(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 905 933 A2

(12) EUROPÄISCHE PATENTANMELDUNG

(43) Veröffentlichungstag:
31.03.1999 Patentblatt 1999/13

(51) Int. Cl.[6]: **H04H 7/00**

(21) Anmeldenummer: **97119295.0**

(22) Anmeldetag: **05.11.1997**

(84) Benannte Vertragsstaaten:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(30) Priorität: **24.09.1997 CH 2248/97**

(71) Anmelder:
STUDER Professional Audio AG
CH-8105 Regensdorf ZH (CH)

(72) Erfinder: **Horbach, Ulrich Dr.**
8114 Dänikon (CH)

(54) Verfahren und Vorrichtung zum Mischen von Tonsignalen

(57) Die Erfindung betrifft ein Verfahren und eine Vorrichtung zum Mischen von Tonsignalen. Um dabei mit möglichst geringem technischem Aufwand eine möglichst natürliche Wiedergabe von Tonsignalen über eine Mehrzahl Lautsprecher zu erreiochen, wenn eine dazu abweichende Zahl Schallquellen vorliegt, soll jedes Tonsignal aufgeteilt und wahlweise verzögert, jedes aufgeteilte und wahlweise verzögerte Tonsignal wahlweise gewichtet und mit entsprechenden weiteren Tonsignalen zu einem Zwischensignal (1 bis K) addiert werden. Dann wird jedes Zwischensignal aufgeteilt, gefiltert und anschliessend mit den anderen Zwischensignalen summiert und die summierten Zwischensignale ergeben zusammen ein Ausgangssignal für einen Lautsprecher.

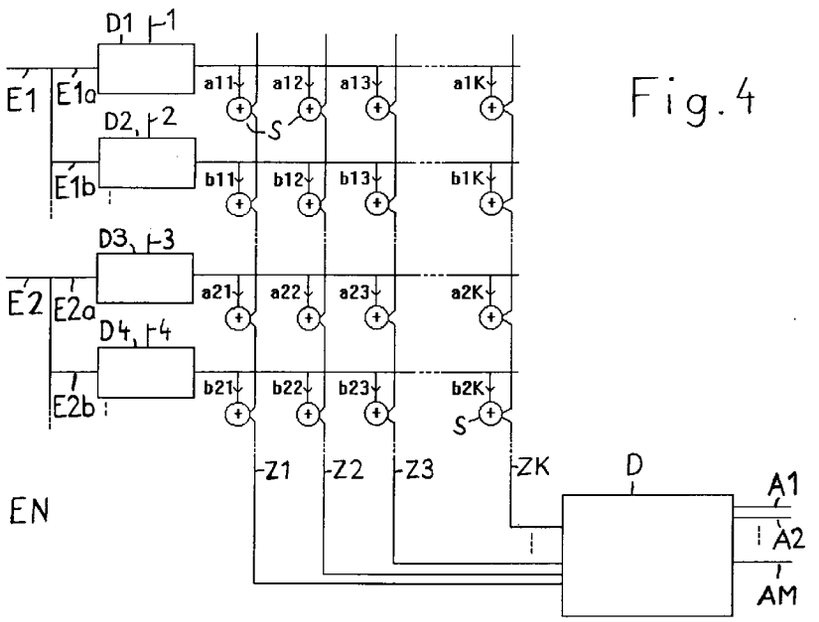

Fig. 4

EP 0 905 933 A2

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren und eine Vorrichtung zum Mischen von Tonsignalen.

[0002] Vorrichtungen dieser Art werden gemeinhin als Mischpulte bezeichnet und dienen der parallelen Verarbeitung von mehreren Tonsignalen. Im Zuge der Einführung neuer Medien (HDTV, Heimkino, DVD) wird die Stereotechnik in Zukunft durch mehrkanalige, sog. „Surround"-Wedergabeverfahren abgelöst werden. Momentan auf dem Markt befindliche Surround-Tonmischpulte enthalten lediglich eine auf mehrere Ausgangskanäle erweiterte Surnmenmatrix. Die von Mono-Mikrofonen generierten N Eingangssignale (typ. N=8..256) werden in den Einzelkanälen 1..N bearbeitet und, jeweils mit Faktoren gewichtet, auf Summenschienen aufgeschaltet. Die Steuerung dieser Faktoren erfolgt mit sog. „Panpots" (Panoramapotentiometern), derart dass eine akustische Positionierung der Tonquelle im Raum erzielt wird. Erlebt der Hörer die Illusion, der Klang enstünde im Raum ausserhalb der Lautsprecher, so spricht man von „Phantomschallquellen".

[0003] Psychoakustische Forschungen und Erfahrungen der letzten Jahre haben gezeigt, dass mit den oben beschriebenen, hier kurz als „Amplituden-Panning" bezeichneten Verfahren nur eine unzureichende Raumabbildung oder Wedergabe eines Schallfeldes in einem Raum in zwei Dimensionen zu erzielen ist. Die Phantomschallquellen können nur auf Verbindungslinien zwischen den Lautsprechern auftreten. Sie sind des weiteren nicht sehr stabil. Ihr Ort ändert sich mit der Position des Hörers. Eine wesentlich natürlichere Wedergabe empfindet man, wenn zwei Aspekte berücksichtigt werden:

a) Die Lautsprechersignale müssen so beschaffen sein, dass der Hörer die gleichen relativen Laufzeitunterschiede und frequenzabhängigen Dämpfungsverläufe zwischen linkem und rechtem Ohrsignal empfängt, wie es beim Hören natürlicher Schallquellen der Fall wäre. Die Ohrsignale müssen in ähnlicher Weise korreliert sein. Bei niedrigen Frequenzen sind in erster Linie Laufzeit-(Phasen-)unterschiede für die Ortbarkeit von Schallereignissen (Lokalisation) wirksam, bei höheren Frequenzen (>1000Hz) hauptsächlich Intensitätsunterschiede. Beim bisher üblichen Amplituden-Panning werden alle Frequenzen gleich gedämpft. Laufzeitunterschiede werden nicht berücksichtigt. Ersetzt man nun die Gewichtsfaktoren durch geeignet dimensionierte variable Filter, so kann beiden Lokalisationsmechanismen Rechnung getragen werden. Wir bezeichnen das Verfahren im folgenden als „Pan-Filterung" (Panoramasteller mittels Filterung).

b) Befindet sich das Schallereignis in einem Raum, so wirken erste bis max. etwa 80msec nach dem Direktschall eintreffende Reflektionen unterstützend für die Lokalisation der Schallquelle. Insbesondere hängt das Entfernungsempfinden davon ab, wie stark der Anteil der Reflektionen bezogen auf den Direktanteil ist. Solche Reflektionen können im Mischpult simuliert oder synthetisch erzeugt werden, indem man das Signal mehrfach verzögert und die so entstandenen Signale durch die oben beschriebenen Pan-Filter unterschiedlichen Richtungen zuordnet. Die Problemstellung bestand nun darin, ein Mischpult zu konstruieren, das die oben beschriebenen Merkmale gemäss a) und b) besitzt, bei vergleichsweise geringem, bezahlbarem technischen Mehraufwand.

[0004] Eine der ersten digitalen Realisierungen wurde von F. Richter und A. Persterer unter dem Titel „Design and applications of a creative audio processor" an der 86. AES Convention in Hamburg 1989 vorgestellt und im Preprint 2782 veröffentlicht. Bei dieser Vorrichtung sind als Pan-Filter direkt Paare sogenannter „Kopfübertragungsfunktionen" (Head Related Transfer Functions = HRTFs) verwendet, d.h. die Filterfunktionen, die man am rechten bzw. linken Ohr misst, wenn man ein Testsignal aus einer bestimmten Raumrichtung aussendet. Jedem Kanalausgangssignal und dessen Echos, erzeugt mittels Verzögerern wird gemäss der jeweiligen Raumrichtung ein passendes HRTF-Paar zugeordnet. Die entstehenden Stereosignale werden auf einen Zweikanal- Summenbus aufgeschaltet. Diese Vorrichtung hat folgende Nachteile:

a) Die Nachbildung einer einzigen Kopfübertragungsfunktion ist bereits sehr aufwendig, wenn eine ausreichende Genauigkeit erzielt werden soll. Es werden nichtrekursive Digitalfilter vom Grad 50..150 oder rekursive vom Grad 10..30 benötigt. Dies macht bereits einen beträchtlichen Teil der verfügbaren Rechenleistung eines modernen digitalen Signalprozessors (DSPs) aus. Für eine natürliche Wiedergabe müssem mehrere Echos pro Einzelkanal simuliert werden (z. B. 5..30), so dass bei hoher Kanalanzahl das Gesamtsystem aufgrund der hohen Anzahl an Filtern nahezu „unerschwinglich" wird.

b) Das binaurale Mischpult liefert am Ausgang lediglich ein für Kopfhörerwiedergabe geeignetes Stereosignal. Man könnte sich eine Anpassung an Lautsprecher- Mehrkanaltechnik vorstellen, indem man die Filter modifizierte und die Anzahl der Summenschienen erhöhte. Der Aufwand würde aber nochmals wesentlich höher.

[0005] D. S. McGrath und A. Reilly haben unter dem Titel „A Suite of DSP Tools for Creation, Manipulation and Playback of Soundfields in the Huron Digital Audio Convolution Workstation" an der 100. AES Convention, Kopenhagen 1996 und im Preprint 4233 eine weitere

Vorrichtung vorgestellt. Bei dieser ist die Anzahl der Summenschienen reduziert, indem man ein Zwischenformat zur Darstellung des Schallfeldes verwendet. Dieses ist unabhängig von der Anzahl bzw. Anordnung der Lautsprecher. Die Umsetzung auf das jeweilige Ausgangsformat erfolgt mit Hilfe eines Dekoders am Summenausgang. Es ist auch das sog. „B-Format" zur Schallfelddarstellung vorgeschlagen, das im zweidimensionalen Fall aus drei Kanälen besteht. Das Signal wird gewichtet mit den Faktoren w, $x = \sin \varphi$, $y = \cos \varphi$ auf den Summenbus aufgeschaltet, wobei w den Signalpegel und φ die Raumrichtung bezeichnen. Der B-Format Dekoder hat die Aufgabe, die Lautsprecher so anzusteuern, dass das Schallfeld in einem Punkt im Raum, dem des Hörers, in optimaler Weise rekonstruiert wird. Folgende Nachteile bezogen auf die hier gegebene Problemstellung sind anzuführen:

Es zeigt sich, dass die erzielbare Lokalisationsschärfe zu gering ist. Benachbarte und gegenüberliegende Lautsprecher strahlen bis auf einen kleinen Pegelunterschied das gleiche Signal ab. Zur Erzielung „diskreter Effekte" wird jedoch eine hohe Kanaltrennung gefordert. Bei einer Filmabmischung soll z.B. ein Geräusch exakt aus einer bestimmten Richtung kommen. Dieses Problem ist einerseits auf das gewählte Schallfeldformat zurückzuführen (zu geringe Kanalanzahl), andererseits auf die Auslegung des Dekoders, der auf die Reproduktion des Schallfelds hin optimiert wurde, nicht hingegen auf optimale Kanaltrennung. Dann ist im Dekoder lediglich eine passive Matrixschaltung vorgesehen. Der Einsatz der eingangs geforderten richtungsabhängigen „Pan-Filter" würde eine wesentlich höhere Anzahl diskret übertragener Richtungen erfordern, wie im folgenden noch näher erläutert wird.

[0006] Die Erfindung, wie sie in den Patentansprüchen wiedergegeben ist, löst somit die Aufgabe, ein Verfahren und eine Vorrichtung zu schaffen, mit denen bei möglichst geringem technischem Aufwand eine möglichst natürliche Wiedergabe von Tonsignalen über eine Mehrzahl Lautsprecher ermöglicht wird, wenn eine dazu abweichende Zahl Schallquellen vorliegt.

[0007] Dies wird dadurch erreicht, dass zum Mischen von 1 bis N Tonsignalen zu 1 bis M Ausgangssignalen, das Tonsignal aus jedem Eingangskanal aufgeteilt und wahlweise verzögert wird, dass jedes aufgeteilte und wahlweise verzögerte Ton- oder Eingangssignal wahlweise gewichtet und mit entsprechenden weiteren Eingangssignalen aus anderen Eingangskanälen zu je einem Zwischensignal 1 bis K addiert wird und dass jedes Zwischensignal in Ausgangskanäle 1 bis M aufgeteilt, gefiltert und anschliessend mit den anderen Zwischensignalen summiert wird. Summierte Zwischensignale ergeben zusammen ein Ausgangssignal für einen Lautsprecher.

[0008] Die Vorrichtung zum Mischen von Tonsignalen aus Eingangskanälen E1 bis EN auf Ausgangskanäle A1 bis AM weist Zwischenkanäle Z1 bis ZK auf, die einerseits über je einen Summierer S und je einen Multiplizierer M mit 1 bis n Teilkanälen jedes Eingangskanales und andererseits mit einem Dekodierer D verbunden sind, der A1 bis AM Ausgangskanäle aufweist. Im Dekodierer D ist jeder Zwischenkanal in eine, der Zahl der Ausgangskanäle entsprechende Anzahl Filterkanäle mit Filtern aufgeteilt und jeder Filterkanal ist über einen Summierer mit einem Filterkanal jedes anderen Zwischenkanales verbunden.

[0009] Die dadurch erreichten Vorteile sind insbesondere darin zu sehen, dass die eingangs definierte Aufgabenstellung in allen Punkten gelöst ist. Insbesondere ist der Aufwand minimal, weil die rechenintensiven Filter nur einmal im System, am Ausgang, benötigt werden.Das vorgeschlagene Schallfeldformat eignet sich hervorragend für die Archivierung von Musikmaterial, da alle gängigen Mehrkanalformate durch Wahl des passenden Dekoders erzeugt werden können. Ebenso können bewegte Quellen in einfacher Weise simuliert werden, da keine Umschaltung von Filtern erforderlich ist.

[0010] Im folgenden wird die Erfindung anhand von einen Ausführungsweg darstellenden Figuren näher erläutert. Es zeigen:

Figur 1, 2 und 3 je ein Schema des Aufbaus einer Vorrichtung gemäss dem Stand der Technik,

Figur 4 ein Schema des Aufbaus einer erfindungsgemässen Vorrichtung,

Figur 5 und 6 je einen Teil der Vorrichtung gemäss Figur 4

Figur 7 und 8 ein Schallfeldformat oder eine Anordnung von Lautsprechern und

Figuren 9, 10 und 11 Frequenzgänge, die mit der Erfindung erzielt werden können.

[0011] Fig. 1 zeigt eine bekannte und in der Einleitung bereits erwähnte Vorrichtung mit Kanälen K1, K2 bis KN für Eingangssignale beispielsweise aus einem Mikrophon und Kanälen A1, A2, A3, A4, A5 usw. für Ausgangssignale beispielsweise für eine entsprechende Anzahl Lautsprecher. Die Kanäle K1 bis KN sind über hier nicht gezeigte Multiplikatoren für Faktoren a11 bis aN5 und Summierer S mit den Kanälen oder Summenschienen A1, A2, A3, A4, A5, usw. verbunden. Wie bereits eingangs beschrieben stellt diese Vorrichtung eine sog. Summenmatrix-Schaltung dar, bei der Eingangssignale direkt durch die Multiplizierer gewichtet und den Summenschienen A1, A2, A3, A4, A5 zugeleitet werden. Damit ist für jeden Lautsprecher ein Signal verfügbar, das aus mehreren Eingangssignalen zusammengesetzt ist, wobei der Anteil eines Eingangssignales im Ausgangssignal der Summenschiene A1, A2 usw. durch einen Multiplikationsfaktor a11 bis aN5 bemessen wird.

[0012] Fig. 2 zeigt eine weitere bekannte und in der Einleitung bereits erwähnte Vorrichtung bei der hier nur einer von vielen möglichen Eingangskanälen E1 gezeigt ist. Dieser ist aufgeteilt in weitere Kanäle e11, e12 usw. in die eine Verzögerungsschaltung V1, V2 usw. eingesetzt ist. Ausgänge jeder dieser Verzögerungsschaltungen V1, V2 münden in je eine Schaltung HRTF1 - 4 zur Realisierung einer Kopfübertragungsfunktion. Ausgänge dieser HRTF-Schaltungen wiederum sind über Summierer S an zwei Summenbusse B1, B2 für zweikanal Stereowiedergabe angeschlossen. Dies entspricht dem eingangs erwähnten binauralen Mischpult gemäss Richter und Persterer.

[0013] Fig. 3 zeigt eine dritte bekannte Vorrichtung nach D. McGrath bei der, wie ebenfalls eingangs erwähnt, ein Eingangssignal aus einem Kanal E mehrfach aufgeteilt und in Verzögerungsschaltungen Ve verzögert, bekannterweise mit Faktoren w1 bis y2 multipliziert oder abgeschwächt und mit Summierern S auf drei Kanäle Kw, Kx und Ky gelangen und dort die bereits erwähnten Signale w, x und y bilden. Ein Dekoder BD wandelt diese Signale w, x und y zu Eingangssignalen für z.B. fünf Lautsprecher um.

[0014] Fig. 4 zeigt schematisch die erfindungsgemässe Vorrichtung mit zwei Eingangskanälen E1, E2, die aber bis zu EN Eingangskanälen erweitert sein kann. Jeder Eingangskanal E1, E2 usw. ist wieder aufgeteilt in mehrere Kanäle E1a, E1b, E2a, E2b, usw. wobei auch hier die Aufteilung in n Kanäle erfolgen kann. In jeden dieser Kanäle ist eine Verzögerungsschaltung D1, D2, D3, D4 usw. eingesetzt, die über Steuereingänge 1, 2, 3, 4 ansteuerbar sind. Zwischenkanäle Z1 bis ZK sind über Summierer S mit jedem Eingangskanal E1a, E1b, E2a, E2b bis ENn verbunden, wobei jedem Summierer noch ein hier nicht gezeigter Multiplizierer vorgeschaltet ist (siehe auch Fig. 6). Alle Zwischenkanäle Z1 bis ZK münden in einen Dekodierer D, dessen Ausgänge Ausgangskanäle A1, A2 bis AM bilden.

[0015] Fig. 5 zeigt ein Schema für den Aufbau des Dekoders D gemäss Fig. 4. Dieser weist die gleiche Anzahl Eingänge wie Zwischenkanäle Z1 bis ZK auf. Hier ist lediglich ein einziger Eingang oder Zwischenkanal Z1 dargestellt. Jeder Zwischenkanal ist wiederum in soviele Filterkanäle aufgeteilt, wie es Ausgänge geben soll. Deshalb sind diese Filterkanäle mit denselben Bezugszeichen A1, A2, AM wie die Ausgangskanäle in Fig. 4 versehen. In jedem Filterkanal oder Ausgangskanal A1 bis AM ist im Dekoder D je ein IIR-Filter und ein FIR-Filter eingesetzt und in Serie geschaltet. In jedem Filter- oder Ausgangskanal A1 bis AM ist vor dem Ausgang des Dekoders D ein Summierer S1, S2, SM vorgesehen. Die Summierer S weisen jeweils soviele Eingänge auf, wie es Zwischenkanäle Z1 bis ZK gibt.

[0016] Fig. 6 zeigt einen Summierer S, der hier beispielsweise an den Zwischenkanal Z1 angeschlossen ist, mit vorgeschaltetem Multiplizierer M, der einen Eingang für Faktoren a11, a12, usw., wie aus der Fig. 4 bekannt, aufweist und der an einen Eingangskanal, hier E1a, angeschlossen ist.

[0017] Fig. 7 zeigt das zur Zeit wichtigste standardisierte Surround-Format. Es besteht aus einem, direkt vor einem als Kreis dargestellten Hörer 15 angeordneten „Center - Lautsprecher" 20 (Aufstellungswinkel 0°), zwei im gleichen Abstand zum Hörer befindlichen Stereolautsprechern 21, 22 im Winkel vom +/- 30°, sowie zwei rückwärtigen Surroundlautsprechern 23, 24 im Winkel von +/- (110..130)°. Bei Musikwiedergabe dienen die Frontlautsprecher 20, 21, 22 zur Übertragung der Schallereignisse, so dass eine Bühne entsteht. Die rückwärtigen Systeme 23, 24 strahlen vorwiegend diffuse Raumechos ab.

[0018] Vorne wird also eine wesentlich präzisere Abbildungsfähigkeit gefordert. Dieser Tatsache kann bei der Auswahl der Raumrichtungen Rechnung getragen werden, indem man die Auflösung entsprechend unterschiedlich wählt. Sehr gute Ergebnisse erzielt man bereits mit K = 9 Kanälen, mit den folgenden Intervallgrenzen:

Kanal 1: hinten links
Kanal 2: -37.5°..-52.5°
Kanal 3: -22.5°..-37.5°
Kanal 4: -7.5° .. -22.5°
Kanal 5: -7.5° .. 7.5°
Kanal 6: 7.5° .. 22.5°
Kanal 7: 22.5° .. 37.5°
Kanal 8: 37.5° .. 52.5°
Kanal 9: hinten rechts.

[0019] Fig. 8 zeigt einen Kopf eines Hörers 25, hier als Kreis dargestellt und einen, aus einer Quelle in einem Winkel α einfallenden Strahl 26.

[0020] Fig. 9 zeigt resultierende Amplitudenfrequenzgänge eines auf 30° normierten kopfbezogenen Filterpaares für unterschiedliche Schalleinfallswinkel. Je nach Einfallswinkel des Schalls, der auf einen Hörer (Kopf) trifft, ergeben sich unterschiedliche Frequenzgänge 10 bis 14 für die Amplituden eines Signals, das von einem Lautsprecher abgestrahlt wird. Der Lautsprecher, der sich in der gleichen Halbebene wie das einfallende Schallsignal befindet, strahlt „Direktanteile" ab, der gegenüberliegende dagegen „Indirektanteile". Wegen der Normierung ergibt sich für ein aus dem Winkel 30° kommendes Signal direkt abgestraht der lineare Frequenzgang 9. Mit 10 ist ein Frequenzgang für Schall der in einem Einfallswinkel von 15° direkt, mit 11 von 0°, mit 12 von 15° indirekt, mit 13 von 30° indirekt und mit 14 von 60° indirekt abgestrahlt gezeigt.

[0021] Fig. 10 zeigt einen Frequenzgang für die Laufzeit für ein Signal aus drei festgelegten Raumrichtungen mit 15°, 22.5° und 30° Einfallswinkel, wobei längs der Abszisse Werte für Frequenzen von 10 - 100000 Hertz und längs der Ordinate Werte für Zeitverzögerungen aufgetragen sind.

[0022] Fig. 11 zeigt resultierende Amplitudenfre-

quenzgänge des Indirektanteils für ein Signal aus drei Raumrichtungen, wobei längs der Abszisse Werte für Frequenzen und längs der Ordinate Werte für die Dämpfung der Amplituden in dB aufgetragen sind. Dies für Signale aus den Raumrichtungen 15°, 22.5° und 30°.

[0023] Die Wirkungsweise der Erfindung ist dabei wie folgt:

[0024] Sollen beispielsweise zwei Eingangssignale über die Vorrichtung gemäss Fig. 4 in M = 5 Ausgangssignale für fünf Lautsprecher umgewandelt werden, so werden die beiden Eingangssignale E1 und E2 zunächst beispielsweise in je zwei Eingangssignale E1a, E1b und E2a, E2b aufgeteilt. Die Eingangssignale E1a und E2a sollen nicht verzögert werden, weil sie zur direkten, nicht reflektierten Abstrahlung an den Hörer vorgesehen sind. Sie erhalten deshalb eine Verzögerung mit dem Betrag Null. Die Eingangssignale E1b und E2b sollen reflektierte und deshalb eine grössere Laufzeit aufweisende Signale bilden oder simulieren. Deshalb werden sie in den Verzögerungsschaltungen D2 und D4 mit einer bestimmten Verzögerung versehen. Entsprechend dem in Fig. 7 gezeigten Format, sind beispielsweise neun Zwischenkanäle Z1 bis Z9 vorgesehen. Der Bediener der Vorrichtung, also des beispielsweise so aufgebauten Mischpultes legt nun einerseits die vorgenannte Verzögerung sowie die Faktoren a11 bis b2K fest, wobei er sich durch weiter unten wiedergegebene Gesichtspunkte leiten lässt.

[0025] Dann stehen am Dekodierer D (siehe Beisp. Fig. 7) neun Zwischensignale Z1 bis ZK an, von denen jedes wieder in M = 5 Signale A1 - AM aufgeteilt und zuerst im IIR und dann im FIR-Filter gefiltert werden. Anschliessend werden die aufgeteilten Signale A1 bis AM mit dem entsprechenden Signal A1 bis AM aus den anderen Zwischenkanälen Z2 bis ZK summiert, was in unserem Beispiel aus zunächst 5 X 9 = 45 Signalen wieder fünf Ausgangssignale A1 bis AM ergibt.

[0026] Mit anderen Worten werden von N Eingangskanälen mit Verzögerungsgliedern Echos erzeugt und die direkten Signalanteile (im allgemeinen ist Delay 1 = 0) werden gewichtet mit den Faktoren a11, b11, ... auf K Summenschienen aufgeschaltet, welche unmittelbar bestimmten, beliebig wählbaren Raumrichtungen zugeordnet sind. Ebenso werden Echos mit Faktoren b11 .... b1K usw. auf die Summenschienen aufgeschaltet. Ein Dekoder D setzt das entstandene Summensignal Z1 bis ZK in das jeweils gewünschte Lautsprecherformat um. Die Auflösung vorn beträgt dabei 15°. Die Gewichtsfaktoren a11....b2K werden wie folgt bestimmt: Gemäss der Zuordnung zu einer bestimmten Raumrichtung sind maximal zwei der K Faktoren von Null verschieden. Soll das Signal aus einem Winkel φ (Fig. 7) kommen, der nicht exakt in der Mitte der definierten Winkelintervalle liegt, so wird eine Gewichtung gemäss den Funktionen $0.5\,(1-\cos \pi x)$ und $0.5(1+\cos \pi x)$, $x \varepsilon (0, 1)$ vorgenommen. Diese entsprechen konventionellen Amplituden-Panning Funktionen, mit dem Unterschied, dass die Summe der Funktionen konstant eins ist (nicht die Summe der Quadrate). Im oben angegebenen Fall 1 würden z. B. bei $\varphi = 22.5°$ (genau an der Grenze der Intervalle der Kanäle 6 und 7, d.h. x = 0.5) folgende Werte resultieren (w entspricht dem gewünschte Pegel):

$a_1 = 0,\ a_2 = 0,\ a_3 = 0,\ a_4 = 0,\ a_5 = 0,\ a_6 = 0.5\,w,\ a_7 = 0.5\,w,\ a_8 = 0,\ a_9 = 0.$

[0027] Es ist insbesondere darauf hinzuweisen, dass der Dekoder D (Fig. 5) nur einmal im System (am Summenausgang) benötigt wird. Alle i Summensignale (i = 1..K) werden jeweils über M Filterstrecken geführt, deren Ausgangssignale die Lautsprecher L_1 .. L_m ansteuern. Entsprechend gefilterte Einzelsignale werden dabei aufaddiert. Die Filter sind als kopfbezogene Filter ausgelegt, wobei die Abschattung des Kopfes relativ zu einer Referenzrichtung (z.B. 0° oder 30°) simuliert wird. Dies berücksichtigt die eingangs erläuterte Regel, dass die Lautsprecher der Natur entsprechend korrelierte Signale abstrahlen sollen. Realisiert werden also, bezogen auf diese Richtung, normierte Kopfübertragungsfunktionen. Man erhält die in Fig. 9 dargestellten typischen Frequenzgänge. Gezeigt sind die zugewandte Seite des Kopfes („direkt") sowie die abgewandte Seite („indirekt"). Mit zunehmender Abschattung nimmt die Dämpfung hoher Frequenzen zu. Die Filter basieren auf einem einfachen Kopfmodell (Kugel). Der Vorteil dieser Wahl besteht darin, dass die empfundene Klangfarbe unabhängig vom individuellen Hörer und der exakten Hörposition weitgehend neutral bleibt.

[0028] Wesentlicher Bestandteil der Erfindung ist, dass die Filter, wie in Fig. 5 dargestellt, aufgeteilt werden. Ein rekursives Filter (IIR-Allpass) modelliert die interauralen Laufzeitunterschiede bis zu einer bestimmten oberen Grenzfrequenz (siehe Fig. 10). Ein linearphasiges FIR-Filter modelliert unabhängig davon die Intensitätsunterschiede, wie z. B. in Fig. 9 dargestellt. Mit dieser Anordnung vermeidet man unerwünschte Kammfilter-effekte, die entstehen, wenn man zwei unterschiedlich verzögerte Signale addiert. Man würde dann oberhalb einer bestimmten Frequenzgrenze Auslöschungen an den Stellen bekommen, wo die Phasendifferenz 180° erreicht. Daher wird hier keine konstante, sondern eine frequenzabhängige Laufzeit realisiert, die bei hohen Frequenzen gegen Null strebt. Ordnet man einem Signal einen Raumwinkel zu, der genau an der Grenze zwischen zwei Intervallen liegt, wie im Beispiel oben, so erhält man die in Fig. 10 bzw. Fig. 11 dargestellten Frequenzgänge. Man erkennt, dass trotz der zugrundeliegenden, relativ geringen Kanalanzahl von 9 eine sehr gute Interpolation erzielt wird. Dies bedeutet, dass eine Schallquelle quasi kontinuierlich im Raum bewegt werden kann, trotz der niedrigen Anzahl fest implementierter Kopfübertragungsfunktionen.

[0029] Die Auslegung der Filter im Dekoder hat vorzugsweise folgendermassen zu geschehen:

Der Entwurf soll am gegebenen Beispiel mit 9 Schall-

feldsignalen und 5 Lautsprechern erklärt werden (siehe Fig. 7). Mit Ausnahme der Kanäle 1 und 9, die hier direkt, ohne Filter zu durchlaufen, mit den entsprechenden rückwärtigen Lautsprechern verbunden werden, sind die in Fig. 5 gezeigten Filter von Kopfübertragungsfunktionen abgeleitet, welche gemäss Fig. 8 definiert sind. Die Filterfunktion $H_{(D, a)}$ bezeichnet die an dem der Schallquelle zugewandten Ohr auftretende Übertragungsfunktion, $H_{(I, a)}$ die der gegenüberliegenden Seite. Die Funktionen sind von der Einfallsrichtung α abhängig, die gezählt wird beginnend vom rechten Ohr gegen den Uhrzeigersinn. Solche Funktionen erhält man z. B. durch Messungen an Versuchspersonen, Kunstköpfen oder durch Berechnung an einfachen Kopfmodellen wie von D. H. Cooper in: „Calculator Program for Head-Related Transfer Function" Journal Audio Engineering Society (AES), No 37, 1989, pp 3-17 oder von B. Gardner, K. Martin in: „Measurements of a KEMAR dummy head" Internet http://sound.media.mit.edu/KEMAR.html dargelegt. Letzteres ist für die hier vorliegende Anwendung der Lautsprecherwiedergabe besonders zu empfehlen, da eine vom jeweiligen Hörer unabhängige Wiedergabequalität resultiert. Beim Entwurf der Filter wird nach folgendem Schema vorgegangen:

1) Wahl einer Referenzrichtung $\alpha 0$ zur Normierung. Man erhält für jede Einfallsrichtung α das Filterpaar $H_1 = H_{(D, a)}/H_{(D, a0)}$ und $H_2 = H_{(I, a)}/H_{(I, a0)}$ Sinnvoll sind z. B. die Wahl von $\alpha 0=30°$ (Normierung auf den Winkel der Stereolautsprecher vorn) oder $\alpha 0=0°$ (Normierung auf frontalen Schalleinfall).

2) Approximation der Beträge von H_1 und H_2 durch Übertragungsfunktionen von rekursiven Filtern niedrigen Grades, z. B. Grad 1-6. Man kaskadiert hierfür eine ausreichende Anzahl von Filtern 1. und 2. Grades, für welche man vorab geeignete Typen wählt (Peak-Notch, Shelving). Mittels einschlägig erhältlicher nichtlinearer Optimierungsprogramme variiert man die Parameter (z.B. Güte, Grenzfrequenz, Verstärkung), bis sich eine optimale Annäherung an endlich vielen Punkten auf einer logarithmischen Frequenzskala ergibt. Werte für die Güte sind dabei nach oben auf Werte bis max. ca. 4 zu beschränken. Sinn dieser Massnahme ist die Gewinnung geglätteter, von Resonanzen hoher Güte befreiter Filter. Dies resultiert in neutraler, verfärbungsarmer Wiedergabe. Die für das Hören wichtige Korrelation der links und rechtsseitig abgestrahlten Lautsprechersignale bleibt dabei erhalten. Die Entwürfe sind für alle Raumwinkel in der Intervallmitte der Schallfeldkanäle durchzuführen, d.h. im vorliegenden Beispiel (Fig. 7) α = +/- (0°, 15°, 30°, 45°).

3) Die linearphasigen FIR Filter (nichtrekursiv) erhält man, indem man die Impulsantworten der in (2) erhaltenen rekursiven Filter mit einem Zeitfenster bewertet (z.B. Rechteckfenster der Länge 100) und symmetrisch fortsetzt.

4) Die IIR-Allpässe approximieren die Schall-Laufzeiten des Direktanteils, t_D, zum rechten bzw. Indirektanteils, t_I, zum linken Ohr beim Schalleinfallswinkel α. Abhängig vom effektiven Kopfdurchmesser h erhält man aufgrund einfacher geometrischer Überlegungen $t_I - t_D = h \sin(90° - \alpha)$. Die IIR-Filter sind kaskadierte Allpässe 2. Grades, die man aus dem Nennerpolynom eines Bessel-Tiefpasses konstruiert. Grenzfrequenz und Filtergrad sind so zu optimieren, dass sich günstige Verläufe der in Fig. 11 dargestellten Interpolationsfunktionen ergeben, die dem Frequenzgang von einem Mischpult-Eingangssignal (Fig. 4) zum Lautsprecherausgang entspricht, wenn man einen Raumwinkel an der Grenze zweier Intervalle der Schallfeldkanäle wählt.

5) Mit je einem nach 1) - 4) gewonnenen Filterpaar werden die vorderen Stereolautsprecher gemäss Fig. 5 angesteuert. Der in der Mitte vorn aufgestellte „Center Lautsprecher" wird, je nach gewählter Normierung, ungefiltert angesteuert (im Fall der 0° Normierung) oder über ein festes Filter $H_{(D, 0)}/H_{(D, 30)}$.

Patentansprüche

1. Verfahren zum Mischen von (1 bis N) Tonsignalen, dadurch gekennzeichnet, dass

   jedes Tonsignal aufgeteilt und wahlweise verzögert wird, dass
   jedes aufgeteilte und wahlweise verzögerte Tonsignal wahlweise gewichtet und mit entsprechenden weiteren Tonsignalen zu einem Zwischensignal (1 bis K) addiert wird, dass
   jedes Zwischensignal aufgeteilt, gefiltert und anschliessend mit den anderen Zwischensignalen summiert wird und dass
   summierte Zwischensignale zusammen ein Ausgangssignal ergeben.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass beim Filtern interaurale Laufzeitunterschiede modelliert werden.

3. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass beim Filtern interaurale Intensitätsunterschiede modelliert werden.

4. Verfahren nach Anspruch 2 und 3, dadurch gekennzeichnet, dass Intensitätsunterschiede und Laufzeitunterschiede unabhängig voneinander modelliert werden.

5. Vorrichtung zum Mischen von Tonsignalen aus Eingangskanälen (E1 bis EN) auf Ausgangskanäle (A1 bis AM), gekennzeichnet durch Zwischenkanäle

(Z1 bis ZK), die einerseits mit (1 bis n) Teilkanälen jedes Eingangskanales und andererseits mit einem Dekodierer (D) verbunden sind, der (A1 bis AM) Ausgangskanäle aufweist.

6. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, dass im Dekodierer jeder Zwischenkanal in eine, der Zahl der Ausgangskanäle entsprechende Anzahl Filterkanäle (A1 bis AM) mit Filtern (IIR, FIR), aufgeteilt ist und dass jeder Filterkanal über einen Summierer (S1 bis SM) mit einem Filterkanal jedes anderen Zwischenkanales verbunden ist.

7. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, dass die Zwischenkanäle über je einen Summierer (S) und je einen Multiplizierer (M) mit Teilkanälen verbunden sind.

8. Vorrichtung nach Anspruch 6, dadurch gekennzeichnet, dass als Filter eine Serieschaltung eines IIR-Filters und eines FIR-Filters vorgesehen ist.

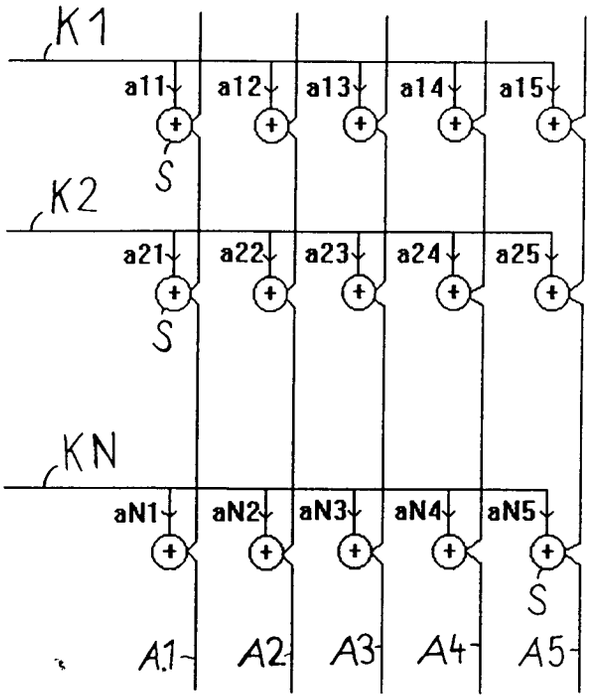

Fig.1

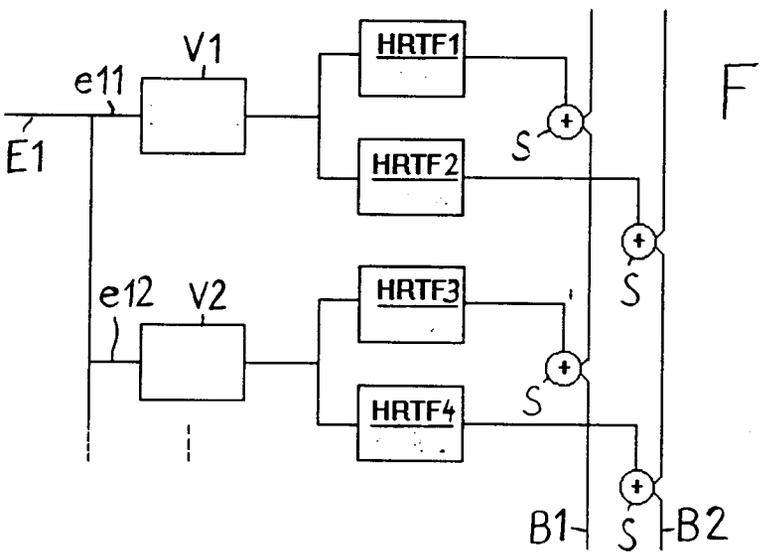

Fig.2

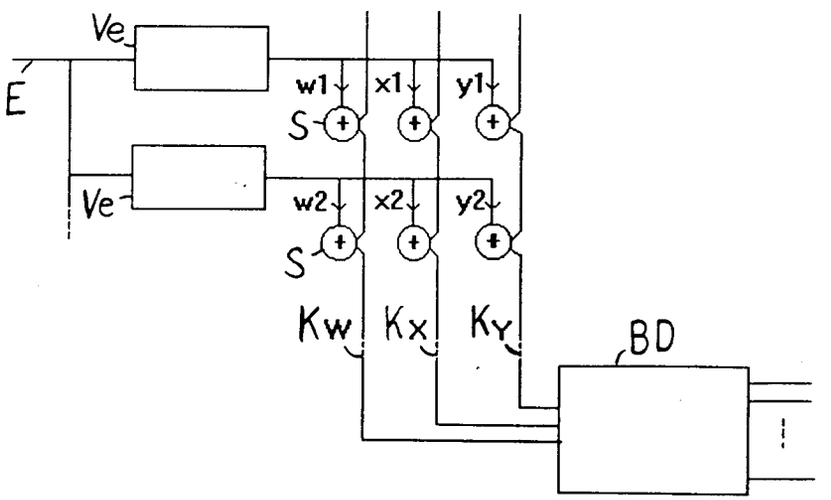

Fig.3

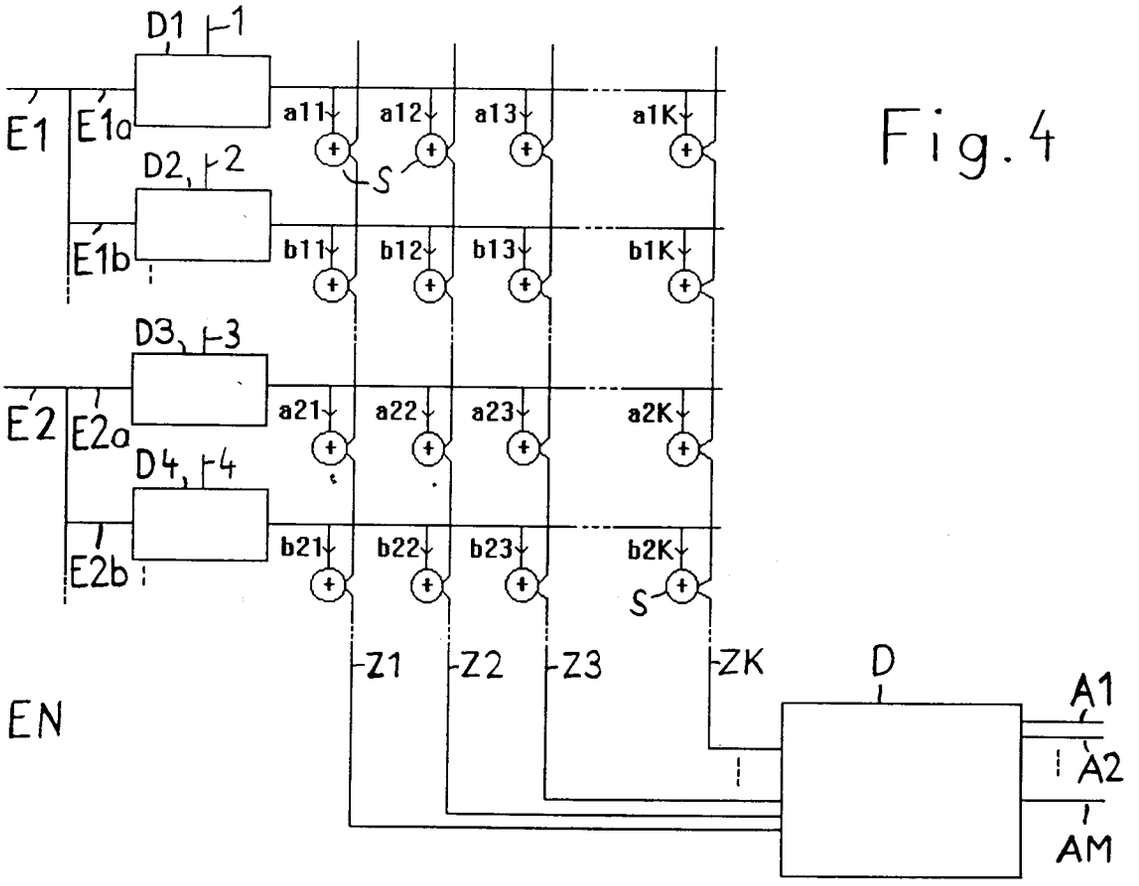
D1
1
E1
E1a
a11
a12
a13
a1K
D2
2
S
E1b
b11
b12
b13
b1K
D3
3
E2
E2a
a21
a22
a23
a2K
D4
4
E2b
b21
b22
b23
b2K
S
Z1
Z2
Z3
ZK
D
A1
A2
AM
EN

Fig.4

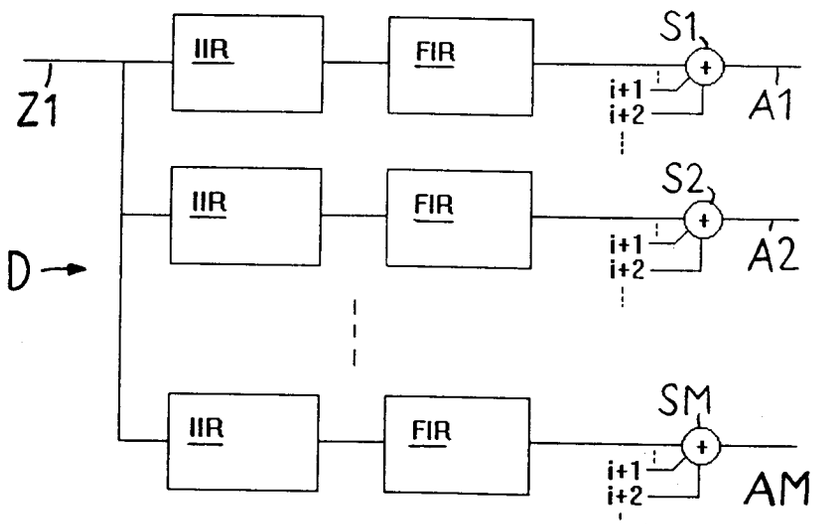
IIR
FIR
S1
i+1
i+2
A1
Z1
S2
A2
D
SM
AM

Fig.5

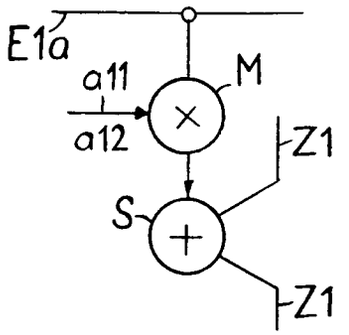

Fig.6

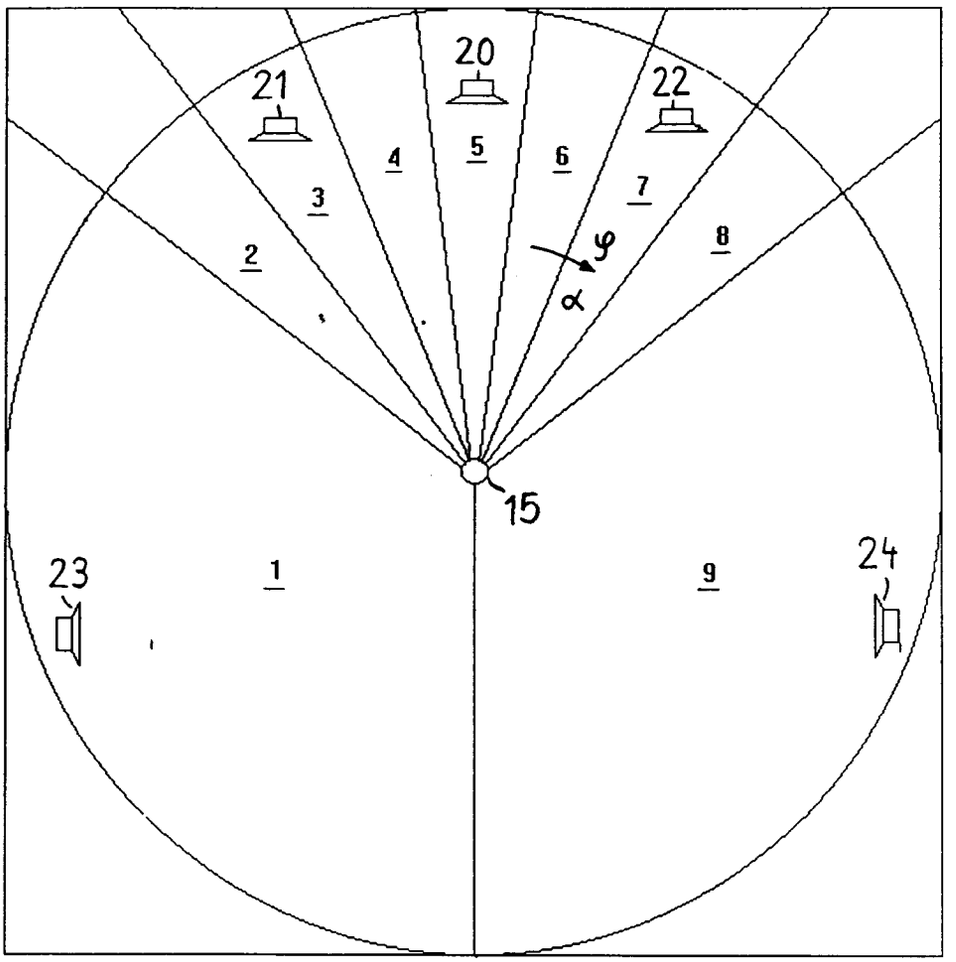

Fig.7

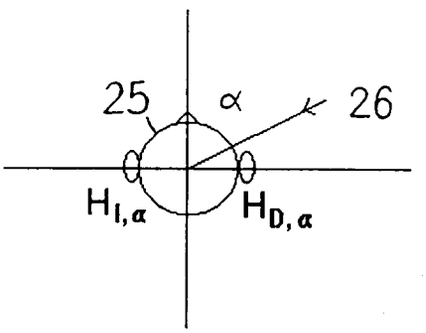

Fig.8

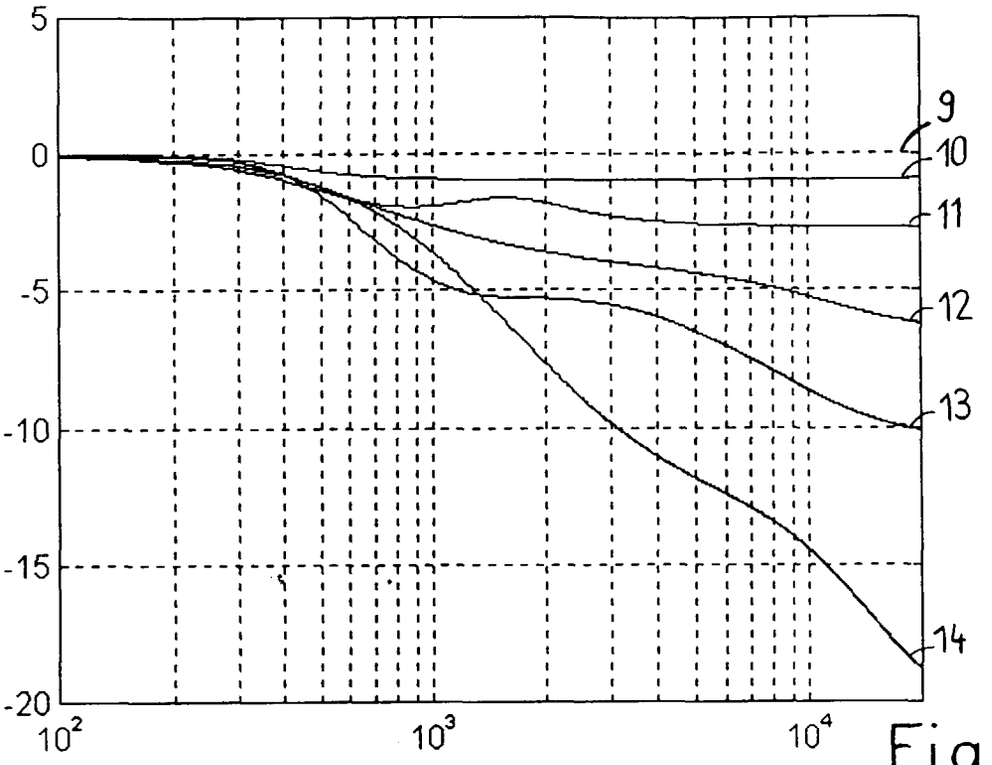
5
0
-5
-10
-15
-20
10^2
10^3
10^4
9
10
11
12
13
14

Fig.9

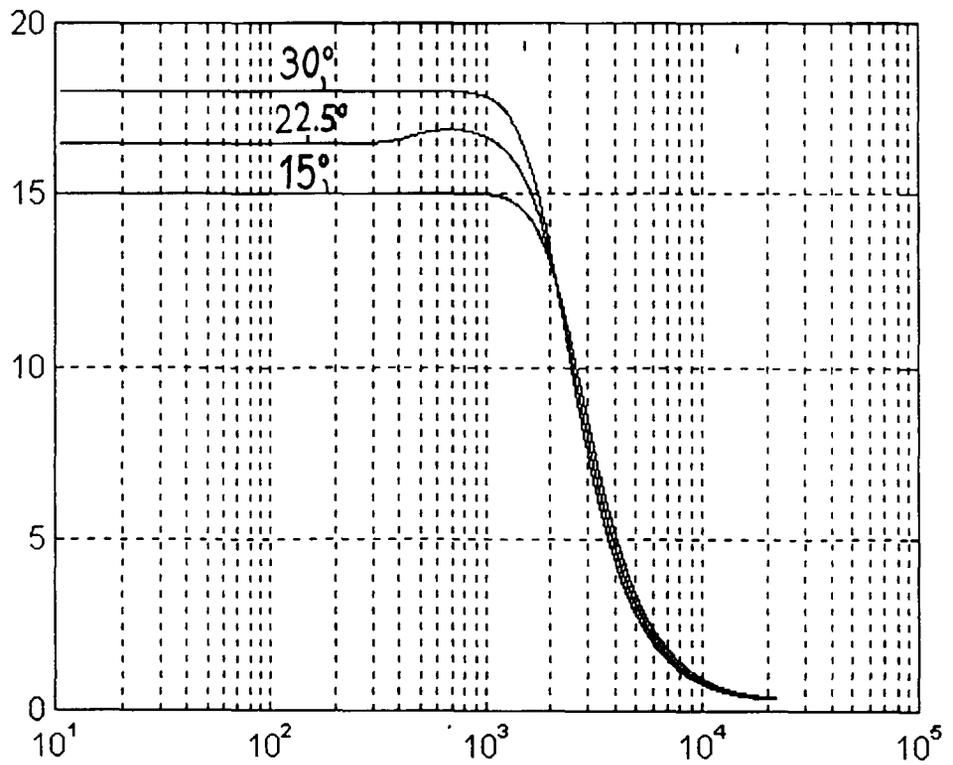
20
15
10
5
0
10^1
10^2
10^3
10^4
10^5
30°
22.5°
15°

Fig.10

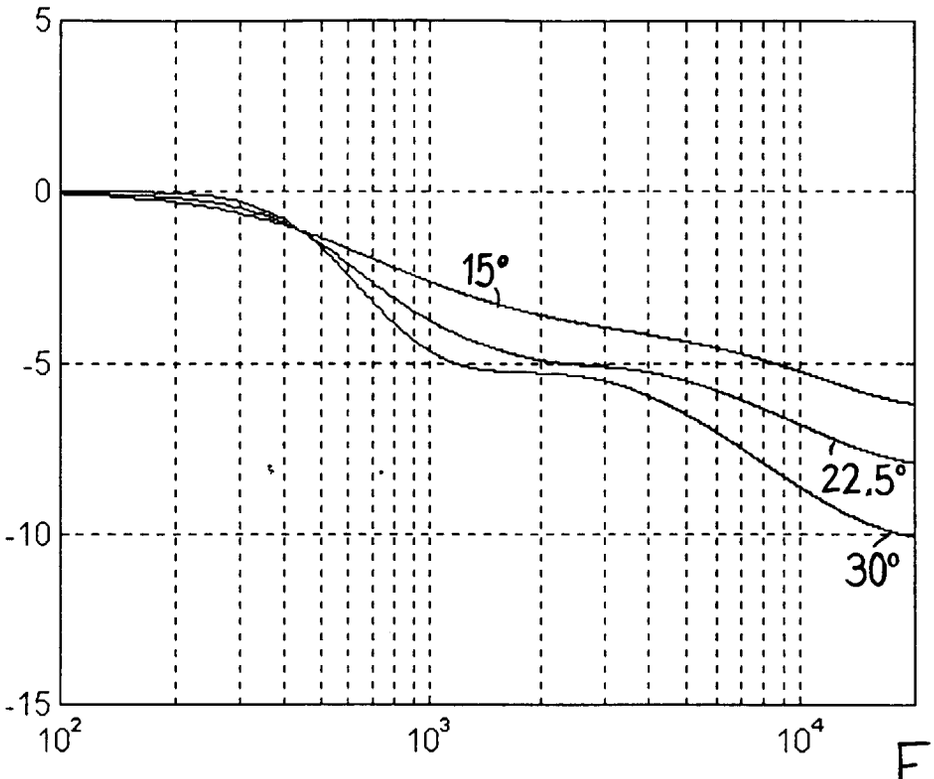

Fig.11