

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 1 595 427 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des
Hinweises auf die Patenterteilung:
13.12.2006 Patentblatt 2006/50

(51) Int Cl.:
H04R 3/00 (2006.01)

(21) Anmeldenummer: **05707893.3**

(86) Internationale Anmeldenummer:
PCT/EP2005/050386

(22) Anmeldetag: **31.01.2005**

(87) Internationale Veröffentlichungsnummer:
WO 2005/076659 (18.08.2005 Gazette 2005/33)

(54) VERFAHREN UND VORRICHTUNG ZUR SEPARIERUNG VON SCHALLSIGNALEN

METHOD AND DEVICE FOR THE SEPARATION OF SOUND SIGNALS

PROCEDE ET DISPOSITIF DE SEPARATION DE SIGNAUX SONORES

(84) Benannte Vertragsstaaten:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU MC NL PL PT RO SE SI SK TR**

(30) Priorität: **06.02.2004 DE 102004005998**

(43) Veröffentlichungstag der Anmeldung:
16.11.2005 Patentblatt 2005/46

(73) Patentinhaber: **Ruwisch, Dietmar, Dr.
12557 Berlin (DE)**

(72) Erfinder: **Ruwisch, Dietmar, Dr.
12557 Berlin (DE)**

(74) Vertreter: **Betten & Resch
Patentanwälte,
Theatinerstrasse 8
80333 München (DE)**

(56) Entgegenhaltungen:
**EP-A- 0 831 458 WO-A-02/061732
US-A- 5 539 859**

EP 1 595 427 B1

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

Beschreibung

[0001] Die vorliegende Erfindung betrifft ein Verfahren und eine Vorrichtung zur Separierung von Schallsignalen.

[0002] Die Erfindung liegt auf dem Gebiet der digitalen Signalverarbeitung zum Entmischen verschiedener akustischer Signale aus unterschiedlichen Raumrichtungen, welche mit zwei Mikrofonen in bekanntem Abstand stereo aufgenommen werden.

[0003] Das Gebiet der Quellentrennung, auch "beamforming" genannt, erfährt wachsende Bedeutung durch die Zunahme der mobilen Kommunikation sowie der automatischen Verarbeitung menschlicher Sprache. In sehr vielen Anwendungen tritt das Problem auf, dass das gewünschte Sprachsignal (Nutzsignal) durch verschiedene Störeinflüsse beeinträchtigt ist. Hier sind hauptsächlich Störungen durch Hintergrundgeräusche, Störungen durch andere Sprecher sowie Störungen durch Lautsprecherabgaben von Musik oder Sprache zu nennen. Die verschiedenen Störeinflüsse erfordern je nach ihrer Art und nach der Vorkenntnis über das Nutzsignal unterschiedliche Behandlungen.

[0004] Beispielhafte Anwendungen der Erfindung finden sich also in Kommunikationseinrichtungen, in denen die Position eines Sprechers bekannt ist, und in denen Störungen durch Hintergrundgeräusche oder andere Sprecher sowie Lautsprecherabgaben vorhanden sind. Anwendungsbeispiele sind Kfz-Freisprecheinrichtungen, in denen die Mikrofone z.B. im Rückspiegel untergebracht sind und eine sogenannte Richthyperbel auf den Fahrer gerichtet wird. In dieser Anwendung kann eine zweite Richthyperbel auf den Beifahrer gerichtet werden, so dass während eines Telefongesprächs gezielt zwischen Fahrer und Beifahrer hin- und hergeschaltet werden kann.

[0005] In Fällen, in denen die geometrische Position der Nutzsignalquelle zu den aufnehmenden Mikrofonen bekannt ist, ist die geometrische Quellenseparation ein mächtiges Werkzeug. Das Standardverfahren dieser Klasse von "beamforming"-Algorithmen ist das sog. "shift and add" Verfahren, bei welchem auf eines der Mikrofonensignale ein Filter angewendet wird, und das gefilterte Signal sodann zum zweiten Mikrofonensignal hinzuaddiert wird (siehe z.B. Haddad und Benoit, "Capabilities of a beamforming technique for acoustic measurements inside a moving car", The 2002 International Congress and Exposition On Noise Control Engineering, Dearborn, Mi, USA, August 19-21, 2002).

[0006] Eine Erweiterung dieses Verfahrens beschäftigt sich mit adaptiven beam forming" bzw. "adaptiver Quellenseparation", wo die Lage der Quellen im Raum a priori unbekannt ist und durch die Algorithmen erst ermittelt werden muss (WO 02/061732, US6,654,719). Hier ist es das Ziel, die Lage der Quellen im Raum aus den Mikrofonensignalen zu bestimmen und nicht, wie beim "geometrischen" beam forming, fest vorzugeben. Adaptive Verfahren erweisen sich zwar als nützlich, allerdings ist auch hier gewöhnlich a-priori-Information erforderlich, da ein Algorithmus in der Regel nicht entscheiden kann, welche der detektierten Sprachquellen Nutz- und welche Störsignal ist. Nachteilig bei allen bekannten adaptiven Verfahren ist die Tatsache, dass die Algorithmen eine gewisse Adaptionszeit benötigen, bevor ausreichende Konvergenz besteht und die Quellentrennung gelingt. Außerdem sind adaptive Verfahren prinzipiell anfälliger für diffuse Hintergrundstörungen, da diese die Konvergenz erheblich beeinträchtigen können. Ein gravierender Nachteil beim klassischen "shift and add"-Verfahren ist die Tatsache, dass sich mit zwei Mikrofonen lediglich zwei Signalquellen voneinander separieren lassen und die Dämpfung von diffusem Hintergrundschall in der Regel nicht in ausreichendem Maße gelingt.

[0007] Aus der DE 69314514 T2 ist ein Verfahren zur Separierung von Schallsignalen gemäß dem Oberbegriff von Anspruch 1 bekannt. Das in diesem Dokument vorgeschlagene Verfahren führt eine Separation der Schallsignale dergestalt durch, dass ein gewünschtes Nutzschallsignal von Umgebungsrauschen befreit wird, und nennt als Anwendungsbeispiele die Sprachsignale eines Fahrzeuginsassen, die auf Grund des allgemeinen und nicht lokalisierten Fahrzeuglärms nur schwer verständlich sind.

[0008] Zum Herausfiltern des Sprachsignals schlägt dieses Dokument des Stands der Technik vor, mit Hilfe von zwei Mikrofonen jeweils ein Gesamtschallsignal zu messen, jedes der beiden Mikrofonensignale zur Ermittlung seines Frequenzspektrums einer Fouriertransformation zu unterziehen, in mehreren Frequenzbändern basierend auf der jeweiligen Phasendifferenz einen Einfallswinkel des jeweiligen Signals zu bestimmen, und schließlich die eigentliche "Filterung" vorzunehmen. Hierzu wird ein bevorzugter Einfallswinkel bestimmt, und dann eine Filterfunktion, nämlich ein Rauschspektrum, von einem der beiden Frequenzspektren subtrahiert, wobei dieses Rauschspektrum derart gewählt ist, dass Schallsignale aus der Umgebung des bevorzugten Einfallswinkels, der dem Sprecher zugeordnet ist, relativ zu den anderen Schallsignalen, die im wesentlichen Hintergrundlärm des Fahrzeugs darstellen, verstärkt werden. Das derart gefilterte Frequenzspektrum wird anschließend einer inversen Fourier-Transformation unterzogen und als gefiltertes Schallsignal ausgegeben.

[0009] Das in der DE 69314514 T2 offenbarte Verfahren leidet an mehreren Nachteilen:

- a) Die Schallsignalseparation gemäß diesem Dokument des Stands der Technik basiert auf dem vollständigen Entfernen eines Anteils des ursprünglich gemessenen Gesamtschallsignals, nämlich demjenigen Anteil, der als Rauschen bezeichnet wird. Dieses Dokument geht nämlich von einem akustischen Szenario aus, bei dem nur eine einzige Nutzschallquelle vorhanden ist, deren Signale gleichsam eingebettet sind in Störsignale von nicht beziehungsweise weniger lokalisierten Quellen, insbesondere Fahrzeuglärm. Das Verfahren gemäß diesem Dokument

des Stands der Technik erlaubt daher ausschließlich das Herausfiltern dieses einen Nutzsignals durch vollständiges Eliminieren aller Rauschsignale.

In Fällen mit einem einzigen Nutzschallsignal mag das Verfahren gemäß dieses Dokuments zufriedenstellende Ergebnisse liefern. Es kann jedoch auf Grund seines Grundprinzips nicht sinnvoll in Situationen eingesetzt werden, in denen nicht nur eine Nutzschallquelle, sondern mehrere derartige Quellen zum Gesamtschallsignal beitragen. Dies liegt insbesondere daran, dass gemäß dieser Lehre nur ein einziger sog. dominanter Ankunftswinkel verarbeitet werden kann, nämlich derjenige Einfallswinkel, unter dem das energiereichste Schallsignal einfällt. Alle Signale, die unter anderen Ankunftswinkeln auf die Mikrophone fallen, werden zwangsläufig als Rauschen behandelt.

b) Darüber hinaus scheint dieses Dokument selbst davon auszugehen, dass die dort vorgeschlagene Filterung in Form einer Subtraktion des Rauschspektrums von einem der beiden Frequenzspektren noch keine zufriedenstellenden Ergebnisse liefert. Daher sieht dieses Dokument zusätzlich, nämlich unmittelbar vor dieser eigentlichen Filterung, noch eine weitere Signalverarbeitung vor: Es werden nämlich in allen Frequenzbändern, nachdem der dominante Einfallswinkel bestimmt worden ist, durch entsprechende Phasenverschiebung eines der beiden fouriertransformierten Schallsignale in diesem Frequenzband die Rauschanteile im jeweiligen Frequenzband relativ zu den in diesem Frequenzband möglicherweise ebenfalls enthaltenen Nutzschallsignalen abgeschwächt. Somit sieht dieses Dokument die in ihr offenbarte Filterung in Form einer Subtraktion des Rauschspektrums offenbar selbst als ungenügend an, so dass sie selbst weitere, nämlich unmittelbar vorhergehende Signalverarbeitungsschritte vorschlägt, die durch hierfür gesondert bereitgestellte Bauteile vorgenommen werden. Insbesondere benötigt das System zusätzlich zu einer Rauschspektrumsabtraktionsvorrichtung (Vorrichtung 24 in der einzigen Figur dieses Dokuments) vorgeschaltete Mittel 20 zur Phasenverschiebung sowie Mittel 21 zur phasenrichtigen Addition von Spektren in den einzelnen Frequenzbändern (vergleiche die entsprechenden Bauteile in der einzigen Figur dieses Dokuments). Hierdurch werden das Verfahren und die zu seiner Durchführung erforderliche Vorrichtung aufwendig.

[0010] Es ist daher Aufgabe der vorliegenden Erfindung, ein Verfahren zur Separierung von Schallsignalen von einer Mehrzahl von Schallquellen sowie eine entsprechende Vorrichtung vorzuschlagen, die durch den reinen Filterschritt eine ausreichende Qualität der Ausgangssignale erzeugen, ohne zuvor eine phasenrichtige Addition von Schallspektren in verschiedenen Frequenzbändern durchführen zu müssen, um eine zufriedenstellende Separierung zu erzielen, und die es ferner erlaubt, nicht nur Signale einer einzigen Nutzschallquelle von allen anderen Schallsignalen zu befreien, sondern grundsätzlich in der Lage ist, Schallsignale von einer Mehrzahl von Schallquellen ohne Eliminierung separat auszugeben.

[0011] Erfindungsgemäß wird diese Aufgabe durch ein Verfahren nach Anspruch 1 bzw. eine Vorrichtung nach Anspruch 7 gelöst. Vorteilhafte Weiterbildungen der Erfindung sind in den jeweiligen Unteransprüchen definiert.

[0012] Das erfindungsgemäße Verfahren benötigt keine Konvergenzzeit und kann mit zwei Mikrofonen mehr als zwei Schallquellen im Raum separierbaren, sofern diese in ausreichendem Maße räumlich getrennt sind. Das Verfahren stellt nur geringe Anforderungen an Speicherbedarf und Rechenleistung, und es ist sehr stabil gegenüber diffusen Störsignalen. Anders als beim herkömmlichen beam forming lassen sich solche diffusen Störungen effektiv dämpfen. Wie bei allen Zwei-Mikrofon-Verfahren sind die Raumbereiche, zwischen denen das Verfahren differenzieren kann, rotations-symmetrisch zur Mikrofon-Achse, d.h. zu der Geraden, welche durch die beiden Mikrofonpositionen definiert ist. In einem Schnitt durch den Raum, welcher die Symmetrieachse enthält, entspricht der Raumbereich, in dem sich eine Schallquelle befinden muss um als Nutzsignal betrachtet zu werden, einer Hyperbel. Der Winkel ϑ_0 , den der Scheitel der Hyperbel zur Symmetrieachse einnimmt, ist frei wählbar, und die Breite der Hyperbel, welche durch einen Winkel γ_{3db} bestimmt wird, ist ebenfalls ein wählbarer Parameter. Mit nur zwei Mikrofonen lassen sich gleichzeitig Ausgangssignale zu beliebigen, verschiedenen Winkeln ϑ_0 erzeugen, wobei die Trennschärfe zwischen den Bereichen mit dem Überlappingsgrad der entsprechenden Hyperbeln abnimmt. Schallquellen innerhalb einer Hyperbel werden als Nutzsignale betrachtet und mit weniger als ϑ db gedämpft. Störsignale werden in Abhängigkeit ihres Einfallswinkels ϑ eliminiert, wobei eine Dämpfung von >25 db für Einfallswinkel ϑ außerhalb der Akzeptanzhyperbel erreichbar ist.

[0013] Das Verfahren arbeitet im Frequenzbereich. Das einer Richthyperbel zuzuordnende Signalspektrum entsteht durch Multiplikation einer Korrekturfunktion $K_2(x_1)$ sowie einer Filterfunktion $F(f,T)$ mit dem Signalspektrum $M(f,T)$ eines der Mikrofone. Die Filterfunktion entsteht durch spektrale Glättung (z.B. durch Diffusion) einer Zuordnungsfunktion $Z(\vartheta - \vartheta_0)$, wobei im Argument der Zuordnungsfunktion der berechnete Einfallswinkel ϑ einer spektralen Signalkomponente steht. Dieser Einfallswinkel ϑ wird aus dem Phasenwinkel φ des komplexen Quotienten der Spektren der beiden Mikrofon-signale, $M_2(f,T)/M_1(f,T)$, ermittelt, indem man φ mit der Schallgeschwindigkeit c multipliziert und durch $2\pi f d$ dividiert, wobei d den Mikrofonabstand bezeichnet. Das Ergebnis $x_1 = \varphi c / 2\pi f d$, welches zugleich das Argument der Korrekturfunktion $K_2(x_1)$ ist, liefert nach Beschränkung $x = K_1(x_1)$ auf einen Betrag kleiner oder gleich eins den Kosinus des Einfallswinkels ϑ , welcher im Argument der Zuordnungsfunktion $Z(\vartheta - \vartheta_0)$ steht; $K_1(x_1)$ bezeichnet dabei eine weitere Korrekturfunktion.

Kurzbeschreibung der Abbildungen:

[0014]

5 Fig. 1 zeigt die Definition des Einfallswinkels ϑ durch die Positionen beider Mikrofone, deren Signale verarbeitet werden.

Fig. 2 zeigt beispielhaft eine Zuordnungsfunktion $Z(\vartheta)$ mit Halbwertsbreite $2\gamma_{3db}$, aus welcher eine Hyperbel mit Scheitel bei $\vartheta=0$ resultiert.

10 Fig. 3 zeigt eine Hyperbel mit Scheitel bei $\vartheta=\vartheta_0$, die die Richtcharakteristik der Quellenseparation bestimmt. Signale innerhalb des durch die Hyperbel definierten Raumbereichs werden mit einer Dämpfung $<3db$ als Nutzsignal ausgegeben

15 Fig. 4 zeigt den Aufbau des Quellenseparators, in welchem die Zeitsignale zweier Mikrofone, $m_1(t)$ und $m_2(t)$, in einer Stereo-Abtast- und -Fourier-Transformator-Einheit (20) zu Spektren $M_1(f,T)$ und $M_2(f,T)$ transformiert werden, wobei T den Zeitpunkt der Entstehung der Spektren bezeichnet. Aus den Spektren wird in der ϑ -Berechnungseinheit (30) der frequenzabhängige Einfallswinkel $\vartheta(f,T)$ sowie das korrigierte Mikrofonspektrum $M(f,T)$ berechnet, woraus in Signalgeneratoren (40) für verschiedene Richtwinkel ϑ_0 Ausgangssignale $s_{\vartheta_0}(t)$ entstehen.

20 Fig. 5 zeigt den Aufbau der ϑ -Berechnungseinheit (30), in welcher der Phasenwinkel $\phi(f,T)$ einer spektralen Komponente des komplexen Quotienten der beiden Mikrofonspektren $M_1(f,T)$ und $M_2(f,T)$ berechnet wird, welcher sodann mit der Schallgeschwindigkeit c zu multiplizieren und durch $2\pi fd$ zu dividieren ist, wobei d den Mikrofonabstand bezeichnet. Bei dieser Operation entsteht die Größe $x_1(f,T)$, welche das Argument der beiden Korrekturfunktionen K_2 und K_1 darstellt. Mit diesen Korrekturfunktionen entsteht das korrigierte Mikrofonspektrum $M(f,T)=M_1(f,T)*K_2(x_1(f,T))$ sowie die Größe $x(f,T)=K_1(x_1(f,T))$, aus welcher durch Anwendung der Arcuskosinus-Funktion der Einfallswinkel $\vartheta(f,T)$ zu berechnen ist.

25 Fig. 6 zeigt einen Signalgenerator, in welcher eine Zuordnungsfunktion $Z(\vartheta-\vartheta_0)$ mit einem einstellbaren Winkel ϑ_0 durch spektrale Diffusion zu einer Filterfunktion $F(f,T)$ geglättet wird, welche mit dem korrigierten Mikrofonspektrum $M(f,T)$ zu multiplizieren ist. Daraus resultiert ein Ausgangsspektrum $S_{\vartheta_0}(f,T)$, aus welchem durch inverse Fouriertransformation ein Ausgangssignal $s_{\vartheta_0}(t)$ entsteht, welches die Schallsignale innerhalb des durch die Zuordnungsfunktion Z und den Winkel ϑ_0 festgelegten Raumbereichs enthält.

35 Fig. 7 zeigt exemplarisch die beiden Korrekturfunktionen $K_2(x_1)$ und $K_1(x_1)$.

[0015] Ein Grundgedanke der Erfindung ist es, jeder spektralen Komponente des einfallenden Signals zu jedem Zeitpunkt T einen Einfallswinkel ϑ zuzuordnen und allein anhand des berechneten Einfallswinkels zu entscheiden, ob die entsprechende Schallquelle innerhalb einer gewünschten Richthyperbel liegt, oder nicht. Um die Zugehörigkeitsentscheidung etwas abzumildern, wird anstatt einer harten Ja/Nein-Entscheidung eine "weiche" Zuordnungsfunktion $Z(\vartheta)$ (Fig. 2) benutzt, die einen kontinuierlichen Übergang zwischen erwünschten und unerwünschten Einfallsrichtungen erlaubt, was sich vorteilhaft auf die Integrität der Signale auswirkt. Die Breite der Zuordnungsfunktion entspricht dann der Breite der Richthyperbel (Fig. 3). Durch Division der komplexen Spektren der beiden Mikrofonsignale wird zunächst für jede Frequenz f zu einem Zeitpunkt T die Phasendifferenz ϕ berechnet. Mit Hilfe der Schallgeschwindigkeit c und der Frequenz f der entsprechenden Signalkomponente lässt sich aus der Phasendifferenz ein Wegunterschied berechnen, der zwischen den beiden Mikrofonen liegt, wenn das Signal von einer Punktquelle ausgesandt wurde. Ist der Mikrofonabstand d bekannt, ergibt eine einfache geometrische Überlegung, dass der Quotient x_1 aus Wegunterschied und Mikrofonabstand dem Kosinus des gesuchten Einfallswinkels entspricht. In der Praxis ist aufgrund von Störungen wie diffusem Störschall oder Raumhall die Annahme einer Punktquelle selten erfüllt, weshalb x_1 gewöhnlich nicht auf den erwarteten Wertebereich $[-1,1]$ beschränkt ist. Bevor der Einfallswinkel ϑ berechnet werden kann, ist daher noch eine Korrektur erforderlich, die x_1 auf das genannte Intervall beschränkt. Wurde zum Zeitpunkt T für jede Frequenz f der Einfallswinkel $\vartheta(f,T)$ bestimmt, ergibt sich das Spektrum des gewünschten Signals innerhalb einer Richthyperbel mit Scheitel beim Winkel $\vartheta=\vartheta_0$ durch einfache frequenzweise Multiplikation mit dem Spektrum eines der Mikrofone, also $M_1(f,T)K(\vartheta(f,T)-\vartheta_0)$. Unter Umständen ist es vorteilhaft, $K(\vartheta(f,T)-\vartheta_0)$ vor Ausführung der Multiplikation spektral zu glätten. Eine Glättung, deren Ergebnis als $F_{\vartheta_0}(f,T)$ bezeichnet sei, erhält man z.B. durch Anwendung eines Diffusionsoperators. In Fällen, in denen durch Störeinflüsse die Größe x , die zur Berechnung des Einfallswinkels dient, außerhalb ihres Wertebereichs liegt, ist es vorteilhaft, die entsprechende spektrale Komponente des Mikrofonsignals abzuschwächen, da zu vermuten ist, dass sich Störsignale überlagert haben. Dies geschieht z.B. durch Anwendung einer Korrektur

turfunktion, deren Argument die Größe x_1 ist. Sei $M(f, T)$ das korrigierte Mikrofonsignal, dann schreibt sich die Erzeugung des gewünschten Signalspektrums inklusive spektraler Glättung und Korrektur als $S_{\vartheta 0}(f, T) = F_{\vartheta 0}(f, T) M(f, T)$. Aus $S_{\vartheta 0}(f, T)$ entsteht durch inverse Fouriertransformation das Zeitsignal $S_{\vartheta 0}(t)$ für die entsprechende Richthyperbel mit Scheitelswinkel ϑ_0 .

[0016] Anders ausgedrückt ist es eine Grundidee der Erfindung, verschiedene Schallquellen, beispielsweise den Fahrer und den Beifahrer in einem Kraftfahrzeug, räumlich voneinander zu unterscheiden und somit beispielsweise das Nutz-Sprachsignal des Fahrers vom Stör-Sprachsignal des Beifahrers zu separieren, indem man die Tatsache ausnutzt, dass diese beiden Sprachsignale, also Schallsignale, in der Regel auch bei unterschiedlichen Frequenzen vorliegen. Die erfindungsgemäß vorgesehene Frequenzanalyse erlaubt also zunächst, das Gesamt-Schallsignal in die zwei Einzel-Schallsignale (nämlich vom Fahrer und vom Beifahrer) aufzuspalten. Es muß dann "nur noch" mit Hilfe geometrischer Überlegungen anhand der jeweiligen Frequenz jedes der beiden Schallsignale und der zu ermittelnden Phasendifferenz zwischen dem Ausgangssignal des Mikrofons 1 und des Mikrofons 2, die jeweils diesem Schallsignal zugeordnet sind, die Einfallswinkel jedes der beiden Schallsignale berechnet werden. Da die Geometrie zwischen beispielsweise der Position des Fahrers, der Position des Beifahrers und der Position der Mikrofone, etwa in einer Freisprecheinrichtung im Kraftfahrzeug, bekannt ist, kann dann das weiter zu verarbeitende Nutz-Schallsignal aufgrund seines anderen Einfallswinkels vom Stör-Schallsignal separiert werden.

[0017] Es folgt ein detailliertes Ausführungsbeispiel der Erfindung, das anhand der Abbildungen beschrieben wird.

[0018] Die Zeitsignale $m_1(t)$ und $m_2(t)$ zweier Mikrofone, die einen festen Abstand d zueinander haben, werden einem Rechenwerk (10) zugeführt (Fig. 4), wo sie in einer Stereo-Abtast- und -Fourier-Transformator-Einheit (20) mit einer Abtastrate f_A diskretisiert und digitalisiert werden. Eine Folge von a Abtastwerten jeweils eines der Mikrofonsignale $m_1(t)$ und $m_2(t)$ wird durch Fourier-Transformation zum komplexwertigen Spektrum $M_1(f, T)$ bzw. $M_2(f, T)$ transformiert, wobei f die Frequenz der jeweiligen Signalkomponente bezeichnet, und T den Zeitpunkt der Entstehung eines Spektrums angibt. Für die praktische Anwendung ist folgende Parameterwahl geeignet: $f_A = 11025$ Hz, $a = 256$, $T = a/2 = 128$. Wenn Rechenleistung und Speicherplatz es erlauben, ist jedoch $a = 1024$ zu bevorzugen. Der Mikrofonabstand d sollte kleiner sein als die halbe Wellenlänge der höchsten zu verarbeitenden Frequenz, welche sich aus der Abtastfrequenz ergibt, d.h. $d < c/4f_A$. Für die oben angegebene Parameterwahl eignet sich ein Mikrofonabstand $d = 20$ mm.

[0019] Die Spektren $M_1(f, T)$ und $M_2(f, T)$ werden einer ϑ -Berechnungseinheit mit Spektrum-Korrektur (30) zugeführt, die aus den Spektren $M_1(f, T)$ und $M_2(f, T)$ einen Einfallswinkel $\vartheta(f, T)$ berechnet, der angibt, aus welcher Richtung relativ zur Mikrofonachse eine Signalkomponente mit Frequenz f zum Zeitpunkt T in die Mikrofone einfällt (Fig. 1). Dazu wird $M_2(f, T)$ durch $M_1(f, T)$ komplex dividiert. $\varphi(f, T)$ bezeichne den Phasenwinkel dieses Quotienten. Wo Verwechslungen ausgeschlossen sind, wird im folgenden das Argument (f, T) der zeit- und frequenzabhängigen Größen fortgelassen. Die genaue Rechenvorschrift zur Bestimmung von φ lautet gemäß der Eulerschen Formel und den Rechenregeln für komplexe Zahlen:

$$\varphi = \arctan((\operatorname{Re}1 \cdot \operatorname{Im}2 - \operatorname{Im}1 \cdot \operatorname{Re}2) / (\operatorname{Re}1 \cdot \operatorname{Re}2 + \operatorname{Im}1 \cdot \operatorname{Im}2)),$$

wobei $\operatorname{Re}1$ und $\operatorname{Re}2$ die Realteile und $\operatorname{Im}1$ und $\operatorname{Im}2$ die Imaginärteile von M_1 bzw. M_2 bezeichnen. Die Größe $x_1 = \varphi c / 2\pi f d$ entsteht mit Hilfe der Schallgeschwindigkeit c aus dem Winkel φ , auch x_1 ist frequenz- und zeitabhängig: $x_1 = x_1(f, T)$. Der Wertebereich von x_1 muss in der Praxis mit Hilfe einer Korrekturfunktion $x = K_1(x_1)$ (Fig. 7) auf das Intervall $[-1, 1]$ beschränkt werden. Auf die so berechnete Größe x wird durch Anwendung der Arcuskosinus-Funktion ein Einfallswinkel ϑ der betrachteten Signalkomponente errechnet, welcher von der Mikrofonachse zu messen ist, d.h. von der durch die Positionen der beiden Mikrofone definierten Geraden (Fig. 1). Unter Berücksichtigung aller Abhängigkeiten lautet damit der Einfallswinkel einer Signalkomponente mit Frequenz f zum Zeitpunkt T : $\vartheta(f, T) = \arccos(x(f, T))$. Des weiteren wird mit Hilfe einer zweiten Korrekturfunktion $K_2(x_1)$ das Mikrofonspektrum korrigiert (Fig. 7): $M(f, T) = K_2(x_1) M_1(f, T)$. Diese Korrektur dient dazu, in Fällen, in denen die erste Korrekturfunktion greift, die entsprechende Signalkomponente zu reduzieren, da zu vermuten ist, dass sich Störungen überlagert haben, die das Signal verfälschen. Die zweite Korrektur ist optional, alternativ kann auch $M(f, T) = M_1(f, T)$ gewählt werden; $M(f, T) = M_2(f, T)$ ist ebenfalls möglich.

[0020] Das Spektrum $M(f, T)$ wird zusammen mit dem Winkel $\vartheta(f, T)$ einem oder mehreren Signalgeneratoren (40) zugeführt, wo mit Hilfe einer Zuordnungsfunktion $Z(\vartheta)$ (Fig. 2) und einem wählbaren Winkel ϑ_0 jeweils ein auszugebendes Signal $s_{\vartheta 0}(t)$ entsteht. Dies geschieht, indem zu einem Zeitpunkt T jede spektrale Komponente des Spektrums $M(f, T)$ mit der entsprechenden Komponente eines ϑ_0 -spezifischen Filters $F_{\vartheta 0}(f, T)$ multipliziert wird. $F_{\vartheta 0}(f, T)$ entsteht durch spektrale Glättung von $Z(\vartheta - \vartheta_0)$. Diese Glättung erfolgt z.B. durch spektrale Diffusion:

$$F_{\vartheta 0}(f, T) = Z(\vartheta(f, T) - \vartheta_0) + D \Delta_f^2 Z(\vartheta(f, T) - \vartheta_0).$$

5 **[0021]** Dabei bezeichnet D die Diffusionskonstante, welche ein frei wählbarer Parameter größer oder gleich null ist. Der diskrete Diffusionsoperator Δ_f^2 ist eine Abkürzung für

$$10 \Delta_f^2 Z(\vartheta(f, T) - \vartheta_0) = (Z(\vartheta(f - f_A/a, T) - \vartheta_0) - 2Z(\vartheta(f, T) - \vartheta_0) + Z(\vartheta(f + f_A/a, T) - \vartheta_0)) / (f_A/a)^2.$$

15 **[0022]** Der auftretende Quotient f_A/a aus Abtastrate f_A und Anzahl a der Abtastwerte entspricht dem Abstand zweier Frequenzen im diskreten Spektrum. Durch Anwendung des so erzeugten Filters $F_{\vartheta 0}(f, T)$ entsteht ein Spektrum $S_{\vartheta 0}(f, T) = F_{\vartheta 0}(f, T)M(f, T)$, welches durch inverse Fouriertransformation in das Zeitsignal $s_{\vartheta 0}(t)$ übergeht.

[0023] Das von einem Signalgenerator (40) auszugebene Signal $s_{\vartheta 0}(t)$ entspricht dem Schallsignal innerhalb desjenigen Raumbereichs, der durch die Zuordnungsfunktion $Z(\vartheta)$ und den Winkel ϑ_0 definiert ist. Der Einfachheit halber wird in der gewählten Nomenklatur für verschiedene Signalgeneratoren nur von einer Zuordnungsfunktion $Z(\vartheta)$ ausgegangen, verschiedene Signalgeneratoren benutzen lediglich verschiedene Winkel ϑ_0 . Praktisch spricht natürlich nichts dagegen, in jedem Signalgenerator auch eine eigene Form der Zuordnungsfunktion zu wählen. Die Anwendung von Zuordnungsfunktionen, welche über die Zugehörigkeit von Signalkomponenten zu verschiedenen Raumbereichen entscheiden, ist einer der zentralen Gedanken der Erfindung. Eine Zuordnungsfunktion muss eine gerade Funktion sein, geeignete Funktionen sind z.B. $Z(\vartheta) = ((1 + \cos \vartheta)/2)^n$ mit einem Parameter $n > 0$. Der Raumbereich, in welchem Signale mit weniger als 3db gedämpft werden, entspricht einer Hyperbel mit Öffnungswinkel $2\gamma_{3db}$ (Fig. 3) und Scheitel bei dem Winkel ϑ_0 . Hierbei entspricht $2\gamma_{3db}$ dem Halbwertswinkel der Zuordnungsfunktion $Z(\vartheta)$ (Fig. 2), mit der angegebenen Formel für die Zuordnungsfunktion gilt $\gamma_{3db} = \arccos(2^{1-1/n} - 1)$. Bei diesen zweidimensionalen geometrischen Überlegungen ist zu beachten, dass der tatsächliche Bereich des dreidimensionalen Raums, aus welchem mit dem beschriebenen Verfahren Schallsignale extrahiert werden, ein Rotationshyperboloid ist, der durch Rotation der beschriebene Hyperbel um die Mikrofonachse entsteht.

30 **[0024]** Selbstverständlich ist die vorliegende Erfindung nicht auf den Einsatz in Kraftfahrzeugen und Freisprecheinrichtungen beschränkt: Weitere Anwendungen sind Konferenz-Telefonanlagen, bei denen mehrere Richthyperbeln in verschiedene Raumrichtungen gelegt werden, um die Sprachsignale einzelner Personen zu extrahieren und Rückkopplungen bzw. Echo-Effekte zu vermeiden. Des weiteren lässt sich das Verfahren mit einer Kamera kombinieren, wobei die Richthyperbel stets in die gleiche Richtung blickt wie die Kamera, und so nur aus dem Bildbereich kommende Schallsignale aufgezeichnet werden. In Bildtelefonsystemen ist mit der Kamera zugleich ein Monitor verbunden, in den die Mikrofonanordnung ebenfalls eingebaut werden kann, um eine Richthyperbel senkrecht zur Monitor-Oberfläche zu generieren, denn es ist zu erwarten, dass sich der Sprecher vor dem Monitor befindet.

35 **[0025]** Eine ganz andere Klasse von Anwendungen ergibt sich, wenn man anstatt des auszugebenden Signals den ermittelten Einfallswinkel ϑ auswertet, indem man z.B. zu einem Zeitpunkt T über Frequenzen f mittelt. Ein solche $\vartheta(T)$ -Auswertung kann zu Überwachungszwecken benutzt werden, wenn innerhalb eines ansonsten ruhigen Raums die Position einer Schallquelle geortet werden soll.

40 **[0026]** Das richtige "Ausschneiden" des gewünschten Bereichs entsprechend dem zu separierenden Nutz-Schallsignal aus einem Mikrofonspektrum muss nicht, wie in Figur 6 beispielhaft gezeigt, durch Multiplikation mit einer Filterfunktion erfolgen, deren Zuordnungsfunktion den in Figur 2 gezeigten beispielhaften Verlauf hat. Jede andere Art der Verknüpfung des Mikrofonspektrums mit einer Filterfunktion ist geeignet, solange diese Filterfunktion und diese Verknüpfung dazu führen, dass Werte im Mikrofonspektrum umso stärker "gedämpft" werden, je weiter ihr zugeordneter Einfallswinkel ϑ vom bevorzugten Einfallswinkel ϑ_0 (beispielsweise der Richtung des Fahrers im Kraftfahrzeug) entfernt ist.

50 Bezugszeichenliste:

[0027]

- 10 Rechenwerk zur Durchführung der erfindungsgemäßen Verfahrensschritte
- 55 20 Stereo-Abtast- und -Fourier-Transformator Einheit
- 30 ϑ -Berechnungseinheit
- 40 Signalgenerator
- a Anzahl der Abtastwerte, die zu Spektren M1 bzw. M2 transformiert werden

d	Mikrofonabstand
D	Diffusionskonstante, wählbarer Parameter größer oder gleich Null
A_f^2	Diffusionsoperator
f	Frequenz
5 f _A	Abtastrate
K1	erste Korrekturfunktion
K2	zweite Korrekturfunktion
m1(t)	Zeitsignal des ersten Mikrofons
m2(t)	Zeitsignal des zweiten Mikrofons
10 M1(f,T)	Spektrum zum Zeitpunkt T des ersten Mikrofonsignals
M2(f,T)	Spektrum zum Zeitpunkt T des zweiten Mikrofonsignals
M(f,t)	Spektrum zum Zeitpunkt T des korrigierten Mikrofonsignals
S _{ϑ0} (t)	erzeugtes Zeitsignal, entsprechend einem Winkel ϑ ₀ der Richthyperbel
S _{ϑ0} (f,T)	Spektrum des Signals s _{ϑ0} (t)
15 γ _{3db}	Winkel, welcher die Halbwertsbreite einer Zuordnungsfunktion Z(ϑ) bestimmt
φ	Phasenwinkel des komplexen Quotienten M2/M1
ϑ(f,T)	Einfallswinkel einer Signalkomponente, gemessen von der Mikrofonachse
ϑ ₀	Winkel des Scheitels einer Richthyperbel, Parameter in Z(ϑ-ϑ ₀)
x, x1	Zwischengrößen bei der ϑ-Berechnung
20 t	Zeitbasis der Signalabtastung
T	Zeitbasis der Spektrumerzeugung
Z(ϑ)	Zuordnungsfunktion

25 Patentansprüche

1. Verfahren zur Separierung von Schallsignalen von einer Mehrzahl von Schallquellen (S1, S2), umfassend die Schritte:

- 30 - Anordnen von zwei Mikrofonen (MIK1, MIK2) in einem vorbestimmten Abstand (d) zueinander;
 - Erfassen der Schallsignale mit beiden Mikrofonen (MIK1, MIK2) und Erzeugen zugeordneter Mikrofonsignale (m1, m2); und
 - Separieren des Schallsignals einer der Schallquellen (S1) von den Schallsignalen der anderen Schallquellen (S2) basierend auf den Mikrofonsignalen (m1, m2),

35 wobei der Schritt des Separierens die Schritte umfaßt:

- Fourier-Transformieren der Mikrofonsignale zur Ermittlung ihrer Frequenzspektren (M1, M2);
 - Bestimmen der Phasendifferenz (φ) zwischen den beiden Mikrofonsignalen (m1, m2) für jede Frequenzkomponente ihrer Frequenzspektren (M1, M2);
 40 - Bestimmen des Einfallswinkels (ϑ) jedes einer Frequenz der Frequenzspektren (M1, M2) zugeordneten Schallsignals basierend auf der Phasendifferenz (φ) und der Frequenz;
 - Erzeugen eines Signalspektrums (S) eines auszugebenden Signals durch Verknüpfen eines der beiden Frequenzspektren (M1, M2) mit einer Filterfunktion (F_{ϑ0}), die derart gewählt ist, dass Schallsignale aus einer Umgebung (γ_{3dB}) um einen bevorzugten Einfallswinkel (ϑ₀) herum relativ zu Schallsignalen von außerhalb dieser Umgebung (γ_{3dB}) verstärkt werden; und
 45 - inverses Fourier-Transformieren des derart erzeugten Signalspektrums, **dadurch gekennzeichnet, dass** die Filterfunktion (F_{ϑ0}) ϑ-abhängig ist und

50 unter Variation von ϑ ein Maximum bei dem bevorzugten Einfallswinkel (ϑ₀) aufweist, und die Verknüpfung der Filterfunktion (F_{ϑ0}) mit einem der beiden Frequenzspektren eine Multiplikation derselben umfasst.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** die Filterfunktion (F_{ϑ0}) die Form aufweist:

55

$$F_{\vartheta_0}(f,T)=Z(\vartheta-\vartheta_0)+D\Delta^2fZ(\vartheta-\vartheta_0)$$

wobei

f die jeweilige Frequenz

T der Zeitpunkt der Ermittlung der Frequenzspektren (M1, M2)

$Z(\vartheta - \vartheta_0)$ eine gerade Zuordnungsfunktion mit Maximum bei ϑ_0

$D \geq 0$ eine Diffusionskonstante, und

Δ_f^2 ein diskreter Diffusionsoperator derart ist, dass

$$\Delta_f^2 Z(\vartheta(f, T) - \vartheta_0) = (Z(\vartheta(f - f_A/a, T) - \vartheta_0) - 2Z(\vartheta(f, T) - \vartheta_0) + Z(\vartheta(f + f_A/a, T) - \vartheta_0)) / (f_A/a)^2$$

wobei f_A die Abtastrate beim Fourier-Transformieren, und a die Anzahl der Abtastwerte darstellen.

3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet, dass** die Zuordnungsfunktion (Z) die Form aufweist:

$$Z(\vartheta - \vartheta_0) = \left(\frac{1 + \cos(\vartheta - \vartheta_0)}{2} \right)^n$$

wobei $n > 0$ ist.

4. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet, dass** die Bestimmung des Einfallswinkels ϑ über die Beziehung

$$\vartheta = \arccos(x_1(f, T))$$

erfolgt mit

$$x_1(f, T) = \varphi c / 2\pi f d$$

für Werte von x_1 im Bereich $[-1, 1]$

wobei

φ die Phasendifferenz zwischen den beiden Mikrofonsignalkomponenten (M1, M2)

c die Schallgeschwindigkeit

f die Frequenz der Schallsignalkomponente und

d der vorbestimmte Abstand der beiden Mikrofone (MIK1, MIK2) ist.

5. Verfahren nach Anspruch 4, **dadurch gekennzeichnet, dass** es ferner den Schritt umfasst:

Begrenzen des Werts von $x_1(f, T)$ auf das Intervall $[-1, 1]$.

6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet, dass** es ferner den Schritt umfasst:

Reduzieren von Signalkomponenten, für die der Wert von $x_1(f, T)$ vor der Begrenzung außerhalb des Intervalls $[-1, 1]$ lag.

7. Vorrichtung zur Durchführung des Verfahrens nach einem der Ansprüche 1 bis 6, umfassend:

- zwei Mikrofone (MIK1, MIK2);

- eine an die Mikrofone angeschlossene Abtast- und Fourier-Transformationseinheit (20) zum Diskretisieren, Digitalisieren und Fourier-Transformieren der Mikrofonsignale (m_1, m_2);

- eine an die Abtast- und Fourier-Transformationseinheit (20) angeschlossene Berechnungseinheit (30) zur Berechnung des Einfallswinkels (ϑ) jeder Schallsignalkomponente; und

- wenigstens einen an die Berechnungseinheit (30) angeschlossenen Signalgenerator (40) zur Ausgabe des separierten Schallsignals, wobei der wenigstens eine Signalgenerator (40) Mittel zum Multiplizieren einer der Fourier-Transformierten (M1, M2) mit einer Filterfunktion (F_{ϑ_0}) umfasst, die ϑ -abhängig ist und unter Variation von ϑ ein Maximum bei einem bevorzugten Einfallswinkel (ϑ_0) aufweist.

8. Vorrichtung nach Anspruch 7, **dadurch gekennzeichnet, dass** der Abstand (d) zwischen den Mikrofonen der Relation genügt:

$$d < c/4f_A$$

wobei c die Schallgeschwindigkeit und f_A die Abtastfrequenz der Abtast- und Fourier-Transformationseinheit (20) ist.

9. Vorrichtung nach Anspruch 7 oder 8, **dadurch gekennzeichnet, dass** die Vorrichtung für jede zu separierende Schallquelle (S1, S2) einen Signalgenerator (40) umfasst.

Claims

1. Method of separating acoustic signals from a plurality of sound sources (S1, S2), comprising the following steps:

- disposing two microphones (MIK1, MIK2) at a predefined distance (d) from one another;
- picking up the acoustic signals with both microphones (MIK1, MIK2) and generating associated microphone signals (m1, m2); and
- separating the acoustic signal of one of the sound sources (S1) from the acoustic signals of the other sound sources (S2) on the basis of the microphone signals (m1, m2),

in which the separation step comprises the following steps:

- applying a Fourier transform to the microphone signals in order to determine their frequency spectra (M1, M2);
- determining the phase difference (φ) between the two microphone signals (m1, m2) for every frequency component of their frequency spectra (M1, M2);
- determining the angle of incidence (ϑ) of every acoustic signal allocated to a frequency of the frequency spectra (M1, M2) on the basis of the phase difference (φ) and the frequency;
- generating a signal spectrum (S) of a signal to be output by correlating one of the two frequency spectra (M1, M2) with a filter function (F_{ϑ_0}) which is selected so that acoustic signals from an area (γ_{3db}) around a preferred angle of incidence (ϑ_0) are amplified relative to acoustic signals from outside this area (γ_{3db}); and
- applying an inverse Fourier transform to the resultant signal spectrum,

characterised in that the filter function (F_{ϑ_0}) is dependent on ϑ and has a maximum at the preferred angle of incidence (ϑ_0) when ϑ is varied, and the correlation of the filter function (F_{ϑ_0}) with one of the two frequency spectra comprises multiplying the same.

2. Method as claimed in claim 1, **characterised in that** the filter function (F_{ϑ_0}) is expressed as follows:

$$F_{\vartheta_0}(\vartheta, T) = Z(\vartheta - \vartheta_0) + D\Delta^2_{\vartheta} Z(\vartheta - \vartheta_0)$$

in which

f is the respective frequency

T is the instant at which the frequency spectra (M1, M2) are determined

$Z(\vartheta - \vartheta_0)$ is an even allocation function with a maximum at ϑ_0

$D > 0$ is a diffusion constant and

Δ^2_{ϑ} is a discrete diffusion operator, such that

$$\Delta^2_{\mathbf{f}} Z(\mathcal{G}(\mathbf{f}, T) - \mathcal{G}_0) = (Z(\mathcal{G}(\mathbf{f} - \mathbf{f}_A/a, T) - \mathcal{G}_0) - 2Z(\mathcal{G}(\mathbf{f}, T) - \mathcal{G}_0, T) + Z(\mathcal{G}(\mathbf{f} + \mathbf{f}_A/a, T) - \mathcal{G}_0)) / (\mathbf{f}_A/a)^2$$

where $\mathbf{f}_A$ is the sampling rate in applying the Fourier transform, and a is the number of sampling values.

3. Method as claimed in claim 2, **characterised in that** the allocation function (Z) is expressed as follows:

$$Z(\mathcal{G} - \mathcal{G}_0) = \left(\frac{1 + \cos(\mathcal{G} - \mathcal{G}_0)}{2} \right)^n$$

where $n > 0$.

4. Method as claimed in one of claims 1 to 3, **characterised in that** the angle of incidence ϑ is determined by the equation

$$\mathcal{G} = \arccos(x_1(\mathbf{f}, T))$$

with

$$x_1(\mathbf{f}, T) = \varphi c / 2\pi f d$$

for values of x_1 in the interval $[-1, 1]$,

where

φ is the phase difference between the two microphone signal components (M1, M2)

c is the acoustic velocity

f is the frequency of the acoustic signal component and

d is the predefined distance of the two microphones (MIK1, MIK2).

5. Method as claimed in claim 4, **characterised in that** it additionally incorporates the following step:

limiting the value of $x_1(f, T)$ to the interval $[-1, 1]$.

6. Method as claimed in claim 5, **characterised in that** it additionally incorporates the following step:

reducing signal components whose value of $x_1(f, T)$ lay outside of the interval $[-1, 1]$ prior to limitation.

7. Device for implementing the method as claimed in one of claims 1 to 6, comprising:

- two microphones (MIK1, MIK2);
- a sampling and Fourier transform unit (20) connected to the microphones for discretizing and digitising the microphone signals (m_1 , m_2) and applying a Fourier transform to them;
- a calculating unit (30) connected to the sampling and Fourier transform unit (20) for calculating the angle of incidence (ϑ) of every acoustic signal component; and
- at least one signal generator (40) connected to the calculating unit (30) for outputting the separated acoustic signal, at least one signal generator (40) having means for multiplying one of the Fourier transformed frequency spectra (M1, M2) by a filter function (F_{ϑ_0}) which is dependent on ϑ and has a maximum at a preferred angle of incidence (ϑ_0) when ϑ is varied.

8. Device as claimed in claim 7, **characterised in that** the distance (d) between the microphones satisfies the equation:

$$d < c / 4 f_A$$

where c is the acoustic velocity and f_A is the sampling frequency of the stereo sampling and Fourier transform unit (20).

9. Device as claimed in claim 7 or 8, **characterised in that** the device has a signal generator (40) for every sound source (S1, S2) to be separated.

Revendications

1. Procédé de séparation de signaux sonores de plusieurs sources sonores (S1, S2), comprenant les étapes suivantes :

- agencement de deux microphones (MIK1, MIK2) à une distance prédéfinie (d) de l'un de l'autre ;
- détection des signaux sonores avec les deux microphones (MIK1, MIK2) et de production de signaux de microphones associés ($m1$, $m2$) ; et
- séparation entre le signal sonore d'une des sources sonores (S1) et des signaux sonores des autres sources sonores (S2) en se basant sur les signaux de microphones ($m1$, $m2$),

dans lequel l'étape de séparation comprend les étapes suivantes :

- transformation de Fourier des signaux de microphones pour déterminer leurs spectres de fréquences ($M1$, $M2$) ;
- définition du déphasage (φ) entre les deux signaux de microphones ($m1$, $m2$) pour chaque composante de fréquence de leurs spectres de fréquences ($M1$, $M2$) ;
- définition de l'angle d'incidence (ϑ) de chaque signal sonore associé à une fréquence des spectres de fréquences ($M1$, $M2$) en se basant sur le déphasage (φ) et la fréquence ;
- production d'un spectre de signal (S) d'un signal à émettre en associant l'un des deux spectres de fréquences ($M1$, $M2$) avec une fonction de filtrage ($F_{\vartheta 0}$) qui est choisie de manière à amplifier des signaux sonores d'une zone (γ_{3dB}) autour d'un angle d'incidence préféré (ϑ_0) par rapport à des signaux sonores de l'extérieur de cette zone (γ_{3dB}) ; et
- transformation de Fourier inverse du spectre de signal ainsi produit,

caractérisé en ce que :

la fonction de filtrage ($F_{\vartheta 0}$) est dépendante de ϑ et présente, en faisant varier ϑ , un maximum pour l'angle d'incidence (ϑ_0) préféré, et l'association de la fonction de filtrage ($F_{\vartheta 0}$) avec l'un des deux spectres de fréquences comprend une multiplication de ceux-ci.

2. Procédé selon la revendication 1, **caractérisé en ce que** la fonction de filtrage ($F_{\vartheta 0}$) a la forme suivante :

$$F_{\vartheta 0}(f, T) = Z(\vartheta - \vartheta_0) + D \Delta_f^2 Z(\vartheta - \vartheta_0)$$

où

f représente la fréquence respective

T représente le moment de la détermination des spectres de fréquences ($M1$, $M2$)

$Z(\vartheta - \vartheta_0)$ représente une fonction d'affectation paire avec un maximum en ϑ_0

$D \geq 0$ une constante de diffusion, et

Δ^2 représente un opérateur de diffusion discret, de telle sorte que

$$\Delta_f^2 Z(\vartheta(f, T) - \vartheta_0) = (Z(\vartheta(f - f_A/a, T) - \vartheta_0) - 2Z(\vartheta(f, T) - \vartheta_0, T) + Z(\vartheta(f + f_A/a, T) - \vartheta_0)) / (f_A/a)^2$$

où

f_A représente la fréquence d'échantillonnage pour la transformation de Fourier, et
 a représente le nombre d'échantillons.

3. Procédé selon la revendication 2, **caractérisé en ce que**
la fonction d'affectation (Z) a la forme suivante :

$$Z(\vartheta - \vartheta_0) = \left(\frac{1 + \cos(\vartheta - \vartheta_0)}{2} \right)^n$$

où n est > 0.

4. Procédé selon l'une quelconque des revendications 1 à 3, **caractérisé en ce que** :
la définition de l'angle d'incidence ϑ est donnée par la relation :

$$\vartheta = \arccos(x_1(f, T))$$

avec

$$x_1(f, T) = \varphi c / 2\pi f d$$

pour des valeurs de x_1 comprises dans la plage [-1, 1]

où :

φ représente le déphasage entre les deux composantes de signaux de microphones (M1, M2),
 c représente la vitesse du son
 f représente la fréquence des composantes sonores, et
 d représente la distance prédéfinie entre les deux microphones (MIK1, MIK2).

5. Procédé selon la revendication 4, **caractérisé en ce qu'il** comprend en outre une étape de limitation de la valeur
de $x_1(f, T)$ sur l'intervalle [-1, 1].

6. Procédé selon la revendication 5, **caractérisé en ce qu'il** comprend en outre une étape de réduction de composantes
de signaux, pour lesquelles la valeur de $x_1(f, T)$ se situait à l'extérieur de l'intervalle [-1, 1] avant la limitation.

7. Dispositif de mise en oeuvre du procédé selon l'une quelconque des revendications 1 à 6, comprenant :

- deux microphones (MIK1, MIK2) ;
- une unité d'échantillonnage et de transformation de Fourier (20), raccordée aux microphones, servant à effectuer la discrétisation, la numérisation et la transformation de Fourier des signaux des microphones (m1, m2) ;
- une unité de calcul (30) raccordée à l'unité d'échantillonnage et de transformation de Fourier (20) pour le calcul de l'angle d'incidence (ϑ) de chaque composante de signal sonore ; et
- au moins un générateur de signaux (40), raccordé à l'unité de calcul (30), pour fournir le signal sonore séparé, le générateur de signaux (40) comprenant un moyen de multiplication d'une transformation de Fourier (M1, M2) avec une fonction de filtrage (F_{ϑ_0}) qui est dépendante de ϑ et présente, en faisant varier ϑ , un maximum pour l'angle d'incidence (ϑ_0) préféré.

8. Dispositif selon la revendication 7, **caractérisé en ce que** la distance (d) entre les microphones satisfait à la relation
suivante :

$$d < c/4f_A$$

5 où c représente la vitesse du son et f_A la fréquence d'échantillonnage de l'unité d'échantillonnage et de transformation de Fourier (20).

9. Dispositif selon la revendication 7 ou 8, **caractérisé en ce que** le dispositif comprend un générateur de signaux (40) pour chaque source sonore (S1, S2) à séparer.

10

15

20

25

30

35

40

45

50

55

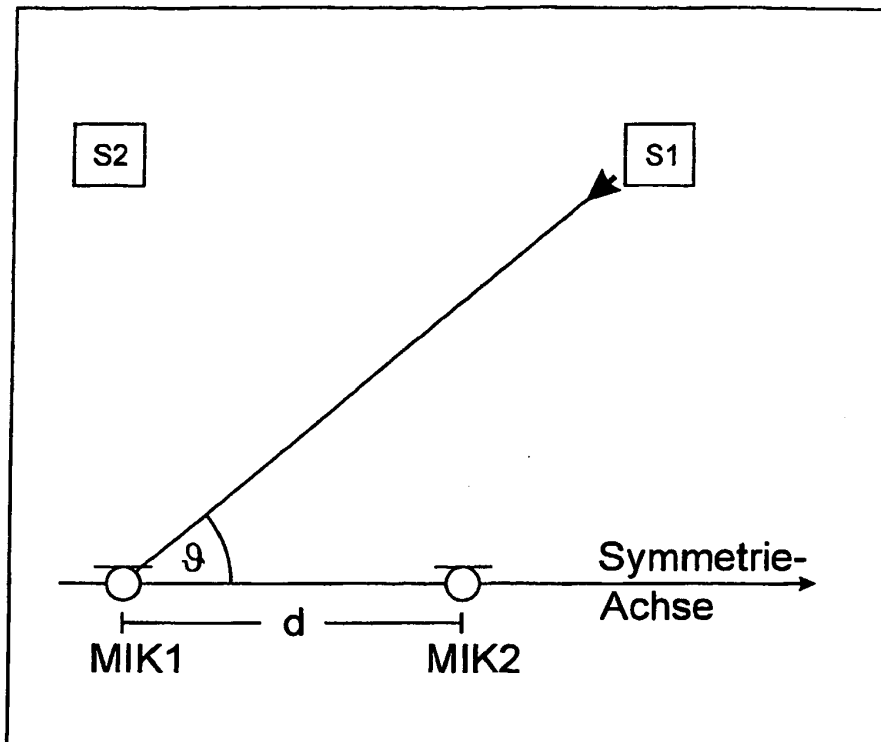


Fig. 1

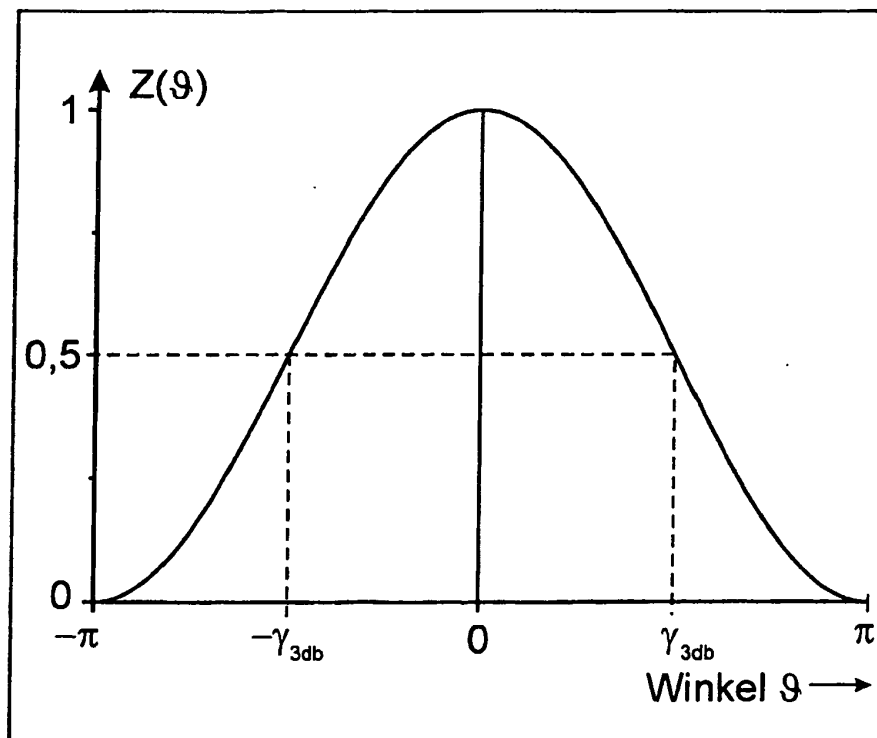


Fig. 2

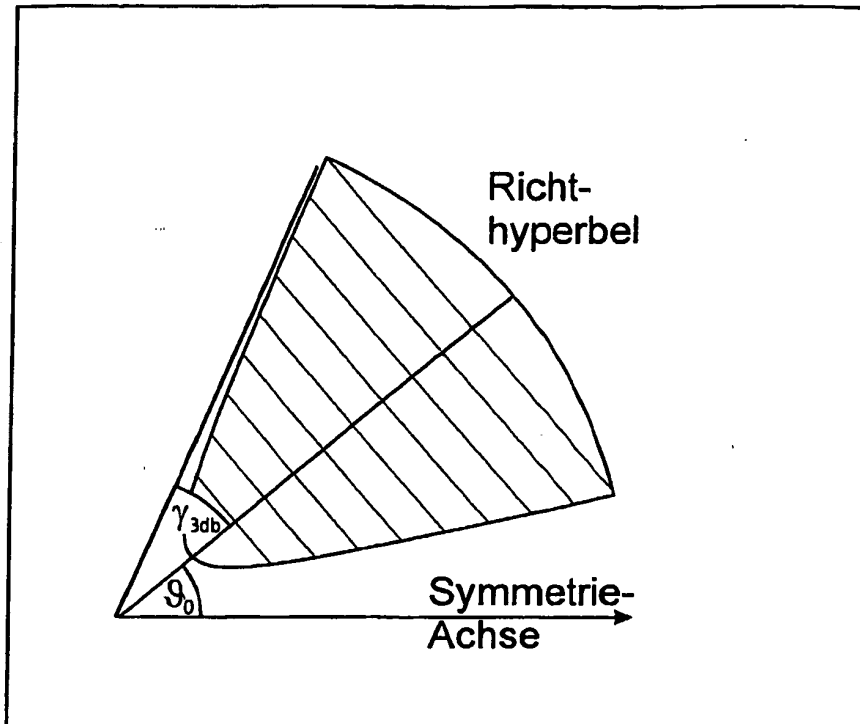


Fig. 3

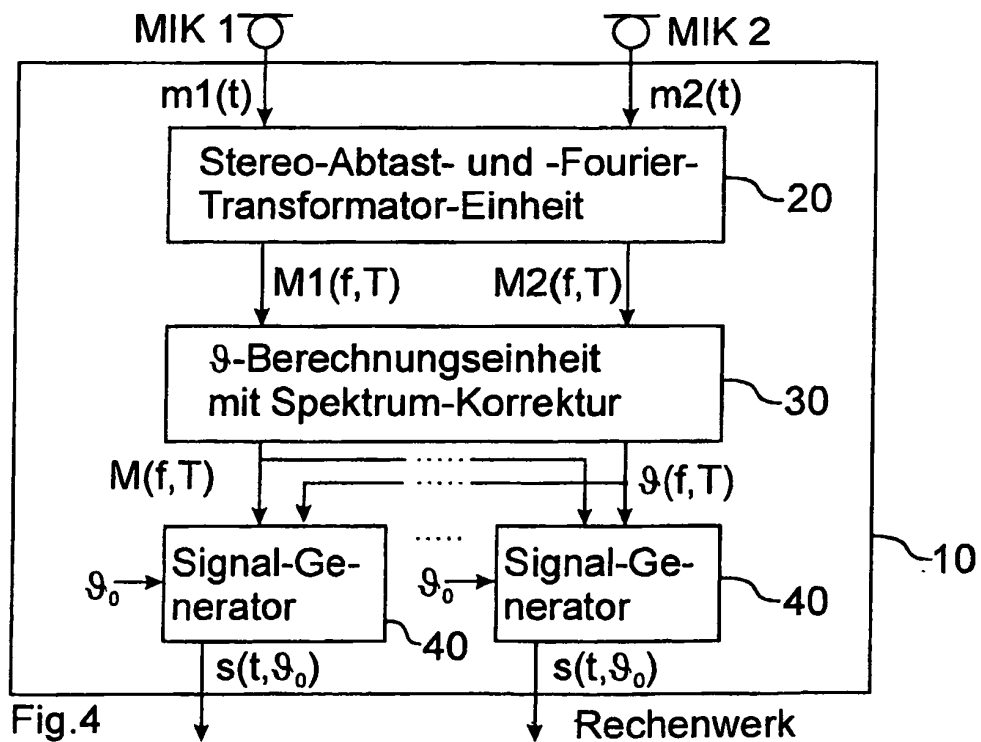
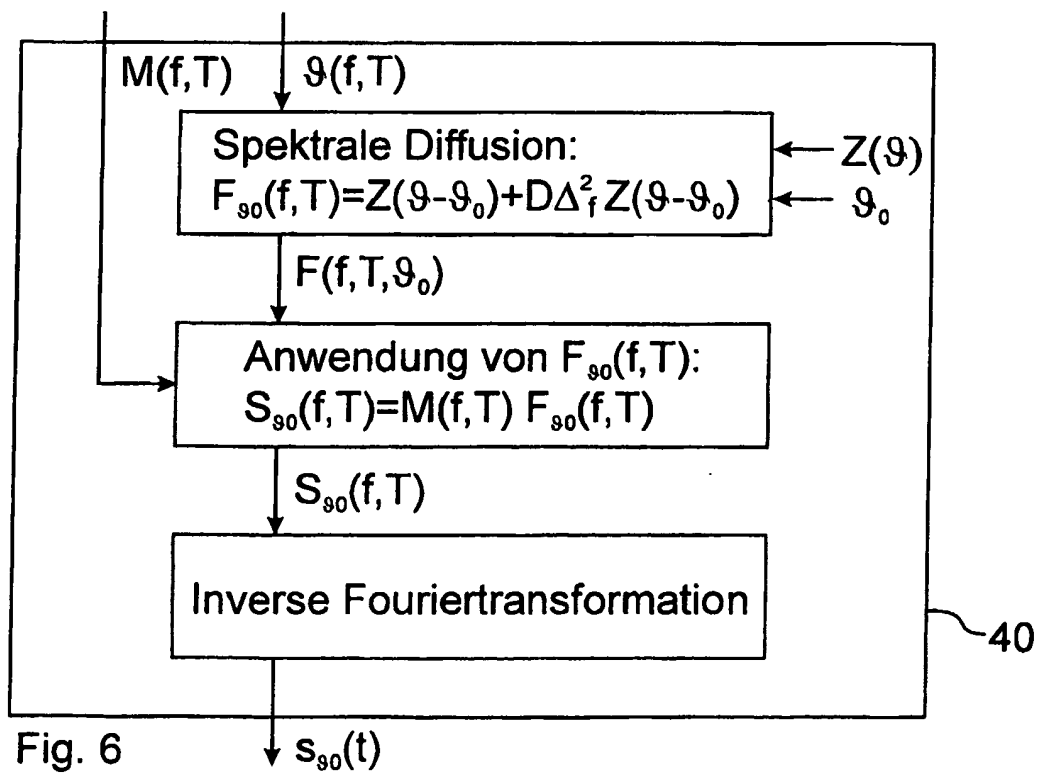
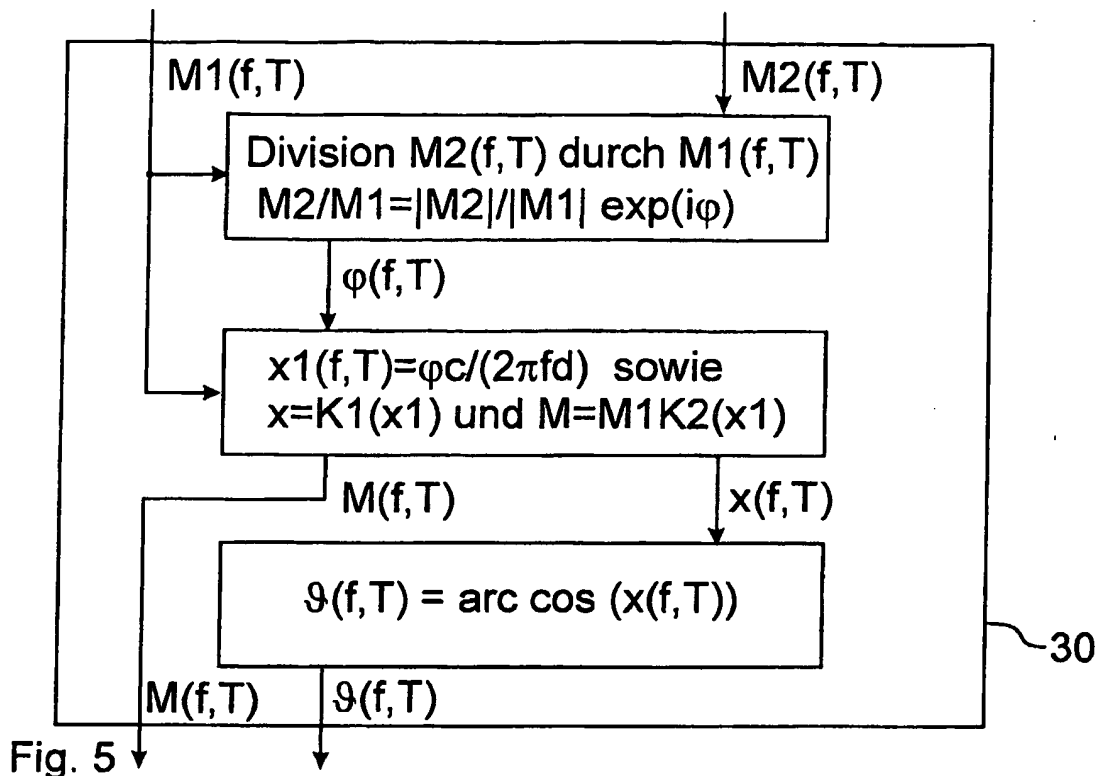


Fig. 4



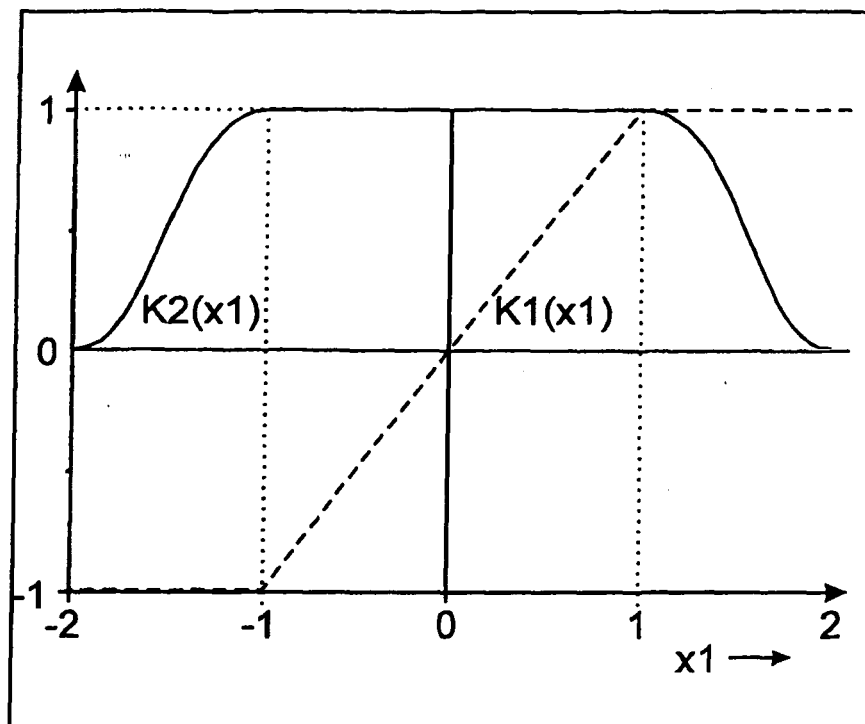


Fig. 7