



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) DE 600 30 137 T2 2007.07.05

(12)

Übersetzung der europäischen Patentschrift

(97) EP 1 210 814 B1

(51) Int Cl.⁸: **H04M 9/08** (2006.01)

(21) Deutsches Aktenzeichen: **600 30 137.0**

(86) PCT-Aktenzeichen: **PCT/EP00/04666**

(96) Europäisches Aktenzeichen: **00 938 672.3**

(87) PCT-Veröffentlichungs-Nr.: **WO 2000/074362**

(86) PCT-Anmeldetag: **23.05.2000**

(87) Veröffentlichungstag

der PCT-Anmeldung: **07.12.2000**

(97) Erstveröffentlichung durch das EPA: **05.06.2002**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **16.08.2006**

(47) Veröffentlichungstag im Patentblatt: **05.07.2007**

(30) Unionspriorität:

320659 27.05.1999 US

(84) Benannte Vertragsstaaten:

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

(73) Patentinhaber:

**Telefonaktiebolaget LM Ericsson (publ),
Stockholm, SE**

(72) Erfinder:

RASMUSSON, Jim, S-235 41 Vellinge, SE

(74) Vertreter:

HOFFMANN & EITLE, 81925 München

(54) Bezeichnung: **VERFAHREN UND VORRICHTUNG ZUR LEISTUNGSVERBESSERUNG EINES ADAPTIVEN FILTERS MITTELS NICHT HÖRBARER INFORMATION**

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 eingelebt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

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

Beschreibung**Gebiet der Erfindung**

[0001] Die vorliegende Erfindung bezieht sich auf Kommunikationssysteme und genauer gesagt auf das adaptive Filtern von Audiosignalen in Kommunikationssystemen.

Hintergrund der Erfindung

[0002] In heutigen Kommunikationssystemen überwiegen adaptive Filteranordnungen. Solche Anordnungen werden typischerweise verwendet, um unerwünschte Signalkomponenten zu reduzieren oder zu entfernen und/oder gewünschte Signalkomponenten zu kontrollieren oder zu verstärken.

[0003] Ein übliches Beispiel einer solchen Filteranordnung bezieht sich auf Freisprechtelefone, bei denen ein eingebauter Kopfhörer und ein Mikrofon eines konventionellen Telefonhörers durch einen externen Lautsprecher bzw. ein externes Mikrofon ersetzt sind, so dass der Telefonbenutzer kommunizieren kann, ohne die Telefoneinheit physisch in der Hand zu halten. Da aus dem externen Lautsprecher dringendes Geräusch vom externen Mikrofon aufgenommen werden kann, wird üblicherweise eine adaptive Filterung durchgeführt, um zu verhindern, dass die Lautsprecherausgabe zurück echot und den entfernten Verwender am anderen Ende der Konversation stört. Diese Art adaptiver Filterung oder Echo-Unterdrückung ist ein grundlegendes Merkmal von Voll duplex-Freisprech-Kommunikationsvorrichtungen von heute geworden.

[0004] EP 0,708,536 offenbart einen bekannten Echo-Unterdrücker zum Trainieren einer Echo-Pfadabschätzung ohne für Sprache ein Hindernis darzustellen.

[0005] Typischerweise wird eine Echo-Unterdrückung erzielt, indem das Lautsprechersignal durch einen adaptiven Finite Impulse Reaktions (FIR)-Filter geleitet wird, der den akustischen Echopfad zwischen dem Freisprech-Lautsprecher und dem Freisprech-Mikrofon annähert oder modelliert (z. B. ein Passagier-Innenraum bei einer Automobil-Freisprech-Telefonanwendung). Der FIR-Filter stellt daher eine Echoabschätzung bereit, die von dem Mikrofon-Ausgangssignal vor der Übertragung an den entfernten Nutzer beseitigt werden kann. Die Filterungs-Charakteristik (das heißt, der Satz von FIR-Koeffizienten) des adaptiven FIR-Filters wird dynamisch und kontinuierlich eingestellt, basierend sowohl auf der Lautsprechereingabe als auch der echo-unterdrückten Mikrofonausgabe, um eine enge Annäherung an den Echopfad bereitzustellen und Änderungen im Echopfad zu verfolgen (z. B. wenn ein hiesiger Nutzer eines Automobil-Freisprech-Telefons in-

nerhalb des Passagierraums seine Position verschiebt).

[0006] Die Einstellung der Filter-Charakteristik wird üblicherweise unter Verwendung einer Art des wohlbekannten Adoptions-Algorithmus des kleinsten mittleren Quadrats (LMS, Least Mean Square) erreicht, der von Widrow und Hoff 1960 entwickelt worden ist. Der LMS-Algorithmus ist das kleinste quadratisch stochastische Gradienten-Schrittverfahren, welches, da es sowohl effizient als auch robust ist, oft in vielen Echtzeitanwendungen verwendet wird. Der LMS-Algorithmus und seine bekannten Variationen (z. B. der normalisierte LMS oder NLMS-Algorithmus) haben jedoch gewisse Nachteile. Beispielsweise kann der LMS-Algorithmus manchmal etwas langsam beim Konvergieren (das heißt beim Annähern der Ziel-Filter-Charakteristik, wie etwa dem akustischen Echopfad in einer Freisprech-Telefonanwendung) sein, insbesondere wenn der Algorithmus basierend auf nicht-weißem oder getöntem Eingangssignal adaptiert oder trainiert wird.

[0007] Eine langsame LMS-Adaptation ist ein besonderes Problem beim Freisprech-Telefonkontext, insoweit als das Trainingssignal (das heißt das Lautsprechersignal) menschliche Sprache enthält, die nur einen relativ kleinen Teil des gesamten möglichen Signalraums anregt und die langsam abnehmende Autokorrelations-Eigenschaften aufweist, insbesondere in Bezug auf stimmhafte (das heißt Vokal-) Geräusche. Darüber hinaus kann der Nah-End-Hintergrundlärm (z. B. Automobil-Innenraum und Straßenlärm) den LMS-Adoptionsprozess stören und weiter verlangsamen. Folglich gibt es einen Bedarf an verbesserten adaptiven Filterungstechniken bei der Freisprech-Telefonie und bei anderen Kontexten.

Zusammenfassung der Erfindung

[0008] Die vorliegende Erfindung erfüllt die oben beschriebenen und anderen Notwendigkeiten durch Bereitstellen von schnell konvergierenden adaptiven Filter-Algorithmen. Gemäß exemplarischen Ausführungsformen wird die Konvergenz-Geschwindigkeit eines konventionellen adaptiven Algorithmus durch Kombinieren eines adaptionsverstärkenden Signals mit einem konventionellen Algorithmen-Trainingssignal verbessert, um ein verstärktes oder optimiertes Trainingssignal bereitzustellen. Das adaptionsverstärkende Signal wird so ausgewählt, dass es Eigenschaften (z. B. spektrale Weißheit und rasch abnehmender Autokorrelation zwischen Proben) aufweist, die dem Adoptions-Algorithmus gestatten, rascher zu konvergieren. Vorteilhafterweise kann das adaptionsverstärkende Signal mit dem konventionellen Trainingssignal so kombiniert werden, dass das verbesserte Trainingssignal nicht wahrnehmbar anders als das konventionelle Trainingssignal ist. Beispielsweise kann im Freisprech-Telefonkontext das adaptions-

verbessernde Signal so maßgeschneidert werden, dass es vom Lautsprechersignal (das heißt dem konventionellen Trainingssignal) maskiert ist und somit für den Telefonnutzer unhörbar ist.

[0009] Eine beispielhafte Signalverarbeitungsvorrichtung gemäß der Erfindung beinhaltet einen adaptiven Filter, der dafür konfiguriert ist, ein Eingangssignal zu filtern und dadurch ein gefiltertes Ausgangssignal bereitzustellen, wobei eine Filtereigenschaft des adaptiven Filters dynamisch, basierend auf dem gefilterten Ausgabesignal und auf einem Trainingssignal, justiert wird. Die beispielhafte Signalverarbeitungsvorrichtung enthält auch einen Adoptionsverstärkungs-Prozessor, der ein adaptionsverstärkendes Signal mit einem Referenzsignal kombiniert, um das Trainingssignal bereitzustellen, wobei das adaptionsverstärkende Signal basierend auf dem Referenzsignal dynamisch eingestellt wird. Vorteilhaftweise kann das adaptionsverstärkende Signal eine unhörbare Komponente des Trainingssignals sein, die vom Referenzsignal maskiert ist. Beispielsweise kann das adaptionsverstärkende Signal durch Filtern einer weißen Pseudo-Rauschen-Sequenz oder einer reproduzierbaren Maximallängen-Sequenz unter Verwendung der Frequenzmaske des Referenzsignals erzeugt werden. Alternativ kann das adaptionsverstärkende Signal erzeugt werden, indem zuerst Audiocodierung und dann Audiodecodierung des Referenzsignales durchgeführt wird.

[0010] Ein beispielhaftes Verfahren gemäß der Erfindung beinhaltet die Schritte des Berechnens einer Frequenzmaske eines Referenzsignals; Berechnens eines adaptionsverstärkenden Signals basierend auf der Frequenzmaske des Referenzsignals; Berechnens eines Trainingssignals basierend auf dem Referenzsignal und dem adaptionsverstärkenden Signal; und Einstellens einer Filter-Charakteristik des adaptiven Filters, basierend auf dem Trainingssignal. Wie oben kann das adaptionsverstärkende Signal eine unhörbare Komponente des Trainingssignals sein, die vom Referenzsignal maskiert ist. Beispielsweise kann der Schritt des Berechnens des adaptionsverstärkenden Signals den Schritt des Filterns einer weißen Pseudo-Rausch-Sequenz oder einer Maximallängen-Sequenz unter Verwendung der Frequenzmaske des Referenzsignals beinhalten. Alternativ kann der Schritt des Berechnens des adaptionsverstärkenden Signales den Schritt der Audiocodierung des Referenzsignals und dann der Audiodecodierung des Referenzsignals beinhalten.

[0011] Die oben beschriebenen und anderen Merkmale und Vorteile der Erfindung werden nachstehend detailliert und unter Bezugnahme auf die illustrativen Beispiele, die in den beigefügten Zeichnungen gezeigt werden, erläutert. Fachleute werden erkennen, dass die beschriebenen Ausführungsformen für illustrative Zwecke und zum Verständnis bereitgestellt

werden und dass verschiedene äquivalente Ausführungsformen hier erwogen werden.

Kurze Beschreibung der Zeichnungen

[0012] [Fig. 1](#) ist ein Blockdiagramm eines beispielhaften Freisprech-Telefonsystems, in dem adaptionsverstärkende Techniken der Erfindung implementiert sind.

[0013] [Fig. 2](#) ist ein Spektralplot, der eine beispielhafte sofortige Frequenzmaske entsprechend einem einzelnen hörbaren Ton darstellt.

[0014] [Fig. 3](#) ist ein Blockdiagramm eines beispielhaften Adoptionsverstärkungs-Prozessors gemäß der Erfindung.

[0015] [Fig. 4](#) ist ein Blockdiagramm eines alternativen Adoptionsverstärkungs-Prozessors gemäß der Erfindung.

Detaillierte Beschreibung der Erfindung

[0016] [Fig. 1](#) zeigt ein beispielhaftes Freisprech-Telefonsystem **100**, das Adoptionsverstärkungstechniken gemäß der Erfindung beinhaltet. Wie gezeigt, beinhaltet das beispielhafte System **100** ein Mikrofon **110**, einen Lautsprecher **120**, eine Summievorrichtung **130**, einen finiten Impuls-Reaktions (FIR)-Filter **140**, einen Kleinste-mittlere-Quadrat (LMS)-Kreuzkorrelator **150** und einen Adoptionsverstärkungs-Prozessor **160**. Fachleute werden erkennen, dass die unten beschriebene Funktionalität der in [Fig. 1](#) dargestellten Komponenten unter Verwendung bekannter digitaler Signalverarbeitungs-Hardware und/oder eines Universal-Digital-Computers implementiert werden kann. Fachleute werden auch erkennen, dass in der Praxis das beispielhafte System **100** Komponenten (z. B. einen Analog/Digital-Wandler am Ausgang des Mikrofons **110** und einen Digital/Analog-Wandler am Eingang des Lautsprechers **120**) beinhaltet, die in [Fig. 1](#) weggelassen sind, da sie für ein Verständnis der vorliegenden Erfindung nicht kritisch sind.

[0017] Im Betrieb wird ein entferntes Audiosignal einschließlich der Sprache eines entfernten Anwenders (nicht gezeigt) am Lautsprecher **120** für die Präsentation an den Nah-End-Anwender (ebenfalls nicht gezeigt) eingegeben. Die Lautsprecherausgabe wird dann über einen unbekannten und sich manchmal ändernden Echopfad zurück zum Mikrofon **110** geschot, wie durch die variable Übertragungsfunktion $H(z)$ in [Fig. 1](#) angezeigt ist. Somit beinhaltet die Audioausgabe aus dem Mikrofon **110** das Lautsprecher-Echo, wie auch die Nah-End-Anwendersprache und Nah-End-Hintergrundgeräusche. Um zu verhindern, dass das Lautsprecher-Echo den entfernten Anwender erreicht und stört, filtert der FIR-Filter **140**

das Lautsprechersignal, um eine Abschätzung des am Mikrofon **110** empfangenen Lautsprecher-Echos bereitzustellen und die resultierende Echo-Abschätzung wird von der Mikrofonausgabe über die Summierzvorrichtung **130** subtrahiert. Die echo-unterdrückte Ausgabe aus der Summierzvorrichtung **130** wird dann an den entfernten Anwender übertragen und zum LMS-Kreuzkorrelator **150** zur Verwendung beim Adaptieren der Filter-Koeffizienten oder Abgriffe des FIR-Filters **140** zurückgekoppelt, so dass sie zum echten Echopfad $H(z)$ konvergieren und ihm folgen. Wie im Stand der Technik bekannt, berechnet der LMS-Kreuzkorrelator **150** Filter-Koeffizienten-Aktualisierungen basierend sowohl auf der echo-unterdrückten Ausgabe, oder dem Fehlersignal, als auch dem Lautsprecher-Eingangs- oder -Trainingssignal.

[0018] Bei konventionellen Systemen wird das entfernte Audiosignal direkt zum Lautsprecher zur Präsentation an dem Nah-End-Anwender gereicht und eine digitalisierte Version des entfernten Audiosignals wird direkt als Trainings- oder Referenzsignal zum Entwickeln der Echoabschätzung verwendet (das heißt das digitale entfernte Audio wird direkt sowohl am FIR-Filter **140** als auch dem LMS-Kreuzkorrelator **150** eingegeben). Folglich können die Filter-Koeffizienten des FIR-Filters **140** langsam damit sein, zum wahren Echopfad $H(z)$ zu konvergieren und ihm zu folgen, wie oben beschrieben. Mit anderen Worten ist, da die Sprache des entfernten Anwenders typischerweise langsam abnehmende Autokorrelations-Eigenschaften aufweist und nur einen relativ kleinen Teil des gesamten möglichen Signalraums anregt, das entfernte Audiosignal nicht ideal zur Verwendung beim Trainieren eines LMS-basierten Algorithmus.

[0019] Gemäß der vorliegenden Erfindung wird jedoch das entfernte Audiosignal nicht direkt als ein Trainingssignal für den LMS-Prozess verwendet. Stattdessen verwendet der adaptionsverstärkende Prozessor **160** das entfernte Audiosignal als eine Referenz beim Entwickeln eines adaptionsverstärkenden Signals (das heißt eines Signals, welches Eigenschaften aufweist, die zum Trainieren eines adaptiven Algorithmus geeigneter sind), das mit dem entfernten Audio-Referenzsignal kombiniert wird, um ein optimales, adaptionsverstärkendes Trainingssignal bereitzustellen. Vorteilhafterweise können Unterschiede zwischen dem optimierten Trainingssignal und dem entfernten Audio-Referenzsignal (das heißt dem konventionellen Trainingssignal) für das menschliche Ohr unwahrnehmbar gemacht werden. Genauer gesagt, kann das Adoptionsverstärkungssignal so maßgeschneidert werden, dass es vom entfernten Audiosignal maskiert wird und somit für den Nah-End-Anwender unhörbar ist. Als Ergebnis wird eine Adoptionsverstärkung erzielt, ohne die Systemleistung aus der Sicht des Nah-End-Anwenders zu beeinträchtigen.

[0020] Um das Adoptionsverstärkungssignal für den Nah-End-Anwender, der einem entfernten Audiosignal zuhört, unwahrnehmbar zu machen, werden bekannte Maskierungseigenschaften des menschlichen Hör-Mechanismus eingesetzt. Speziell wird eine dynamische Frequenzmaske des entfernten Audiosignals in Echtzeit berechnet und das Adoptionsverstärkungssignal wird dynamisch übereinstimmend mit der Frequenzmaske eingestellt, so dass das Adoptionsverstärkungssignal unhörbar bleibt. Anders ausgedrückt, wird das Adoptionsverstärkungssignal kontinuierlich maßgeschneidert, so dass es effektiv durch das entfernte Audiosignal verborgen wird.

[0021] Das Maskieren eines Signals durch ein anderes ist in [Fig. 2](#) dargestellt. In der Figur ist eine verzögerungsfreie Frequenzmaske **210** eines Einzeltons **230** gezeigt. Die verzögerungsfreie Frequenzmaske **210** definiert einen verzögerungsfreien Bereich **220**, der vom Einzelton **230** maskiert (das heißt unhörbar gemacht) ist. Mit anderen Worten kann kein Signal, das einen Spektralinhalt vollständig innerhalb des Bereichs **220** aufweist, von einem dem Ton **230** zuhörenden menschlichen Ohr wahrgenommen werden. In der Praxis wird die Frequenzmaske eines Audiosignals dynamisch und kontinuierlich verändert, und Fachleute werden erkennen, dass die verzögerungsfreie Maske **210** nur für einen festen Zeitpunkt gilt.

[0022] Die Maskereigenschaften des menschlichen Hör-Mechanismus werden beispielsweise in Zwicker and Fastl, "Psychoacoustics, Facts and Models", Springer, Heidelberg, 1990 beschrieben. Zudem sind Algorithmen zum dynamischen Berechnen der Frequenzmaske eines Audiosignals in Echtzeit im Detail in den Spezifikationen beschrieben, die von der bekannten Moving Pictures Experts Group (MPEG) herausgegeben werden. Man siehe beispielsweise K. Brandenburg und Marini Bosi, "Overview of MPEG-Audio: Current and future standards for low bit-rate audio coding", 99. AES, New York, 6.-9. Oktober 1995, Vorabdruck #4130 (29 Seiten) und Nikil Jayant, James Johnston und Robert Safranek, "Signal Compression Based on Models of Human Perception", Proceedings of the IEEE, Oktober 1993, Band 81, Nummer 10, Seiten 1385-1421, die alle unter Bezugnahme hier inkorporiert sind.

[0023] Solche Algorithmen werden beispielsweise bei MPEG-Audiocodierern verwendet, um die Bit-Rate zu reduzieren, die benötigt wird, um ein interessierendes Signal zu repräsentieren. Mit anderen Worten werden die Frequenzmasken-Algorithmen verwendet, um zu bestimmen, welche Anteile eines Signals entfernt werden können, ohne das Signal hörbar zu stören. Im Gegensatz dazu werden dynamische Frequenzmasken-Algorithmen im Kontext der vorliegenden Erfindung verwendet, um eine Frequenzverteilung des Adoptionsverstärkungssignals zu etablieren.

ren. Anders ausgedrückt wird, bei gegebener dynamischer Frequenzmaske des entfernten Audiosignals, das Adoptionsverstärkungssignal in Echtzeit so konstruiert, dass es einen Spektralinhalt aufweist, der vollständig innerhalb des unhörbaren Bereichs liegt (das heißt innerhalb des Bereichs, der vom entfernten Audiosignal maskiert wird). Beispielsweise kann das Adoptionsverstärkungssignal so eingestellt werden, dass es eine Spektralverteilung aufweist, die sich der der dynamisch berechneten Frequenzmaske annähert.

[0024] Die Erzeugung des Adoptionsverstärkungssignals anhand der Frequenzmaske des entfernten Audiosignals kann in einer Reihe von Weisen bewerkstelligt werden. Beispielsweise kann ein Modellsignal mit Eigenschaften, die eine rasche adaptive Algorithmus-Konvergenz (z. B. spektrale Weißheit und rasches Abklingen oder nichtexistierende Auto-korrelation zwischen Proben) fördern, unter Verwendung der entfernten Audio-Frequenzmaske gefiltert werden, um das Adoptionsverstärkungssignal bereitzustellen.

[0025] Beispielhafte Modellsignale enthalten weiße Pesudo-Rausch(PN)-Signale und reproduzierbare Maximallängen-Sequenz (MLS)-Signale. Verfahren zum Erzeugen solcher Signale in Echtzeit sind bekannt. Weißes Rauschen und MLS-Signale werden oft beispielsweise beim Echtzeittesten von Frequenz- und Raumreaktionen für High-Fidelity-Audio-Ausstattung verwendet. Vorteilhafterweise, wenn das sich ergebende Adoptionsverstärkungssignal mit dem entfernten Audiosignal kombiniert (z. B. dazu-ad-diert) wird, veranlasst das resultierende optimierte Trainingssignal die Filterungseigenschaft des FIR-Filters **140** dazu, rascher zu konvergieren, während der Nah-End-Anwender sich nicht bewusst wird, dass das Adoptionsverstärkungssignal vorhanden ist.

[0026] [Fig. 3](#) stellt einen beispielhaften Prozessor **300** zum Erzeugen eines verbesserten adaptiven Algorithmus-Trainingssignals in der oben beschriebenen Art und Weise dar. Der beispielhafte Prozessor **300** kann beispielsweise verwendet werden, um den Adoptionsverstärkungs-Prozessor **160** von [Fig. 1](#) zu implementieren. Wie gezeigt, beinhaltet der beispielhafte Prozessor **300** einen Frequenzmasken-Berechnungsprozessor **310**, einen Pseudo-Rauschen-Generator **320**, einen Frequenzmaskenfilter **330** und eine Summivorrichtung **340**. Fachleute werden erkennen, dass die unten beschriebene Funktionalität der Komponenten von [Fig. 3](#) unter Verwendung jeglicher aus einer Vielzahl von bekannten Hardware-Konfigurationen implementiert werden kann, einschließlich standardmäßiger digitaler Signalverarbeitungskomponenten, einem Universal-Computer und/oder einer oder mehrerer applikations-spezifischer integrierter Schaltungen (ASICs).

[0027] In [Fig. 3](#) wird ein Referenzsignal (z. B. ein entferntes Audiosignal) mit einem Eingang des Frequenzmasken-Berechnungsprozessors **310** und mit einem ersten additiven Eingang der Summivorrichtung **340** gekoppelt. Zusätzlich wird ein Ausgang des Frequenzmasken-Berechnungsprozessors **310** mit einem Kontrolleingang des Frequenzmaskenfilters **330** gekoppelt und ein Ausgang des Pseudo-Rauschen-Generators **320** wird mit einem Audioeingang des Frequenzmaskenfilters **330** gekoppelt. Weiterhin wird ein Audioausgang des Frequenzmaskenfilters **330** mit einem zweiten additiven Eingang der Summivorrichtung **340** gekoppelt und ein Ausgang der Summivorrichtung **330** dient als adaptionsverstärkendes Trainingssignal (z. B. zur Eingabe in den adaptiven Filter **140** und den LMS-Kreuzkorrelator **150** von [Fig. 1](#)).

[0028] Im Betrieb berechnet der Frequenzmasken-Berechnungsprozessor **310** die dynamische Frequenzmaske des entfernten Audiosignals (z. B. einmal für jeden Block an Proben des entfernten Audiosignals). Wie oben beschrieben, wird die Frequenzmaske unter Verwendung jeglichen geeigneten Algorithmus berechnet, wie etwa den in den oben inkorporierten MPEG-Dokumenten beschriebenen Algorithmen. Die sich ergebende Frequenzmaske wird dann verwendet, um die Filterübertragungsfunktion des Frequenzmaskenfilters **330** zu aktualisieren.

[0029] Zum gleichen Zeitpunkt stellt der Generator **320** ein Modellsignal (z. B. eine Pseudo-Weiß-Rauschen-Sequenz oder eine Maximallängen-Sequenz) dem Audioeingang des Frequenzmaskenfilters **330** bereit und der Filter **330** bearbeitet das Modellsignal, um ein gefiltertes modelliertes Signal bereitzustellen, das dem entfernten Audiosignal hinzugefügt wird (über die Summivorrichtung **340**), um das verbesserte adaptive Algorithmen-Referenzsignal bereitzustellen. Wie oben beschrieben, fördert der addierte Signalinhalt die rasche Konvergenz des echo-unterdrückenden adaptiven Filters (z. B. Filter **130** von [Fig. 1](#)) und ist für den Nah-End-Anwender unhörbar.

[0030] Fachleute werden erkennen, dass die vom Frequenzmaskenfilter **330** durchgeführte Filterung entweder in der Frequenz- oder in der Zeitdomäne ausgeführt werden kann. Bei Anwendungen, bei denen eine Frequenzdomänenfilterung (das heißt Filterung durch Vektor-Multiplikation eines Satzes von Frequenzdomänen-Koeffizienten des Filters **330** mit einer Frequenzdomänen-Repräsentation eines Probenblocks des Modellsignals) bevorzugt wird, kann der Filter **330** aktualisiert werden, indem periodisch (das heißt einmal für jeden Probenblock) die Spektralwerte, die sich aus der Frequenzmaskenberechnung ergeben, direkt zum Filter **330** kopiert werden. Alternativ kann bei Anwendungen, bei denen eine Zeitdomänenfilterung (das heißt, Filterung durch probeweise Konvolution von Proben des Modellsignals

mit Zeitdomänen-Koeffizienten des Frequenzmaskenfilters **330**) bevorzugt wird, der Filter **330** durch Umwandeln der sich aus der Frequenzmaskenberechnung der Zeitdomäne (z. B. unter Verwendung einer Inversen Fast Fourier Transformation oder IFFT) ergebenden Spektralwerte und dann Kopieren der sich ergebenden Zeitdomänen-Koeffizienten zum Filter **300** aktualisiert werden. Die Auswahl zwischen Zeitdomänen und Frequenzdomänenfilterung ist eine Frage der Designauswahl.

[0031] [Fig. 4](#) stellt ein alternatives System **400** zum Erzeugen des Adoptionsverstärkungssignals anhand der Frequenzmaske des entfernten Audiosignals dar. Wie der Prozessor **300** von [Fig. 3](#) kann der alternative Prozessor **400** beispielsweise verwendet werden, um den Adoptionsverstärkungsprozessor **160** von [Fig. 1](#) zu implementieren. Wie gezeigt, beinhaltet der alternative Prozessor **400** einen MPEG-Audio-Codierer **410** und einen MPEG-Audio-Decodierer **420**. Fachleute werden erkennen, dass die unten beschriebene Funktionalität der Komponenten von [Fig. 4](#) unter Verwendung jeglicher aus einer Vielzahl von bekannten Hardware-Konfigurationen implementiert werden kann, einschließlich standardisierter digitaler Signalverarbeitungs-Komponenten, einem Universal-Digital-Computer und/oder einem oder mehreren applikationsspezifischen integrierten Schaltung (ASICs).

[0032] In [Fig. 4](#) wird ein Referenzsignal (z. B. das entfernte Audiosignal) mit einem Eingang des Audio-Codierers **410** gekoppelt und ein Ausgang des Audio-Codierers **410** ist mit einem Eingang des Audio-Decodierers **420** gekoppelt. Ein Ausgang des Audio-Decodierers **420** dient als das adaptionsverstärkende Trainingssignal (z. B. zur Eingabe an den adaptiven Filter **140** und den LMS-Kreuzkorrelator **150** von [Fig. 1](#)).

[0033] Im Betrieb codiert der Audio-Codierer **410** das entfernte Audiosignal (z. B. unter Verwendung eines MPEG-Codierungs-Algorithmus, wie in den oben inkorporierten MPEG-Dokumenten beschrieben) und der Audio-Decodierer **420** decodiert das resultierende Signal unmittelbar (z. B. unter Verwendung eines komplementären MPEG-Decodier-Algorithmus, wie in den oben inkorporierten MPEG-Artikeln beschrieben), um das adaptionsverstärkende Referenzsignal bereitzustellen. Da der MPEG-Audio-Codierer unhörbare Informationen im entfernten Audiosignal unter Verwendung sehr weniger Bits codiert, wird die unhörbare Information schwer quantisiert und es wird unhörbares Quantisierungs-Rauschen nach Wunsch eingeführt. Wie beim Prozessor **300** von [Fig. 3](#) fördert der im entfernten Audiosignal in [Fig. 4](#) eingeführte Signalthalt rasche Konvergenz des echo-unterdrückenden adaptiven Filters (z. B. Filter **130** von [Fig. 1](#)) und ist für Nah-End-Anwender unhörbar.

[0034] Allgemein stellt die vorliegende Erfindung Verfahren und Vorrichtungen zum Verbessern der Konvergenzgeschwindigkeit eines adaptiven Filterungs-Algorithmus bereit. Gemäß beispielhafter Ausführungsformen wird die Konvergenzgeschwindigkeit eines adaptiven Algorithmus durch Kombinieren eines adaptionsverstärkenden Signals mit einem konventionellen Algorithmus-Trainingssignal verbessert, um ein verstärktes oder optimiertes Trainingssignal bereitzustellen. Das adaptionsverstärkende Signal ist so ausgelegt, dass es Eigenschaften (z. B. spektrale Weißheit und rasch abklingende Autokorrelation zwischen Proben) aufweist, die dem Adaptions-Algorithmus gestatten, rascher zu konvergieren. Vorteilhaftweise kann das Adaptionsverstärkungssignal mit dem konventionellen Trainingssignal so kombiniert werden, dass das verstärkte Trainingssignal nicht wahrnehmbar anders als das konventionelle Trainingssignal ist. Beispielsweise kann im Freisprech-Telefonkontext das Adaptionsverstärkungssignal so maßgeschneidert werden, dass es vom Lautsprechersignal (das heißt dem konventionellen Trainingssignal) maskiert und damit für den Telefon-Anwender unhörbar ist.

[0035] Fachleute werden erkennen, dass die vorliegende Erfindung nicht auf die spezifischen beispielhaften Ausführungsformen beschränkt ist, die hier aus illustrativen Gründen beschrieben worden sind, und dass zahlreiche alternative Ausführungsformen ebenfalls erwogen werden. Beispielsweise sind, obwohl die beispielhaften Ausführungsformen in Bezug auf akustische Echo-Unterdrückung im Freisprech-Telefonie-Kontext beschrieben worden sind, die offenbarten Adaptionsverstärkungstechniken gleichermaßen auf alle adaptiven Filterungsanordnungen anwendbar (z. B. adaptive Geräuschunterdrückung und Netzwerk-Echo-Unterdrückung). Darüber hinaus können die offenbarten Maskierungstechniken beispielsweise ebenfalls verwendet werden, um unhörbare Daten hinter einem hörbaren, interessierenden Signal zu übertragen, wie etwa Sprache oder Musik. Der Umfang der Erfindung ist daher durch die hier angehängten Ansprüche definiert, und nicht durch die vorstehende Beschreibung, und alle mit der Bedeutung der Ansprüche konsistenten Äquivalente sollen hiervon erfasst sein.

Patentansprüche

1. Signalverarbeitungs-Vorrichtung, umfassend: einen adaptiven Filter (**130**, **140**), der dafür konfiguriert ist, ein Eingangssignal zu filtern und dadurch ein gefiltertes Ausgangssignal bereitzustellen, wobei eine Filtereigenschaft des adaptiven Filters (**130**, **140**) dynamisch justiert wird, basierend auf dem gefilterten Ausgangssignal und einem Trainingssignal; und
einen Adaptionsverstärkungsprozessor (**300**, **400**) der ein adaptionsverstärkendes Signal mit einem

Referenzsignal kombiniert, um das Trainingssignal dem adaptiven Filter (130, 140) bereitzustellen, wobei das adaptionsverstärkende Signal dynamisch justiert wird, basierend auf dem Referenzsignal.

2. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei das Eingangssignal ein Nahend-Audiosignal ist und das Referenzsignal ein Fernend-Audiosignal ist.

3. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei der adaptive Filter (130, 140) ein Echokompensationsfilter (130) ist und wobei das gefilterte Ausgangssignal ein Nahend-echokompensiertes Audiosignal ist.

4. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei der Adoptionsverstärkungsprozessor (300) dafür adaptiert ist, das adaptionsverstärkende Signal zum Referenzsignal hinzuzuaddieren (340), um das Trainingssignal für den adaptiven Filter (130, 140) bereitzustellen.

5. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei das adaptionsverstärkende Signal basierend auf einer dynamisch berechneten Frequenzmaske (310) des Referenzsignals erzeugt wird.

6. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei das adaptionsverstärkende Signal eine unhörbare Komponente des Trainingssignals ist, die vom Referenzsignal maskiert ist.

7. Signalverarbeitungs-Vorrichtung gemäß Anspruch 6, wobei eine Frequenzverteilung des adaptionsverstärkenden Signals ungefähr zur Frequenzmaske des Referenzsignals passt.

8. Signalverarbeitungs-Vorrichtung gemäß Anspruch 6, wobei das adaptionsverstärkende Signal durch Filtern (330) einer weißen Pseudo-Rausch-Sequenz 320 unter Verwendung der Frequenzmaske (310) des Referenzsignals erzeugt wird.

9. Signalverarbeitungs-Vorrichtung gemäß Anspruch 6, wobei das adaptionsverstärkende Signal durch Filtern (330) einer reproduzierbaren Maximallängensequenz unter Verwendung der Frequenzmaske (310) des Referenzsignals erzeugt wird.

10. Signalverarbeitungs-Vorrichtung gemäß Anspruch 6, wobei das adaptionsverstärkende Signal erzeugt wird, indem zuerst das Referenzsignal audiocodiert (410) und dann audiodecodiert (420) wird.

11. Signalverarbeitungs-Vorrichtung gemäß Anspruch 1, wobei die Filterungs-Charakteristik des adaptiven Filters gemäß einem Algorithmus (150) der kleinsten mittleren Quadrate (LMS) justierbar ist.

12. Kommunikationsvorrichtung zum Bereitstellen bidirektonaler Kommunikationen zwischen einem Nahend-Anwender und einem Fernend-Anwender über ein bidirektonalen Kommunikationskanal, umfassend:

ein Nahend-Mikrofon (110), das Nahend-Geräusche empfängt und ein entsprechendes Nahend-Audiosignal bereitstellt;

einen adaptiven Echokompensator (130), der das Nahend-Audiosignal empfängt und ein echokompensierte Nahend-Signal zur Übertragung an den Fernend-Anwender über den Kommunikationskanal bereitstellt, wobei adaptive Filterkoeffizienten des adaptiven Echokompensators (130) dynamisch in Abhängigkeit von dem echokompensierten Nahend-Signal und von einem Audio-Trainingssignal justierbar sind; einen Adoptionsverstärkungsprozessor (300, 400), der ein Fernend-Audiosignal über den Kommunikationskanal empfängt und das Audio-Trainingssignal dem adaptiven Echokompensator bereitstellt, wobei der Prozessor dafür ausgelegt ist, das Audio-Trainingssignal durch Kombinieren des Fernend-Audiosignals mit einem adaptionsverstärkenden Signal dynamisch zu berechnen, wobei das adaptionsverstärkende Signal dynamisch basierend auf einer Frequenzmaske (310) des Fernend-Audiosignals justierbar ist; und

einen Nahend-Lautsprecher (120), der das Audio-Trainingssignal empfängt und das entsprechende Fernend-Geräusch dem Nahend-Anwender bereitstellt.

13. Kommunikationsvorrichtung gemäß Anspruch 12, wobei der Prozessor dafür ausgelegt ist, das Audio-Trainingssignal durch Summieren (340) von Proben des Fernend-Audiosignals und des adaptionsverstärkenden Signals zu berechnen.

14. Kommunikationsvorrichtung gemäß Anspruch 12, wobei das adaptionsverstärkende Signal eine nicht hörbare Komponente des Audio-Trainingssignals ist.

15. Kommunikationsvorrichtung gemäß Anspruch 14, wobei das adaptionsverstärkende Signal so einstellbar ist, dass eine Frequenzverteilung des adaptionsverstärkenden Signals ungefähr die Frequenzmaske des Fernend-Audiosignals verfolgt.

16. Kommunikationsvorrichtung gemäß Anspruch 14, wobei das adaptionsverstärkende Signal durch Filtern (330) einer weißen Pseudo-Rausch-Sequenz (320) unter Verwendung der Frequenzmaske (310) des Fernend-Audiosignales hergestellt wird.

17. Kommunikationsvorrichtung gemäß Anspruch 14, wobei das adaptionsverstärkende Signal durch Filtern einer reproduzierbaren Maximallängen-Sequenz unter Verwendung der Frequenzmaske des Fernend-Audiosignals hergestellt wird.

18. Kommunikationsvorrichtung gemäß Anspruch 14, wobei das adaptionsverstärkende Signal hergestellt wird, indem das Fernend-Audiosignal zuerst audiocodiert (**410**) und dann audiodecodiert (**420**) wird.

19. Kommunikationsvorrichtung gemäß Anspruch 12, wobei die adaptiven Filterkoeffizienten des adaptiven Echokompensators gemäß einem Algorithmus kleinsten mittleren Quadrat (LMS) justierbar sind.

20. Verfahren zum Trainieren eines adaptiven Filters, umfassend die Schritte:

Berechnen einer Frequenzmaske (**310**) eines Referenzsignals;

Berechnen eines adaptionsverstärkenden Signals basierend auf der Frequenzmaske (**310**) des Referenzsignals;

Berechnen eines Trainingssignals basierend auf dem Referenzsignal und dem adaptionsverstärkenden Signal; und

Einstellen einer Filter-Charakteristik des adaptiven Filters, basierend auf dem Trainingssignal.

21. Verfahren gemäß Anspruch 20, wobei der adaptive Filter ein Echokompensations-Filter ist und wobei das Referenzsignal ein Audiosignal ist.

22. Verfahren gemäß Anspruch 20, wobei der Schritt des Berechnens eines Trainingssignals den Schritt des Addierens (**340**) des adaptionsverstärkenden Signals zum Referenzsignal umfasst.

23. Verfahren gemäß Anspruch 20, wobei das adaptionsverstärkende Signal eine unhörbare Komponente des Trainingssignals ist, die durch das Referenzsignal maskiert ist.

24. Verfahren gemäß Anspruch 23, wobei der Schritt des Berechnens eines adaptionsverstärkenden Signals den Schritt des Veranlassens einer Frequenzverteilung des adaptionsverstärkenden Signals zur Annäherung an die berechnete Frequenzmaske umfasst.

25. Verfahren gemäß Anspruch 23, wobei das adaptionsverstärkende Signal durch Filtern (**330**) einer weißen Pseudo-Rauschen-Sequenz (**320**) unter Verwendung der Frequenzmaske (**310**) des Referenzsignals hergestellt wird.

26. Verfahren gemäß Anspruch 23, wobei das adaptionsverstärkende Signal durch Filtern einer reproduzierbaren Maximallängen-Sequenz unter Verwendung der Frequenzmaske des Referenzsignals hergestellt wird.

27. Verfahren gemäß Anspruch 23, wobei das adaptionsverstärkende Signal hergestellt wird, indem

das Referenzsignal zuerst audiocodiert (**410**) und dann audiodecodiert (**420**) wird.

28. Verfahren gemäß Anspruch 20, wobei die Frequenz-Charakteristik des adaptiven Filters anhand eines Algorithmus (**150**) der kleinsten mittleren Quadrat (LMS) eingestellt wird.

Es folgen 2 Blatt Zeichnungen

Anhängende Zeichnungen

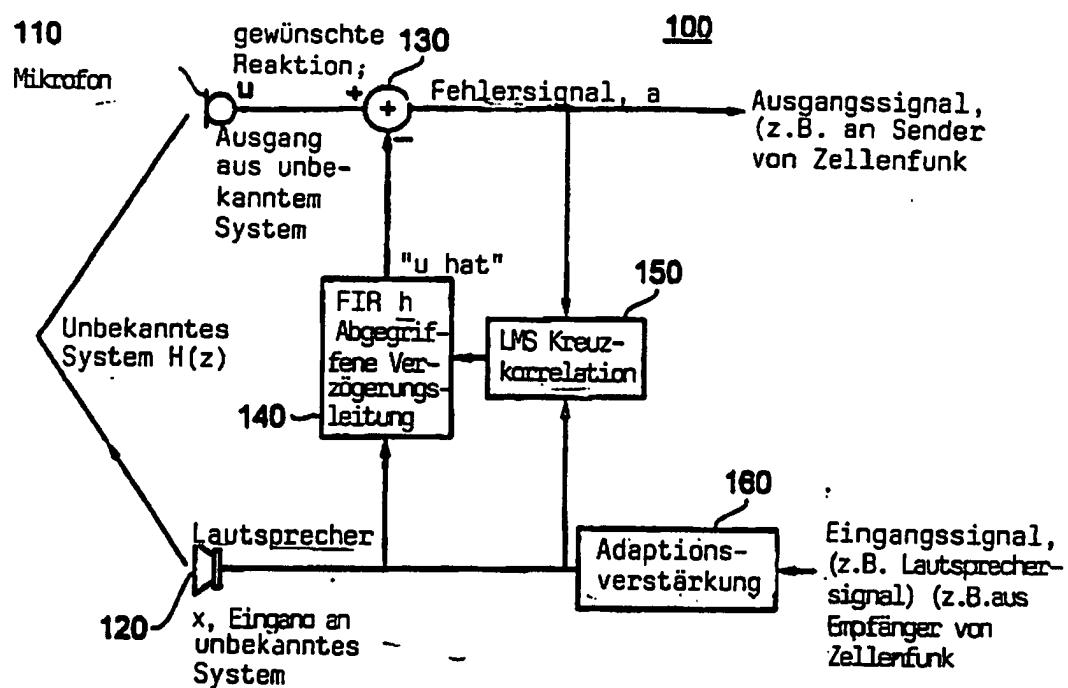


FIG. 1

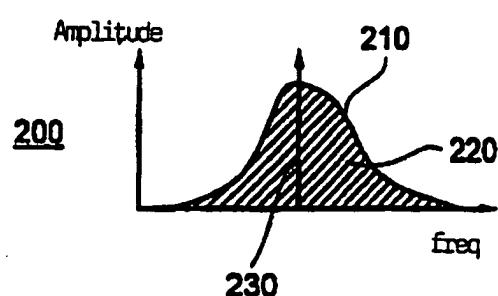


FIG. 2

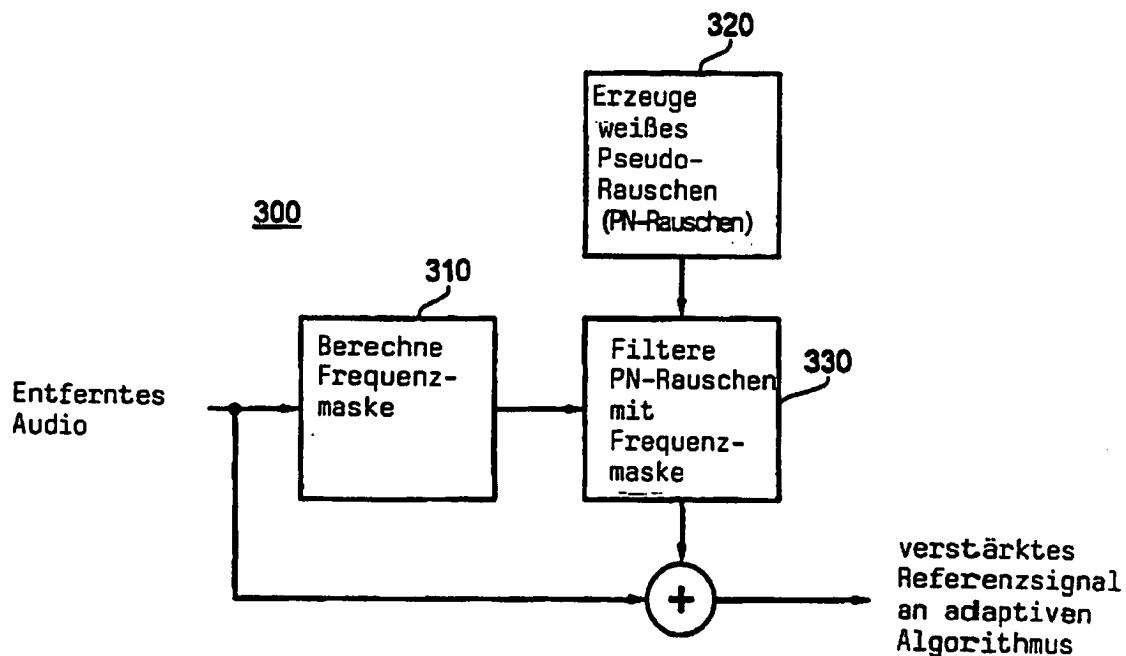


FIG. 3

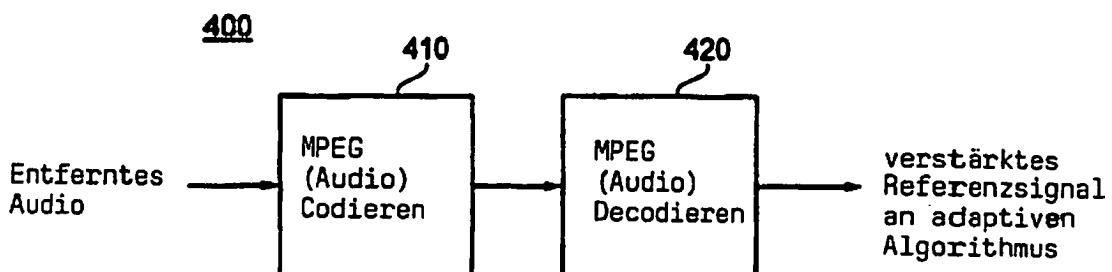


FIG. 4