



(11)

EP 2 982 136 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
13.06.2018 Patentblatt 2018/24

(51) Int Cl.:
H04R 25/00 (2006.01)

(21) Anmeldenummer: **14710644.7**

(86) Internationale Anmeldenummer:
PCT/IB2014/059290

(22) Anmeldetag: **27.02.2014**

(87) Internationale Veröffentlichungsnummer:
WO 2014/162214 (09.10.2014 Gazette 2014/41)

(54) VERFAHREN ZUM SCHÄTZEN EINES NUTZSIGNALS UND HÖRVORRICHTUNG

METHOD FOR EVALUATING A USEFUL SIGNAL AND AUDIO DEVICE

PROCÉDÉ D'ESTIMATION D'UN SIGNAL UTILE ET DISPOSITIF AUDITIF

(84) Benannte Vertragsstaaten:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(30) Priorität: **02.04.2013 DE 102013205790**

(43) Veröffentlichungstag der Anmeldung:
10.02.2016 Patentblatt 2016/06

(73) Patentinhaber: **Sivantos Pte. Ltd. Singapore 539775 (SG)**

(72) Erfinder:
• **KELLERMANN, Walter**
90542 Eckental (DE)
• **REINDL, Klaus**
01069 Dresden (DE)
• **ZHENG, Yuanhang**
72766 Reutlingen (DE)

(74) Vertreter: **FDST Patentanwälte Nordostpark 16 90411 Nürnberg (DE)**

(56) Entgegenhaltungen:
EP-A1- 2 395 506

- **SOUDEN M ET AL:** "On Optimal Frequency-Domain Multichannel Linear Filtering for Noise Reduction", IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, USA, Bd. 18, Nr. 2, 1. Februar 2010 (2010-02-01), Seiten 260-276, XP011329126, ISSN: 1558-7916, DOI: 10.1109/TASL.2009.2025790
- **MOONEN M ET AL:** "Robustness Analysis of Multichannel Wiener Filtering and Generalized Sidelobe Cancellation for Multimicrophone Noise Reduction in Hearing Aid Applications", IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, Bd. 13, Nr. 4, 1. Juli 2005 (2005-07-01), Seiten 487-503, XP011134931, ISSN: 1063-6676, DOI: 10.1109/TSA.2005.845821
- **SPRIET A ET AL:** "The impact of speech detection errors on the noise reduction performance of multi-channel Wiener filtering and Generalized Sidelobe Cancellation", SIGNAL PROCESSING, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, Bd. 85, Nr. 6, 1. Juni 2005 (2005-06-01), Seiten 1073-1088, XP027670885, ISSN: 0165-1684 [gefunden am 2005-06-01]

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 2 982 136 B1

Beschreibung

[0001] Die vorliegende Erfindung betrifft ein Verfahren zum Schätzen eines Nutzsignals einer Hörvorrichtung durch Gewinnen von mindestens zwei Mikrofonsignalen aus einem Schallsignal, Gewinnen eines Restsignals aus den Mikrofonsignalen, bei dem ein Anteil der Mikrofonsignale aus einer vorgebbaren Richtung blockiert ist, und Filtern der Mikrofonsignale mit einem Filter, wodurch eine Schätzung für das Nutzsignal erhalten wird. Darüber hinaus betrifft die vorliegende Erfindung eine Hörvorrichtung mit einer entsprechenden Mikrofoneinrichtung, Blockiereinrichtung und einem Filter. Unter einer Hörvorrichtung wird hier jedes im oder am Ohr tragbare, einen Schallreiz produzierende Gerät verstanden, insbesondere ein Hörgerät, ein Headset, Kopfhörer und dergleichen.

[0002] Hörgeräte sind tragbare Hörvorrichtungen, die zur Versorgung von Schwerhörenden dienen. Um den zahlreichen individuellen Bedürfnissen entgegenzukommen, werden unterschiedliche Bauformen von Hörgeräten wie Hinter-dem-Ohr-Hörgeräte (HdO), Hörgerät mit externem Hörer (RIC: receiver in the canal) und In-dem-Ohr-Hörgeräte (IdO), z.B. auch Concha-Hörgeräte oder Kanal-Hörgeräte (ITE, CIC), bereitgestellt. Die beispielhaft aufgeführten Hörgeräte werden am Außenohr oder im Gehörgang getragen. Darüber hinaus stehen auf dem Markt aber auch Knochenleitungshörhilfen, implantierbare oder vibrotaktile Hörhilfen zur Verfügung. Dabei erfolgt die Stimulation des geschädigten Gehörs entweder mechanisch oder elektrisch.

[0003] Hörgeräte besitzen prinzipiell als wesentliche Komponenten einen Eingangswandler, einen Verstärker und einen Ausgangswandler. Der Eingangswandler ist in der Regel ein Schallempfänger, z. B. ein Mikrofon, und/oder ein elektromagnetischer Empfänger, z. B. eine Induktionsspule. Der Ausgangswandler ist meist als elektroakustischer Wandler, z. B. Miniaturlautsprecher, oder als elektromechanischer Wandler, z. B. Knochenleitungshörer, realisiert. Der Verstärker ist üblicherweise in eine Signalverarbeitungseinheit integriert. Dieser prinzipielle Aufbau ist in FIG 1 am Beispiel eines Hinter-dem-Ohr-Hörgeräts dargestellt. In ein Hörgerätegehäuse 1 zum Tragen hinter dem Ohr sind ein oder mehrere Mikrofone 2 zur Aufnahme des Schalls aus der Umgebung eingebaut. Eine Signalverarbeitungseinheit 3, die ebenfalls in das Hörgerätegehäuse 1 integriert ist, verarbeitet die Mikrofonsignale und verstärkt sie. Das Ausgangssignal der Signalverarbeitungseinheit 3 wird an einen Lautsprecher bzw. Hörer 4 übertragen, der ein akustisches Signal ausgibt. Der Schall wird gegebenenfalls über einen Schallschlauch, der mit einer Otoplastik im Gehörgang fixiert ist, zum Trommelfell des Geräteträgers übertragen. Die Energieversorgung des Hörgeräts und insbesondere die der Signalverarbeitungseinheit 3 erfolgt durch eine ebenfalls ins Hörgerätegehäuse 1 integrierte Batterie 5.

[0004] Eine besondere Herausforderung bei der Nut-

zung eines Hörgeräts oder einer anderen Hörvorrichtung besteht in ihrem Einsatz in einem sogenannten Cafeteria-Szenario. Dabei unterhält sich der Träger des Hörgeräts oder der Hörvorrichtung mit einem Gesprächspartner. Die akustische Umgebung ist weiterhin geprägt durch andere sprechende Personen sowie durch undefiniertes Hintergrundgeräusch. In einem solchen Szenario ist es besonders schwierig, die Sprache des Gesprächspartners aus dem gesamten Schallsignal zu extrahieren, d.h. das tatsächliche Nutzsignal zu ermitteln bzw. zu schätzen. Das Störsignal bzw. Störgeräusch besteht hierbei in der Regel also aus Hintergrundrauschen und/oder Störsprachanteilen bzw. Interferenzen.

[0005] Um mehrkanalige Störreduktionstechniken zu realisieren, müssen Statistikgrößen zweiter Ordnung (insbesondere spektrale Leistungsdichte PSD; Power Spectral Density) der Störgeräuschkomponenten geschätzt werden. Typischerweise werden diese Komponenten während der Zielsprachpausen geschätzt. Damit verlässliche Schätzungen nur während der Zielsprachpausen durchgeführt werden, müssen die Störkomponenten zeitlich ausreichend stationär sein, sodass die erhaltene Schätzung auch dann gültig ist, wenn der Zielsprecher nach einer gewissen Pause wieder aktiv ist. In der Realität sind die Störsignale jedoch nicht immer stationär. Daher sind wirkungsvolle Mehrkanal-Störreduktionstechniken in ihrer Anwendung begrenzt, da sie in Szenarien mit nicht stationären Signalen (z. B. sprachähnlichen Interferenzen) kaum ausführbar sind.

[0006] Die Schätzung von Störgeräuschstatistikgrößen für mehrkanalige Störreduktionstechniken basiert typischerweise auf einer sogenannten Ziel-Sprachaktivitätsdetektion (VAD; Voice Activity Detection). Dies bedeutet, dass eine Schätzung der gesamten Störgeräusch-PSD-Matrix nur in Perioden möglich ist, in denen der Zielsprecher inaktiv ist. Wenn die Störgeräusch-PSD-Matrix nur während der Zielsprachpausen geschätzt werden kann, ist es wichtig, dass die PSD der Störgeräuschkomponenten sich über der Zeit nicht stark ändert, d.h. die Störgeräuschsignale müssen (zeitlich) hinreichend stationär sein. Der größte Nachteil dieser Strategie besteht daher darin, dass für (zeitlich) sehr instationäre Signale (z. B. sprachähnliche Interferenzen) die Schätzungen der Störgeräusch-PSD-Matrix, die nur während der Zielsprachpausen erhalten werden können, nicht verlässlich sind, da nicht angenommen werden kann, dass die während einer Sprachpause erhaltene Schätzung auch noch gültig ist, nachdem der Zielsprecher bereits wieder eine Zeit lang aktiv ist.

[0007] Aus der EP 2 395 506 A1 ist ein Verfahren und ein Schallsignalverarbeitungssystem zur Unterdrückung von Interferenzen und Rauschen in binauralen Mikrofonkonfigurationen bekannt. Bei dem Verfahren wird anhand von Referenzsignalen, bei welchen ein Anteil der von der Mikrofonkonfiguration empfangenen Mikrofonsignale aus einer vorgebbaren Richtung blockiert ist, eine Kohärenzgröße zur Schätzung einer Störgeräusch-PSD-Matrix bestimmt.

[0008] Die Aufgabe der vorliegenden Erfindung besteht somit darin, ein Verfahren zum Schätzen eines Nutzsignals einer Hörvorrichtung bereitzustellen, das auch bei zeitlich nichtstationären Signalen, wie etwa bei Sprache, eingesetzt werden kann. Darüber hinaus soll eine entsprechende Hörvorrichtung bereitgestellt werden.

[0009] Erfindungsgemäß wird diese Aufgabe gelöst durch ein Verfahren zum Schätzen eines Nutzsignals einer Hörvorrichtung durch

- Gewinnen von mindestens zwei Mikrofonsignalen aus jeweils einem Schallsignal, wobei die Mikrofon-signale einen Mikrofon-signalvektor bilden,
- Gewinnen eines Referenzsignalvektors aus dem Mikrofon-signalvektor, bei welchem Referenzsignal-vektor ein Anteil der Mikrofon-signale aus einer vorgebbaren Richtung blockiert ist, und
- Filtern des Mikrofon-signalvektors mit einem Filter, wodurch ein Schätzsignal für das Nutzsignal erhalten wird, sowie
- Ermitteln einer Kohärenzgröße aus dem Referenz-signalvektor und dem Mikrofon-signalvektor,
- Ermitteln einer Leistungsdichtegröße aus der Kohärenzgröße und
- Parametrieren des Filters anhand der Leistungs-dichtegröße.

[0010] Darüber hinaus wird erfindungsgemäß bereitgestellt eine Hörvorrichtung mit

- einer Mikrofoneinrichtung zum Gewinnen von mindestens zwei Mikrofonsignalen aus jeweils einem Schallsignal, wobei die Mikrofon-signale einen Mikrofon-signalvektor bilden,
- einer Blockiereinrichtung zum Gewinnen eines Referenzsignalvektors aus dem Mikrofon-signalvektor, bei welchem Referenzsignalvektor ein Anteil der Mikrofon-signale aus einer vorgebbaren Richtung blockiert ist, und
- einem Filter zum Filtern des Mikrofon-signalvektors, wodurch ein Schätzsignal für das Nutzsignal erhalten wird, sowie mit
- einer Recheneinrichtung zum Ermitteln einer Kohärenzgröße aus dem Referenzsignalvektor und dem Mikrofon-signalvektor und zum Ermitteln einer Leistungsdichtegröße aus der Kohärenzgröße sowie zum Parametrieren des Filters anhand der Leistungsdichtegröße.

[0011] Der Referenzsignalvektor kann auch eindimensional sein, d.h. aus einem einzigen Referenzsignal bestehen. In der Regel wird er aber aus mehreren Referenzsignalen bestehen.

[0012] In vorteilhafter Weise wird also aus dem Referenzsignalvektor, d.h. aus Anteilen des Restsignals eine Kohärenzgröße und insbesondere eine Kohärenzmatrix

gewonnen, aus der sich eine Leistungsdichtegröße, und insbesondere eine Leistungsdichtematrix, des Restsignals (d.h. der Störgeräuschanteile) ermitteln lässt. Mit dieser Leistungsdichtegröße wird das Filter parametrisiert, sodass aus den Mikrofonsignalen bzw. dem Mikrofon-signalvektor eine spezifische Nutzsignalquelle herausgefiltert bzw. geschätzt werden kann. Mit dem vorgeschlagenen Konzept können also auch spektrale Leistungsdichten von Störgeräuschkomponenten für zeitlich nicht stationäre Signale (z. B. Sprache) geschätzt werden, sodass mehrkanalige Störgeräuschreduktionstechniken in praktisch beliebigen Szenarien angewendet bzw. realisiert werden können.

[0013] Vorzugsweise wird zum Gewinnen des Referenzsignalvektors die vorgebbare Richtung des Nutzsignals aus dem Mikrofon-signalvektor geschätzt. Damit ist es möglich, aus dem gesamten Einfallsraum des Schalls das Nutzsignal auszublenden.

[0014] Insbesondere ist es von Vorteil, zum Gewinnen des Referenzsignalvektors einen direktionalen Blinde-Quellentrennungs-Algorithmus zu verwenden. Ein derartiger Blinde-Quellentrennungs-Algorithmus hat sich in der Störgeräuschunterdrückung bewährt, und er ist aufgrund einer zuvor durchgeführten Quellenlokalisierung sehr leistungstark.

[0015] Zum Gewinnen des Referenzsignalvektors können jeweils eine Nutzsignalkomponente jedes Mikrofon-signals untereinander angeglichen und anschließend voneinander subtrahiert werden. Dadurch lassen sich die Signalkanäle (je ein Kanal für ein Mikrofon bzw. ein Mikrofon-signal) wirksam von Ziel- bzw. Nutzsignalkomponenten befreien. Dabei ist es besonders günstig, wenn die Nutzsignalkomponenten sowohl hinsichtlich der Verzögerung als auch hinsichtlich ihrer Spektren aneinander angeglichen werden. Somit können die Nutzsignalkomponenten nahezu restlos aus den Signalkanälen entfernt werden.

[0016] Zum Ermitteln der Leistungsdichtegröße und insbesondere der Leistungsdichtematrix des (mehrkanaligen) Restsignalvektors kann neben der Kohärenzgröße auch der Restsignalvektor selbst herangezogen werden. Damit kann für das Filter auf der Basis der Kohärenzgröße und des Referenzsignalvektors eine Steuerung anhand der Leistungsdichte bereitgestellt werden.

[0017] Das Nutzsignal kann insbesondere ein Sprachsignal sein. Damit lässt sich das erfindungsgemäße Verfahren bzw. die erfindungsgemäße Hörvorrichtung insbesondere dafür nutzen, die Sprachverständlichkeit zu erhöhen.

[0018] Des Weiteren kann der Referenzsignalvektor Sprachsignalanteile umfassen, die nicht Teil des Nutzsignals sind. Der Referenzsignalvektor beinhaltet beispielsweise Sprachanteile von Sprechern, die von dem Zielsprecher verschieden sind.

[0019] Die oben geschilderten Verfahrensmerkmale können auch in den Einrichtungen der Hörvorrichtung umgesetzt werden, wodurch diese Einrichtungen die jeweilige Funktionalität erhalten.

[0020] Die vorliegende Erfindung wird anhand der beigefügten Zeichnungen näher erläutert, in denen zeigen:

FIG 1 den prinzipiellen Aufbau einer Hörvorrichtung gemäß dem Stand der Technik und

FIG 2 ein Blockschaltdiagramm zum erfindungsgemäßen Schätzen eines Nutzsignals.

[0021] Die nachfolgend näher geschilderten Ausführungsbeispiele stellen bevorzugte Ausführungsformen der vorliegenden Erfindung dar.

[0022] Das in FIG 2 dargestellte Blockdiagramm stellt zum einen ein Verfahren dar, welches in einem Hörgerät gemäß FIG 1 oder in einer anderen Hörvorrichtung implementiert werden kann. Zum anderen können die in FIG 2 dargestellten Blöcke entsprechende Einrichtungen einer Hörvorrichtung darstellen.

[0023] Eine beispielhafte Hörvorrichtung bzw. ein beispielhaftes Hörgerät umfasst eine Sensor- bzw. Mikrofonanordnung mit mindestens zwei Sensoren bzw. zwei Mikrofonen M1, M_p. Nachfolgend wird stellvertretend immer von Mikrofonen gesprochen.

[0024] Jedes Mikrofon M1, M_p wandelt das jeweils anliegende Schallsignal in ein korrespondierendes Mikrofonsignal. Die Schallsignale sind Komponenten eines Schallfelds, welches die akustische Situation beispielsweise eines Hörgeräteträgers repräsentiert. Eine solche typische Situation wäre die eines "Cafeteria-Szenarios", in welcher der Hörgeräteträger mit einem Gesprächspartner spricht, im Hintergrund eine oder mehrere andere Personen sprechen und sonstiges Hintergrundrauschen vorliegt. Es kann aber auch eine andere akustische Situation vorliegen, in der nichtstationäre Störgeräusche gegeben sind.

[0025] Die Mikrofonsignale, die zusammen einen Mikrofonsignalvektor x bilden, werden jeweils in eigenen Kanälen weiterverarbeitet, d.h. in jedem Kanal wird ein Mikrofonsignal verarbeitet. In FIG 2 ist diese mehrkanalige Verarbeitung durch dicke Pfeile dargestellt. Der Mikrofonsignalvektor x wird in dem mehrkanaligen System 10 einer Quellenlokalisierungseinheit LOC (source localisation) zugeführt. Diese gewinnt aus dem Mikrofonsignalvektor x Positionsdaten Φ_q einer Quelle S_q . Insbesondere wird die Positionsinformation Φ_q der Nutzsignalquelle S_q im dreidimensionalen Raum oder einfach nur als Winkel bzw. Winkel und Entfernung ermittelt. Diese Positionsinformation Φ_q wird als grobe Referenzinformation zum Erstellen einer Blockiermatrix BM genutzt. Mit Hilfe der Blockiermatrix BM werden aus den Mikrofonsignalen bzw. dem Mikrofonsignalvektor x diejenigen Anteile räumlich ausgeblendet, die aus dem räumlichen Bereich der Nutzsignalquelle stammen. Eine derartige Blockiermatrix BM kann beispielsweise auf einem direktionalen Blinde-Quellentrennungs-Algorithmus beruhen, wie er in Y. Zheng, K. Reindl, and W. Kellermann "BSS for improved interference estimation for blind speech signal extraction with two microphones," in

IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) Aruba, Dutch Antilles, Dez. 2009 beschrieben ist. Es können aber auch beliebige andere Algorithmen zur Ermittlung der Blockiermatrix BM eingesetzt werden.

[0026] Aus dem Mikrofonsignalvektor x ergibt sich durch Anwendung der Blockiermatrix BM damit ein mehrkanaliges Referenzsignal bzw. ein Referenzsignalvektor n . Werden die Signale beispielsweise paarweise in der Blockiermatrix subtrahiert, so kann die Anzahl der Signale des mehrdimensionalen Referenzsignalvektors n der Hälfte der Anzahl der Mikrofonsignale bzw. -kanäle entsprechen. Bei ungerader Anzahl der Mikrofonsignale wird vorzugsweise aufgerundet. Der Referenzsignalvektor ist also in der Regel ein mehrdimensionaler Vektor aus mehreren Einzelsignalen.

[0027] Der Referenzsignalvektor n wird zusammen mit dem Mikrofonsignalvektor x , der aus den einzelnen Mikrofonsignalen besteht, einer Kohärenzschätzeinheit COH zugeführt. Diese schätzt aus den beiden Vektoren n und x eine Kohärenzmatrix r . Die Kohärenzmatrix r wird einer PSD-Schätzeinheit PSD zugeführt. Die PSD-Schätzeinheit schätzt aus der Kohärenzmatrix r und dem Referenzvektor n eine mehrdimensionale Leistungsdichteschätzgröße S wie beispielsweise beschrieben in I. McCowan and H. Bourlard, "Microphone array post-filter for diffuse noise field," in IEEE Int. conf. Acoustics, Speech, Signal Processing (ICASSP), 2002, Seiten 905-908 oder in K. Reindl., Y. Zheng, A. Schwarz, S. Meier, R. Maas, A. Sehr, and W. Kellermann, "A stereophonic acoustic signal extraction scheme for noise and reverberant environments," Computer Speech and Language, 2012.

[0028] Ein mehrkanaliges Filter FILT schätzt aus der Leistungsdichteschätzgröße S Filterparameter. Diese werden im Filter FILT auf die Mikrofonsignale bzw. auf den Mikrofonsignalvektor x angewandt, wodurch sich das Schätzsignal $\hat{s}_q$ für die bestimmte Nutzquelle bzw. das Nutzsignal ergibt.

[0029] Somit kann vor allem eine Schätzung einer Störkomponenten betreffenden nichtstationären Statistikgröße zweiter Ordnung mittels PSD erreicht werden, indem die Kohärenz der entsprechenden Störkomponenten genutzt wird. Dabei können insbesondere die Zielsprachkomponenten zunächst in allen Kanälen gleichgesetzt werden (Verzögerungskompensation sowie spektrale Angleichung), sodass nahezu identische Zielsprachkomponenten in den verfügbaren Kanälen enthalten sind. Zu dieser Angleichung kann ein direktonaler Blinde-Quellentrennungs-Algorithmus der oben genannten Art verwendet werden. Aus den resultierenden Signalen können, wie oben im Detail dargestellt wurde, die Störsignalkohärenzmatrix geschätzt werden, welche ihrerseits zur Schätzung der Stör-PSD-Matrix S eingesetzt wird. Zur Schätzung des Nutzsignals bedarf es erfindungsgemäß also keiner Einschränkungen der zeitlichen Signalcharakteristiken. Im Gegensatz zu bekannten und typischerweise eingesetzten Konzepten, die nur bei hinreichend (zeitlich) stationären Störgeräuschsigna-

len anwendbar sind, nutzt die vorliegende Erfindung, dass das jeweilige akustische Szenario räumlich stationär ist, um die Störgeräusch-PSD-Matrix zu schätzen. Dabei kann angenommen werden, dass die räumliche Domäne für beliebige Szenarien hinreichend stationär ist im Gegensatz zur zeitlichen Domäne. Dies liegt daran, dass die Änderungen der Kohärenzfunktion hauptsächlich von den räumlichen Eigenschaften abhängen, d.h. von der geometrischen Anordnung der Quellen und Objekte in der akustischen Szene. Die Änderungen der Kohärenzfunktion hängen hingegen nur wenig von den zeitlichen Eigenschaften der Signale ab.

[0030] Zusammenfassend bedeutet dies also, dass das erfindungsgemäße Verfahren bzw. die erfindungsgemäße Hörvorrichtung nicht auf spezielle Szenarien beschränkt ist, die sich auf zeitlich stationäre Störgeräusche beziehen. Dementsprechend macht das erfindungsgemäße Konzept leistungsstarke, mehrkanalige Störreduktionstechniken für beliebige Szenarien, in denen Störgeräuschunterdrückung notwendig ist, einsetzbar oder realisierbar. Ein wesentlicher Bestandteil der Erfindung basiert also auf der Erkenntnis, die Schätzung der räumlichen Kohärenz von Störsignalen von der Schätzung der zeitlichen Statistikgrößen zweiter Ordnung (PSD der Störgeräuschkomponenten) zu trennen. Dabei können die Raum-Zeit-Kohärenzmatrizen auch für Szenarien mit (zeitlich) instationären Sprachsignalen kontinuierlich geschätzt werden.

[0031] In einem konkreten Beispiel kann als Filter ein mehrkanaliges Wiener-Filter eingesetzt werden. Prinzipiell kann aber auch ein einkanaliges Filter Verwendung finden. Ein derartiges Filter lässt sich beispielsweise bei der Störgeräuschunterdrückung in einem binauralen Hörgerät einsetzen.

[0032] Die PSD-Störgeräuschschätzung zusammen mit dem mehrkanaligen Wiener-Filter kann im Verbund mit einer Polyphasenfilterbank implementiert werden, wie sie typischerweise in Hörgeräten eingesetzt wird. Das erfindungsgemäße Konzept kann auf der Grundlage einer in SIR/SINR-Verstärkung (signal to interference ratio/signal to interference and noise ratio) realisiert werden. Darüber hinaus wird für die Berechnung beispielsweise ein ideales Blinde-Quellentrennung-Schema angenommen, d.h. die Zielsprachkomponenten sind in allen verfügbaren Kanälen in etwa die gleichen. Darüber hinaus kann im konkreten Fall eine ideale blockbasierte Sprachaktivitätsdetektion (VAD) verwendet werden, um die Störgeräuschkohärenzmatrix zu schätzen.

[0033] In Experimenten konnte gezeigt werden, dass gegebenenfalls mehrere Interferenz- bzw. Sprachsignale deutlich reduziert werden können (SIR mindestens 10dB). Auch wenn zusätzliches (diffuses) Hintergrundgeplapper vorhanden war, konnte ein SINR von 8dB erreicht werden. Dabei waren Verarbeitungsartefakte (Störung der einzelnen Signale) nicht hörbar.

Patentansprüche

1. Verfahren zum Schätzen eines Nutzsignals einer Hörvorrichtung durch
 - Gewinnen von mindestens zwei Mikrofonsignalen aus jeweils einem Schallsignal, wobei die Mikrofonsignale einen Mikrofonsignalvektor (x) bilden,
 - Gewinnen eines Referenzsignalvektors aus dem Mikrofonsignalvektor (x), bei welchem Referenzsignalvektor (n) ein Anteil der Mikrofonsignale aus einer vorgebbaren Richtung blockiert ist, und
 - Filtern des Mikrofonsignalvektors (x) mit einem Filter (FILT), wodurch ein Schätzsignal ($\hat{s}_q$) für ein Nutzsignal erhalten wird,

gekennzeichnet durch

 - Ermitteln einer Kohärenzgröße (r) aus dem Referenzsignalvektor (n) und dem Mikrofonsignalvektor (x),
 - Ermitteln einer Leistungsdichtegröße (S) aus der Kohärenzgröße (r) und
 - Parametrieren des Filters (FILT) anhand der Leistungsdichtegröße (S).
2. Verfahren nach Anspruch 1, wobei zum Gewinnen des Referenzsignalvektors (n) die vorgebbare Richtung des Nutzsignals aus dem Mikrofonsignalvektor (x) geschätzt wird.
3. Verfahren nach Anspruch 2, wobei das Gewinnen des Referenzsignalvektors (n) durch einen direkten Blinde-Quellentrennungs-Algorithmus erfolgt.
4. Verfahren nach einem der vorhergehenden Ansprüche, wobei zum Gewinnen des Referenzsignalvektors (n) jeweils eine Nutzsignalkomponente jedes Mikrofonsignals untereinander angeglichen und anschließend voneinander subtrahiert werden.
5. Verfahren nach Anspruch 4, wobei die Nutzsignalkomponenten sowohl hinsichtlich der Verzögerung als auch hinsichtlich ihrer Spektren aneinander angeglichen werden.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei die Kohärenzgröße (r) eine Kohärenzmatrix ist.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei zum Ermitteln der Leistungsdichtegröße (S) auch der Referenzsignalvektor (n) herangezogen wird.
8. Verfahren nach einem der vorhergehenden Ansprüche, wobei das Nutzsignal ein Sprachsignal ist.

9. Verfahren nach einem der vorhergehenden Ansprüche, wobei der Referenzsignalvektor (n) Sprachsignalanteile umfasst, die nicht Teil des Nutzsignals sind.

10. Hörvorrichtung mit

- einer Mikrofoneinrichtung (M1,...,Mp) zum Gewinnen von mindestens zwei Mikrofonsignalen aus jeweils einem Schallsignal, wobei die Mikrofonsignale einen Mikrofonsignalvektor (x) bilden,
- einer Blockiereinrichtung zum Gewinnen eines Referenzsignalvektors (n) aus dem Mikrofonsignalvektor (x), bei dem ein Anteil der Mikrofonsignale aus einer vorgebbaren Richtung blockiert ist, und
- einem Filter (FILT) zum Filtern des Mikrofonsignalvektors (x), wodurch ein Schätzsignal ($\tilde{s}_q$) für ein Nutzsignal erhalten wird,

gekennzeichnet durch

- eine Recheneinrichtung zum Ermitteln einer Kohärenzgröße (Γ) aus dem Referenzsignalvektor (n) und dem Mikrofonsignalvektor (x) und zum Ermitteln einer Leistungsdichtegröße (S) aus der Kohärenzgröße (Γ) sowie zum Parametrieren des Filters (FILT) anhand der Leistungsdichtegröße (S).

Claims

1. Method for estimating a useful signal from a hearing apparatus by

- obtaining at least two microphone signals from a respective sound signal, wherein the microphone signals form a microphone signal vector (x),
 - obtaining a reference signal vector from the microphone signal vector (x), which reference signal vector (x) has a portion of the microphone signals from a prescribable direction in a blocked state, and
 - filtering the microphone signal vector (x) using a filter (FILT), as a result of which an estimation signal ($\tilde{s}_q$) is obtained for a useful signal,
- characterized by**
- ascertainment of a coherence variable (Γ) from the reference signal vector (n) and the microphone signal vector (x),
 - ascertainment of a power density variable (S) from the coherence variable (Γ) and
 - parameterization of the filter (FILT) on the basis of the power density variable (S).

2. Method according to Claim 1, wherein obtaining the reference signal vector (n) involves the prescribable

direction of the useful signal being estimated from the microphone signal vector (x).

3. Method according to Claim 2, wherein the reference signal vector (n) is obtained by a directional blind source separation algorithm.

4. Method according to one of the preceding claims, wherein obtaining the reference signal vector (n) involves a respective useful signal component of each microphone signal being aligned with one another and then subtracted from one another.

5. Method according to Claim 4, wherein the useful signal components are aligned with one another both in terms of delay and in terms of their spectra.

6. Method according to one of the preceding claims, wherein the coherence variable (Γ) is a coherence matrix.

7. Method according to one of the preceding claims, wherein ascertaining the power density variable (S) also involves the use of the reference signal vector (n).

8. Method according to one of the preceding claims, wherein the useful signal is a voice signal.

9. Method according to one of the preceding claims, wherein the reference signal vector (n) comprises voice signal portions that are not part of the useful signal.

10. Hearing apparatus having

- a microphone device (M1, ..., Mp) for obtaining at least two microphone signals from a respective sound signal, wherein the microphone signals form a microphone signal vector (x),
 - a blocking device for obtaining a reference signal vector (n) from the microphone signal vector (x), which has a portion of the microphone signals from a prescribable direction in a blocked state, and
 - a filter (FILT) for filtering the microphone signal vector (x), as a result of which an estimation signal ($\tilde{s}_q$) is obtained for a useful signal,
- characterized by**
- a computation device for ascertaining a coherence variable (Γ) from the reference signal vector (n) and the microphone signal vector (x) and for ascertaining a power density variable (S) from the coherence variable (Γ) and also for parameterizing the filter (FILT) on the basis of the power density variable (S).

Revendications

1. Procédé d'évaluation d'un signal utile d'un appareil auditif par

- acquisition d'au moins deux signaux de microphone respectivement à partir d'un signal sonore, les signaux de microphone formant un vecteur de signal de microphone (x),

- acquisition d'un vecteur de signal de référence à partir du vecteur de signal de microphone (x), vecteur de signal de référence (n) avec lequel une part des signaux de microphone provenant d'une direction pouvant être prédéfinie est bloquée, et

- filtrage du vecteur de signal de microphone (x) avec un filtre (FILT), un signal d'évaluation ($\tilde{s}_q$) pour un signal utile étant ainsi obtenu,

caractérisé par

- détermination d'une grandeur de cohérence (Γ) à partir du vecteur de signal de référence (n) et du vecteur de signal de microphone (x),

- détermination d'une grandeur de densité de puissance (S) à partir de la grandeur de cohérence (Γ) et

- paramétrage du filtre (FILT) à l'aide de la grandeur de densité de puissance (S).

2. Procédé selon la revendication 1, la direction pouvant être prédéfinie du signal utile étant estimée à partir du vecteur de signal de microphone (x) en vue d'acquérir le vecteur de signal de référence (n).

3. Procédé selon la revendication 2, l'acquisition du vecteur de signal de référence (n) étant effectuée par un algorithme de séparation aveugle de source directionnel.

4. Procédé selon l'une des revendications précédentes, l'acquisition du vecteur de signal de référence (n) étant effectuée en alignant l'une sur l'autre une composante de signal utile de chaque signal de microphone puis en les soustrayant ensuite l'une de l'autre.

5. Procédé selon la revendication 4, les composantes de signal utile étant alignées l'une sur l'autre aussi bien du point de vue du retard que du point de vue de leurs spectres.

6. Procédé selon l'une des revendications précédentes, la grandeur de cohérence (Γ) étant une matrice de cohérence.

7. Procédé selon l'une des revendications précédentes, le vecteur de signal de référence (n) étant également utilisé pour la détermination de la grandeur de densité de puissance (S).

8. Procédé selon l'une des revendications précédentes, le signal utile étant un signal vocal.

9. Procédé selon l'une des revendications précédentes, le vecteur de signal de référence (n) comprenant des parts de signal vocal qui ne font pas partie du signal utile.

10. Appareil auditif comprenant

- un dispositif à microphone (M1, ..., Mp) destiné à acquérir au moins deux signaux de microphone respectivement à partir d'un signal sonore, les signaux de microphone formant un vecteur de signal de microphone (x),

- un dispositif de blocage destiné à acquérir un vecteur de signal de référence (n) à partir du vecteur de signal de microphone (x), avec lequel une part des signaux de microphone provenant d'une direction pouvant être prédéfinie est bloquée, et

- un filtre (FILT) destiné à filtrer le vecteur de signal de microphone (x), un signal d'évaluation ($\tilde{s}_q$) pour un signal utile étant ainsi obtenu,

caractérisé par

- un dispositif de calcul destiné à déterminer une grandeur de cohérence (Γ) à partir du vecteur de signal de référence (n) et du vecteur de signal de microphone (x) et à déterminer une grandeur de densité de puissance (S) à partir de la grandeur de cohérence (Γ) ainsi qu'à paramétrer le filtre (FILT) à l'aide de la grandeur de densité de puissance (S).

FIG 1
(Stand der Technik)

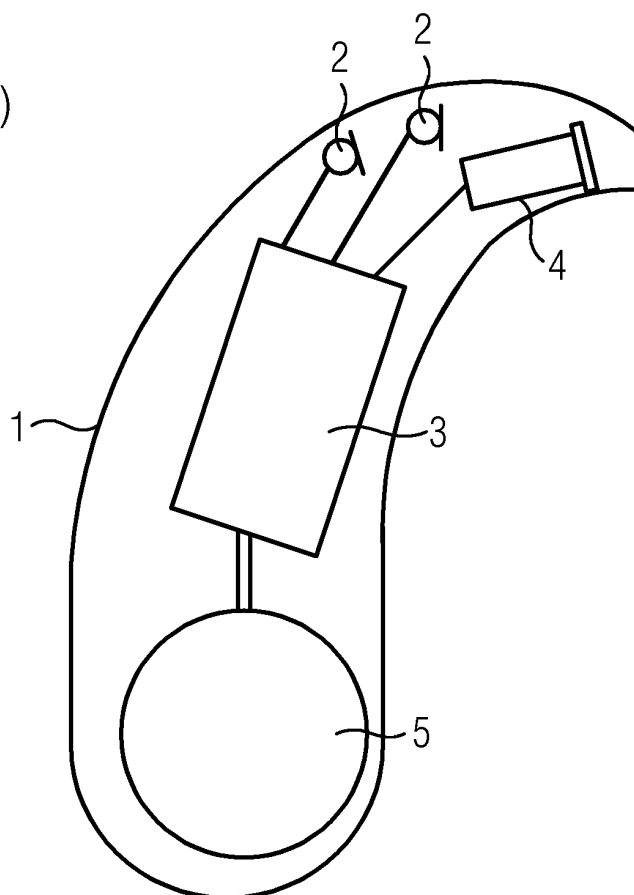
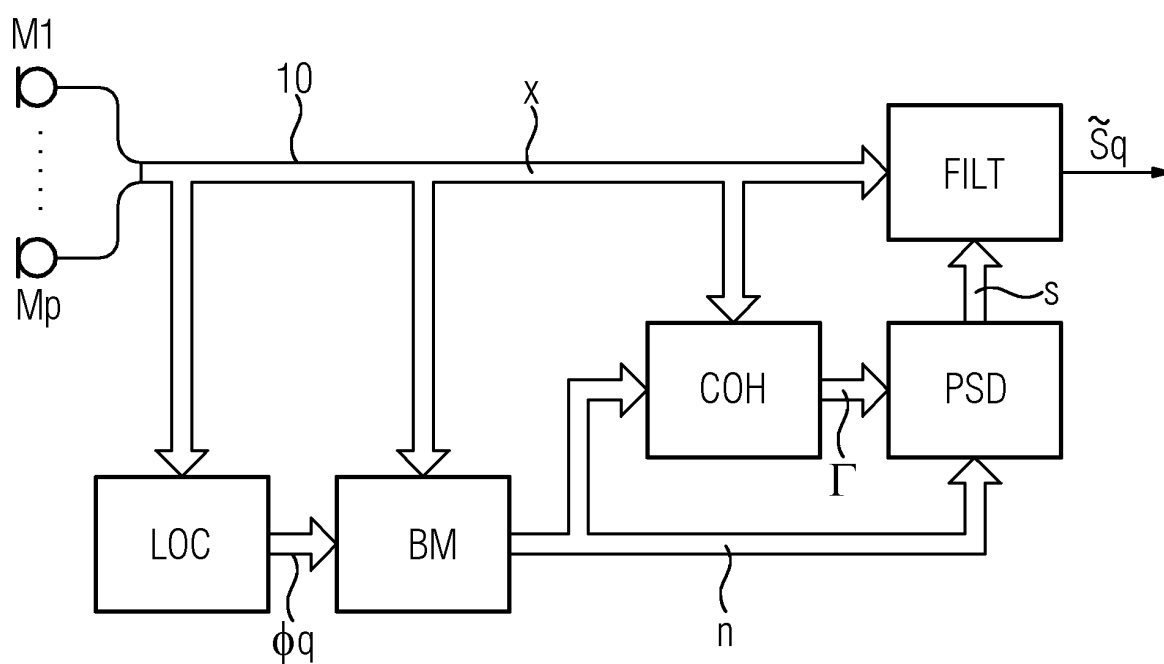


FIG 2



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 2395506 A1 [0007]

In der Beschreibung aufgeführte Nicht-Patentliteratur

- **Y. ZHENG ; K. REINDL ; W. KELLERMANN.** BSS for improved interference estimation for blind speech signal extraction with two microphones. *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)* Aruba, Oktober 2009 [0025]
- **I. MCCOWAN ; H. BOURLARD.** Microphone array post-filter for diffuse noise field. *IEEE Int. conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2002, 905-908 [0027]
- **K. REINDL. ; Y. ZHENG ; A. SCHWARZ ; S. MEIER ; R. MAAS ; A. SEHR ; W. KELLERMANN.** A stereophonic acoustic signal extraction scheme for noise and reverberant environments. *Computer Speech and Language*, 2012 [0027]