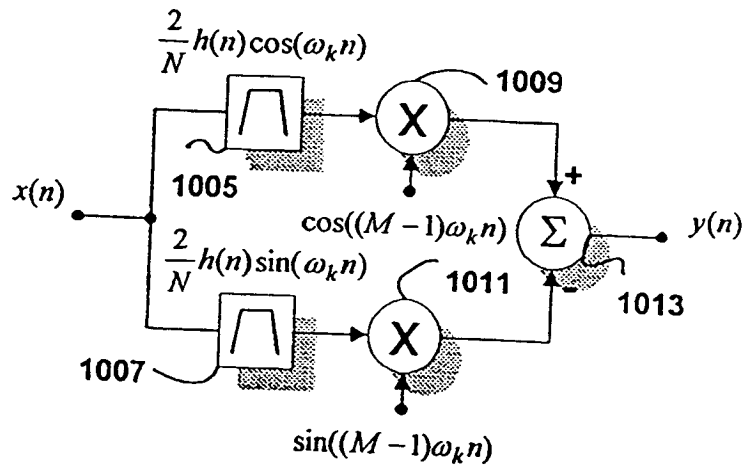


Fig. 10b

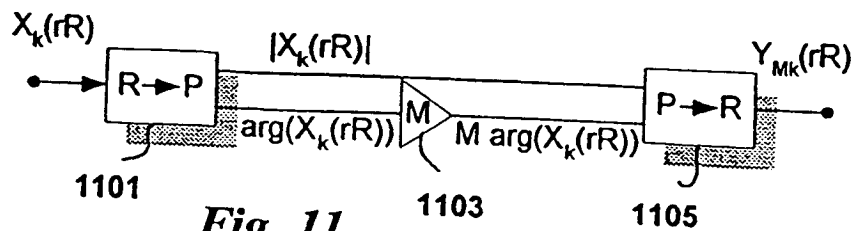


Fig. 11

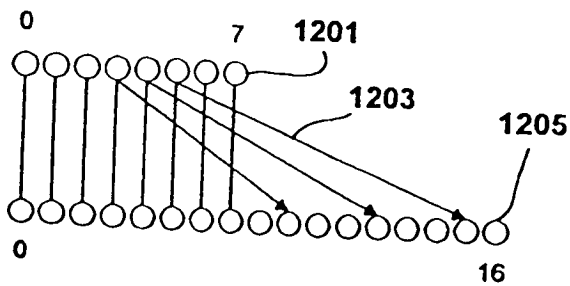


Fig. 12

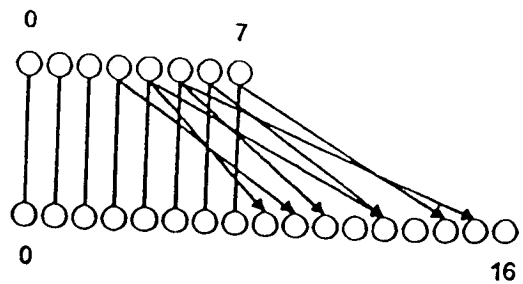


Fig. 13

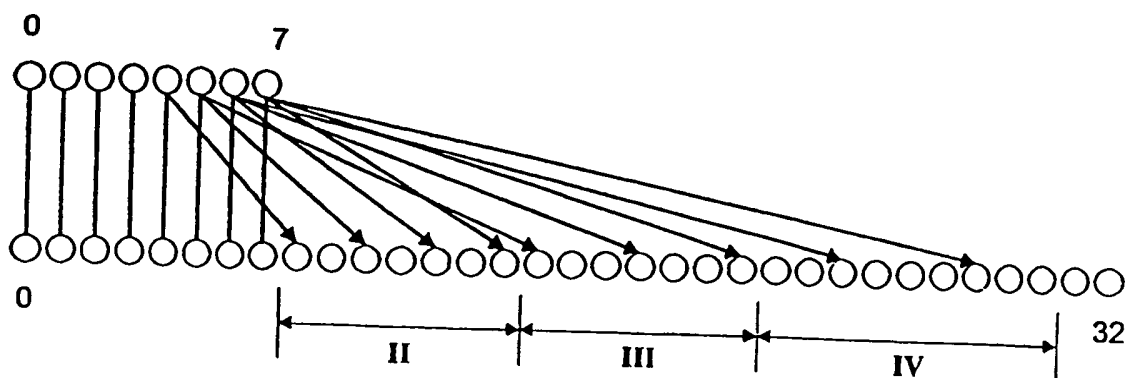


Fig. 14

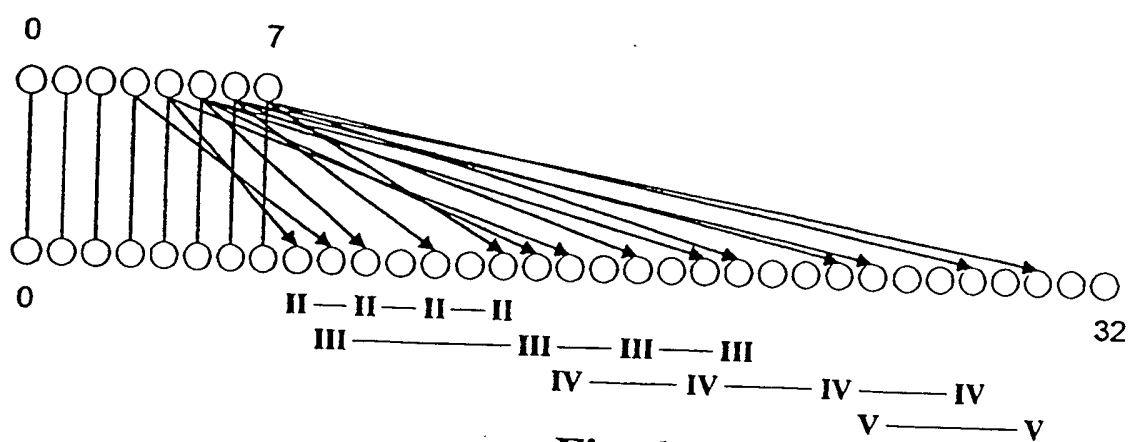


Fig. 15

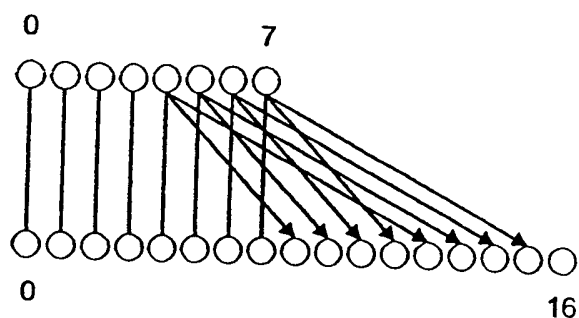


Fig. 16

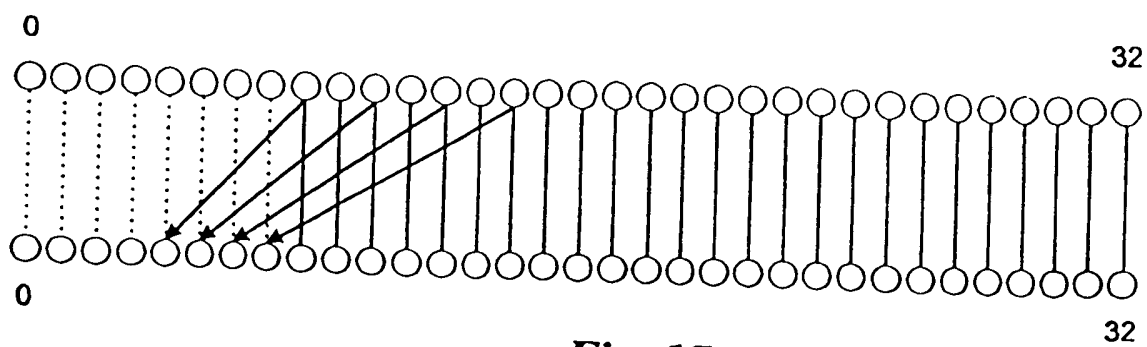
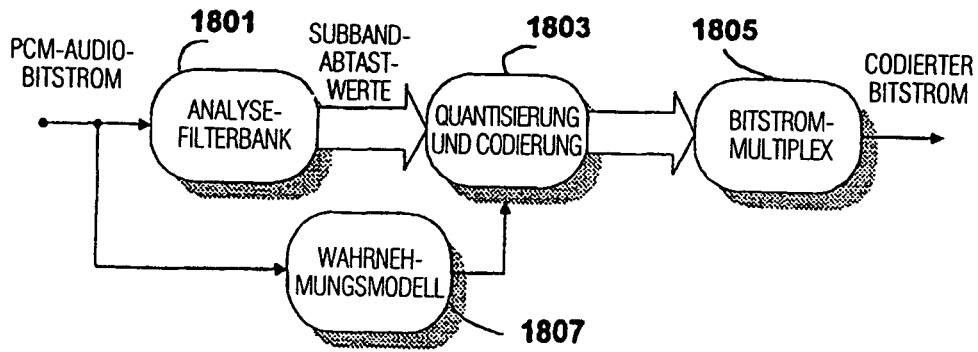
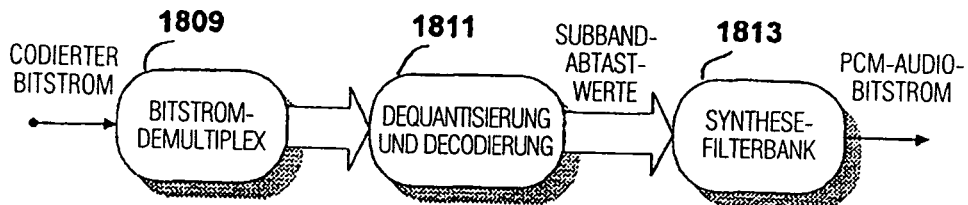
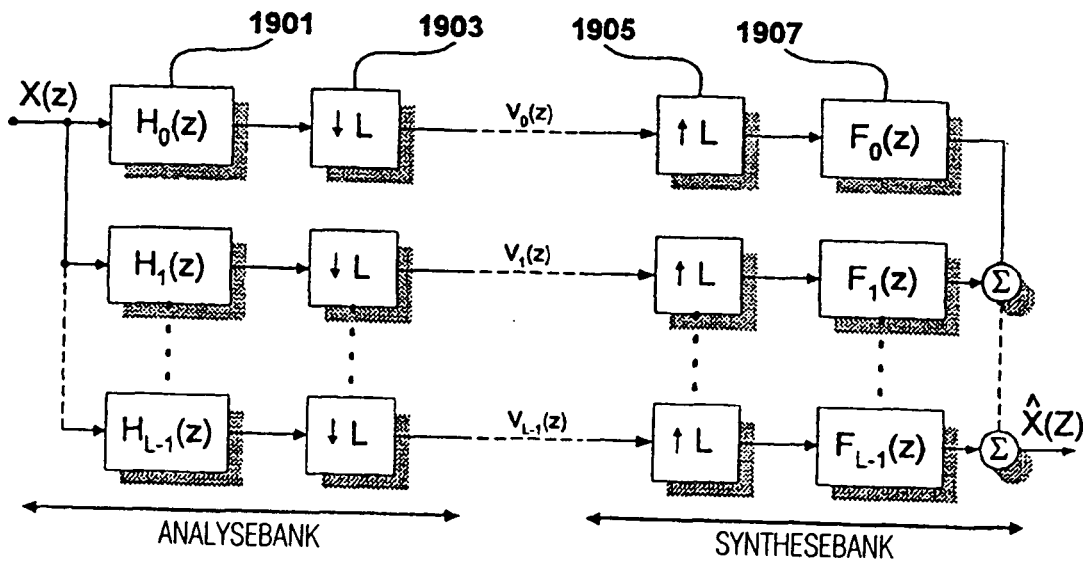


Fig. 17

*Fig. 18a**Fig. 18b**Fig. 19*

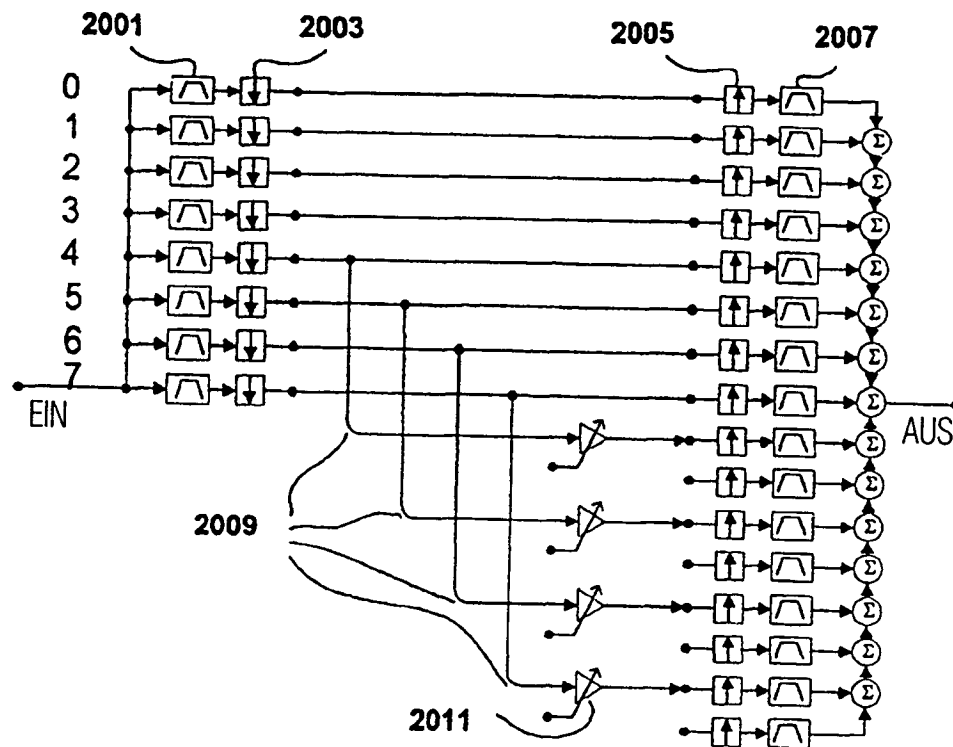


Fig. 20

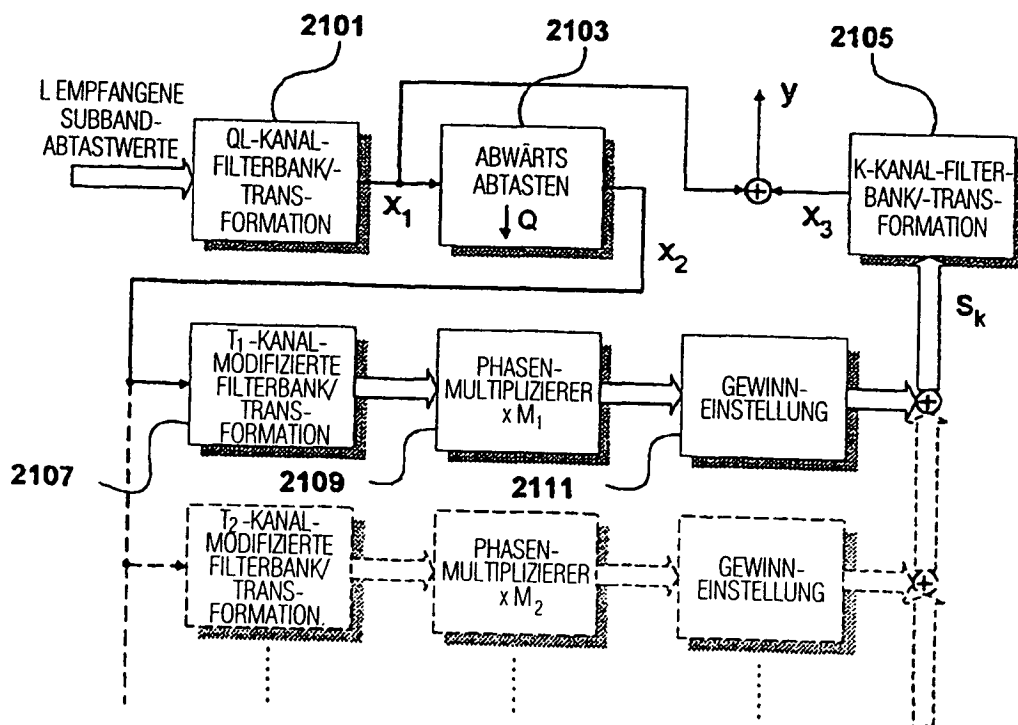


Fig. 21

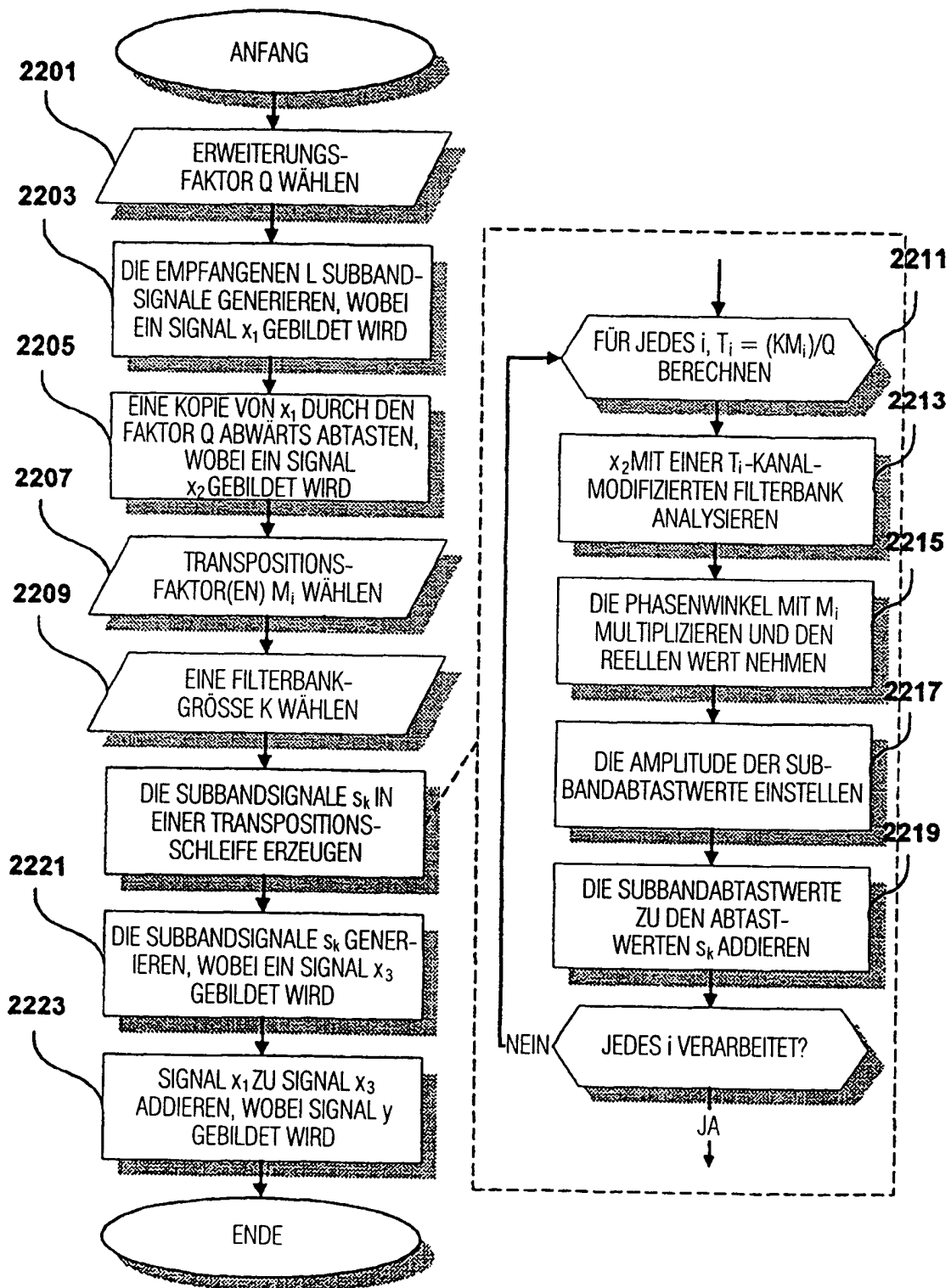
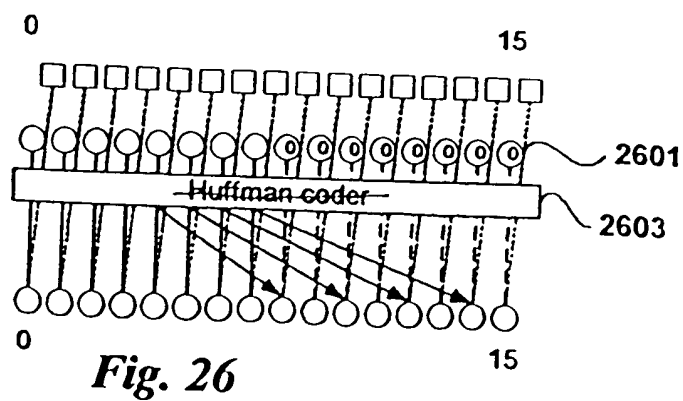
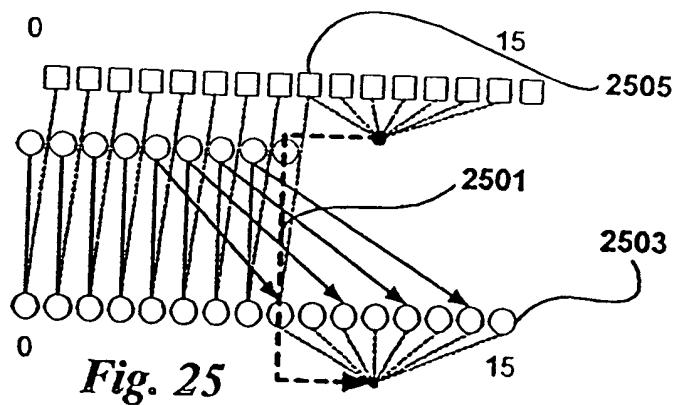
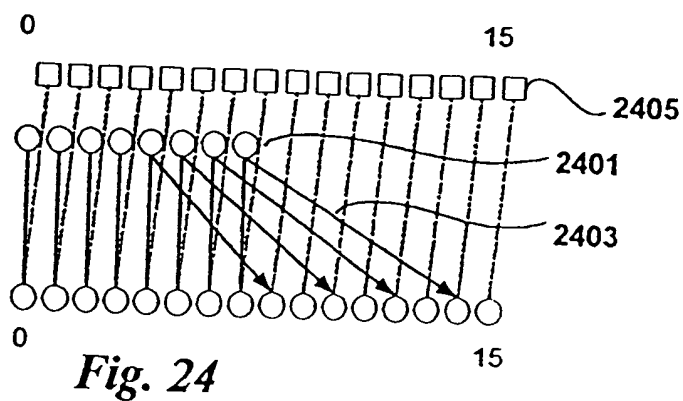
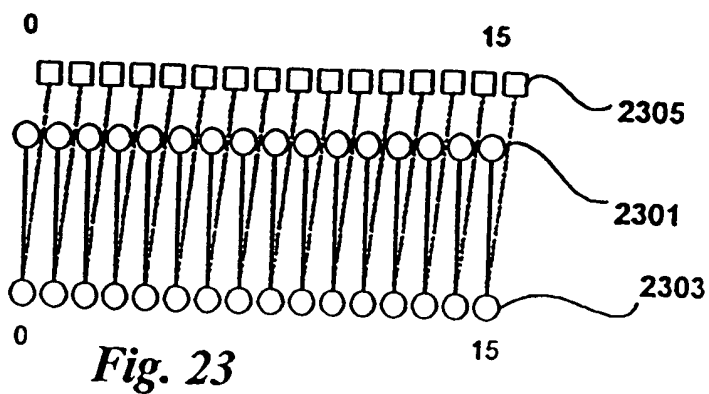


Fig. 22



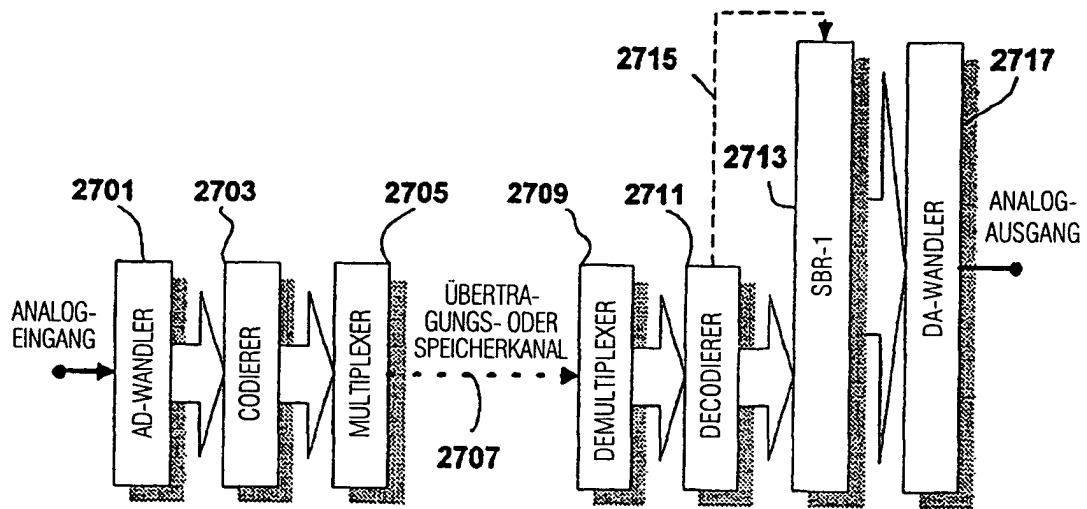


Fig. 27

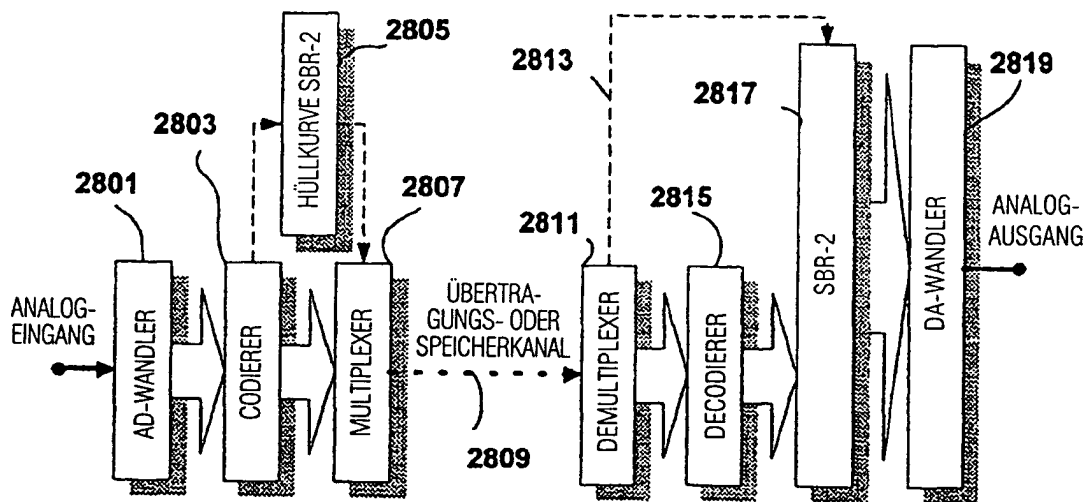


Fig. 28

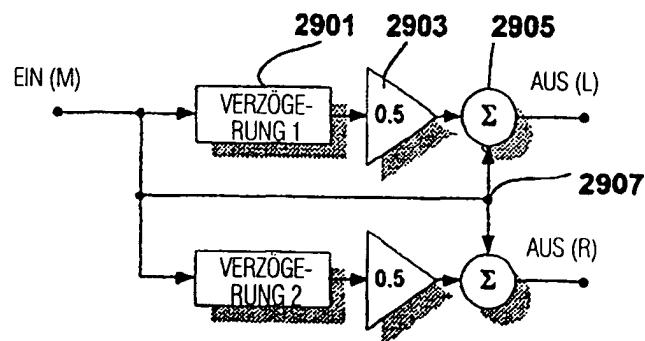


Fig. 29