



(11)

EP 1 979 899 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
11.03.2015 Patentblatt 2015/11

(51) Int Cl.:
G10L 19/12^(2013.01) G10L 21/02^(2013.01)

(21) Anmeldenummer: **06706507.8**

(86) Internationale Anmeldenummer:
PCT/EP2006/000811

(22) Anmeldetag: **31.01.2006**

(87) Internationale Veröffentlichungsnummer:
WO 2007/087823 (09.08.2007 Gazette 2007/32)

(54) VERFAHREN UND ANORDNUNGEN ZUR AUDIOSIGNALKODIERUNG

METHOD AND ARRANGEMENTS FOR ENCODING AUDIO SIGNALS

PROCÉDÉ ET DISPOSITIFS POUR CODER UN SIGNAL AUDIO

(84) Benannte Vertragsstaaten:
DE FR GB IT SE

(56) Entgegenhaltungen:
EP-A- 0 883 107

(43) Veröffentlichungstag der Anmeldung:
15.10.2008 Patentblatt 2008/42

(73) Patentinhaber: **Unify GmbH & Co. KG**
81379 München (DE)

(72) Erfinder:
• **GEISER, Bernd**
52072 Aachen (DE)
• **JAX, Peter**
30539 Hannover (DE)
• **SCHANDL, Stefan**
A-1130 Wien (AT)
• **TADDEI, Hervé**
53115 Bonn (DE)

(74) Vertreter: **Fritzsche, Thomas et al**
Fritzsche Patent
Naupliastraße 110
81545 München (DE)

- **HONGMEI AI ET AL: "A 6.6 kb/s CELP speech coder: high performance for GSM half-rate system" ISSIPNN '94. 1994 INTERNATIONAL SYMPOSIUM ON SPEECH, IMAGE PROCESSING AND NEURAL NETWORKS PROCEEDINGS (CAT. NO.94TH0638-7) IEEE NEW YORK, NY, USA, Bd. 2, 1994, Seiten 555-558 vol.2, XP002382323 ISBN: 0-7803-1865-X**
- **CHOI J: "A FAST DETERMINATION OF STOCHASTIC EXCITATION WITHOUT CODEBOOK SEARCH IN CELP CODER" IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, Bd. 3, Nr. 6, November 1995 (1995-11), Seiten 473-480, XP000730633 ISSN: 1063-6676**
- **SALAMI R A: "BINARY CODE EXCITED LINEAR PREDICTION (BCELP): NEW APPROACH TO CELPCODING OF SPEECH WITHOUT CODEBOOKS" ELECTRONICS LETTERS, IEE STEVENAGE, GB, Bd. 25, Nr. 6, 16. März 1989 (1989-03-16), Seiten 401-403, XP000096828 ISSN: 0013-5194**

Anmerkung: Innerhalb von neun Monaten nach Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents im Europäischen Patentblatt kann jedermann nach Maßgabe der Ausführungsordnung beim Europäischen Patentamt gegen dieses Patent Einspruch einlegen. Der Einspruch gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist. (Art. 99(1) Europäisches Patentübereinkommen).

EP 1 979 899 B1

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren und Anordnungen zur Audiosignalkodierung. Insbesondere betrifft die Erfindung ein Verfahren und einen Anregungssignalgenerator zum Bilden eines Anregungssignals zum Anregen eines Audiosynthesefilters sowie einen Audiosignalenkodierer und einen Audiosignaldekodierer.

[0002] In vielen zeitgemäßen Kommunikationssystemen und insbesondere in mobilen Kommunikationssystemen stehen für Echtzeit-Audioübertragungen, wie z.B. Sprach- oder Musikübertragungen, nur begrenzte Übertragungsbandbreiten zur Verfügung. Um über eine Übertragungsstrecke mit begrenzter Bandbreite, wie z.B. über ein Funknetz, möglichst viele Audio- oder Sprachkanäle in Echtzeit zu übertragen, ist deshalb häufig vorgesehen, die zu übertragenden Audiosignale durch echtzeit- oder quasiechtzeitfähige Audio-Kodierungsverfahren zu komprimieren.

[0003] Bei derartigen Audio-Kodierungsverfahren wird in der Regel angestrebt, die zu übertragende Datenmenge und damit die Übertragungsrate möglichst zu reduzieren ohne den subjektiven Höreindruck bzw. bei Sprachübertragungen die Verständlichkeit, zu sehr zu beeinträchtigen.

[0004] Eine effiziente Komprimierung von Audiosignalen ist auch im Zusammenhang mit einer Speicherung oder Archivierung von Audiosignalen ein wesentlicher Gesichtspunkt.

[0005] Als besonders effizient erweisen sich Kodierungsverfahren, bei denen ein zu übertragendes Audiosignal zeitraumenweise an ein durch ein Audiosynthesefilter synthetisiertes Audiosignal durch Optimierung von Filterparametern angeglichen wird. Eine derartige Verfahrensweise wird häufig auch als Analysis-by-Synthesis bezeichnet. Das Audiosynthesefilter wird dabei durch ein vorzugsweise ebenfalls zu optimierendes Anregungssignal angeregt. Die Filterung wird häufig auch als FormantSynthese bezeichnet. Als Filterparameter können z.B. sog. LPC-Koeffizienten (LPC: Linear Predictive Coding) und/oder Parameter, die eine spektrale und/oder zeitliche Einhüllende des Audiosignals spezifizieren, verwendet werden. Die optimierten Filterparameter sowie das Anregungssignal spezifizierende Parameter werden dann zeitraumenweise zum Empfänger übertragen, um dort mittels eines empfängerseitig vorgesehenen Audiosynthesefilters ein synthetisches Audiosignal zu formen, das dem ursprünglichen Audiosignal hinsichtlich des subjektiven Höreindrucks möglichst ähnlich ist.

[0006] Ein derartiges Audio-Kodierungsverfahren ist aus der ITU-T Empfehlung G.729 bekannt. Mittels des dort beschriebenen Audio-Kodierungsverfahren kann ein Echtzeit-Audiosignal mit einer Bandbreite von 4 kHz auf eine Übertragungsrate von 8 kbit/s reduziert werden. Gemäß der G.729-Empfehlung wird das Anregungssignal mittels eines sog. adaptiven Kodebuchs im Zusammenwirken mit einem sog. fixen Kodebuch erzeugt. Im fixen Kodebuch ist eine Vielzahl von vorbestimmten Anregungssignalsequenzen fest gespeichert, die anhand eines Kodebuchindex abrufbar sind. Dagegen werden im adaptiven Kodebuch bereits erzeugte Anregungssignalsequenzen gespeichert. Eine jeweilige Sequenz des Anregungssignals wird durch Mischen einer Sequenz aus dem adaptiven Kodebuch mit einer Sequenz aus dem fixen Kodebuch erzeugt. Zur Optimierung des Anregungssignals werden für jeden Zeitrahmen sowohl das fixe als auch das adaptive Kodebuch nach Anregungssignalsequenzen durchsucht, die eine möglichst gute Angleichung des synthetischen Audiosignals an das zu übertragende Audiosignal erlauben. Als das Anregungssignal spezifizierende Parameter werden schließlich Zugriffsinformationen auf die als optimal gefundenen Sequenzen aus dem fixen und dem adaptiven Kodebuch zum Empfänger übertragen. Beim Empfänger werden diese Parameter dazu genutzt, um mittels eines fixen und eines adaptiven Kodebuchs des Empfängers ein Anregungssignal zu rekonstruieren.

[0007] Das gemäß der G.729-Empfehlung für jeden Zeitrahmen in Echtzeit auszuführende Durchsuchen zweier Kodebücher bedingt jedoch einen erheblichen Rechenaufwand, der aufwändige Prozessoren erfordert.

[0008] In EP 0 883 107 A1 ist Sprachkodierer / Dekodierer vom CELP-Typ (CELP: Code-Excited Linear Prediction) beschrieben, wobei ein Zufallskodevektor-Leseabschnitt und ein Zufallskodebuch eines herkömmlichen Sprachkodierers / Dekodierers jeweils durch einen Oszillator zum Ausgeben unterschiedlicher Vektor-Streams gemäß Werten von Eingangsseeds, und einen Seedspeicherabschnitt zum Speichern einer Vielzahl von Seeds ersetzt werden.

[0009] Darüber hinaus wird derzeit angestrebt, ein zu übertragendes Audiosignal zur Verbesserung des Höreindrucks mit höherer Bandbreite zu synthetisieren. Bei der gegenwärtig diskutierten Erweiterung G.729EV der G.729-Empfehlung wird versucht, die Audio-Bandbreite von 4 kHz auf 8 kHz zu erweitern.

[0010] Eine derartige Bandbreitenerweiterung des synthetisierten Audiosignals kann dadurch erzielt werden, dass aus einem schmalbandigen Anregungssignal, z.B. mit einer Bandbreite von 4 kHz, ein geeignetes Anregungssignal höherer Bandbreite, beispielsweise 8 kHz Bandbreite, konstruiert wird, um den Audiosynthesefilter breitbandig anzuregen. In diesem Zusammenhang werden gegenwärtig verschiedene Verfahrensweisen zum Bilden eines solchen breitbandigen Anregungssignals diskutiert. Danach kann das breitbandige Anregungssignal durch Quadrieren des schmalbandigen Anregungssignals im Zeitbereich oder durch Erzeugen eines Erweiterungsbandes durch Verschieben oder Spiegeln des Frequenzspektrums des schmalbandigen Anregungssignals erzeugt werden. Durch die genannten Verfahrensweisen wird jedoch das Spektrum des Anregungssignals anharmonisch verzerrt und/oder es wird ein erheblicher, hörbarer Phasenfehler im Spektrum verursacht.

[0011] Es ist Aufgabe der vorliegenden Erfindung ein Verfahren zum Bilden eines Anregungssignals für ein Audiosynthesefilter anzugeben, das bei Audiosignalübertragungen eine weitere Verringerung der Übertragungsrate und/oder

eine Verbesserung des Höreindrucks sowie eine Verringerung des zur Audiokodierung erforderlichen Rechenaufwandes erlaubt. Es ist weiterhin Aufgabe der Erfindung einen Anregungssignalgenerator zur Durchführung des Verfahrens sowie einen Audiosignalenkodierer und einen Audiosignaldekodierer anzugeben.

[0012] Gelöst wird diese Aufgabe durch ein Verfahren mit den Merkmalen des Patentanspruchs 1, einen Anregungssignalgenerator mit den Merkmalen des Patentanspruchs 7, einen Audiosignalenkodierer mit den Merkmalen des Patentanspruchs 9 sowie einen Audiosignaldekodierer mit den Merkmalen des Patentanspruchs 10.

[0013] Vorteilhafte Ausführungsformen und Weiterbildungen der Erfindung sind in den abhängigen Ansprüchen angegeben.

[0014] Beim erfindungsgemäßen Verfahren zum Bilden eines Anregungssignals zum Anregen eines Audiosynthesefilters wird das Anregungssignal als Folge von Anregungs-Abtastwerten gebildet. Bereits gebildete Anregungs-Abtastwerte werden hierbei zeitlich fortlaufend in einem adaptiven Kodebuch gespeichert. Weiterhin ist ein Rauschgenerator vorgesehen, durch den fortlaufend Zufalls-Abtastwerte generiert werden. Aus dem adaptiven Kodebuch wird eine Sequenz der gespeicherten Anregungs-Abtastwerte an Hand eines zugeführten Audiogrundfrequenzparameters ausgewählt, durch den ein Zeitabstand der auszuwählenden Sequenz zum aktuellen Zeitbezug vorgegeben wird. Falls der Audiogrundfrequenzparameter einen Zeitabstand vorgibt, der kein ganzzahliges Vielfaches eines vorgegebenen Abtastabstandes eines separat zu generierenden, schmalbandigen Anregungssignals ist, werden zwischen die Anregungs-Abtastwerte und/oder zwischen die Zufalls-Abtastwerte abhängig vom Audiogrundfrequenzparameter Zwischen-Abtastwerte derart eingefügt, dass ein Abtastabstand der Abtastwerte geringer ist als der Abtastabstand des schmalbandigen Anregungssignals, wodurch das Anregungssignal gegenüber dem schmalbandigen Anregungssignal zusätzliche Frequenzkomponenten eines Erweiterungsbandes aufweist. Das Anregungssignal wird durch Mischen der ausgewählten Sequenz mit einer aktuellen Zufalls-Abtastwerte des Rauschgenerators umfassenden Zufalls-Sequenz gebildet.

[0015] Durch die Verwendung des Rauschgenerators als Quelle von Zufalls-Abtastwerten kann auf ein fixes Kodebuch zum Befüllen des adaptiven Kodebuchs verzichtet werden. Dementsprechend ist es nicht erforderlich, Kodebuch-Indizes zum Auswählen von vorgegebenen in einem fixen Kodebuch gespeicherten Abtastwertsequenzen bereitzustellen oder zu übertragen. Da derartige Kodebuch-Indizes für ein fixes Kodebuch bei bekannten Verfahren einen erheblichen Anteil an den zu übertragenden Audiodaten einnehmen, kann durch die Erfindung die Übertragungsrate in der Regel erheblich verringert werden. Die eingesparte Übertragungsbandbreite kann entsprechend für andere Zwecke oder zur Erhöhung der Übertragungsqualität genutzt werden.

[0016] Durch den Rauschgenerator, der vorzugsweise ein im Wesentlichen weißes, spektral flaches Rauschen erzeugt, kann ein in Audiosignalen oder Sprachsignalen enthaltener Rauschanteil in der Regel besser modelliert werden, als mittels eines fixen Kodebuchs, das nur fest vorgegebene Abtastwertsequenzen enthält. Eine harmonische Feinstruktur der Audio- oder Sprachsignale kann dagegen durch die vom Audiogrundfrequenzparameter abhängige Auswahl einer Abtastwertsequenz aus dem adaptiven Kodebuch gut nachgebildet werden.

[0017] Da ein Rauschgenerator naturgemäß gut auf unterschiedliche Frequenzbereiche skalierbar ist, lassen sich Bandbreitenerweiterungen mit geringem Aufwand realisieren. Zudem lässt sich durch die Erfindung vermeiden, dass ein Kodierungsrestfehler bei einer Bandbreitenerweiterung in ein Erweiterungsband übertragen wird.

[0018] Erfindungsgemäß ist für den Fall, dass der Audiogrundfrequenzparameter einen Zeitabstand vorgibt, der kein ganzzahliges Vielfaches eines vorgegebenen Abtastabstandes eines separat zu generierenden, schmalbandigen Anregungssignals ist, vorgesehen, zwischen die Anregungs-Abtastwerte und/oder zwischen die Zufalls-Abtastwerte abhängig vom Audiogrundfrequenzparameter Zwischen-Abtastwerte einzufügen. Die Einfügung erfolgt vorzugsweise so, dass ein Abtastabstand der resultierenden Abtastwerte geringer ist als der Abtastabstand des schmalbandigen Anregungssignals. Auf diese Weise kann ein Anregungssignal erzeugt werden, das gegenüber einem schmalbandigen Anregungssignal, z.B. im Frequenzbereich von 0-4 kHz, zusätzliche Frequenzkomponenten eines Erweiterungsbandes, z.B. von 4-8 kHz, aufweist. Das so erzeugte Anregungssignal weist anders als durch bekannte Bandbreiten-Erweiterungsverfahren erzeugte Anregungssignale keine wesentlichen anharmonischen Verzerrungen auf.

[0019] Die Erfindung ist sowohl beim Enkodieren als auch bei Dekodieren eines Audiosignals vorteilhaft einsetzbar. Bei einem Audiosignalenkodierer kann durch einen erfindungsgemäßen Anregungssignalgenerator ein Audiosynthesefilter angeregt werden, dessen Ausgangs-Audiosignal mit einem jeweils aktuellen Rahmen des zu übertragenden Audiosignals verglichen wird. Der Vergleich des aktuellen Rahmens wird vorzugsweise für unterschiedliche Auswahlen von im adaptiven Kodebuch gespeicherten Sequenzen von früheren Anregungs-Abtastwerten durchgeführt. Die zeitliche Position derjenigen Abtastwertsequenz innerhalb des adaptiven Kodebuchs, bei der der Vergleich eine optimale Übereinstimmung anzeigt, kann durch einen entsprechenden Audiogrundfrequenzparameter ausgedrückt werden, der dann zu einem Empfänger übertragen werden kann. Eine Durchsuchung eines weiteren, fixen Kodebuchs sowie eine zusätzliche Übertragung von Kodebuchindizes sind nicht erforderlich.

[0020] Bei einem Audiosignaldekodierer kann durch einen jeweils empfangenen Audiogrundfrequenzparameter ein erfindungsgemäßer Anregungssignalgenerator so gesteuert werden, dass er ein dem Audiogrundfrequenzparameter harmonisch entsprechendes Anregungssignal erzeugt, ohne auf zusätzlich zu übertragende Kodebuchindizes angewiesen zu sein. Durch das so erzeugte Anregungssignal kann ein Audiosynthesefilter angeregt werden, um ein synthe-

tisches Audiosignal zu erzeugen, dass dem ursprünglichen Audiosignal hinsichtlich des Höreindrucks sehr ähnelt.

[0021] Damit verringern sich sowohl der erforderliche Rechenaufwand beim Audiosignalenkodierer als auch die erforderliche Übertragungsrate. Entsprechend kann bei gleich bleibender Übertragungsrate in der Regel eine höhere Übertragungsqualität und damit ein verbesserter Höreindruck erzielt werden.

[0022] Die Audiosynthesefilter beim Audiosignalenkodierer und/oder Audiosignaldekodierer können z.B. als LPC-Filter, Wiener FIR-Filter, als Filter zur Formung einer zeitlichen oder spektralen Einhüllenden des Audiosignals oder als eine Kombination dieser Filter realisiert werden.

[0023] Das erfindungsgemäße Verfahren kann vorzugsweise durch einen Signalprozessor ausgeführt werden.

[0024] Nach einer vorteilhaften Ausführungsform der Erfindung können die Anregungs-Abtastwerte und/oder die Zufalls-Abtastwerte zeitraumenweise verarbeitet werden, wobei die Länge der ausgewählten Sequenz und/oder die Länge der Zufalls-Sequenz einer vorgegebenen Länge eines Zeitrahmens entsprechen.

[0025] Gemäß einer weiteren Ausführungsform der Erfindung können beim Mischen die ausgewählte Sequenz gemäß einem ersten Intensitätsparameter und/oder die Zufalls-Sequenz gemäß einem zweiten Intensitätsparameter verstärkt werden. Der erste und der zweite Intensitätsparameter, wie auch der Audiogrundfrequenzparameter, können vorzugsweise zeitraumenweise, aus dem zu übertragenden Audiosignal abgeleitet und übertragen werden.

[0026] Weiterhin kann das Anregungssignal mit einem gegenüber einem separat zu generierenden, schmalbandigen Anregungssignal geringeren Abtastabstand gebildet werden, wodurch das Anregungssignal gegenüber dem schmalbandigen Anregungssignal zusätzliche Frequenzkomponenten eines Erweiterungsbandes aufweist. In diesem Fall kann der Audiogrundfrequenzparameter sowie der erste und/oder zweite Intensitätsparameter aus Audiosynthese-Parametern abgeleitet werden, die eigentlich zur Generierung des schmalbandigen Anregungssignals vorgesehen sind. Analog dazu können der Audiogrundfrequenzparameter sowie der erste und/oder der zweite Intensitätsparameter aus einem Schmalbandanteil eines zu übertragenden Audiosignals abgeleitet werden.

[0027] Der Audiogrundfrequenzparameter sowie der erste und/oder der zweite Intensitätsparameter können mithin aus Schmalband-Audioparametern abgeleitet, aber auf das Erweiterungsband angewandt werden. Dies ist insofern vorteilhaft, da zur Bandbreitenerweiterung des Anregungssignals außer den zur Generierung des schmalbandigen Anregungssignals vorgesehenen Audiosyntheseparametern keine zusätzlichen Audiosynthese-Parameter erforderlich sind. Die zur Generierung des schmalbandigen Anregungssignals vorgesehenen Audiosynthese-Parameter können in der Regel von bestehenden, schmalbandigen Audiocodern, wie z.B. gemäß G.729-Empfehlung bereitgestellt werden.

[0028] Bei bekannten Schmalband-Übertragungsverfahren, wie z.B. gemäß G.729-Empfehlung, wird der Audiogrundfrequenzparameter häufig genauer ermittelt, als dem Abtastabstand des schmalbandigen Anregungssignals entspricht. Häufig ist eine Genauigkeit von z.B. einem halben oder drittel Abtastabstand vorgesehen. Somit kann der für das schmalbandige Anregungssignal vorgesehene Audiogrundfrequenzparameter in der Regel direkt bzw. im Wesentlichen unverändert zur Erzeugung des bandbreitenerweiterten Anregungssignals verwendet werden.

[0029] Der erste und/oder der zweite Intensitätsparameter kann jeweils durch Anwendung einer vorgegebenen Funktion aus den entsprechenden schmalbandigen Intensitätsparametern abgeleitet werden, um z.B. einen Rauschanteil gegenüber einem harmonischen Anteil im Erweiterungsband eines Audiosignals zu betonen.

[0030] Vorzugsweise kann ein auf das Erweiterungsband entfallender Anteil des Anregungssignals mit dem separat generierten, schmalbandigen Anregungssignal kombiniert werden, um ein breitbandiges Anregungssignal, z.B. im Frequenzbereich von 0 bis 8 kHz, zur Anregung des Audiosynthesefilters zu erzeugern.

[0031] Ein Ausführungsbeispiel der Erfindung wird nachfolgend anhand der Zeichnung näher erläutert.

[0032] Dabei zeigen jeweils in schematischer Darstellung:

Figur 1 ein mit unterschiedlichen Abtastraten abgetastetes Audiosignal,

Figuren 2a und 2b verschiedene Ausführungsbeispiele eines Anregungssignalgenerators,

Figur 3. die Veranschaulichung eines Auswahlvorgangs einer Abtastwert-Sequenz aus einem adaptiven Kodebuch, und

Figur 4 einen Audiosignaldekodierer.

[0033] Figur 1 veranschaulicht ein mit unterschiedlichen, beispielhaften Abtastraten abgetastetes Audiosignal. Einzelne Abtastwerte sind hierbei als Punkte dargestellt, die verschiedene, durch senkrechte Linien veranschaulichte Amplituden aufweisen. Die verschiedenen Abtastraten werden durch unterschiedliche zeitliche Abtastabstände zwischen den Abtastwerten verdeutlicht. Beide Teilfiguren weisen eine gemeinsame Zeitachse T auf.

[0034] Die obere Teilfigur veranschaulicht das mit einer Abtastrate von beispielhaft 8 kHz abgetastete Audiosignal. Die Abtastrate von 8 kHz entspricht einem Abtastabstand $DT1$ von $1/8000$ s. Durch die mit einer Abtastrate von 8 kHz abgetasteten Abtastwerte können nach einem fundamentalen Abtasttheorem Audiosignale im Wesentlichen bis zu einer

Frequenz von 4 kHz dargestellt werden. Dieser Frequenzbereich wird im Folgenden als schmalbandig bezeichnet.

[0035] Die untere Teilfigur veranschaulicht das mit einer Abtastrate von 16 kHz abgetastete Audiosignal. Entsprechend der gegenüber der oberen Teilfigur verdoppelten Abtastrate beträgt der Abtastabstand DT2 in der unteren Teilfigur die Hälfte des Abtastabstandes DT1, d.h. hier 1/16000 s. Durch die mit einer Abtastrate von 16 kHz abgetasteten Abtastwerte kann ein Audiosignal im Wesentlichen bis zu einer Frequenz von 8 kHz dargestellt werden. Der vorstehende Frequenzbereich wird im Folgenden auch als breitbandig bezeichnet. Es versteht sich, dass die Begriffe schmalbandig und breitbandig nicht auf die nur beispielhaft genannten Frequenzbereiche beschränkt sind, sondern auf beliebige Frequenzbereiche insoweit verallgemeinerbar sind, als der Begriff breitbandig einen größeren Frequenzbereich spezifizieren soll, als der Begriff schmalbandig.

[0036] Die Figuren 2a und 2b zeigen in schematischer Darstellung verschiedene Ausführungsbeispiele eines Anregungssignalgenerators. Die dargestellten Anregungssignalgeneratoren umfassen als Funktionskomponenten jeweils einen Rauschgenerator NOISE, ein adaptives Kodebuch ACB sowie eine Mischeinrichtung MIX. Der Zufallsgenerator NOISE dient zum zeitlich fortlaufenden Generieren von Zufalls-Abtastwerten mit einem jeweils vorgegebenen Abtastabstand. Für beide in Figur 2a und 2b dargestellten Ausführungsbeispiele sei beispielhaft angenommen, dass der jeweilige Rauschgenerator NOISE Zufalls-Abtastwerte mit einer schmalbandigen Abtastrate, also z.B. 8 kHz generiert. Unter Zufalls-Abtastwerten seien hierbei Abtastwerte verstanden, die vom Rauschgenerator zeitlich fortlaufend, zufällig oder quasizufällig erzeugt werden und insbesondere nicht vorgegeben sind oder aus vorgegebenen Werten ausgewählt werden. Insbesondere werden die Zufalls-Abtastwerte unabhängig von einem mittels des jeweiligen Anregungssignalgenerators zu enkodierenden oder zu dekodierenden Audiosignals erzeugt. Mithin ist zum Betrieb des Rauschgenerators NOISE kein Zuführen oder Übermitteln spezifischer Zugriffsparameter erforderlich wie bei einem fixen Kodebuch gemäß dem Stand der Technik. In einem solchen fixen Kodebuch sind fest vorgegebene, deterministische Abtastsequenzen gespeichert, zu deren zeitrahenweisem Abruf fortlaufend Kodebuch-Indizes zuzuführen sind, was in der Regel einen erheblichen Anteil an der Übertragungsbandbreite beansprucht.

[0037] Vorzugsweise weist ein durch die Zufalls-Abtastwerte gebildetes Rauschsignal ein im Wesentlichen weißes bzw. flaches Frequenzspektrum auf.

[0038] Im Folgenden wird zunächst das in Figur 2a dargestellte Ausführungsbeispiel des Anregungssignalgenerators betrachtet. Der dort dargestellte Anregungssignalgenerator kann allgemein zur Audio- und/oder Sprachkodierung eingesetzt werden. Sowohl der Rauschgenerator NOISE, als auch das adaptive Kodebuch ACB geben Abtastwerte zeitrahenweise, d.h. als Abfolge von Abtastwerte enthaltenden Zeitrahmen vorgegebener Länge aus. Ein Zeitrahmen von z.B. von 5 ms Länge enthält bei einer Abtastrate von z.B. 8 kHz entsprechend 40 Abtastwerte. Bei einer Abtastrate von 16 kHz enthält einer derartiger Zeitrahmen entsprechend 80 Abtastwerte.

[0039] Während der Rauschgenerator NOISE fortlaufend Zufalls-Sequenzen EXC_N, d.h. Zeitrahmen mit Zufalls-Abtastwerten erzeugt, gibt das adaptive Kodebuch ACB fortlaufend Sequenzen d.h. Zeitrahmen EXC_P von gespeicherten Anregungs-Abtastwerten aus. Die Zufalls-Sequenzen EXC_N und die vom adaptiven Kodebuch ACB ausgegebenen Sequenzen EXC_P werden der Mischeinrichtung MIX zugeleitet, der auch Intensitätsparameter G_N zur Pegelsteuerung der Zufalls-Sequenzen EXC_N sowie Intensitätsparameter G_P zur Pegelsteuerung der vom adaptiven Kodebuch ACB kommenden Sequenzen EXC_P zeitrahenweise zugeleitet werden. In der Mischeinrichtung MIX werden die Zufalls-Abtastwerte einer jeweiligen Zufalls-Sequenz EXC_N mit einem jeweiligen Intensitätsparameter G_N und die Abtastwerte einer jeweiligen vom adaptiven Kodebuch ACB ausgegebenen Sequenz EXC_P mit einem jeweiligen Intensitätsparameter G_P zeitrahenweise multipliziert, d.h. verstärkt. Die Multiplikationen sind in Figur 2a durch mit Multiplikationszeichen versehene Kreise angedeutet. Die gemäß den Intensitätsparametern G_N und G_P verstärkten Abtastwertsequenzen werden durch die Mischeinrichtung MIX zeitrahenweise addiert und das resultierende Summensignal als Anregungssignal EXC in Form einer Folge von Anregungs-Abtastwerten ausgegeben. Die Addition ist in Figur 2a durch einen mit einem Pluszeichen versehenen Kreis veranschaulicht. Das gebildete Anregungssignal EXC wird ausgegeben und parallel dazu zeitlich fortlaufend im adaptiven Kodebuch ACB gespeichert. Das Anregungssignal EXC wird also gewissermaßen vom Ausgang der Mischeinrichtung MIX zum adaptiven Kodebuch ACB rückgekoppelt.

[0040] Hinsichtlich des Anregungssignals EXC wirkt das adaptive Kodebuch ACB ähnlich wie ein Schieberegister, in das aktuell gebildete Sequenzen des Anregungssignals EXC eingespeichert werden und dabei früher gebildete Sequenzen des Anregungssignals unter Wahrung der zeitlichen Reihenfolge sukzessive nach hinten verschieben.

[0041] Die Ausgabe der Sequenzen EXC_P von gespeicherten Anregungs-Abtastwerten wird durch dem adaptiven Kodebuch ACB zeitrahenweise zugeführte Audiogrundfrequenzparameter PITCH gesteuert. An Hand der Audiogrundfrequenzparameter PITCH werden die vom adaptiven Kodebuch ACB auszugebenden Sequenzen EXP aus den gespeicherten Anregungs-Abtastwerten ausgewählt. Die Auswahl erfolgt mittels einer Auswahlrichtung SEL des adaptiven Kodebuchs ACB. Ein derartiger Audiogrundfrequenzparameter PITCH wird in der Fachwelt häufig auch als "pitch lag" bezeichnet.

[0042] Es sei im Folgenden angenommen, dass die Audiogrundfrequenzparameter PITCH jeweils in Einheiten eines schmalbandigen Abtastabstandes, hier z.B. 1/8000 s bei einer schmalbandigen Abtastrate von 8 kHz, vorgegeben sind. Durch die Audiogrundfrequenzparameter PITCH wird jeweils zeitrahenweise eine Periode einer Grundfrequenz des

zu übertragenden oder zu synthetisierenden Audiosignals spezifiziert. Bei zeitgemäßen Audiokodierverfahren, z.B. gemäß der G.729-Empfehlung, werden die Grundfrequenz-Perioden eines Audiosignals häufig mit höherer Auflösung gemessen oder bereitgestellt, als einem jeweils verwendeten Abtastabstand entspricht. Ein solcher, bis auf Bruchteile, von Abtastabständen genauer Audiogrundfrequenzparameter kann also in Einheiten des Abtastabstandes auch nicht-ganzzahlige Werte annehmen. Eine solcher nichtganzzahliger Audiogrundfrequenzparameter PITCH enthält Informationen über höhere Frequenzanteile, als dem Abtastabstand eigentlich entspricht. Während solche höheren Frequenzanteile bei bekannten Audiokodierern, z.B. gemäß der G.729-Empfehlung herausgefiltert werden, kann die Information über die höheren Frequenzanteile bei erfindungsgemäßen Audiosignalgeneratoren auf einfache Weise zur Verbesserung der Audiosynthesequalität genutzt werden.

[0043] Figur 3 veranschaulicht die Auswahl einer Abtastwertsequenz EXC_P aus dem adaptiven Kodebuch ACB anhand des der Auswahlrichtung SEL zugeführten Audiogrundfrequenzparameters PITCH. Figur 3 zeigt einen Ausschnitt der im adaptiven Kodebuch ACB zeitlich fortlaufend gespeicherten Anregungs-Abtastwerte. Die gespeicherten Anregungs-Abtastwerte sind durch mit senkrechten Linien versehene Punkte angedeutet, wobei die Länge einer jeweiligen Linie eine jeweilige Amplitude eines Anregungs-Abtastwerts veranschaulicht. Der zeitliche Verlauf wird durch eine Zeitachse angezeigt.

[0044] Ein aktueller Zeitbezug T0 wird in Figur 3 durch eine senkrechte Linie angedeutet, die diejenige Stelle im adaptiven Kodebuch angibt, an der ein jeweiliger aktuell gebildeter Zeitrahmen des Anregungssignals im adaptiven Kodebuch ACB neu eingespeichert wird. Die Speicherung erfolgt hierbei zeitlich oder logisch angrenzend an einen unmittelbar vorher abgespeicherten Zeitrahmen des Anregungssignals. Aus Übersichtlichkeitsgründen umfasst ein Zeitrahmen in Figur 3 nur vier Abtastwerte. Eine Verallgemeinerung der durch Figur 3 veranschaulichten Zusammenhänge auf Zeitrahmen mit beliebiger vorgegebener Länge ist evident.

[0045] Aus dem adaptiven Kodebuch ACB wird diejenige Sequenz EXC_P von gespeicherten Anregungs-Abtastwerten zur Ausgabe ausgewählt, deren Beginn einen dem Audiogrundfrequenzparameter PITCH entsprechenden Zeitabstand vom aktuellen Zeitbezug T0 aufweist und deren Länge der vorgegebenen Länge eines Zeitrahmens entspricht. Der Zeitabstand wird hierbei vom aktuellen Zeitbezug T0 aus zeitlich rückwärts gerechnet. Es sei darauf hingewiesen, dass der Beginn der ausgewählten Sequenz EXC_P nicht auf eine Zeitrahmengrenze fallen muss, sondern ggf. innerhalb vorgegebener Grenzen auf einen beliebigen gespeicherten Anregungs-Abtastwert fallen kann.

[0046] In Figur 3 ist beispielhaft angenommen, dass durch den mit dem aktuellen Zeitrahmen übermittelten Audiogrundfrequenzparameter PITCH ein Zeitabstand von sechs Abtastabständen spezifiziert wird. Somit wird als ausgewählte Sequenz EXC_P ein Zeitrahmen vom sechstletzten gespeicherten Anregungs-Abtastwert bis zum drittletzten gespeicherten Anregungs-Abtastwert, gerechnet vom aktuellen Zeitbezug T0, ausgegeben. Der ausgegebene Zeitrahmen EXC_P ist in Figur 3 durch ein strichliertes Rechteck angedeutet.

[0047] Beim Einschalten des Anregungssignalgenerators ist das adaptive Kodebuch ACB zunächst leer, um dann sukzessive mit gebildeten Anregungs-Abtastwerten des ausgegebenen Anregungssignals EXC gefüllt zu werden. Da das adaptive Kodebuch ACB zu Beginn leer ist, wird das Anregungssignal EXC zunächst nur durch den Rauschgenerator NOISE als einziger Signalquelle gespeist. Dies bedeutet, dass das adaptive Kodebuch ACB zunächst mit nicht-periodischen Zufalls-Abtastwerten befüllt wird. In diesem Szenario stellt sich die Frage, wie sich mittels des adaptiven Kodebuchs ACB periodische Signalkomponenten gewinnen lassen, da als originäre Signalquelle lediglich ein nicht-periodischer Rauschgenerator NOISE zur Verfügung steht. Tatsächlich wurde nach bisherigen Vorstellungen für nötig erachtet, neben einem adaptiven Kodebuch, auch ein fixes Kodebuch vorzusehen, um das adaptive Kodebuch ACB mit determinierten, im fixen Kodebuch gespeicherten Signalsequenzen, zu befüllen.

[0048] Gemäß einer Untersuchung der Erfinder ist ein derartiges fixes Kodebuch jedoch nicht erforderlich. Tatsächlich kann auch ohne fixes Kodebuch ein Anregungssignal mit harmonischer Feinstruktur durch fortlaufende geeignete Auswahl von Abtastwertsequenzen EXC_P aus dem adaptiven Kodebuch ACB erzeugt werden. Zur Verdeutlichung des zugrunde liegenden Wirkungsprinzips sei ein Fall betrachtet, bei dem der Audiogrundfrequenzparameter PITCH über mehrere Zeitrahmen hinweg konstant bleibt. In einem solchen Fall wird aus dem adaptiven Kodebuch ACB mehrmals hintereinander ein Zeitrahmen mit der gleichen Zeitlage relativ zu T0 aus dem adaptiven Kodebuch ACB ausgelesen, mit einer Zufalls-Sequenz EXC_N des Rauschgenerators NOISE gemischt und als aktueller Zeitrahmen des Anregungssignals EXC wieder im aktuellen Kodebuch ACB gespeichert. Der aktuelle Zeitrahmen wird dabei mit einem durch den Audiogrundfrequenzparameter PITCH spezifizierten Abstand zur vorher ausgegebenen Sequenz EXC_P gespeichert. Auf diese Weise bildet sich im adaptiven Kodebuch ACB sukzessive ein periodischer Signalanteil, dessen Periode durch den Audiogrundfrequenzparameter PITCH bestimmt wird. Der periodische Anteil am Gesamtanregungssignal EXC wird dabei durch die Intensitätsparameter G_N und G_P gesteuert.

[0049] Aufgrund der Verwendung des Rauschgenerators NOISE an Stelle eines fixen Kodebuchs kann eine Übertragung von Kodebuch-Indizes für ein fixes Kodebuch entfallen. Auf diese Weise kann die Übertragungsrate bzw. Bandbreite für die Übertragung von Audiosignalen erheblich reduziert werden. Zudem lässt sich durch die Verwendung des Rauschgenerators NOISE ein besserer Höreindruck, insbesondere bei einer Wiedergabe nicht harmonischer oder rauschartiger Audioanteile erzielen.

[0050] Anhand von Figur 2b wird im Folgenden eine Ausführungsform des erfindungsgemäßen Anregungssignalgenerators zum Erzeugen eines bandbreitenerweiterten Anregungssignals EXC erläutert. Das ausgegebene Anregungssignal EXC wird mit einer um einen Bandbreitenerweiterungsfaktor N erweiterten Bandbreite erzeugt. Die auch in Figur 2a verwendeten Bezugszeichen behalten in Figur 2b ihre Bedeutung.

[0051] Es sei beispielhaft angenommen, dass der Bandbreitenerweiterungsfaktor N einen Wert von 2 aufweist, und dass bei einer schmalbandigen Abtastrate von 8 kHz die Abtastrate des auszugebenden Anregungssignals EXC entsprechend $N \times 8 \text{ kHz} = 16 \text{ kHz}$ beträgt.

[0052] Während der Rauschgenerator NOISE Zufalls-Abtastwerte mit der schmalbandigen Abtastrate von 8 kHz ausgibt, verwenden das adaptive Kodebuch ACB und die Mischeinrichtung MIX die breitbandige Abtastrate von 16 kHz. Zur Anpassung der schmalbandigen Abtastrate des Rauschgenerators NOISE an die breitbandige Abtastrate der Mischeinrichtung MIX ist zwischen diese und dem Rauschgenerator NOISE ein Interpolator INT_N geschaltet. Der Interpolator INT_N empfängt die vom Rauschgenerator NOISE mit schmalbandiger Abtastrate ausgegebenen Zufalls-Abtastwerte und setzt zwischen jeweils zwei dieser Zufalls-Abtastwerte einen Zwischen-Abtastwert mit der Amplitude 0. Für andere Werte des Bandbreitenerweiterungsfaktors N werden analog dazu N-1 Zwischen-Abtastwerte, jeweils mit Amplitude 0, zwischen jeweils zwei Zufalls-Abtastwerten eingefügt. Auf diese Weise wird ein schmalbandiges weißes Rauschspektrum des Rauschgenerators NOISE auf ein breitbandiges weißes Spektrum umgesetzt.

[0053] Es sei angenommen, dass der Audiogrundfrequenzparameter PITCH in Einheiten des schmalbandigen Abtastabstandes zugeführt wird. Es sei weiterhin angenommen, dass der Audiogrundfrequenzparameter PITCH in diesen Einheiten bis mindestens auf einen Bruchteil $1/N$ genau, das heißt hier bis auf $1/2$ genau bereitgestellt wird. Der nicht-ganzzahlige Audiogrundfrequenzparameter PITCH enthält Informationen über Frequenzanteile außerhalb des schmalbandigen Frequenzbereiches. Ein derartiger, nicht ganzzahliger Audiogrundfrequenzparameter PITCH wird häufig auch durch $\text{pitch} = p + p_{\text{frac}}/N$ dargestellt, wobei p und p_{frac} ganzzahlige Parameter mit $p_{\text{frac}} = 0, \dots, N-1$ sind. Da das adaptive Kodebuch ACB einen gegenüber dem schmalbandigen Abtastabstand halbierten, bzw. durch N geteilten, Abtastabstand verwendet, wird der Audiogrundfrequenzparameter PITCH zunächst mit N multipliziert. Das resultierende Produkt $\text{PITCH} \times N = p \times N + p_{\text{frac}}$ wird dann zur Auswahl der gespeicherten Abtastwertsequenz EXC_P - wie schon im Zusammenhang mit Figur 3 erläutert - verwendet.

[0054] Durch den in Figur 2b dargestellten Anregungssignalgenerator kann auf einfache Weise ein bandbreitenerweitertes Anregungssignal EXC erzeugt werden, dessen harmonische Feinstruktur im Erweiterungsband durch Nutzung des gebrochen-zahligen Anteils des Audiogrundfrequenzparameters PITCH besser modelliert werden kann. Insbesondere kann die harmonische Feinstruktur des Anregungssignals im schmalbandigen Frequenzbereich harmonisch und konsistent in das Erweiterungsband fortgesetzt werden.

[0055] In Figur 4. ist ein erfindungsgemäßer Audiosignaldekodierer zum Empfangen eines zu übertragenden Audiosignals schematisch dargestellt. Der Audiosignaldekodierer umfasst ein Audiosynthesefilter ASYN, das durch ein breitbandiges Anregungssignal S_{EXC} , z.B. im Frequenzbereich von 0 bis 8 kHz angeregt wird und durch Filterung ein synthetisches Audiosignal SAS generiert. Dem Audiosynthesefilter ASYN werden spektrale Parameter F_{ENV} , die eine spektrale Einhüllende des zu übertragenden Audiosignals spezifizieren, sowie Zeitverlauf-Parameter T_{ENV} , die eine zeitliche Einhüllende des Audiosignals spezifizieren, zugeführt. Das Audiosynthesefilter ASYN formt anhand der zugeführten Parameter F_{ENV} und T_{ENV} die spektrale und zeitliche Einhüllende des zu synthetisierenden Audiosignals SAS. Die Parameter F_{ENV} und T_{ENV} werden vom Sender des zu übertragenden Audiosignals zeitrahenweise ermittelt und zum Empfänger bzw. Audiosignaldekodierer übertragen.

[0056] Die Erzeugung des breitbandigen Anregungssignals S_{EXC} ist in unterschiedliche Schichten, nämlich eine Schicht für den schmalbandigen Frequenzbereich, hier von 0 bis 4 kHz, und eine Schicht für das Erweiterungsband, hier von 4 bis 8 kHz, aufgeteilt. Der Audiosignaldekodierer weist zur Erzeugung eines schmalbandigen Anregungssignals N_{EXC} , hier im Frequenzbereich von 0 bis 4 kHz, einen schmalbandigen Anregungssignalgenerator NBC und zur Erzeugung eines frequenzerweiterten Anregungssignals E_{EXC} , hier im Frequenzbereich von 4 bis 8 kHz, einen Anregungssignalgenerator EBC gemäß Figur 2b für das Erweiterungsband auf. Der schmalbandige Anregungssignalgenerator NBC kann wie der in Figur 2a dargestellte Anregungssignalgenerator oder wie ein herkömmlicher, mit adaptivem und fixem Kodebuch ausgestatteter Anregungssignalgenerator, z.B. gemäß G.729-Empfehlung, ausgestaltet sein.

[0057] Dem schmalbandigen Anregungssignalgenerator NBC werden zeitrahenweise jeweils der Audiogrundfrequenzparameter PITCH sowie die Intensitätsparameter G_N und G_P zugeführt. An Stelle der Intensitätsparameter G_N und G_P können auch ein Summenparameter $G_S + G_N$ und ein Verhältnisparameter G_S/G_N bzw. dessen Kehrwert zugeführt werden.

[0058] Der Audiogrundfrequenzparameter PITCH ist wie im Zusammenhang mit Figur 2b bereits erläutert nicht-ganzzahlig und gegeben durch $\text{pitch} = p + p_{\text{frac}}/N$. Der Bandbreitenerweiterungsfaktor N hat entsprechend dem Bandbreitenverhältnis zwischen dem breitbandigen Frequenzbereich von 0 bis 8 kHz und dem schmalbandigen Frequenzbereich von 0 bis 4 kHz beispielhaft einen Wert von $N=2$. Anhand der zugeführten Parameter PITCH, G_S und G_N erzeugt der schmalbandige Anregungssignalgenerator NBC das schmalbandige Anregungssignal N_{EXC} .

[0059] Dem gemäß Figur 2b ausgestalteten Anregungssignalgenerator EBC werden die vom schmalbandigen Anre-

gungssignalgenerator NBC verwendeten Parameter PITCH, G_S und G_N zugeleitet. Die Intensitätsparameter G_S und G_N werden ggf. durch eine vorgegebene Funktion umgesetzt, bevor sie in der Mischeinrichtung MIX des Anregungssignalgenerators EBC zur Pegelsteuerung verwendet werden. Die zugeleiteten Audiogrundfrequenzparameter PITCH werden, wie in Figur 2b veranschaulicht, mit N, hier N=2, multipliziert, um aus dem adaptiven Kodebuch des Anregungssignalgenerators EBC eine gespeicherte Anregungssignalsequenz auszuwählen. Anhand der zugeführten Parameter PITCH, G_S und G_N erzeugt der Anregungssignalgenerator EBC, wie bereits in Zusammenhang mit Figur 2b erläutert, das Anregungssignal EXC, das zunächst noch eine Bandbreite von 0 bis 8 kHz aufweist. Da der Anregungssignalgenerator EBC beim dargestellten Audiosignaldekodierer nur für das Erweiterungsband zuständig sein soll, wird das Anregungssignal EXC einem Hochpassfilter HP zugeführt. Dieser lässt im Wesentlichen nur Frequenzen des Erweiterungsbandes von 4 bis 8 kHz passieren und gibt ein frequenzerweitertes Anregungssignal E_EXC aus. Das frequenzerweiterte Anregungssignal E_EXC wird mit dem schmalbandigen Anregungssignal N_EXC, wie in Figur 4 durch ein Pluszeichen angedeutet, kombiniert, um das breitbandige Anregungssignal S_EXC zu bilden. Letzteres wird schließlich dem Audiosynthesefilter ASYN zugeführt.

[0060] Beim erfindungsgemäßen Audiosignaldekodierer werden zur Erzeugung des bandbreitenerweiterten Anregungssignals E_EXC und damit zur Erzeugung des breitbandigen Anregungssignals S_EXC nur die Audioparameter PITCH, G_S und G_N benötigt, die zur Erzeugung des schmalbandigen Anregungssignals ohnehin übertragen werden oder von einem schmalbandigen Anregungssignalgenerator bereitgestellt werden. Die Audioparameter PITCH, G_N und G_P können also vorteilhafterweise aus dem schmalbandigen Frequenzbereich des zu übertragenden Audiosignals oder aus Parametern eines schmalbandigen Codecs abgeleitet werden, um dann auf ein hinzuzufügendes Erweiterungsband angewandt zu werden. Zur Erzeugung des breitbandigen Anregungssignals S_EXC sind gegenüber einer Erzeugung des schmalbandigen Anregungssignals N_EXC keine zusätzlichen Audioparameter zu übertragen. Durch den Verzicht auf ein fixes Kodebuch in den Anregungssignalgeneratoren EBC und/oder NBC kann darüber hinaus eine zusätzliche Übertragung von Kodebuch-Indizes entfallen. Zusatzinformationen über eine Audiostruktur im Erweiterungsband können durch die Parameter F_ENV und T_ENV übertragen werden.

[0061] Der in Figur 4 dargestellte Audiosignaldekodierer kann zu einem Audiosignalenkodierer nach dem Analysis-by-Synthesis-Principle erweitert werden. Hierbei wird das synthetisierte Audiosignal SAS durch eine Vergleichseinrichtung mit dem zu enkodierenden Audiosignal verglichen und durch Variation der Audiosyntheseparameter PITCH, G_S, G_N, F_ENV und T_ENV angeglichen. Eine Kombination aus Audiosignaldekodierer und Audiosignalenkodierer wird häufig auch als Codec bezeichnet.

Patentansprüche

1. Verfahren zum Bilden eines Anregungssignals (EXC) als Folge von Anregungs-Abtastwerten zum Anregen eines Audiosynthesefilters (ASYN), wobei

- a) bereits gebildete der Anregungs-Abtastwerte zeitlich fortlaufend in einem adaptiven Kodebuch (ACB) gespeichert werden,
- b) durch einen Rauschgenerator (NOISE) fortlaufend Zufalls-Abtastwerte generiert werden,
- c) aus dem adaptiven Kodebuch (ACB) eine Sequenz (EXC_P) der gespeicherten Anregungs-Abtastwerte anhand eines zugeführten Audiogrundfrequenzparameters (PITCH) ausgewählt wird, durch den ein Zeitabstand der auszuwählenden Sequenz (EXC_P) zum aktuellen Zeitbezug vorgegeben wird, **dadurch gekennzeichnet, dass**, falls der Audiogrundfrequenzparameter (PITCH) einen Zeitabstand vorgibt, der kein ganzzahliges Vielfaches eines vorgegebenen Abtastabstandes eines separat zu generierenden, schmalbandigen Anregungssignals ist, zwischen die Anregungs-Abtastwerte und/oder zwischen die Zufalls-Abtastwerte abhängig vom Audiogrundfrequenzparameter (PITCH) Zwischen-Abtastwerte derart eingefügt werden, dass ein Abtastabstand der Abtastwerte geringer ist als der Abtastabstand des schmalbandigen Anregungssignals, wodurch das Anregungssignal (EXC) gegenüber dem schmalbandigen Anregungssignal zusätzliche Frequenzkomponenten eines Erweiterungsbandes aufweist, und
- d) das Anregungssignal (EXC) durch Mischen der ausgewählten Sequenz (EXC_P) mit einer aktuellen Zufalls-Abtastwerte des Rauschgenerators (NOISE) umfassenden Zufalls-Sequenz (EXC_N) gebildet wird.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet,**

dass die Anregungs-Abtastwerte und/oder die Zufalls-Abtastwerte zeitraumweise verarbeitet werden, und dass die Länge der ausgewählten Sequenz (EXC_P) und/oder die Länge der Zufalls-Sequenz (EXC_N) einer vorgegebenen Länge eines Zeitrahmens entsprechen.

3. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass beim Mischen die ausgewählte Sequenz (EXC_P) gemäß einem ersten Intensitätsparameter (G_P) und/oder
 die Zufalls-Sequenz (EXC_N) gemäß einem zweiten Intensitätsparameter (G_N) verstärkt werden.
4. Verfahren nach Anspruch 3,
dadurch gekennzeichnet,
dass das Anregungssignal (EXC) mit einem gegenüber einem separat zu generierenden, schmalbandigen Anregungssignal geringeren Abtastabstand gebildet wird, wodurch das Anregungssignal (EXC) gegenüber dem schmalbandigen Anregungssignal zusätzliche Frequenzkomponenten eines Erweiterungsbandes aufweist, und
dass der Audiogrundfrequenzparameter (PITCH) sowie der erste Intensitätsparameter (G_P) und/oder der zweite Intensitätsparameter (G_N) aus zur Generierung des schmalbandigen Anregungssignals vorgesehenen Audiosynthese-Parametern abgeleitet werden.
5. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass ein auf das Erweiterungsband entfallender Anteil des Anregungssignals (EXC) mit dem separat generierten, schmalbandigen Anregungssignal kombiniert wird, um ein breitbandiges Anregungssignal zur Anregung des Audiosynthesefilters (ASYN) zu erzeugen.
6. Verfahren nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass das Anregungssignal (EXC) durch ein LPC-Filter (LPC: Linear Predictive Coding) und/oder durch Formung einer spektralen und/oder zeitlichen Einhüllenden gefiltert wird.
7. Anregungssignalgenerator zum Bilden eines Anregungssignals (EXC) als Folge von Anregungs-Abtastwerten zum Anregen eines Audiosynthesefilters (ASYN), aufweisend
- a) ein adaptives Kodebuch (ACB) zum zeitlich fortlaufenden Speichern von bereits gebildeten Anregungs-Abtastwerten,
 - b) einen Rauschgenerator (NOISE) zum fortlaufenden Generieren von Zufalls-Abtastwerten,
 - c) eine Auswahlrichtung (SEL) zum Zuführen eines Audiogrundfrequenzparameters (PITCH) sowie zum Auswählen einer Sequenz (EXC_P) der gespeicherten Anregungs-Abtastwerte anhand des einen Zeitabstand der auszuwählenden Sequenz (EXC_P) zum aktuellen Zeitbezug vorgebenden Audiogrundfrequenzparameters (PITCH), **gekennzeichnet durch**
 - d) einen Interpolator (INT_N), der derart ausgebildet ist, dass, falls der Audiogrundfrequenzparameter (PITCH) einen Zeitabstand vorgibt, der kein ganzzahliges Vielfaches eines vorgegebenen Abtastabstandes eines separat zu generierenden, schmalbandigen Anregungssignals ist, zwischen die Anregungs-Abtastwerte und/oder zwischen die Zufalls-Abtastwerte abhängig vom Audiogrundfrequenzparameter (PITCH) Zwischen-Abtastwerte derart eingefügt werden, dass ein Abtastabstand der Abtastwerte geringer ist als der Abtastabstand des schmalbandigen Anregungssignals, wodurch das Anregungssignal (EXC) gegenüber dem schmalbandigen Anregungssignal zusätzliche Frequenzkomponenten eines Erweiterungsbandes aufweist, sowie
 - e) eine mit dem Rauschgenerator (NOISE) und dem adaptiven Kodebuch (ACB) gekoppelte Mischeinrichtung (MIX) zum Mischen der ausgewählten Sequenz (EXC_P) mit einer aktuellen Zufalls-Abtastwerte des Rauschgenerators (NOISE) umfassenden Zufalls-Sequenz (EXC_N) und zum Ausgeben des dadurch gebildeten Anregungssignals (EXC).
8. Anregungssignalgenerator nach Anspruch 7,
dadurch gekennzeichnet,
dass der Rauschgenerator (NOISE) mit der Mischeinrichtung (MIX) über einen Interpolator (INT_N) zum Einfügen von Zwischen-Abtastwerten zwischen die Zufalls-Abtastwerte gekoppelt ist.
9. Audiosignalenkodierer mit einem Anregungssignalgenerator nach Anspruch 7 oder 8 zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 6, einem durch den Anregungssignalgenerator anzuregenden Audiosynthesefilter sowie einer Vergleichseinrichtung zum Angleichen eines synthetischen, vom angeregten Audiosynthesefilter erzeugten Audiosignals an ein zu übertragendes Audiosignal.
10. Audiosignaldekodierer mit einem Anregungssignalgenerator nach Anspruch 7 oder 8 zur Durchführung des Ver-

fahrens nach einem der Ansprüche 1 bis 6 und einem durch den Anregungssignalgenerator anzuregenden Audio-synthesefilter zum Erzeugen eines synthetischen Audiosignals.

5 Claims

1. Method for forming an excitation signal (EXC) as a sequence of sampled excitation values for the excitation of an audio synthesis filter (ASYN), wherein

a) already formed sampled excitation values are continuously stored in a codebook (ACB),
 b) a noise generator (NOISE) continuously generates random sampled values,
 c) from the adaptive codebook (ACB) a sequence (EXC_P) of the stored sampled excitation values is selected on the basis of a fed audio fundamental frequency parameter (PITCH), by means of which a time gap between the sequence (EXC_P) that is to be selected and the actual time reference is predefined,

characterised in that

if the audio fundamental frequency parameter (PITCH) defines a time gap which is not a whole-figure multiple of a narrow-band excitation signal which is to be generated separately, then, between the sampled excitation values and/or between the random sampled values, depending on the audio fundamental frequency parameter (PITCH), intermediate sampled values are inserted in such a way that a sampling gap of the sampled values is smaller than the sampling gap of the narrow-band signal, as a result of which the excitation signal (EXC) exhibits additional frequency components of an extension band in relation to the narrow-band excitation signal, and

d) the excitation signal (EXC) is formed by mixing the selected sequence (EXC_P) with a random sequence (EXC_N) comprising actual random sampled value of the noise generator (NOISE).

2. Method according to claim 1,

characterised in that

the sampled excitation values and/or the random sampled values are processed time-frame by time-frame, and that the length of the selected sequence (EXC_P) and/or the length of the random sequence (EXC_N) correspond to the predefined length of the time-frame.

3. Method according to any one of the preceding claims, **characterised in that**, during mixing, the selected sequence (EXC_P) is amplified in accordance with a first intensity parameter (G_P) and/or the random sequence (EXC_N) in accordance with a second intensity parameter (G_N).

4. Method according to claim 3, **characterised in that**

the excitation signal (EXC) is formed with a smaller sampling gap in relation to a narrow-band excitation signal, to be generated separately, as a result of which the excitation signal (EXC) comprises additional frequency components of an extension band in relation to the narrow-band excitation signal, and

that the audio fundamental frequency parameter (PITCH) and the first intensity parameter (G_P) and/or the second intensity parameter (G_N) are derived from audio synthesis parameters provided for the generation of the narrow-band excitation signal.

5. Method according to any one of the preceding claims

characterised in that

a portion of the excitation signal (EXC) allotted to the extension band is combined with the separately generated narrow-band excitation signal, in order to produce a broad-band excitation signal to excite the audio synthesis filter (ASYN).

6. Method according to any one of the preceding claims

characterised in that

the excitation signal (EXC) is filtered through an LPC filter (LPC: Linear Predictive Coding) and/or by forming a spectral and/or temporal envelope.

7. Excitation signal generator for forming an excitation signals (EXC) as a sequence of sampled excitation values for exciting an audio synthesis filter (ASYN), comprising

a) an adaptive codebook (AGB) for the continuous storage of already formed sampled excitation values,

b) a noise generator (NOISE) for the continuous generation of random sampled values,
 c) a selector device (SEL) for feeding audio fundamental frequency parameters (PITCH) and for selecting a sequence (EXC_P) of the stored sampled excitation values on the basis of the audio fundamental frequency parameter (PITCH) predefining a time gap of the selected sequence (EXC_P) to the actual time reference,
 5 **characterised by**
 d) an interpolator (INT_N), which is formed in such a way that, if the audio fundamental frequency parameter (PITCH) predefines a time gap which is not a whole number multiple of a predefined time gap of a narrow-band excitation signal, to be generated separately, intermediate sampling values are inserted between the sampled excitation values and/or between the random sampled values, depending on the audio fundamental frequency
 10 parameter (PITCH), in such a way that a sampling gap of the sampled values is smaller than the sampling gap of the narrow-band excitation signal, as a result of which the excitation signal (EXC) comprises additional frequency component of an extension band in relation to the narrow-band excitation signal, and
 e) a mixing device (MIX) coupled to the noise generator (NOISE) and the adaptive codebook (ACB), for mixing the selected sequence (EXC_P) with a random sequence (EXC_N) comprising an actual random sampled value
 15 of the noise generators (NOISE), and for issuing the excitation signal (EXC) thereby formed.

8. Excitation signal generator according to claim 7, **characterised in that** the noise generator (NOISE) is coupled to the mixing device (MIX) via an interpolator (INT_N) for the insertion of intermediate sampled values between the random sampled values.

9. Audio signal encoder with an excitation signal generator according to claim 7 or 8, for carrying out the method according to one of claims 1 to 6, an audio synthesis filter to be excited by the excitation signal generator, and a comparator device for comparing a synthetic audio signal produced by the excited audio synthesis filter to an audio signal which is to be transferred.

10. Audio signal decoder with an excitation signal generator according to claim 7 or 8 for carrying out the method according to one of claims 1 to 6 and an audio synthesis filter to be excited by the excitation signal generator for the production of a synthetic audio signal.

Revendications

1. Procédé de formation d'un signal d'excitation (EXC) sous forme d'une suite de valeurs d'excitation échantillonnées en vue de l'excitation d'un filtre de synthèse audio (ASYN),

a) des valeurs déjà formées des valeurs d'excitation échantillonnées étant enregistrées continuellement dans le temps dans une table de codage adaptative (ACB),
 b) des valeurs échantillonnées aléatoires étant générées en continu par un générateur de bruit (NOISE),
 c) à partir de la table de codage adaptative (ACB), une séquence (EXC_P) des valeurs d'excitation échantillonnées enregistrées étant sélectionnée à l'aide d'un paramètre de fréquence fondamentale audio (PITCH) fourni par lequel un intervalle de temps entre la séquence à sélectionner (EXC_P) et la référence temporelle courante est prédéfini, **caractérisé en ce que**,
 dans le cas où le paramètre de fréquence fondamentale audio (PITCH) prédéfinit un intervalle de temps qui n'est pas un multiple entier d'un intervalle d'échantillonnage prédéfini d'un signal d'excitation à bande étroite à
 40 générer séparément, des valeurs échantillonnées intermédiaires sont introduites entre les valeurs d'excitation échantillonnées et/ou entre les valeurs échantillonnées aléatoires, en fonction du paramètre de fréquence fondamentale audio (PITCH), d'une manière telle qu'un intervalle d'échantillonnage des valeurs échantillonnées est inférieur à l'intervalle d'échantillonnage du signal d'excitation à bande étroite, moyennant quoi le signal d'excitation (EXC) comprend par rapport au signal d'excitation à bande étroite des composantes fréquentielles
 50 supplémentaires d'une bande d'extension, et
 d) **en ce que** le signal d'excitation (EXC) est formé par mélange de la séquence (EXC_P) sélectionnée avec une séquence aléatoire (EXC_N) comportant des valeurs échantillonnées aléatoires courantes du générateur de bruit (NOISE).

2. Procédé selon la revendication 1, **caractérisé en ce que** les valeurs d'excitation échantillonnées et/ou les valeurs échantillonnées aléatoires sont traitées selon un laps temporel, et
en ce que la longueur de la séquence sélectionnée (EXC_P) et/ou la longueur de la séquence aléatoire (EXC_N) correspondent à une longueur prédéfinie d'un laps de temps.

3. Procédé selon l'une quelconque des revendications précédentes, **caractérisé en ce que**, lors du mélange, la séquence sélectionnée (EXC_P) est amplifiée selon un premier paramètre d'intensité (G_P) et/ou la séquence aléatoire (EXC_N) est amplifiée selon un second paramètre d'intensité (G_N).

4. Procédé selon la revendication 3, **caractérisé en ce que** le signal d'excitation (EXC) est formé avec un intervalle d'échantillonnage plus petit par rapport à un signal d'excitation à bande étroite à générer séparément, moyennant quoi le signal d'excitation (EXC) comprend des composantes fréquentielles supplémentaires d'une bande d'extension par rapport au signal d'excitation à bande étroite, et **en ce que** le paramètre de fréquence fondamentale audio (PITCH) ainsi que le premier paramètre d'intensité (G_P) et/ou le second paramètre d'intensité (G_N) sont déduits de paramètres de synthèse audio prévus pour la génération du signal d'excitation à bande étroite.

5. Procédé selon l'une quelconque des revendications précédentes, **caractérisé en ce qu'**une partie du signal d'excitation (EXC) comprise dans la bande d'extension est combinée avec le signal d'excitation à bande étroite généré séparément afin de générer un signal d'excitation à large bande pour l'excitation du filtre de synthèse audio (ASYN).

6. Procédé selon l'une quelconque des revendications précédentes, **caractérisé en ce que** le signal d'excitation (EXC) est filtré par un filtre LPC (LPC : codage par prédiction linéaire) et/ou par formation d'enveloppantes spectrales et/ou temporelles.

7. Générateur de signaux d'excitation destiné à la formation d'un signal d'excitation (EXC) sous forme d'une suite de valeurs d'excitation échantillonnées en vue de l'excitation d'un filtre de synthèse audio (ASYN), comprenant

a) une table de codage adaptative (ACB) destinée à l'enregistrement continu dans le temps de valeurs d'excitation échantillonnées déjà formées,

b) un générateur de bruit (NOISE) servant à la génération en continu de valeurs échantillonnées aléatoires,

c) un dispositif de sélection (SEL) servant à la fourniture d'un paramètre de fréquence fondamentale audio (PITCH) ainsi qu'à la sélection d'une séquence (EXC_P) des valeurs d'excitation échantillonnées enregistrées à l'aide du paramètre de fréquence fondamentale audio (PITCH) prédéfinissant un intervalle de temps de la séquence (EXC_P) à sélectionner par rapport à la référence temporelle courante,

caractérisé par

d) un interpolateur (INT_N) qui est conçu de manière telle que, dans le cas où le paramètre de fréquence fondamentale audio (PITCH) prédéfinit un intervalle de temps qui n'est pas un multiple entier d'un intervalle d'échantillonnage prédéfini d'un signal d'excitation à bande étroite à générer séparément, des valeurs échantillonnées intermédiaires sont introduites entre les valeurs d'excitation échantillonnées et/ou entre les valeurs échantillonnées aléatoires, en fonction du paramètre de fréquence fondamentale audio (PITCH), d'une manière telle qu'un intervalle d'échantillonnage des valeurs échantillonnées soit inférieur à l'intervalle d'échantillonnage du signal d'excitation à bande étroite, moyennant quoi le signal d'excitation (EXC) comprend par rapport au signal d'excitation à bande étroite des composantes fréquentielles supplémentaires d'une bande d'extension, ainsi que

e) un dispositif de mélange (MIX) couplé au générateur de bruit (NOISE) et à la table de codage adaptative (ACB) et servant au mélange de la séquence sélectionnée (EXC_P) avec une séquence aléatoire (EXC_N) comportant des valeurs échantillonnées aléatoires du générateur de bruit (NOISE) et servant à la sortie du signal d'excitation. (EXC) formé de cette façon.

8. Générateur de signaux d'excitation selon la revendication 7, **caractérisé en ce que** le générateur de bruit (NOISE) est couplé au dispositif de mélange (MIX) par le biais d'un interpolateur (INT_N) en vue de l'introduction de valeurs échantillonnées intermédiaires entre les valeurs échantillonnées aléatoires.

9. Codeur de signaux audio équipé d'un générateur de signaux d'excitation selon la revendication 7 ou 8 en vue de l'exécution du procédé selon l'une quelconque des revendications 1 à 6, d'un filtre de synthèse audio à exciter par le générateur de signaux d'excitation ainsi que d'un dispositif de comparaison servant à l'égalesation d'un signal audio de synthèse, généré par le filtre de synthèse audio excité, avec un signal audio à transmettre.

10. Décodeur de signaux audio équipé d'un générateur de signaux d'excitation selon la revendication 7 ou 8 en vue de l'exécution du procédé selon l'une quelconque des revendications 1 à 6 et d'un filtre de synthèse audio à exciter par le générateur de signaux d'excitation en vue de la génération d'un signal audio de synthèse.

FIG 1

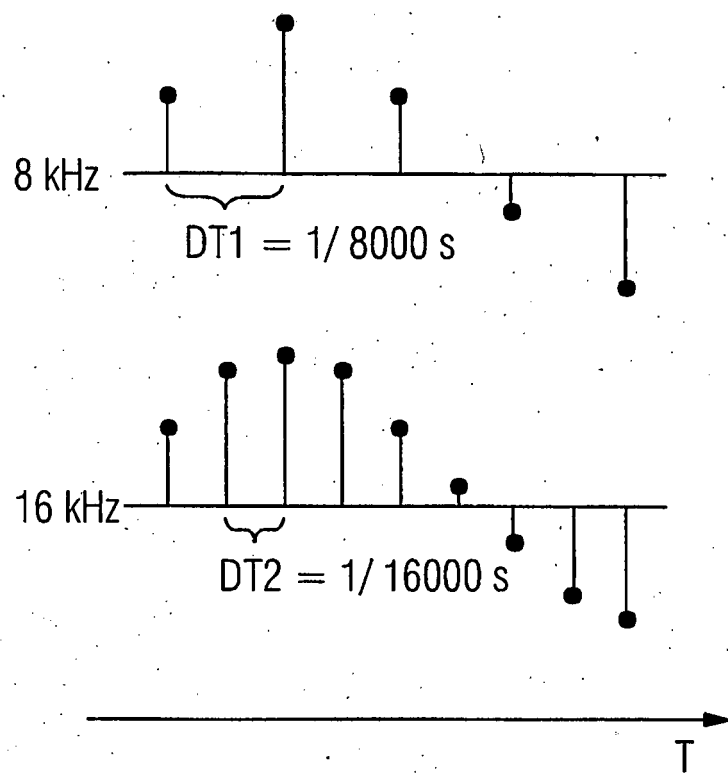


FIG 2a

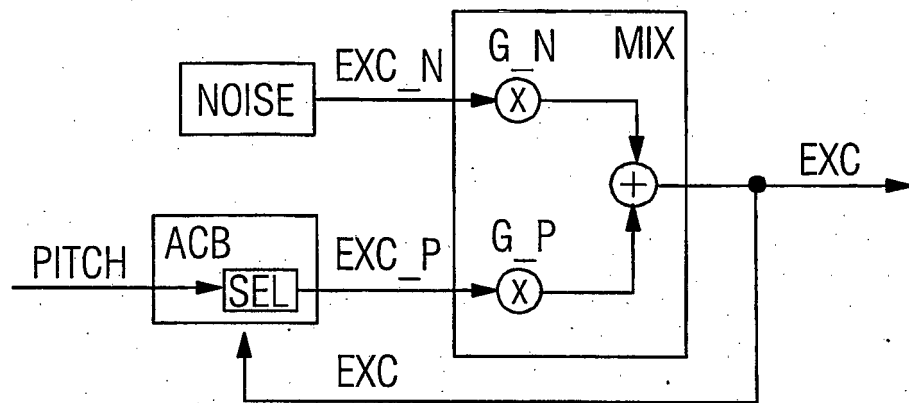


FIG 2b

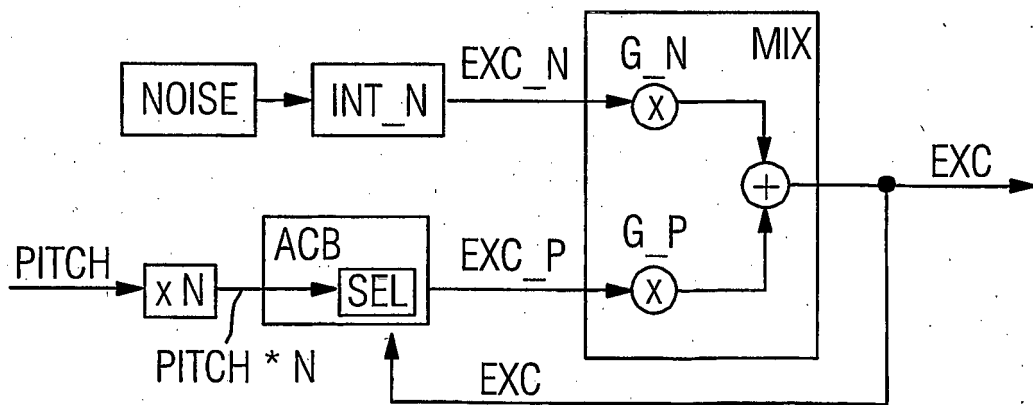


FIG 3

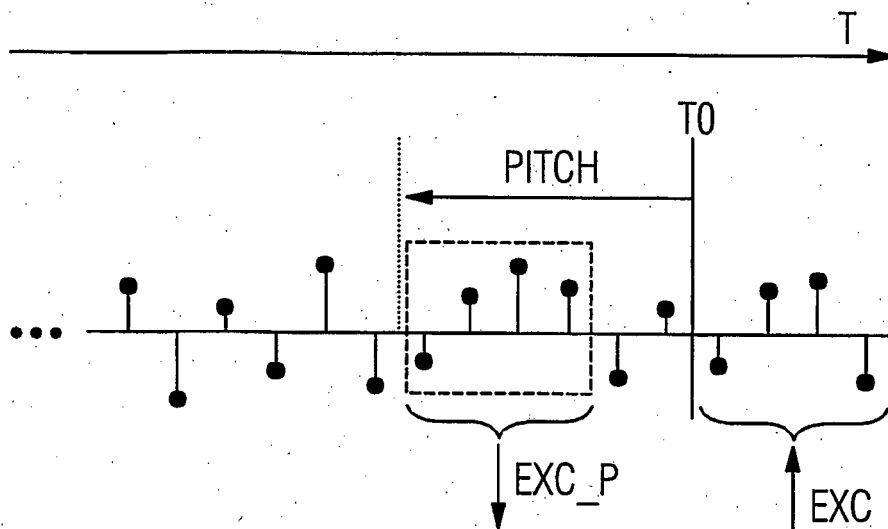
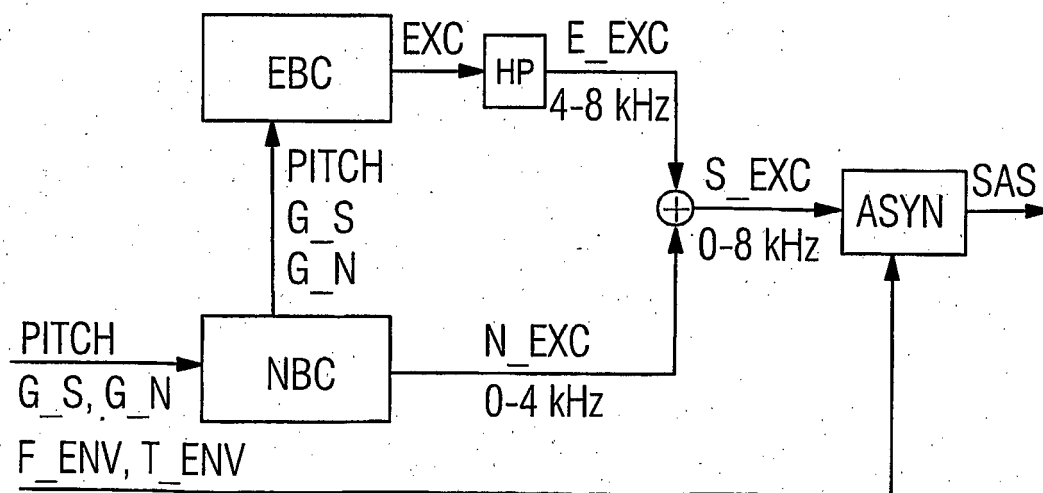


FIG 4



IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE

Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.

In der Beschreibung aufgeführte Patentdokumente

- EP 0883107 A1 [0008]