

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 388 147 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:

29.12.2004 Patentblatt 2004/53

(51) Int Cl.7: **G10L 21/02**

(86) Internationale Anmeldenummer:

PCT/DE2001/001826

(21) Anmeldenummer: **01943072.7**

(87) Internationale Veröffentlichungsnummer:

(22) Anmeldetag: **11.05.2001**

WO 2002/093561 (21.11.2002 Gazette 2002/47)

(54) **VERFAHREN ZUR ERWEITERUNG DER BANDBREITE EINES SCHMALBANDIG GEFILTERTEN SPRACHSIGNALS, INSBESONDERE EINES VON EINEM TELEKOMMUNIKATIONSGERÄT GESENDETEN SPRACHSIGNALS**

METHOD FOR ENLARGING THE BAND WIDTH OF A NARROW-BAND FILTERED VOICE SIGNAL, ESPECIALLY A VOICE SIGNAL EMITTED BY A TELECOMMUNICATION APPLIANCE

PROCEDE D'AGRANDISSEMENT DE LA LARGEUR DE BANDE D'UN SIGNAL VOCAL FILTRE EN BANDE ETROITE, EN PARTICULIER D'UN SIGNAL VOCAL EMIS PAR UN APPAREIL DE TELECOMMUNICATION

(84) Benannte Vertragsstaaten:
DE FR GB

(43) Veröffentlichungstag der Anmeldung:
11.02.2004 Patentblatt 2004/07

(73) Patentinhaber: **SIEMENS
AKTIENGESELLSCHAFT
80333 München (DE)**

(72) Erfinder:
• **AUBAUER, Roland
81669 München (DE)**

• **KLINKE, Stefano, Ambrosius
50169 Kerpen (DE)**
• **LORENZ, Frannk
46397 Bocholt (DE)**

(56) Entgegenhaltungen:

EP-A- 0 945 852	EP-A- 0 994 464
US-A- 4 700 390	US-A- 5 455 888
US-A- 5 933 801	

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).

EP 1 388 147 B1

Beschreibung

[0001] Die vorliegende Erfindung betrifft ein Verfahren zur Erweiterung der Bandbreite eines schmalbandig gefilterten Sprachsignals, insbesondere eines von einem Telekommunikationsgerät gesendeten Sprachsignals gemäß dem Oberbegriff des Patentanspruches 1, dem Oberbegriff des Patentanspruches 3 und dem Oberbegriff des Patentanspruches 12.

[0002] Sprachcodierverfahren sind durch ihre unterschiedlichen Bandbreiten charakterisiert. So gibt es beispielsweise Schmalband-Codierer (engl.: narrow-band coder), welche Sprachsignale, die im Frequenzbereich bis 4000 Hz liegen, in codierte Sprachsignale umsetzen und Breitband-Codierer (engl.: wideband coder), welche Sprachsignale, die typischerweise zwischen 50 und 7000 Hz liegen, in codierte Sprachsignale umsetzen. Die Sprachsignale, die dem Schmalband-Codierer zugeführt werden, werden dabei in der Regel mit einer geringeren Abtastrate abgetastet als die Sprachsignale, die dem Breitband-Codierer zugeführt werden. Dafür ist die Nettobitrate des Schmalband-Codierers in der Regel niedriger als die Nettobitrate des Breitband-Codierers.

[0003] Werden die codierten Sprachsignale verschiedener Bandbreite innerhalb des gleichen Kanalmodus übertragen, so ermöglicht dies die Anwendung verschiedener Raten bei der Kanalcodierung, was zu unterschiedlichem Fehlerschutz führt. So ist es bei Anwendung des gleichen Kanalmodus möglich, bei schlechten Übertragungsbedingungen über den Übertragungskanal den schmalbandigen codierten Sprachsignalen im Zuge der Kanalcodierung mehr redundante Fehler-schutzbits hinzuzufügen als den breitbandigen codierten Sprachsignalen. Daher bietet sich bei variierenden Übertragungsbedingungen die Übertragung von Sprachsignalen über einen Übertragungskanal an, bei der abhängig von den Übertragungsbedingungen die Sprachcodierung zwischen einer breitbandigen und einer schmalbandigen Sprachcodierung umgeschaltet ["Wide-Band" to Narrow-Band"-Switching ("WB/NB"-Switching)] und die Kanalcodierung, insbesondere die Rate der Kanalcodierung, daran angepaßt wird. Empfangsseitig erfolgt eine an die Codierung angepaßte Decodierung der codierten Sprachsignale.

[0004] Bei dem neuen Telekommunikationssystem zur drahtlosen Telekommunikation UMTS (Universal Mobile Telecommunications System) ist beispielsweise eine Breitband-Codierung standardisiert worden, um mit den zukünftigen UMTS-Endgeräten eine sehr gute Sprachqualität zu gewährleisten.

[0005] Nachteilig bei einem derartigen Ansatz ist, dass ein empfangender Teilnehmer insbesondere das plötzliche Umschalten von Breitband-Codierung auf Schmalband-Codierung und den damit verbundenen Qualitätsverlust als äußerst störend empfindet.

[0006] Dieses sogenannte "WB/NB-Switching"-Problem kann auch bei der Handover-Situation in Telekom-

munikationssystemen zur drahtlosen Telekommunikation mit mehreren Basisstationen und Mobilteilen, wobei die Basisstationen unterschiedlichen Telekommunikationsteilsystemen zugeordnet sind und die Mobilteilen innerhalb des Systems für ein teilsystemübergreifendes Roaming als Dual-Mode-Mobilteilen ausgebildet sind, auftreten: Ausgangspunkt der Betrachtungen ist eine bestehende breitbandige Gesprächsverbindung zwischen einer Basisstation und einem Mobilteil. Wenn nun für das Mobilteil bzw. den Gesprächsteilnehmer eine Übergabe (Handover) an eine andere Basisstation durchgeführt wird, kann der Fall eintreten, dass die übernehmende Basisstation zu einem Teilsystem gehört, welches den breitbandigen Sprachservice nicht unterstützt. Aus diesem Grunde wird dann auf die schmalbandige Codierung und Decodierung zurückgeschaltet.

[0007] Auch in diesem Szenario wird der empfangende Teilnehmer insbesondere das plötzliche Umschalten von Breitband-Codierung auf Schmalband-Codierung und den damit verbundenen Qualitätsverlust als äußerst störend empfinden.

[0008] Basisstationen, die wie oben beschrieben keine breitbandige Gesprächsverbindung unterstützen, sowie andere Telekommunikationsendgeräte, welche lediglich Schmalband-Codierung oder analoge Sprachsignalübertragung im Bereich von typisch 300 bis 3400 Hz ermöglichen, sind noch weit verbreitet, da die bisher bekannten Telekommunikationssysteme Sprachsignale bisher im Allgemeinen mit einer Bandbreite von etwa 3,1 kHz zwischen 3400 Hz (erste Grenzfrequenz) und 300 Hz (zweite Grenzfrequenz) übertragen, da die Verständlichkeit der Kommunikation trotz der damit gegebenen Bandbegrenzung der Sprache ausreichend ist. Zur Übertragung der Sprachsignale verwenden die bisher bekannten Telekommunikationssysteme dabei verschiedene digitale und analoge Codierverfahren.

[0009] Um eine Qualitätsverbesserung derart zu erzielen, dass eine Sprachqualität in Telekommunikationssystemen mit der Sprachqualität bei Radio- und Fernsehsignalen vergleichbar ist, wird es erforderlich, Frequenzanteile der Sprache, die über die Bandbreite von 300 Hz bis 3400 Hz hinausgehen, empfängerseitig abzuschätzen und zu synthetisieren.

[0010] Im Stand der Technik sind verschiedene Verfahren bekannt, die eine Erweiterung der Bandbreite eines schmalbandigen Sprachsignals ermöglichen.

[0011] Beispielsweise ist für eine Erweiterung der Bandbreite im unteren Frequenzbereich (<300Hz) aus der EP 0 994 464 eine Wiederherstellung von Signalanteilen des unteren Frequenzbereichs eines durch eine Hochpassfunktion zu tiefen Frequenzen hin begrenzten Sprachsignals bekannt, wobei die beschriebene Hochpass-Filterung z.B. bei der Sprachübertragung über ein Telefon beim fernen Teilnehmer durchgeführt wird (Sende- Charakteristik des Fernsprechers).

[0012] Die Wiederherstellung erfolgt dabei durch Generieren von Frequenzen des unteren Frequenzbe-

reichs durch eine nichtlineare Signalverarbeitung, mittels der subharmonische Frequenzen des Signals erzeugt und zum Hochpasssignal hinzuaddiert werden.

[0013] Des Weiteren ist in der EP 0 994 464 dazu auch eine Weiterbildung bekannt, bei der die nichtlineare Signalverarbeitung durch die Multiplikation des Signals mit einer Funktion des Signals durchgeführt wird.

[0014] Nachteilig an den genannten Verfahren ist es, dass in der Regel die Filtercharakteristik (Sende-Charakteristik des Fernsprechers), mit der das Signal am fernen Teilnehmerendgerät gefiltert wurde, unbekannt ist und für verschiedene Gerätetypen sehr unterschiedlich sein kann. Dies ist in FIGUR 7 dargestellt. Eine Wiederherstellung des Sprachsignals ist deshalb nur dann möglich, wenn die Filtercharakteristiken der beteiligten Teilnehmergeräte jeweils bekannt oder diese Geräte aufeinander abgestimmt sind.

[0015] In vielen Verfahren der digitalen Sprachcodierung werden das digitale Sprachsignal zur Weiterverarbeitung und Übertragung in Koeffizienten, welche die spektrale Grobstruktur eines Signalabschnitts beschreiben, und in ein Anregungs- bzw. Prädiktionsfehlersignal, das sogenannte Restsignal, welches die spektrale Feinstruktur bildet, aufgespalten. Dieses Restsignal enthält nicht mehr die spektrale Einhüllende des Sprachsignals, die durch die Koeffizienten, die die spektrale Grobstruktur beschreiben, repräsentiert wird.

[0016] Auf der Decodierseite werden diese beiden - meist quantisiert übertragenen - Teile, welche die spektrale Grob- und Feinstruktur beschreiben, wieder zusammengefügt und bilden das decodierte Sprachsignal.

[0017] Eine typische Repräsentation für die spektrale Grobstruktur bilden die bei der linearen Prädiktionsanalyse ermittelten LPC-Koeffizienten (Linear Predictive Coding), welche ein rekursives Filter, das sogenannte Synthesefilter, beschreiben, dessen Übertragungsfunktion der spektralen Grobstruktur entspricht. Diese Koeffizienten werden in ihrer eigentlichen oder einer transformierten Form in vielen Sprachcodierern verwendet. Hierbei wird auf Empfängerseite das empfangene Restsignal als Eingangssignal für das Synthesefilter verwendet, so dass am Ausgang des Filters das rekonstruierte Sprachsignal verfügbar ist. Die LPC-Koeffizienten sind folglich eine Repräsentation der spektralen Grobstruktur eines Sprachsignalabschnitts und können unter Verwendung eines passenden Anregungssignals zur Synthese von Sprachsignalen verwendet werden.

[0018] Für eine Erweiterung der Bandbreite im oberen Frequenzbereich sind Verfahren bekannt, die auf besonderen Sprachdatenbüchern, sogenannten Codebüchern (Codebooks) basieren, die eine Relation zwischen den LPC-Koeffizienten eines schmalbandigen Sprachsignalabschnitts und denen eines breitbandigen Sprachsignalabschnitts bilden. Das hat zur Folge, dass die Codebücher gleichzeitig mit schmalbandiger und breitbandiger Sprache trainiert und im Kommunikationsendgerät abgespeichert werden müssen.

[0019] Außerdem wird aus dem schmalbandigen Restsignal, das durch die lineare Prädiktionsanalyse des schmalbandigen Sprachsignals erzeugt wurde, ein breitbandiges Anregungssignal erzeugt, welches Frequenzkomponenten oberhalb der Bandbreite des schmalbandigen Sprachsignals enthält.

[0020] Da die Codebücher im Telekommunikationsgerät gespeichert werden müssen, ist neben dem aufwendigen Training der Codebücher sowohl mit schmalbandiger als auch mit breitbandiger Sprache, auch der hohe Bedarf an Speicher und die Schwierigkeit einer sprecher- und sprachunabhängigen eindeutigen Zuordnung zwischen beiden Codebüchern nachteilig.

[0021] Um den Speicherplatzbedarf bei der Verwendung von Codebüchern zu verringern, ist es gemäß einem von der Technischen Hochschule Aachen entwickelten Verfahren bekannt, nur noch ein Codebuch in Verbindung mit einem Hidden-Markov-Modell, mit dem die statistischen Spracheigenschaften beschrieben werden können, zu benutzen.

[0022] In der Praxis haben diese Verfahren zur Erweiterung der Bandbreite im oberen Frequenzbereich keine Anwendung gefunden, da zudem die Qualität der erzeugten breitbandigen Sprachsignale unzureichend und von dem jeweiligen Sprachsignal abhängig ist.

[0023] Die der Erfindung zugrundeliegende Aufgabe besteht darin, die Bandbreite eines schmalbandig gefilterten Sprachsignals auf einfache und kostengünstige Weise ohne Qualitätseinbußen zu erweitern.

[0024] Diese Aufgabe wird ausgehend von dem im Oberbegriff des Anspruchs 1 definierten Verfahren durch die im Kennzeichen des Anspruchs 1 angegebenen Merkmale, ausgehend von dem im Oberbegriff des Anspruchs 3 definierten Verfahren durch die im Kennzeichen des Anspruchs 3 angegebenen Merkmale sowie ausgehend von dem im Oberbegriff des Anspruchs 12 definierten Verfahren durch die Kennzeichen des Anspruchs 12 angegebenen Merkmale gelöst.

[0025] Zwei Methoden, wie das schmalbandig gefilterte Sprachsignal in bezug auf Frequenzanteile oberhalb einer Grenzfrequenz im Frequenzbereich geschätzt werden kann, sind in den Ansprüchen 1 und 3 angegeben, wonach

zunächst jeweils das schmalbandige Sprachsignal in eine spektrale Struktur aufweisende Sprachsignalzeitabschnitte unterteilt wird, jeder schmalbandige Sprachsignalzeitabschnitt als ein stimmhafter Laut oder als ein stimmloser Laut klassifiziert wird, eine spektrale Struktur aufweisende Ergänzungen zur Erweiterung des schmalbandigen Sprachsignals in bezug auf die vorgenommene lautartbezogene Klassifizierung erzeugt werden, wobei zumindest für den Fall des stimmhaften Lautes die Ergänzung unabhängig von dem jeweiligen Laut ist, die spektrale Struktur des schmalbandigen Sprachsignalzeitabschnittes, die gemäß Anspruch 2 vorzugsweise durch eine FFT-Analyse (Fast Fourier Transformation) berechnet wird, und die spektrale Struktur der erzeugten Ergänzung zeitabschnittsweise derart ver-

knüpft werden, dass jeweils eine erweiterte spektrale Struktur entsteht und anschließend

gemäß Anspruch 1 aus der erweiterten spektralen Struktur, insbesondere durch eine IFFT-Analyse (Inverse Fast Fourier Transformation) gemäß Anspruch 2, jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt wird oder gemäß Anspruch 3 bezüglich der Zeitabschnittdauer den schmalbandigen Sprachsignalzeitabschnitten entsprechende Prädiktionsfehlersignale erzeugt werden und aus der erweiterten spektralen Struktur und dem jeweiligen breitbandigen Prädiktionsfehlersignalzeitabschnitt jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt wird, bevor abschließend aus den einzelnen breitbandigen erweiterten Sprachsignalzeitabschnitten ein breitbandiges erweitertes Sprachsignal erzeugt wird.

[0026] Eine alternative Methode, wie das schmalbandig gefilterte Sprachsignal in Bezug auf Frequenzanteile oberhalb der Grenzfrequenz im Zeitbereich geschätzt werden kann, ist im Anspruch 12 angegeben, wonach zunächst das schmalbandige Sprachsignal in Sprachsignalzeitabschnitten unterteilt wird und jeder schmalbandige Sprachsignalzeitabschnitt als ein stimmhafter Laut oder als ein stimmloser Laut klassifiziert wird und anschließend

die schmalbandigen Sprachsignalzeitabschnitte derart nichtlinear verarbeitet werden, dass jeweils ein modifizierter Sprachsignalzeitabschnitt erzeugt wird, der einerseits den jeweiligen im wesentlichen unveränderten schmalbandigen Sprachsignalzeitabschnitt und andererseits oberhalb der ersten Grenzfrequenz durch die nichtlineare Signalverarbeitung erzeugte Signalanteile enthält und die modifizierten Sprachsignalzeitabschnitte in Bezug auf die vorgenommene lautartbezogene Klassifizierung derart unterschiedlich gefiltert werden, dass aus den modifizierten Sprachsignalzeitabschnitten breitbandige erweiterte Sprachsignalzeitabschnitte und damit ein breitbandiges erweitertes Sprachsignal entsteht.

[0027] Das Schätzen der Frequenzanteile oberhalb der Grenzfrequenz des schmalbandig gefilterten Sprachsignals im Zeitbereich ist von Vorteil, weil keine Begutachtung des Spektrums und daher keine rechenintensive Transformation in den Spektralbereich notwendig ist. Im Übrigen werden die modifizierten Sprachsignalzeitabschnitte derart gefiltert, dass im Fall eines stimmhaften Sprachsignalzeitabschnittes wenig Energie oberhalb von der ersten Grenzfrequenz - z.B. 4 kHz - und im Fall eines stimmlosen Sprachsignalzeitabschnittes mehr Energie oberhalb von der ersten Grenzfrequenz - z.B. 4 kHz - durchgelassen wird.

[0028] Ein wesentlicher Vorteil der vorgestellten erfindungsgemäßen Verfahren zur Erweiterung eines schmalbandig gefilterten Sprachsignals im oberen Frequenzbereich gemäß der Ansprüche 1, 3 und 12 gegenüber den bekannten Verfahren besteht in der Einsparung von Speicherplatz, weil im Wesentlichen auf spei-

cherplatzaufwendige Codebücher verzichtet werden kann. Außerdem erlauben sie die Erweiterung des schmalbandigen Sprachsignals ohne genaue Kenntnis des ursprünglichen breitbandigen Anregungssignals. Zudem zeichnen sich die Verfahren gemäß Anspruch 3 und 12 durch sehr geringen Rechenaufwand aus. Schließlich entfällt bei sämtlichen Verfahren das Training der speicheraufwendigen Codebücher, welches üblicherweise in der Entwicklungsphase von zur Sprachübertragung benutzten Telekommunikationsgeräten durchgeführt werden muss.

[0029] Bei der Weiterbildung gemäß Anspruch 4 wird die für die als stimmhafte Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils erzeugte Ergänzung derart erzeugt, dass die Energie dieser Ergänzung in Bezug auf die Gesamtenergie des schmalbandigen Sprachsignalabschnittes vernachlässigbar ist.

[0030] Diese Ergänzung kann stets die gleiche sein, unabhängig davon, um welchen stimmhaften Laut - z. B.: "a", "e" oder "i" - es sich handelt, so dass eine Bestimmung des Lautes sowie die Anwendung eines Codebuchs für stimmhafte Laute entfällt.

[0031] Durch die Weiterbildung gemäß Anspruch 4 ist eine Qualitätsverbesserung des breitbandigen erweiterten Sprachsignals gewährleistet, da durch diese Art der Weiterbildung berücksichtigt wird, dass bei stimmlosen Lauten im oberen Frequenzbereich ein wesentlicher Teil der Signalenergie fortgesetzt wird, so dass eine Vernachlässigung des genauen Verlaufs dieses Teils verhindert wird, die dadurch erfolgt, daß stets die gleiche Ergänzung vorgenommen wird und somit das synthetisierte Sprachsignal verfälscht würde.

[0032] Bei der Weiterbildung gemäß Anspruch 5 wird die für die als stimmlose Laute klassifizierten schmalbandigen Sprachsignalabschnitte jeweils erzeugte Ergänzung derart erzeugt, daß die Energie dieser Ergänzung in Bezug auf die Gesamtenergie des schmalbandigen Sprachsignalabschnittes nicht vernachlässigbar ist. Auf diese Weise kann einfach ohne genaue Kenntnis des stimmlosen Lautes eine Erweiterung des schmalbandig gefilterten Sprachsignals durchgeführt werden.

[0033] Bei der Weiterbildung gemäß Anspruch 6 wird die für die als stimmlose Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils erzeugte Ergänzung derart erzeugt, dass auf Basis von zumindest einem Breitband-Codebuch aus ersten Filterkoeffizienten des schmalbandigen Sprachsignalzeitabschnittes zweite Filterkoeffizienten eines breitbandigen Sprachsignalzeitabschnittes ermittelt werden. Dadurch kann die Qualität des synthetisierten Sprachsignals gegenüber dem Sprachsignal, wo kein Codebuch verwendet wird, verbessert werden.

[0034] Die Weiterbildung gemäß Anspruch 7 erlaubt die Wiederherstellung eines im oberen Frequenzbereich erweiterten breitbandigen Sprachsignals anhand von ermittelten breitbandigen Filterkoeffizienten.

[0035] Die Weiterbildung gemäß Anspruch 8 erlaubt

die Wiederherstellung eines im oberen Frequenzbereich erweiterten breitbandigen Sprachsignals anhand von ermittelten breitbandigen Filterkoeffizienten und eines breitbandigen Prädiktionsfehlersignalzeitabschnittes.

[0036] Bei dem Verfahren gemäß Anspruch 3 werden für die Schätzung der Filterkoeffizienten für das Synthesefilter keine Codebücher benötigt, wodurch der Speicherplatzbedarf in vorteilhafter Weise verringert werden konnte. Allerdings ist die Schätzung der Frequenzeinhüllenden oberhalb der Grenzfrequenz, z.B. 4 kHz, sehr grob, was manchmal bei gewissen stimmlosen Lauten dazu führt, dass unerwünschte Artefakte entstehen. Um dies zu vermeiden, werden bei der Weiterbildung gemäß Anspruch 9 die breitbandigen Filterkoeffizienten mit den Einträgen aus einem Breitband-Codebuch verglichen und der Eintrag in dem Breitband-Codebuch, der am besten zu den breitbandigen Filterkoeffizienten passt, als Filterkoeffizient der Synthese des breitbandigen erweiterten Sprachsignals zugrundegelegt. Der Vorteil dieses Verfahrens liegt darin, dass durch die Benutzung eines Codebuches die auf der Basis des vorstehenden Codebuchvergleiches gefundenen Filterkoeffizienten eine gute Annäherung der echten Koeffizienten sowohl unterhalb der ersten Grenzfrequenz (z.B. 4 kHz) als auch oberhalb der ersten Grenzfrequenz (z.B. 4 kHz) sind. Das bedeutet, dass die Schätzung der Koeffizienten oberhalb der ersten Grenzfrequenz nicht mehr so grob ist. Darüber hinaus ist es vorteilhaft, dass einerseits nur noch das breitbandige Codebuch und nicht mehr zusätzlich das schmalbandige Codebuch benötigt wird und andererseits auch wie beim Stand der Technik (entwickeltes Verfahren an der TH-Aachen) kein Hidden-Markov-Modell mehr notwendig ist.

[0037] Um die Qualität des breitbandigen erweiterten Sprachsignals gemäß der Ansprüche 1 bis 3 zu verbessern, ist es von Vorteil, wenn gemäß Anspruch 11 der aus der erweiterten spektralen Struktur jeweils erzeugte breitbandige erweiterte Sprachsignalzeitabschnitt hochpassgefiltert wird, der hochpassgefilterte Sprachsignalzeitabschnitt mit dem entsprechenden schmalbandigen Sprachsignalzeitabschnitt verknüpft wird und aus den einzelnen verknüpften Sprachsignalzeitabschnitten das breitbandige erweiterte Sprachsignal erzeugt wird.

[0038] Bei der Weiterbildung gemäß Anspruch 13 werden die für die als stimmhaften Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils durch die nichtlineare Signalverarbeitung erzeugten Signalanteile derart erzeugt, dass die Energie des jeweiligen Signalanteils in bezug auf die Gesamtenergie des schmalbandigen Sprachsignalzeitabschnittes vernachlässigbar ist.

[0039] Bei der Weiterbildung gemäß Anspruch 14 werden die für die als stimmlosen Laute klassifizierten schmalbandigen Sprachsignalabschnitte jeweils durch die nichtlineare Signalverarbeitung erzeugten Signalanteile derart erzeugt, dass die Energie des jeweiligen Si-

gnalanteils in bezug auf die Gesamtenergie des schmalbandigen Sprachsignalzeitabschnittes nicht vernachlässigbar ist.

[0040] Gemäß Anspruch 15 ist es von Vorteil - weil einfach zu realisieren, wenn die Signalanteile durch spektrale Spiegelung erzeugt werden.

[0041] Das Verfahren zur Erweiterung des schmalbandig gefilterten Sprachsignals kann gemäß Anspruch 16 vorteilhaft - im Sinne einer vereinfachten Berechnung und Durchführung des Verfahrens - weitergebildet werden, indem die schmalbandigen Sprachsignalzeitabschnitte gleich lang gewählt werden.

[0042] Weitere vorteilhafte Ausgestaltungen sind in den übrigen Unteransprüchen angegeben.

[0043] Weitere Einzelheiten, Merkmale und Vorteile der Erfindung werden nachfolgend anhand der in den Figuren dargestellten Ausführungsbeispiele näher erläutert. Dabei zeigen:

FIGUR 1 als ein erstes Ausführungsbeispiel ein Ablaufdiagramm zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz des schmalbandig gefilterten Sprachsignals im Frequenzbereich,

FIGUR 2 als ein zweites Ausführungsbeispiel ein Ablaufdiagramm zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz des schmalbandig gefilterten Sprachsignals im Frequenzbereich,

FIGUR 3 als ein drittes Ausführungsbeispiel ein Ablaufdiagramm zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz des schmalbandig gefilterten Sprachsignals im Zeitbereich,

FIGUR 4a das Spektrum eines stimmhaften Lautes (Vokals),

FIGUR 4b das Spektrum eines stimmlosen Lautes (Frikativs),

FIGUR 5a eine mögliche Erweiterung des Spektrums eines Vokals,

FIGUR 5b eine mögliche Erweiterung des Spektrums eines Frikativs,

FIGUR 6a Verlauf eines ersten Sprachsignals,

FIGUR 6b Verlauf eines ersten sich aus dem Sprachsignal ergebenden Restsignals (first residual signal),

FIGUR 6c Kurzzeitspektralanalyse des Sprachsignals,

FIGUR 6d Kurzzeitspektralanalyse des Restsignals.

[0044] FIGUR 1 zeigt anhand eines Ablaufdiagramms einen ersten Prozess (eine erste Methode) zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz - z. B. 4 kHz - des schmalbandig gefilterten Sprachsignals im Frequenzbereich. Gemäß einem Ausgangszustand AZ des dargestellten Prozesses wird von dem Telekommunikationsgerät das Sprachsignal gesendet. Es liegt somit ein schmalbandig gefiltertes Sprachsignal vor.

[0045] In einem ersten Prozessschritt P0.1 wird dieses Sprachsignal in vorzugsweise gleich große schmalbandige Sprachsignalzeitabschnitte unterteilt. Anschließend werden für jeden Sprachsignalzeitabschnitt in einem zweiten Prozessschritt P1.1 die Spektralstruktur durch eine "Fast Fourier Transformation (FFT)" berechnet und in einem dritten Prozessschritt P2.1 eine Klassifizierung derart durchgeführt, dass der jeweilige Sprachsignalzeitabschnitt als ein stimmhafter Laut - wie beispielsweise "a", "e" oder "i", deren Aussprache ein in FIGUR 4a dargestelltes Spektrum aufweist - oder als ein stimmloser Laut - wie beispielsweise "s", "sch" oder "f", deren Aussprache ein in FIGUR 4b dargestelltes Spektrum aufweist - eingestuft bzw. definiert wird.

[0046] Diese Unterscheidung wird beispielsweise anhand der Position der ersten Formanten oder anhand des Verhältnisses von Spektralanteilen oberhalb und unterhalb einer bestimmten Frequenz - beispielsweise 2 kHz - geschehen. Eine Unterscheidung anhand des schmalbandigen Spektrums ist einfach durchzuführen, da wie ein Vergleich des in FIGUR 4a dargestellten Spektrum eines stimmhaften Lautes mit dem in FIGUR 4b dargestellten Spektrum eines stimmlosen Lautes zeigt, stimmhafte und stimmlose Laute in der Regel sehr unterschiedliche Spektren haben.

[0047] Alternativ dazu wird eine Kurzzeitsignalenergie eines ersten schmalbandig gefilterten Sprachsignalzeitabschnittes sowie eine Langzeitsignalenergie anhand weiterer aufeinanderfolgender zum ersten Signal korrelierender schmalbandig gefilterter Sprachsignalzeitabschnitte ermittelt und anschließend das Detektieren durch Vergleich eines Verhältnisses von Kurzzeitsignalenergie zu Langzeitsignalenergie mit einem Schwellwert realisiert.

[0048] Alternativ dazu kann die Unterscheidung durch Vergleich der Kurzzeitsignalenergie - d.h. der Signalenergie in einem kurzen Zeitausschnitt des

Schmalband-Sprachsignals - und der Langzeitsignalenergie - d.h. der Signalenergie über einen längeren Zeitausschnitt betrachtet - und anschließendem Vergleich des Verhältnis Kurzzeit- zu Langzeitenergie mit einem festen Schwellwert durchgeführt werden.

[0049] Im Anschluss daran wird in einem vierten Prozessschritt P3.1 in bezug auf die im dritten Prozessschritt P2.1 vorgenommene lautartbezogene Klassifizierung die im zweiten Prozessschritt P1.1 berechnete Spektralstruktur durch eine "Inverse Fast Fourier Transformation (IFFT)" erweitert. Dies geschieht derart, dass zeitabschnittsweise in bezug auf die im dritten Prozessschritt P2.1 vorgenommene lautartbezogene Klassifizierung Ergänzungen zur Erweiterung des Sprachsignals, die jeweils eine spektrale Struktur aufweisen, erzeugt werden, wobei beispielsweise (insbesondere) für den Fall des stimmhaften Lautes die Ergänzung unabhängig von dem jeweiligen Laut ist (mit Feststellung der Art des Sprachlautes - stimmhaft/stimmlos - wird auch die zur Erweiterung der Bandbreite notwendige Ergänzung bestimmt), die spektrale Struktur des schmalbandigen Sprachsignalzeitabschnittes und die spektrale Struktur der erzeugten Ergänzung zeitabschnittsweise zu einer erweiterten spektralen Struktur verknüpft werden und aus dieser erweiterten spektralen Struktur jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt wird.

[0050] Daran anschließend gibt es zwei Möglichkeiten, das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal zu erhalten.

[0051] Um eine gewisse Qualitätsverbesserung des breitbandigen erweiterten Sprachsignals zu erzielen, ist es möglich, den jeweiligen im vierten Prozessschritt P3.1 erzeugten breitbandigen erweiterten Sprachsignalzeitabschnitt in einem fünften Prozessschritt P4.1 mittels eines Hochpassfilters zu filtern, danach in einem sechsten Prozessschritt P5.1 diesen gefilterten Sprachsignalzeitabschnitt mit dem entsprechenden schmalbandigen Sprachsignalzeitabschnitt aus dem ersten Prozessschritt P0.1 zu verknüpfen, bevor abschließend in einem siebten Prozessschritt P6.1 aus den einzelnen verknüpften Sprachsignalzeitabschnitten durch Zusammenfügen dieser Zeitabschnitte das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal erzeugt wird.

[0052] Kann auf eine derartige Qualitätsverbesserung des breitbandigen erweiterten Sprachsignals verzichtet werden, so ist es stattdessen auch möglich, unmittelbar nach dem vierten Prozessschritt P3.1 aus den in diesem Prozessschritt jeweils erzeugten breitbandigen erweiterten Sprachsignalzeitabschnitten in dem siebten Prozessschritt P6.1 durch Zusammenfügen dieser Zeitabschnitte das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal zu erzeugen.

[0053] Anhand der FIGUR 2 soll zunächst die erfindungsgemäße Erweiterung eines schmalbandig gefilterten Sprachsignals in die Richtung der oberen Frequenzen gemäß eines zweiten Prozesses (einer zwei-

ten Methode) erläutert werden.

[0054] Im Allgemeinen wird ein Sprachsignal durch lineare Prädiktion analysiert. Dabei werden unter der Annahme, dass ein Sprachabtastwert durch die lineare Kombination von vorherigen Sprachabtastwerten angenähert werden kann, lineare Prädiktionskoeffizienten, sogenannte LPC-Koeffizienten, die die Filterkoeffizienten eines Sprachsynthesefilters darstellen, sowie ein Anregungssignal für dieses Synthesefilter berechnet. Durch Anwenden der zu einem Sprachsignalabschnitt gehörenden LPC-Koeffizienten auf diesen Sprachsignalabschnitt mittels Filterung des Abschnitts mit einem durch diese Koeffizienten definierten nichtrekursiven Digitalfilter entsteht das sogenannte Prädiktionsfehlersignal. Dieses Signal beschreibt die Differenz zwischen dem durch die lineare Prädiktion geschätztem Signalwert und dem tatsächlichem Signalwert. Es stellt auch gleichzeitig das Anregungssignal für das durch die LPC-Koeffizienten definierte rein rekursive Synthesefilter dar, mit dem der Original-Sprachsignalabschnitt durch Filtern des Prädiktionsfehler- bzw. Anregungssignals wiedergewonnen wird.

[0055] Um ein Sprachsignal in die Richtung der oberen Frequenzen zu erweitern, ist die Kenntnis eines breitbandigen Anregungssignals und der Filterkoeffizienten, die das (breitbandige) Sprachsignal im Sinne der linearen Prädiktion beschreiben erforderlich.

[0056] Da beispielsweise in Telekommunikationssystemen in denen schmalbandig übertragen wird, das Sprachsignal schmalbandig vorliegt, wird erfindungsgemäß anhand des mittels linearer Prädiktion aus dem Sprachsignal berechneten schmalbandigen Anregungssignal ein breitbandiges Anregungssignal ermittelt.

[0057] Dies erfolgt beispielweise durch Frequenzspiegelung des schmalbandigen Anregungssignals, bei dem die Frequenzanteile zwischen 0 kHz und 4 kHz an der 4 kHz - Spektrallinie in einen Bereich von 4 kHz bis 8 kHz gespiegelt werden.

[0058] Alternativ kann die Berechnung auch durch Addition des schmalbandigen Signals mit Gauß'schem (weißen) oder begrenzten (gefärbtem) Rauschen realisiert werden.

[0059] FIGUR 2 zeigt anhand eines Ablaufdiagramms den zweiten Prozess (die erste Methode) zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz - z. B. 4 kHz - des schmalbandig gefilterten Sprachsignals im Frequenzbereich. Gemäß dem Ausgangszustand AZ des dargestellten Prozesses wird wieder von dem Telekommunikationsgerät das Sprachsignal gesendet. Es liegt somit wider ein schmalbandig gefiltertes Sprachsignal vor.

[0060] In einem ersten Prozessschritt P0.2 wird dieses Sprachsignal in vorzugsweise gleich große schmalbandige Sprachsignalzeitabschnitte unterteilt. Anschließend werden für jeden Sprachsignalzeitabschnitt

in einem zweiten Prozessschritt P1.2 in bekannter Weise im Rahmen einer Prädiktionsanalyse LPC-Koeffizienten und ein schmalbandiges Prädiktionsfehlersignal berechnet, in einem dritten Prozessschritt P2.2 auf der Basis der LPC-Koeffizienten und des schmalbandigen Prädiktionsfehlersignals die Spektralstruktur der schmalbandigen Sprachsignalzeitabschnitte berechnet und in einem vierten Prozessschritt P3.2 eine Klassifizierung derart durchgeführt, dass der jeweilige Sprachsignalzeitabschnitt als ein stimmhafter Laut - wie beispielsweise "a", "e" oder "i", deren Aussprache ein in FIGUR 4a dargestelltes Spektrum aufweist - oder als ein stimmloser Laut - wie beispielsweise "s", "sch" oder "f", deren Aussprache ein in FIGUR 4b dargestelltes Spektrum aufweist - eingestuft bzw. definiert wird.

[0061] Diese Unterscheidung wird beispielsweise anhand der Position der ersten Formanten oder anhand des Verhältnisses von Spektralanteilen oberhalb und unterhalb einer bestimmten Frequenz - beispielsweise 2 kHz - geschehen. Eine Unterscheidung anhand des schmalbandigen Spektrums ist einfach durchzuführen, da wie ein Vergleich des in FIGUR 4a dargestellten Spektrum eines stimmhaften Lautes mit dem in FIGUR 4b dargestellten Spektrum eines stimmlosen Lautes zeigt, stimmhafte und stimmlose Laute in der Regel sehr unterschiedliche Spektren haben.

[0062] Alternativ dazu wird eine Kurzzeitsignalenergie eines ersten schmalbandig gefilterten Sprachsignalzeitabschnittes sowie eine Langzeitsignalenergie anhand weiterer aufeinanderfolgender zum ersten Signal korrelierender schmalbandig gefilterter Sprachsignalzeitabschnitte ermittelt und anschließend das Detektieren durch Vergleich eines Verhältnisses von Kurzzeitsignalenergie zu Langzeitsignalenergie mit einem Schwellwert realisiert.

[0063] Alternativ dazu kann die Unterscheidung durch Vergleich der Kurzzeitsignalenergie - d.h. der Signalenergie in einem kurzen Zeitausschnitt des Schmalband-Sprachsignals - und der Langzeitsignalenergie - d.h. der Signalenergie über einen längeren Zeitausschnitt betrachtet - und anschließendem Vergleich des Verhältnis Kurzzeit- zu Langzeitenergie mit einem festen Schwellwert durchgeführt werden.

[0064] Im Anschluss daran wird in einem fünften Prozessschritt P4.2 in bezug auf die im dritten Prozessschritt P2.1 vorgenommene lautartbezogene Klassifizierung die im dritten Prozessschritt P2.2 berechnete Spektralstruktur erweitert. Dies geschieht derart, dass zeitabschnittsweise in bezug auf die im vierten Prozessschritt P3.2 vorgenommene lautartbezogene Klassifizierung Ergänzungen zur Erweiterung des Sprachsignals, die jeweils eine spektrale Struktur aufweisen, erzeugt werden, wobei für den Fall des stimmhaften Lautes die Ergänzung unabhängig von dem jeweiligen Laut ist (mit Feststellung der Art des Sprachlautes - stimmhaft/stimmlos - wird auch die zur Erweiterung der Bandbreite notwendige Ergänzung bestimmt), die spektrale Struktur des schmalbandigen Sprachsignalzeitab-

schnittes und die spektrale Struktur der erzeugten Ergänzung zeitabschnittsweise zu einer erweiterten spektralen Struktur verknüpft werden.

[0065] Handelt es sich in dem fünften Prozessschritt P4.2 bei dem untersuchten schmalbandigen Sprachsignal um einen stimmhaften Laut, so wird die schmalbandige spektrale Struktur, wie in FIGUR 5a dargestellt, derart durch eine Ergänzung erweitert, dass die erweiterte breitbandige spektrale Struktur oberhalb von 4 kHz wesentlich weniger Energie als unterhalb von 4 kHz besitzt. Es ist z.B. ein Abfall, ein exponentieller Abfall, ein Anstieg, ein gleichbleibendes Nullniveau oder ein gleichbleibendes Niveau der spektralen Struktur zu höheren Frequenzen hin denkbar.

[0066] Alternativ kann auch ganz von einer Erweiterung abgesehen werden, weil in der Regel die Signalenergie eines stimmhaften Lautes oberhalb der Grenzfrequenz des Schmalband-Sprachsignals (z.B. 4 kHz) vernachlässigbar ist (vgl. FIGUR 4a). Der erzeugte breitbandige Frequenzgang entspricht für diesen Fall dem schmalbandigen Frequenzgang des zugrundeliegenden schmalbandigen Sprachsignals.

[0067] Es ist auch möglich, dass die Erweiterung, die nach Detektion eines stimmhaften Lautes vorgenommen wird, unabhängig von der genauen Kenntnis der Laute stets die gleiche ist (angepasst lediglich an die Energie des Schmalband-Sprachsignals), so dass eine einfache, kostengünstige und schnelle Umsetzung dieser Erweiterung erzielt wird.

[0068] Handelt es sich in dem fünften Prozessschritt P4.2 bei dem untersuchten schmalbandigen Sprachsignal um einen stimmlosen Laut, so wird der schmalbandige Frequenzgang, wie in FIGUR 5b dargestellt, derart erweitert, dass er - im Gegensatz zur Erweiterung bei stimmhaften Lauten - im Bereich oberhalb der Grenzfrequenz des Schmalband-Sprachsignals (z.B. 4 kHz) einen nicht vernachlässigbaren Teil seiner Gesamtenergie besitzt.

[0069] Auch hierbei kann die Erweiterung stets, unabhängig von der genauen Kenntnis der Laute, durch eine gleichartige spektrale Erweiterung erfolgen (angepasst lediglich an die Energie des Schmalband-Sprachsignals), so dass hierdurch ebenso eine einfache, kostengünstige und schnelle Umsetzung dieser Erweiterung erzielt wird.

[0070] Als Ergebnis der ersten bis fünften Prozessschritte P0.2...P4.2 in FIGUR 2 wird also eine neue erweiterte breitbandige spektrale Struktur in Abhängigkeit von dem Laut, der der vorhandenen schmalbandigen spektralen Struktur zugrundeliegt, generiert.

[0071] Als alternativen Ansatz zur Durchführung der Erweiterung im fünften Prozessschritt P4.2 kann man auch auf Codebücher zurückgreifen. Voraussetzung hierfür ist, dass mindestens ein Codebuch vorhanden ist, das den Zusammenhang, beispielsweise unter Zuhilfenahme der statistischen Eigenschaften der Sprache, die z.B. in einem Hidden Markov Model (HMM) abgespeichert werden können, zwischen schmalbandigen

und breitbandigen Filterkoeffizienten darstellt und aufgrund der statistischen Beziehung zu den im zweiten Prozessschritt P1.2 berechneten schmalbandigen Filterkoeffizienten, breitbandige Filterkoeffizienten liefert.

[0072] Bei einer alternativen Zuordnung von schmalbandigen zu breitbandigen Filterkoeffizienten, die durch ein oder mehrere Codebücher wiedergegeben wird, werden aus den im zweiten Prozessschritt P1.2 berechneten schmalbandigen Filterkoeffizienten zugehörige breitbandige Filterkoeffizienten ermittelt. Diese Filterkoeffizienten werden dann zur Synthese von Frequenzanteilen oberhalb der Grenzfrequenz des schmalbandigen Sprachsignals (z.B. 4 kHz) verwendet.

[0073] Die Codebücher werden jedoch nur für den Fall benötigt, dass die Untersuchung der im vierten Prozessschritt P3.2 ermittelten schmalbandigen Spektraleinhüllende einen stimmlosen Laut detektiert. Daher können sie auch auf Filterkoeffizienten für stimmlose Laute eingeschränkt und damit sehr klein sein, wodurch sie keine große Speichieranforderung an ein Telekommunikationsendgerät darstellen.

[0074] Außerdem wird in einem sechsten Prozessschritt P5.2 das in dem zweiten Prozessschritt P1.2 berechnete schmalbandige Prädiktionsfehlersignal zum einem breitbandigen Prädiktionsfehlersignal erweitert, so dass bezüglich der Zeitabschnittdauer den schmalbandigen Sprachsignalzeitabschnitten entsprechende Prädiktionsfehlersignalabschnitte des breitbandigen Prädiktionsfehlersignales erzeugt werden.

[0075] Daran anschließend wird aus der im fünften Prozessschritt P4.2 erzeugten erweiterten spektralen Struktur durch die Berechnung von breitbandigen Filterkoeffizienten in einem siebten Prozessschritt P6.2 und dem im sechsten Prozessschritt P5.2 jeweils erzeugten breitbandigen Prädiktionsfehlersignalabschnitt in einem achten Prozessschritt P7.2 mittels eines sogenannten Synthesefilters jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt.

[0076] Daran anschließend gibt es zwei Möglichkeiten, das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal zu erhalten.

[0077] Um eine gewisse Qualitätsverbesserung des breitbandigen erweiterten Sprachsignals zu erzielen, ist es möglich, den jeweiligen im achten Prozessschritt P7.2 erzeugten breitbandigen erweiterten Sprachsignalzeitabschnitt in einem neunten Prozessschritt P8.2 mittels eines Hochpassfilters zu filtern, danach in einem zehnten Prozessschritt P9.2 diesen gefilterten Sprachsignalzeitabschnitt mit dem entsprechenden schmalbandigen Sprachsignalzeitabschnitt aus dem ersten Prozessschritt P0.2 zu verknüpfen, bevor abschließend in einem elften Prozessschritt P10.2 aus den einzelnen verknüpften Sprachsignalzeitabschnitten durch Zusammenfügen dieser Zeitabschnitte das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal erzeugt wird.

[0078] Kann auf eine derartige Qualitätsverbesserung des breitbandigen erweiterten Sprachsignals ver-

zichtet werden, so ist es stattdessen auch möglich, unmittelbar nach dem achten Prozessschritt P7.2 aus den in diesem Prozessschritt jeweils erzeugten breitbandigen erweiterten Sprachsignalzeitabschnitten in dem elften Prozessschritt P10.2 durch Zusammenfügen dieser Zeitabschnitte das breitbandige in Richtung der oberen Frequenzen erweiterte Sprachsignal zu erzeugen.

[0079] Die breitbandigen Filterkoeffizienten beschreiben auf Grund dessen, dass sie aus der Abschätzung der breitbandigen spektralen Struktur berechnet wurden, die spektrale Struktur eines breitbandigen Sprachsignals.

[0080] Diese breitbandigen Filterkoeffizienten stehen dann für die Sprachsynthese zur Verfügung, mit der unter Verwendung des - wie bereits beschrieben - erzeugten breitbandigen Anregungssignals oder Prädiktions-signals die breitbandigen Sprachsignalzeitabschnitte und damit das breitbandige erweiterte Sprachsignal erzeugt wird, dessen Qualität deutlich besser ist als die des schmalbandig gefilterten Sprachsignals.

[0081] Die auf Basis der Codebücher berechneten und dem Synthesefilter zugeführten breitbandigen Filterkoeffizienten werden zur Synthese des oberen Frequenzbands des Sprachsignals verwendet, was zu einer Qualitätsverbesserung des Sprachsignals durch die Bandbreitenerweiterung führt.

[0082] Erfindungsgemäß können daher breitbandige Filterkoeffizienten ohne die Hilfe von Codebüchern bzw. mit sehr kleinen Codebüchern bestimmt werden, wobei eine mögliche Anwendung des erfindungsgemäßen Verfahrens zur Erweiterung der Sprachsignalbandbreite im oberen Frequenzbereich in Telekommunikationssystemen besteht, in denen Sprachcoder mit variabler Bitrate eingesetzt werden, die sowohl breitbandig als auch schmalbandig codieren können, da dort der Fall eintreten kann, dass der Sprachcoder während der Kommunikation zwischen Schmalband (narrow band) und Breitband (wide band) wechselt.

[0083] Die dadurch verursachte deutliche Verschlechterung in der Kommunikationsqualität wird dabei durch die Anwendung des in dieser Erfindung beschriebenen Verfahrens in Kommunikationsendgeräten verhindert.

[0084] In Telekommunikationssystemen, die beispielsweise gemäß dem UMTS-Standard funktionieren, und bei denen die oben beschriebene Problematik auftaucht, ist daher ein erfindungsgemäßes Schätzen der breitbandigen Sprachsignalanteile während der schmalbandigen Übertragung, um eine konstante Qualität zu gewährleisten, vorteilhaft einsetzbar.

[0085] FIGUR 3 zeigt anhand eines Ablaufdiagramms einen dritten Prozess (eine dritte Methode) zur Erweiterung der Bandbreite eines von einem Telekommunikationsgerät gesendeten Sprachsignals in Richtung der oberen Frequenzen oberhalb einer Grenzfrequenz - z. B. 4 kHz - des schmalbandig gefilterten Sprachsignals im Zeitbereich. Gemäß dem Ausgangszustand AZ des dargestellten Prozesses wird wieder von dem Telekom-

munikationsgerät das Sprachsignal gesendet. Es liegt somit wieder ein schmalbandig gefiltertes Sprachsignal vor.

[0086] In einem ersten Prozessschritt P0.3 wird dieses Sprachsignal in vorzugsweise gleich große schmalbandige Sprachsignalzeitabschnitte unterteilt. Anschließend werden für jeden Sprachsignalzeitabschnitt in einem zweiten Prozessschritt P1.3 eine Klassifizierung derart durchgeführt, dass der jeweilige Sprachsignalzeitabschnitt als ein stimmhafter Laut - wie beispielsweise "a", "e" oder "i", deren Aussprache ein in FIGUR 6a dargestelltes Spektrum aufweist - oder als ein stimmloser Laut - wie beispielsweise "s", "sch" oder "f", deren Aussprache ein in FIGUR 4b dargestelltes Spektrum aufweist - eingestuft bzw. definiert wird.

[0087] Diese Unterscheidung wird beispielsweise anhand der Position der ersten Formanten oder anhand des Verhältnisses von Spektralanteilen oberhalb und unterhalb einer bestimmten Frequenz - beispielsweise 2 kHz - geschehen. Eine Unterscheidung anhand des schmalbandigen Spektrums ist einfach durchzuführen, da wie ein Vergleich des in FIGUR 4a dargestellten Spektrum eines stimmhaften Lautes mit dem in FIGUR 4b dargestellten Spektrum eines stimmlosen Lautes zeigt, stimmhafte und stimmlose Laute in der Regel sehr unterschiedliche Spektren haben.

[0088] Alternativ dazu wird eine Kurzzeitsignalenergie eines ersten schmalbandig gefilterten Sprachsignalzeitabschnittes sowie eine Langzeitsignalenergie anhand weiterer aufeinanderfolgender zum ersten Signal korrelierender schmalbandig gefilterter Sprachsignalzeitabschnitte ermittelt und anschließend das Detektieren durch Vergleich eines Verhältnisses von Kurzzeitsignalenergie zu Langzeitsignalenergie mit einem Schwellwert realisiert.

[0089] Alternativ dazu kann die Unterscheidung durch Vergleich der Kurzzeitsignalenergie - d.h. der Signalenergie in einem kurzen Zeitausschnitt des Schmalband-Sprachsignals - und der Langzeitsignalenergie - d.h. der Signalenergie über einen längeren Zeitausschnitt betrachtet - und anschließendem Vergleich des Verhältnis Kurzzeit- zu Langzeitenergie mit einem festen Schwellwert durchgeführt werden.

[0090] Außerdem werden in einem dritten Prozessschritt P2.3 die schmalbandigen Sprachsignalzeitabschnitte derart nichtlinear, vorzugsweise durch spektrale Spiegelung, verarbeitet, dass jeweils ein modifizierter Sprachsignalzeitabschnitt erzeugt wird, der einerseits den jeweiligen im wesentlichen unveränderten schmalbandige Sprachsignalzeitabschnitt und andererseits oberhalb der Grenzfrequenz durch die nichtlineare Signalverarbeitung erzeugte Signalanteile enthält.

[0091] Im Anschluss daran werden in einem vierten Prozessschritt P3.3 die modifizierten Sprachsignalzeitabschnitte in bezug auf die vorgenommene lautartbezogene Klassifizierung derart unterschiedlich gefiltert werden, dass aus den modifizierten Sprachsignalzeitabschnitten breitbandige erweiterte Sprachsignalzeit-

abschnitte und damit ein breitbandiges erweitertes Sprachsignal entsteht, wobei im Fall eines stimmhaften Sprachsignalzeitabschnittes wenig Energie oberhalb von der Grenzfrequenz - z.B. 4 kHz - und im Fall eines stimmlosen Sprachsignalzeitabschnittes mehr Energie oberhalb von der Grenzfrequenz - z.B. 4 kHz - durchgelassen wird.

[0092] Darüber hinaus ist eine Kombination des erfindungsgemäßen Verfahrens zur Erweiterung schmalbandiger Sprachsignale im oberen Frequenzbereich mit einem Verfahren zur Erweiterung schmalbandiger Sprachsignale im unteren Frequenzbereich, die man als "Wideband Speech Extender" bezeichnen kann, besonders vorteilhaft, da sie die Synthese eines breitbandigen Sprachsignals gewährleistet, das dem zugrundeliegenden Sprachsignal am nächsten kommt, so dass ein Nutzer eines Telekommunikationsendgerätes, welches den "Wideband Speech Extender" einsetzt, ein Sprachsignal hoher Qualität, vergleichbar mit der Qualität bei Sprachsignalen in Radio- und Fernsehgeräten, hört.

[0093] Damit kann der "Wideband Speech Extender" in Telekommunikationsgeräten, wo eine bandbegrenzte Übertragung von Sprachsignalen stattfindet, eingesetzt werden, um beim Benutzer den Eindruck einer breitbandigen Übertragung zu erzeugen.

[0094] Neben dem erfindungsgemäßen Verfahren zur Erweiterung eines schmalbandigen Sprachsignals im oberen Frequenzbereich kann auch der "Wideband Speech Extender" in Telekommunikationssystemen eingesetzt werden, wo das "WB/NB-Switching"-Problem auftritt, so dass stets ein breitbandiges Sprachsignal und damit eine weitgehend konstante Qualität gewährleistet ist.

[0095] Die Erweiterung eines bandbegrenzten Sprachsignals in die Richtung der unteren Frequenzen bzw. die Wiederherstellung der unteren Frequenzanteile soll ausgehend von der FIGUR 7 anhand der FIGUREN 6a bis 6d angedeutet werden.

[0096] Wie eingangs diskutiert ist aus der EP 0 994 464 bereits eine spektrale Wiederherstellung von Signalanteilen des unteren Frequenzbereichs eines durch eine Hochpassfunktion zu tiefen Frequenzen hin begrenzten Sprachsignals bekannt, wobei die Wiederherstellung durch Generieren von Frequenzen des unteren Frequenzbereichs durch eine nichtlineare Signalverarbeitung erfolgt, wobei dazu subharmonische Frequenzen des Signals erzeugt und zum Hochpasssignal hinzu addiert werden.

[0097] Bei bestehenden, insbesondere dem aus der EP 0 994 464 bekannten, Verfahren zur Erweiterung der unteren Frequenzen ist es erforderlich, die Filtercharakteristik, mit der ein Signal an einem fernen Telekommunikationsendgerät gefiltert wurde, zu kennen. Im Allgemeinen sind derartige Verfahren nur unter Verwendung von Telekommunikationseinrichtungen mit gleicher Charakteristik, d.h. Telekommunikationsendgeräte gleichen Typs, optimal einsetzbar, da deren Filtercharakteristik gleich bzw. angepasst ist.

[0098] In heterogenen Systemen, wo eine Vielzahl unterschiedlicher Telekommunikationsgeräte sowie unterschiedliche Typen von Telekommunikationsgeräten Verwendung finden, sind diese Verfahren nicht einsetzbar, da unterschiedliche Typen von Telekommunikationsgeräten, z.B. Siemens-Telekommunikationsgeräten wie in FIGUR 7 gezeigt, unterschiedliche Filtercharakteristiken aufweisen.

[0099] Die Erweiterung bandbegrenzter Sprachsignale im unteren Frequenzbereich kann für solche Systeme beispielsweise durch das Schätzen von Filtercharakteristiken erreicht werden, wobei für die Schätzung zunächst von einem, wie in der FIGUR 6a dargestellten, Sprachsignal ein erstes, wie in der FIGUR 6b dargestelltes, Restsignal (first residual signal), auch Prädiktionsfehlersignal genannt, durch die aus der Literatur bekannte Lineare Prädiktionsmethode berechnet wird, wobei die Berechnung des ersten Restsignals entfallen kann, wenn es schon durch andere Verarbeitungsschritte bekannt ist.

[0100] Da, wie aus der Fachliteratur (Vary, Heute, Hess: "Digitale Sprachsignalverarbeitung", Teubner Stuttgart 1998) bekannt ist, die spektrale Form des ersten Restsignals, insbesondere im Vergleich mit dem in FIGUR 6c dargestellten Spektrum des Sprachsignals, wie in FIGUR 6d entnehmbar, im übertragenen Frequenzbereich nahezu eben ist und lediglich an den Flanken des Filters, der das Sprachsignal im entfernten Kommunikationsendgerät bandbegrenzt hat, abfällt, wird mit dieser Kenntnis und dem berechneten Restsignal eine Schätzung der Filtercharakteristik durchgeführt, wobei insbesondere eine Messung der Restsignalenergie in unterschiedlichen Frequenzbändern Informationen über die Filtercharakteristik liefert.

Patentansprüche

1. Verfahren zur Erweiterung der Bandbreite eines schmalbandig gefilterten Sprachsignals, insbesondere eines von einem Telekommunikationsgerät gesendeten Sprachsignals, oberhalb einer Grenzfrequenz des schmalbandigen Sprachsignals, bei dem

a) das schmalbandige Sprachsignal in Sprachsignalzeitabschnitte unterteilt wird (P0.1) und jeweils eine spektrale Struktur der Sprachsignalzeitabschnitte berechnet wird (P1.1),

b) jeder schmalbandige Sprachsignalzeitabschnitt als ein stimmhafter Laut oder als ein stimmloser Laut klassifiziert wird (P2.1),

dadurch gekennzeichnet, dass

c) eine spektrale Struktur aufweisende Ergänzungen zur Erweiterung des schmalbandigen Sprachsignals in bezug auf die in b) vorgenommene lautartbezogene Klassifizierung erzeugt werden (P3.1), wobei zumindest für den Fall

des stimmhaften Lautes die Ergänzung unabhängig von dem jeweiligen Laut ist,

d) die spektrale Struktur des schmalbandigen Sprachsignalzeitabschnittes und die spektrale Struktur der erzeugten Ergänzung zeitabschnittsweise derart verknüpft werden (P3.1), dass jeweils eine erweiterte spektrale Struktur entsteht, 5
e) aus der erweiterten spektralen Struktur jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt wird (P3.1), 10
f) aus den einzelnen breitbandigen erweiterten Sprachsignalzeitabschnitten ein breitbandiges erweitertes Sprachsignal erzeugt wird (P6.1).

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass

die spektrale Struktur des schmalbandigen Sprachsignalzeitabschnittes durch eine FFT-Analyse berechnet wird und aus der erweiterten spektralen Struktur durch eine IFFT-Analyse der breitbandige erweiterte Sprachsignalzeitabschnitt erzeugt wird. 20

3. Verfahren zur Erweiterung der Bandbreite eines schmalbandig gefilterten Sprachsignals, insbesondere eines von einem Telekommunikationsgerät gesendeten Sprachsignals, oberhalb einer Grenzfrequenz des schmalbandigen Sprachsignals, bei dem 25

a) das schmalbandige Sprachsignal in Sprachsignalzeitabschnitte unterteilt wird (P0.2) und jeweils eine spektrale Struktur der schmalbandigen Sprachsignalzeitabschnitte sowie Prädiktionsfehlersignalzeitabschnitte eines schmalbandigen Prädiktionsfehlersignals berechnet werden (P1.2, P2.2), 30 35

b) jeder schmalbandige Sprachsignalzeitabschnitt als ein stimmhafter Laut oder als ein stimmloser Laut klassifiziert wird (P3.2), 40

dadurch gekennzeichnet, dass

c) eine spektrale Struktur aufweisende Ergänzungen zur Erweiterung des schmalbandigen Sprachsignals in Bezug auf die in b) vorgenommene lautartbezogene Klassifizierung erzeugt werden (P4.2), wobei zumindest für den Fall des stimmhaften Lautes die Ergänzung unabhängig von dem jeweiligen Laut ist, 45
d) die spektrale Struktur des schmalbandigen Sprachsignalzeitabschnittes und die spektrale Struktur der erzeugten Ergänzung zeitabschnittsweise derart verknüpft werden (P4.2), dass jeweils eine erweiterte spektrale Struktur entsteht, 50

e) aus den schmalbandigen Prädiktionsfehlersignalzeitabschnitten bezüglich der Zeitabschnittsdauer den schmalbandigen Sprachsignalzeitabschnitten entsprechende Prädiktions-

onsfehlersignalzeitabschnitte eines breitbandigen Prädiktionsfehlersignals erzeugt werden (P5.2) und aus der erweiterten spektralen Struktur und dem jeweiligen breitbandigen Prädiktionsfehlersignalzeitabschnitt jeweils ein breitbandiger erweiterter Sprachsignalzeitabschnitt erzeugt wird (P6.2, P7.2),

f) aus den einzelnen breitbandigen erweiterten Sprachsignalzeitabschnitten ein breitbandiges erweitertes Sprachsignal erzeugt wird (P10.2).

4. Verfahren nach Anspruch 1 oder 3, dadurch gekennzeichnet, dass

die für die als stimmhafte Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils erzeugte Ergänzung derart erzeugt wird (P4.2), dass die Energie dieser Ergänzung in Bezug auf die Gesamtenergie des schmalbandigen Sprachsignalabschnittes vernachlässigbar ist.

5. Verfahren nach einem der Ansprüche 1, 3 oder 4, dadurch gekennzeichnet, dass

die für die als stimmlose Laute klassifizierten schmalbandigen Sprachsignalabschnitte jeweils erzeugte Ergänzung derart erzeugt wird (P4.2), dass die Energie dieser Ergänzung in Bezug auf die Gesamtenergie des schmalbandigen Sprachsignalabschnittes nicht vernachlässigbar ist.

6. Verfahren nach einem der Ansprüche 1, 3 oder 4, dadurch gekennzeichnet, dass

die für die als stimmlose Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils erzeugte Ergänzung derart erzeugt wird (P4.2), dass auf Basis von zumindest einem Breitband-Codebuch aus ersten Filterkoeffizienten des schmalbandigen Sprachsignalzeitabschnittes zweite Filterkoeffizienten eines breitbandigen Sprachsignalzeitabschnittes ermittelt werden.

7. Verfahren nach einem der Ansprüche 3 bis 6, dadurch gekennzeichnet, dass

aus der erweiterten spektralen Struktur jeweils dritte Filterkoeffizienten berechnet werden (P6.2).

8. Verfahren nach Anspruch 6 oder 7, dadurch gekennzeichnet, dass

mit den zweiten oder dritten Filterkoeffizienten und dem breitbandigen Prädiktionsfehlersignalzeitabschnitt breitbandige erweiterte Sprachsignalzeitabschnitte und damit das breitbandige erweiterte Sprachsignal synthetisiert werden (P7.2).

9. Verfahren nach Anspruch 7, dadurch gekennzeichnet, dass

a) die dritten Filterkoeffizienten mit den Einträgen aus einem Breitband-Codebuch verglichen

werden und

b) der Eintrag in dem Breitband-Codebuch, der am besten zu den dritten Filterkoeffizienten passt, als Filterkoeffizient der Synthese des breitbandigen erweiterten Sprachsignals zugrundegelegt wird.

10. Verfahren nach Anspruch 1, 3, 4 oder 5, dadurch gekennzeichnet, dass

die erzeugte Ergänzung abfällt, exponentiell abfällt, ansteigt, gleichbleibendes Nullniveau aufweist oder gleichbleibendes Niveau aufweist.

11. Verfahren nach Anspruch 1 oder 3, dadurch gekennzeichnet, dass

der aus der erweiterten spektralen Struktur jeweils erzeugte breitbandige erweiterte Sprachsignalzeitabschnitt hochpassgefiltert wird (P4.1, P8.2), der hochpassgefilterte Sprachsignalzeitabschnitt mit dem entsprechenden schmalbandigen Sprachsignalzeitabschnitt verknüpft wird (P5.1, P9.2) und aus den einzelnen verknüpften Sprachsignalzeitabschnitten das breitbandige erweiterte Sprachsignal erzeugt wird (P6.1, P10.2).

12. Verfahren zur Erweiterung der Bandbreite eines schmalbandig gefilterten Sprachsignals, insbesondere eines von einem Telekommunikationsgerät gesendeten Sprachsignals, oberhalb einer Grenzfrequenz des schmalbandigen Sprachsignals, bei dem

a) das schmalbandige Sprachsignal in Sprachsignalzeitabschnitte unterteilt wird (P0.3),

b) jeder schmalbandige Sprachsignalzeitabschnitt als ein stimmhafter Laut oder als ein stimmloser Laut klassifiziert wird (P1.3),

dadurch gekennzeichnet, dass

c) die schmalbandigen Sprachsignalzeitabschnitte derart nichtlinear verarbeitet werden (P2.3), dass jeweils ein modifizierter Sprachsignalzeitabschnitt erzeugt wird, der einerseits den jeweiligen im wesentlichen unveränderten schmalbandigen Sprachsignalzeitabschnitt und andererseits oberhalb der Grenzfrequenz durch die nichtlineare Signalverarbeitung erzeugte Signalanteile enthält,

d) die modifizierten Sprachsignalzeitabschnitte in bezug auf die in b) vorgenommene lautartbezogene Klassifizierung derart unterschiedlich gefiltert werden (P3.3), dass im Fall eines stimmhaften Lautes wenig Energie oberhalb von der Grenzfrequenz und im Fall eines stimmlosen Lautes viel Energie oberhalb von der Grenzfrequenz durchgelassen wird und dass aus den modifizierten Sprachsignalzeitabschnitten breitbandige erweiterte Sprachsignalzeitabschnitte und damit ein breitbandiges

erweitertes Sprachsignal entsteht.

13. Verfahren nach Anspruch 12, dadurch gekennzeichnet, dass

die für die als stimmhafte Laute klassifizierten schmalbandigen Sprachsignalzeitabschnitte jeweils durch die nichtlineare Signalverarbeitung erzeugten Signalanteile derart erzeugt werden (P2.3), dass die Energie des jeweiligen Signalanteils in bezug auf die Gesamtenergie des schmalbandigen Sprachsignalzeitabschnittes vernachlässigbar ist.

14. Verfahren nach einem der Ansprüche 12 oder 13, dadurch gekennzeichnet, dass

die für die als stimmlose Laute klassifizierten schmalbandigen Sprachsignalabschnitte jeweils durch die nichtlineare Signalverarbeitung erzeugten Signalanteile derart erzeugt werden (P2.3), dass die Energie des jeweiligen Signalanteils in bezug auf die Gesamtenergie des schmalbandigen Sprachsignalzeitabschnittes nicht vernachlässigbar ist.

15. Verfahren nach einem der Ansprüche 12 bis 14, dadurch gekennzeichnet, dass

die Signalanteile durch spektrale Spiegelung erzeugt werden.

16. Verfahren nach einem der Ansprüche 1 bis 15, dadurch gekennzeichnet, dass

die schmalbandigen Sprachsignalzeitabschnitte gleich lang gewählt werden.

Claims

1. Method for expanding the bandwidth of a narrowband filtered speech signal, in particular a speech signal transmitted by a telecommunications device, above a cut-off frequency of the narrowband speech signal, wherein

a) the narrowband speech signal is subdivided into speech signal time segments (P0.1) and a spectral structure of the speech signal time segment is computed in each case (P1.1),

b) each narrowband speech signal time segment is classified as a voiced sound or as an unvoiced sound (P2.1),

characterized in that

c) enhancements having a spectral structure for expanding the narrowband speech signal in relation to the sound-related classification (P3.1) performed in b), wherein at least for the case of the voiced sound the enhancement is independent of the respective sound,

- d) the spectral structure of the narrowband speech signal time segment and the spectral structure of the generated enhancement are combined (P3.1) in time segment sequence such that an expanded spectral structure is produced in each case, 5
- e) a wideband expanded speech signal time segment is generated in each case from the expanded spectral structure (P3.1), 10
- f) a wideband expanded speech signal time segment is generated from the individual wideband expanded speech signal time segments (P6.1). 15
2. Method according to Claim 1, **characterized in that** the spectral structure of the narrowband speech signal time segment is computed by means of an FFT analysis and the wideband expanded speech signal time segment is generated from the expanded spectral structure by means of an IFFT analysis. 20
3. Method for expanding the bandwidth of a narrowband filtered speech signal, in particular a speech signal transmitted by a telecommunications device, above a cut-off frequency of the narrowband speech signal, 25
- wherein
- a) the narrowband speech signal is subdivided into speech signal time segments (P0.2) and a spectral structure of the narrowband speech signal time segments as well as a prediction error signal time segment of a narrowband prediction signal error signal are computed in each case (P1.2, P2.2), 30
- b) each narrowband speech signal time segment is classified as a voiced sound or as an unvoiced sound (P3.2), **characterized in that** 35
- c) enhancements having a spectral structure for expanding the narrowband speech signal in relation to the sound-related classification (P4.2) performed in b), wherein at least for the case of the voiced sound the enhancement is independent of the respective sound, 40
- d) the spectral structure of the narrowband speech signal time segments and the spectral structure of the generated enhancement are combined (P4.2) in time segment sequence such that an expanded spectral structure is produced in each case, 45
- e) from the narrowband prediction error signal time segments with regard to the time segment duration, prediction error signal time segments of a wideband prediction error signal corresponding to the narrowband speech signal time segments are generated (P5.2) and a wideband expanded speech signal time segment is generated in each case from the expanded 50
- spectral structure and the respective wideband prediction error signal time segment (P6.2, P7.2),
- f) a wideband expanded speech signal is generated from the individual wideband expanded speech signal time segments (P10.2).
4. Method according to Claim 1 or 3, **characterized in that** the enhancement generated in each case for the narrowband speech signal time segments classified as voiced sounds is generated in such a way (P4.2) that the energy of this enhancement is negligible in relation to the total energy of the narrowband speech signal segment.
5. Method according to one of the Claims 1, 3, or 4, **characterized in that** the enhancement generated in each case for the narrowband speech signal time segments classified as unvoiced sounds is generated in such a way (P4.2) that the energy of this enhancement is not negligible in relation to the total energy of the narrowband speech signal segment.
6. Method according to one of the Claims 1, 3 or 4, **characterized in that** the enhancement generated in each case for the narrowband speech signal time segments classified as unvoiced sounds is generated in such a way (P4.2) that second filter coefficients of a wideband speech signal time segment are determined from first filter coefficients of the narrow-band speech signal time segment on the basis of at least one wideband codebook.
7. Method according to one of Claims 3 to 6, **characterized in that** third filter coefficients are computed in each case from the expanded spectral structure (P6.2).
8. Method according to Claim 6 or 7, **characterized in that** wideband expanded speech signal time segments and hence the wideband expanded speech signal are synthesized by means of the second or third filter coefficients and the wideband prediction error signal time segment (P7.2).
9. Method according to Claim 7, **characterized in that**
- a) the third filter coefficients are compared with the entries from a wideband codebook and
- b) the entry in the wideband codebook which best matches the third filter coefficients is taken as the basis for the filter coefficient of the synthesis of the wideband expanded speech signal.
10. Method according to Claim 1, 3, 4 or 5, **characterized in that** the generated enhancement drops, drops exponentially, rises, maintains a constant ze-

ro level or maintains a constant level.

11. Method according to Claim 1 or 3, **characterized in that** the wideband expanded speech signal time segment generated in each case from the expanded spectral structure is high-pass filtered (P4.1, P8.2), the high-pass filtered speech signal time segment is combined with the corresponding narrowband speech signal time segment (P5.1, P9.2) and the wideband expanded speech signal is generated from the individual combined speech signal time segments (P6.1, P10.2).

12. Method for expanding the bandwidth of a narrowband filtered speech signal, in particular a speech signal transmitted by a telecommunications device, above a cut-off frequency of the narrowband speech signal, wherein

- a) the narrowband speech signal is subdivided into speech signal time segments (p0.3),
- b) each narrowband speech signal time segment is classified as a voiced sound or as an unvoiced sound (P1.3) **characterized in that**
- c) the narrowband speech signal time segments are processed non-linearly (P2.3) in such a way that in each case a modified speech signal time segment is generated which on the one hand contains the respective essentially unmodified narrow-band speech signal time segment and on the other hand contains signal components generated by the non-linear signal processing above the cut-off frequency,
- d) the modified speech signal time segments are filtered differently (P3.3), in relation to the sound-related classification performed in b) in such a way that in the case of voiced sounds, less energy is allowed to pass above the cut-off frequency, and in the case of an unvoiced sound, a lot of energy is allowed to pass above the cut-off frequency, and in such a way that wideband expanded speech signal time segments and hence a wideband expanded speech signal are produced from the modified speech signal time segments.

13. Method according to Claim 12, **characterized in that** the signal components generated in each case by the non-linear signal processing for the narrowband speech signal time segments classified as voiced sounds are generated in such a way (P2.3) that the energy of the respective signal component is negligible in relation to the total energy of the narrowband speech signal time segment.

14. Method according to one of the Claims 12 or 13, **characterized in that** the signal components generated in each case by the non-linear signal

processing for the narrowband speech signal time segments classified as unvoiced sounds are generated in such a way (P2.3) that the energy of the respective signal component is not negligible in relation to the total energy of the narrowband speech signal time segment.

15. Method according to one of the Claims 12 to 14, **characterized in that** the signal components are generated by spectral mirroring.

16. Method according to one of the Claims 1 to 15, **characterized in that** the narrowband speech signal time segments are chosen to be of equal length.

Revendications

1. Procédé d'élargissement de la largeur de bande d'un signal vocal filtré en bande étroite, en particulier d'un signal vocal émis par un appareil de télécommunication, au delà d'une fréquence de coupure du signal vocal à bande étroite, dans lequel

- a) le signal vocal à bande étroite est subdivisé en segments temporels de signal vocal (P0.1) et une structure spectrale des segments temporels de signal vocal est respectivement calculée (P1.1),
- b) chaque segment temporel de signal vocal à bande étroite est classé comme son sonore ou comme son non sonore (P2.1),

caractérisé en ce que

- c) des compléments pour élargir le signal vocal à bande étroite par rapport à la classification opérée selon le type de son en b), lesquels présentent une structure spectrale, sont produits (P3.1), le complément étant indépendant du son respectif du moins dans le cas du son sonore,
- d) il est procédé à l'association, par segments temporels, de la structure spectrale du segment temporel de signal vocal à bande étroite et de la structure spectrale du complément produit (P3.1) de manière telle qu'une structure spectrale élargie se crée respectivement,
- e) un segment temporel de signal vocal élargi à large bande est respectivement produit au départ de la structure spectrale élargie (P3.1),
- f) un signal vocal élargi à large bande est produit au départ des différents segments temporels de signal vocal élargi à large bande (P6.1).

2. Procédé selon la revendication 1, **caractérisé en ce que** la structure spectrale du segment temporel de signal vocal à bande étroite est calculée par une analyse FFT et le segment temporel de signal vocal

élargi à large bande est produit par une analyse IFFT au départ de la structure spectrale élargie.

3. Procédé d'élargissement de la largeur de bande d'un signal vocal filtré en bande étroite, en particulier d'un signal vocal émis par un appareil de télécommunication, au delà d'une fréquence de coupure du signal vocal à bande étroite, dans lequel

a) le signal vocal à bande étroite est subdivisé en segments temporels de signal vocal (P0.2) et respectivement une structure spectrale des segments temporels de signal vocal à bande étroite ainsi que des segments temporels de signal d'erreur de prédiction d'un signal d'erreur de prédiction à bande étroite sont calculés (P1.2, P2.2),

b) chaque segment temporel de signal vocal à bande étroite est classé comme son sonore ou comme son non sonore (P3.2),

caractérisé en ce que

c) des compléments pour élargir le signal vocal à bande étroite par rapport à la classification opérée selon le type de son en b), lesquels présentent une structure spectrale, sont produits (P4.2), le complément étant indépendant du son respectif du moins dans le cas du son sonore,

d) il est procédé à l'association, par segments temporels, de la structure spectrale du segment temporel de signal vocal à bande étroite et de la structure spectrale du complément produit, de manière telle (P4.2) qu'une structure spectrale élargie se crée respectivement,

e) des segments temporels de signal d'erreur de prédiction d'un signal d'erreur de prédiction à large bande, lesquels correspondent aux segments temporels de signal vocal à bande étroite pour ce qui est de la durée du segment temporel, sont produits au départ des segments temporels de signal d'erreur de prédiction à bande étroite (P5.2) et un segment temporel de signal vocal élargi à large bande est respectivement produit au départ de la structure spectrale élargie et du segment temporel de signal d'erreur de prédiction à large bande respectif (P6.2, P7.2),

f) un signal vocal élargi à large bande est produit au départ des différents segments temporels de signal vocal élargi à large bande (P10.2).

4. Procédé selon la revendication 1 ou 3, **caractérisé en ce que** le complément respectivement produit pour les segments temporels de signal vocal à bande étroite qui sont classés comme sons sonores est produit de manière telle (P4.2) que l'énergie de ce complément est négligeable par rapport à l'énergie

totale du segment de signal vocal à bande étroite.

5. Procédé selon l'une des revendications 1, 3 ou 4, **caractérisé en ce que** le complément respectivement produit pour les segments de signal vocal à bande étroite qui sont classés comme sons non sonores est produit de manière telle (P4.2) que l'énergie de ce complément n'est pas négligeable par rapport à l'énergie totale du segment de signal vocal à bande étroite.

6. Procédé selon l'une des revendications 1, 3 ou 4, **caractérisé en ce que** le complément respectivement produit pour les segments temporels de signal vocal à bande étroite qui sont classés comme sons non sonores est produit de manière telle (P4.2) que des deuxièmes coefficients de filtrage d'un segment temporel de signal vocal à large bande sont déterminés sur la base d'au moins une table de codes large bande au départ de premiers coefficients de filtrage du segment temporel de signal vocal à bande étroite.

7. Procédé selon l'une des revendication 3 à 6, **caractérisé en ce que** des troisièmes coefficients de filtrage sont respectivement calculés au départ de la structure spectrale élargie (P6.2).

8. Procédé selon la revendication 6 ou 7, **caractérisé en ce que** des segments temporels de signal vocal élargi à large bande et, par conséquent, le signal vocal élargi à large bande sont synthétisés avec les deuxièmes ou troisièmes coefficients de filtrage et le segment temporel de signal d'erreur de prédiction à large bande (P7.2).

9. Procédé selon la revendication 7, **caractérisé en ce que**

a) les troisièmes coefficients de filtrage sont comparés avec les entrées tirées d'une table de codes large bande et

b) l'entrée de la table de codes large bande qui convient le mieux pour les troisièmes coefficients de filtrage est prise comme base en tant que coefficient de filtrage pour la synthèse du signal vocal élargi à large bande.

10. Procédé selon la revendication 1, 3, 4 ou 5, **caractérisé en ce que** le complément produit chute, chute exponentiellement, augmente, présente un niveau zéro constant ou un niveau constant.

11. Procédé selon la revendication 1 ou 3, **caractérisé en ce que** le segment temporel de signal vocal élargi à large bande respectivement produit au départ de la structure spectrale élargie est filtré en passe-haut (P4.1, P8.2), il est procédé à l'association du

segment temporel de signal vocal filtré en passe-haut et du segment temporel de signal vocal à bande étroite correspondant (P5.1, P9.2) et le signal vocal élargi à large bande est produit au départ des différents segments temporels de signal vocal associés (P6.1, P10.2).

12. Procédé d'élargissement de la largeur de bande d'un signal vocal filtré en bande étroite, en particulier d'un signal vocal émis par un appareil de télécommunication, au delà d'une fréquence de coupure du signal vocal à bande étroite, dans lequel

a) le signal vocal à bande étroite est subdivisé en segments temporels de signal vocal (P0.3),
b) chaque segment temporel de signal vocal à bande étroite est classé comme son sonore ou comme son non sonore (P1.3),

caractérisé en ce que

c) les segments temporels de signal vocal à bande étroite sont traités non linéairement de manière telle (P2.3) qu'un segment temporel de signal vocal modifié est respectivement produit, lequel contient, d'une part, le segment temporel de signal vocal à bande étroite respectif, non modifié, pour l'essentiel, et, d'autre part, des parties de signal produites par le traitement non linéaire du signal au delà de la fréquence de coupure,

d) les segments temporels de signal vocal modifiés sont filtrés de manière tellement différente, par rapport à la classification opérée selon le type de son en b) (P3.3), que, dans le cas d'un son sonore, peu d'énergie peut passer au delà de la fréquence de coupure et, dans le cas d'un son non sonore, beaucoup d'énergie peut passer au delà de la fréquence de coupure, et que des segments temporels de signal vocal élargi à large bande et, par conséquent, un signal vocal élargi à large bande se créent au départ des segments temporels de signal vocal modifiés.

13. Procédé selon la revendication 12, **caractérisé en ce que** les parties de signal respectivement produites par le traitement non linéaire du signal pour les segments temporels de signal vocal à bande étroite qui sont classés comme sons sonores sont produites de manière telle (P2.3) que l'énergie de la partie de signal respective est négligeable par rapport à l'énergie totale du segment temporel de signal vocal à bande étroite.

14. Procédé selon l'une des revendications 12 ou 13, **caractérisé en ce que** les parties de signal respectivement produites par le traitement non linéaire du signal pour les segments de signal vocal à bande étroite qui sont classés comme sons non sonores

sont produites de manière telle (P2.3) que l'énergie de la partie de signal respective n'est pas négligeable par rapport à l'énergie totale du segment temporel de signal vocal à bande étroite.

15. Procédé selon l'une des revendications 12 à 14, **caractérisé en ce que** les parties de signal sont produites par réflexion spectrale.

16. Procédé selon l'une des revendications 1 à 15, **caractérisé en ce que** les segments temporels de signal vocal à bande étroite sont choisis de même longueur.

FIG 1

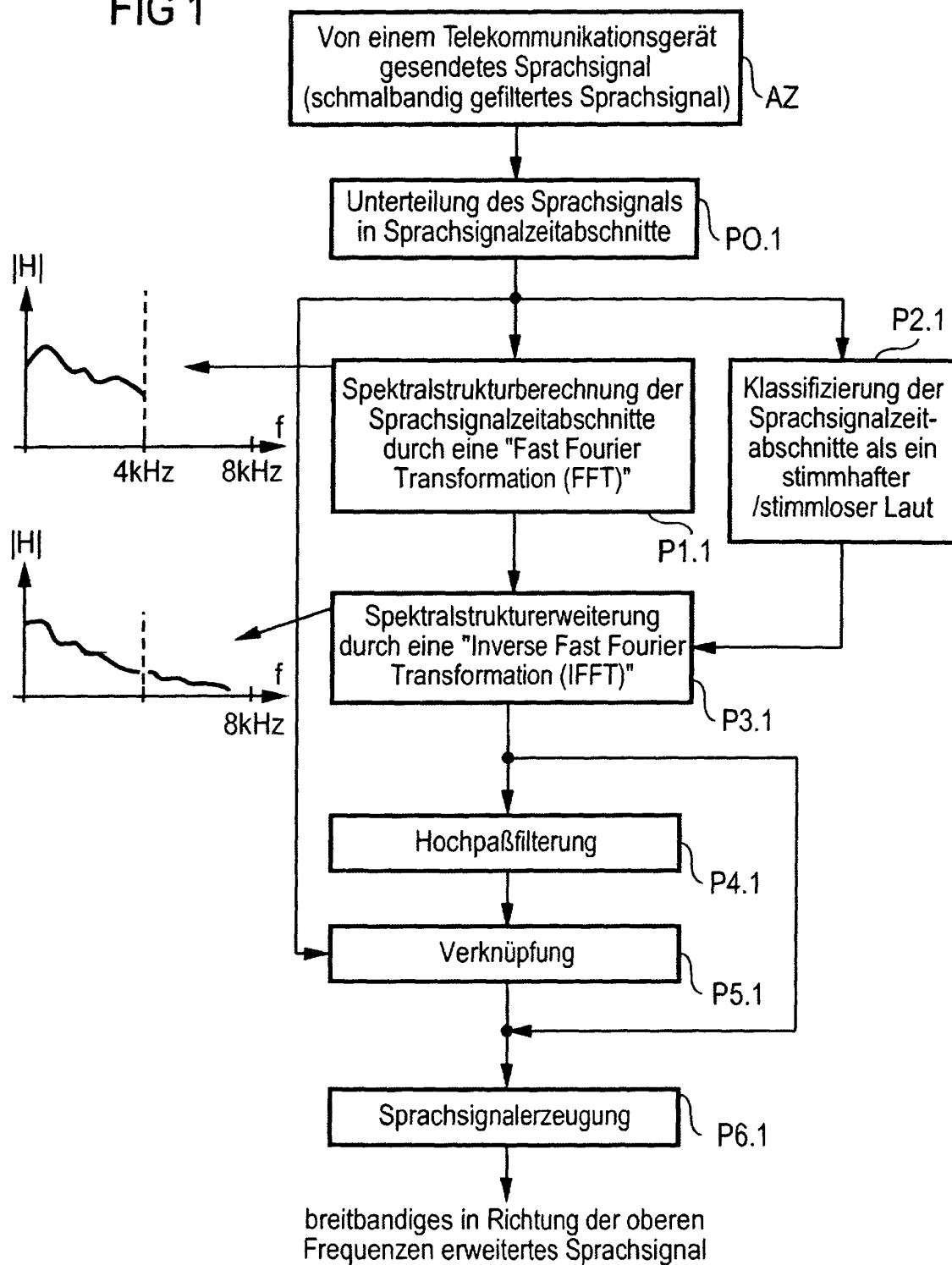


FIG 2

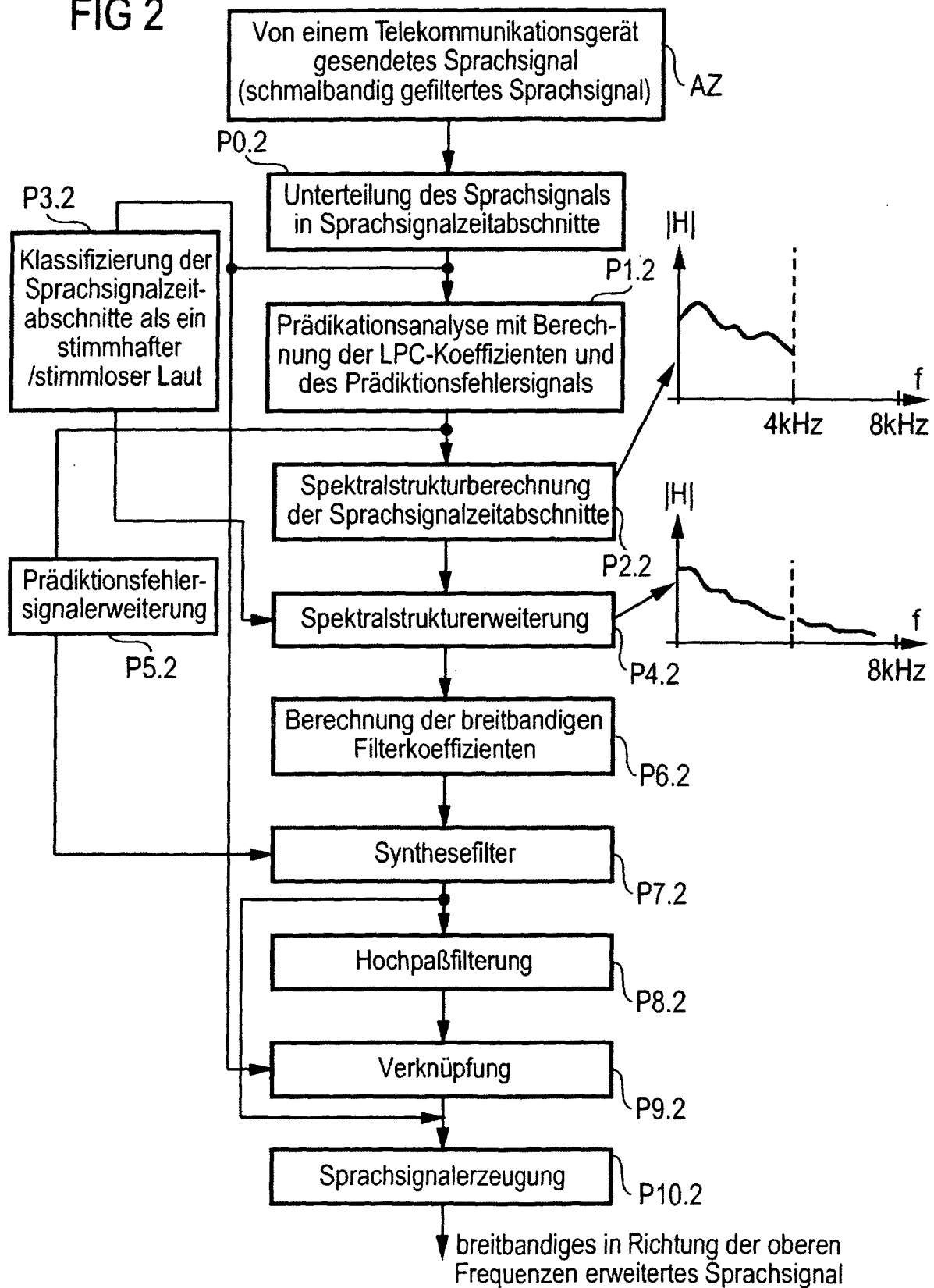


FIG 3

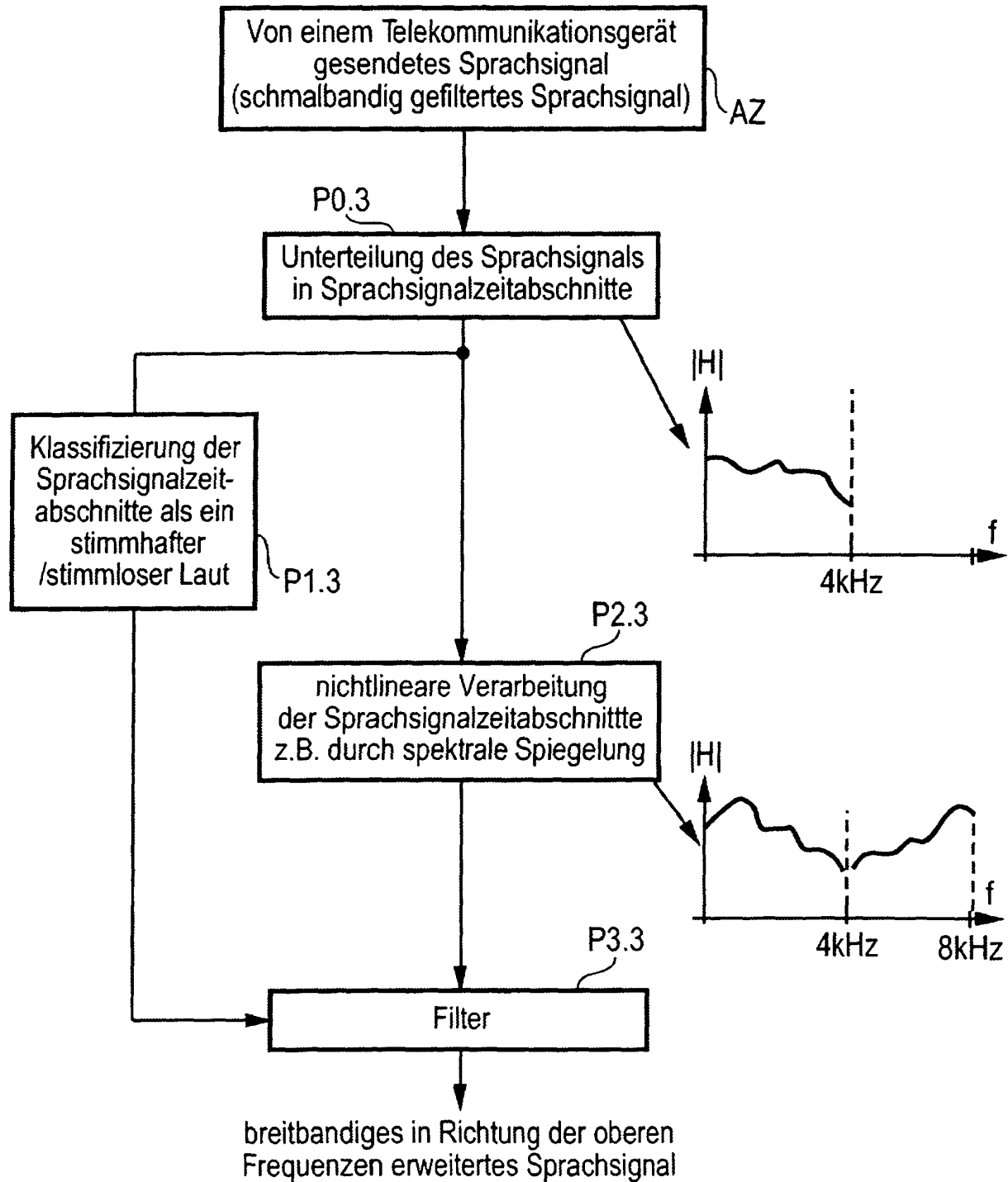


FIG 4A

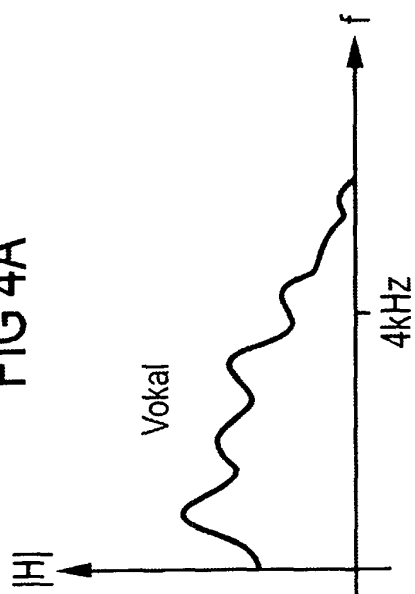


FIG 4B

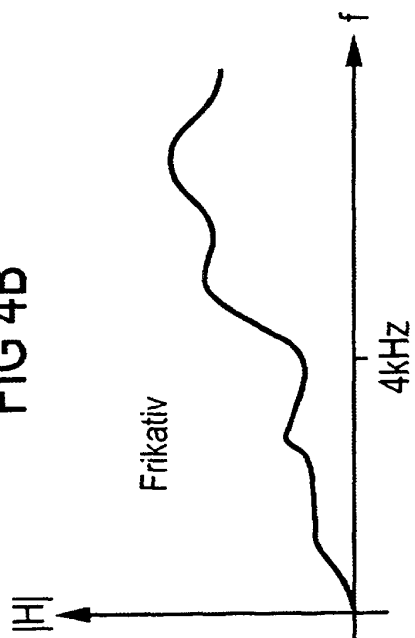


FIG 5A

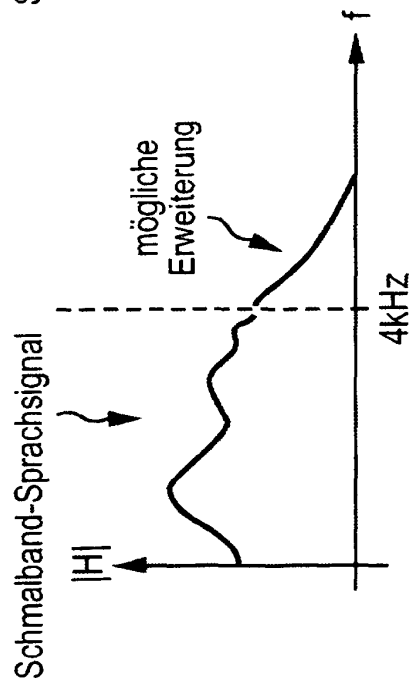
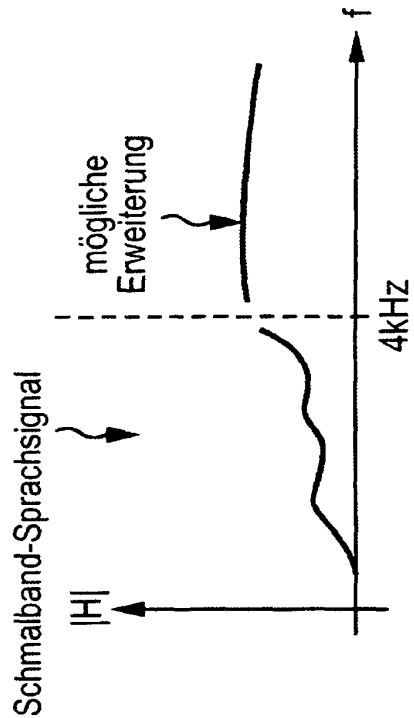


FIG 5B



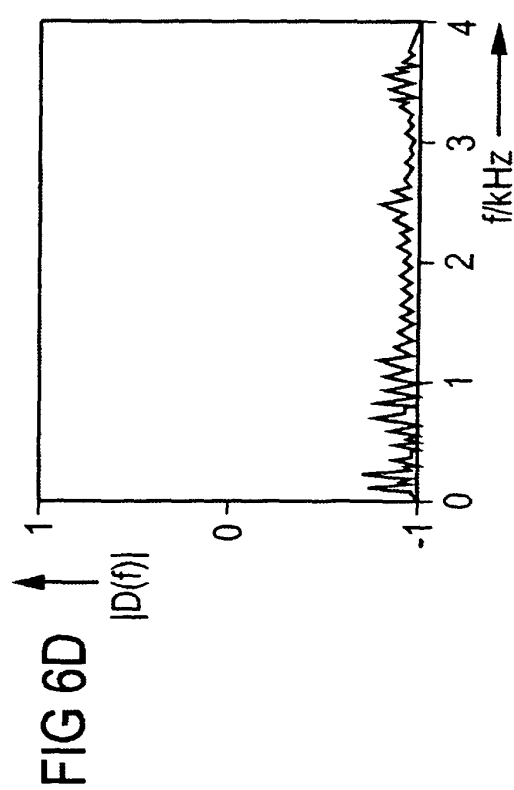
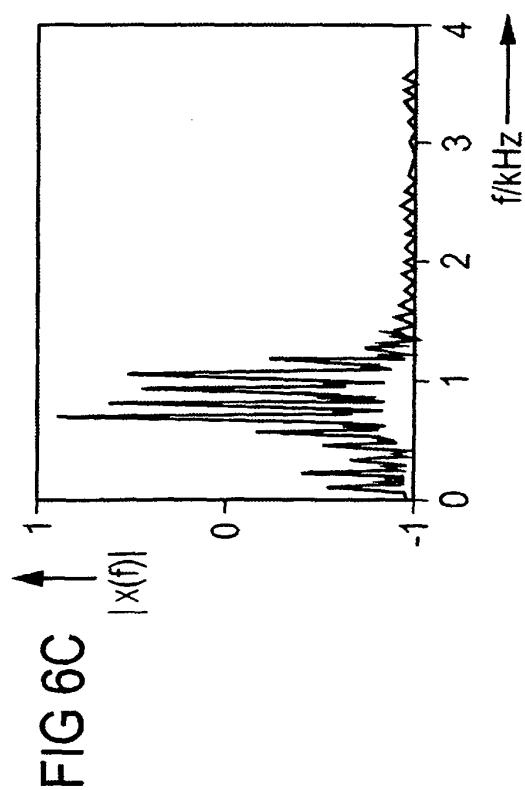
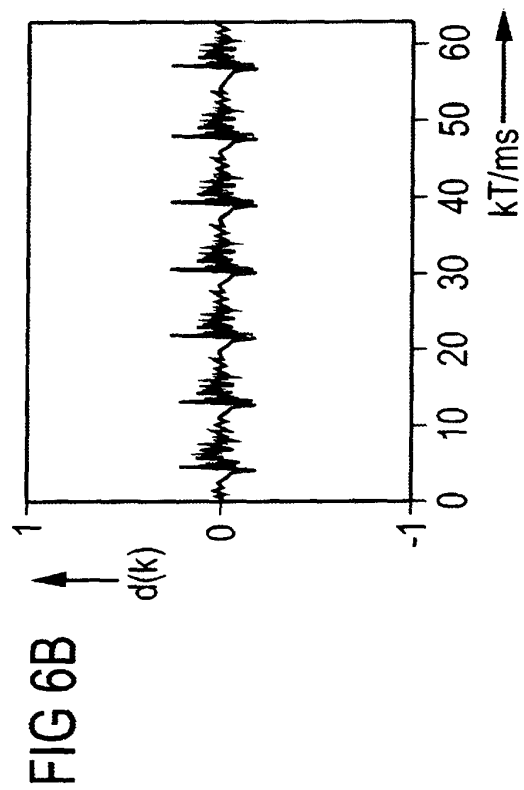
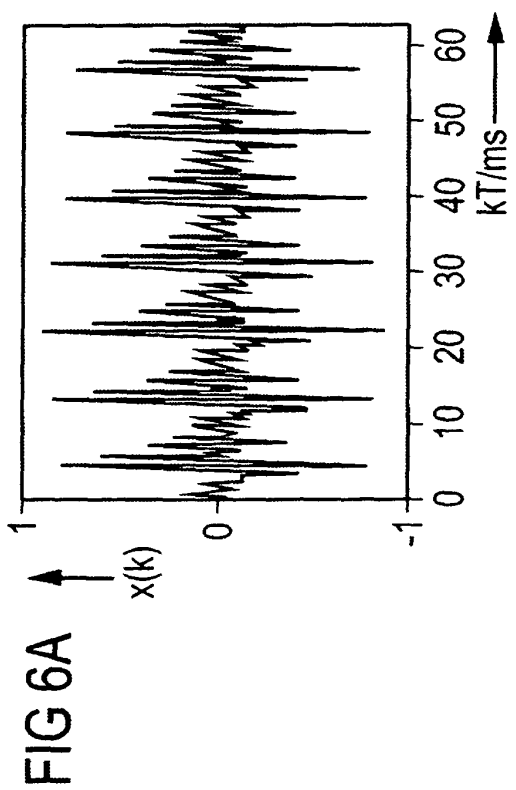


FIG 7 Filtercharakteristiken von verschiedenen Siemens-Telekommunikationsgeräten

