

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
14. Juni 2001 (14.06.2001)

PCT

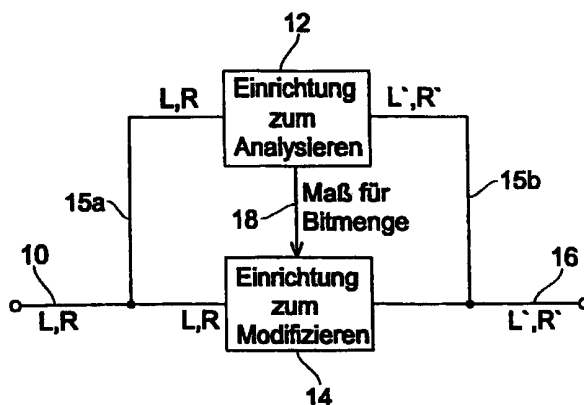
(10) Internationale Veröffentlichungsnummer
WO 01/43503 A2

- (51) Internationale Patentklassifikation⁷: **H04S 1/00** (72) **Erfinder; und**
(21) Internationales Aktenzeichen: **PCT/EP00/12352** (75) **Erfinder/Anmelder (nur für US): TEICHMANN, Bodo**
(22) Internationales Anmeldedatum: **7. Dezember 2000 (07.12.2000)** [DE/DE]; Grossgründlacher Hauptstrasse 29, 90427
Nürnberg (DE). **KUNZ, Oliver** [DE/DE]; Röttenweg 1,
96145 Sesslach (DE). **HERRE, Jürgen** [DE/DE]; Am
Eichengarten 11, 91054 Buckenhof (DE). **PEICHL,**
Klaus [DE/DE]; Dornbergstrasse 10, 91058 Erlangen
(DE). **BEER, Michael** [DE/DE]; Bismarckstrasse 26,
91054 Erlangen (DE).
(25) Einreichungssprache: **Deutsch**
(26) Veröffentlichungssprache: **Deutsch**
(30) Angaben zur Priorität: **199 59 156.3 8. Dezember 1999 (08.12.1999) DE**
(71) **Anmelder (für alle Bestimmungsstaaten mit Aus-**
nahme von US): FRAUNHOFER-GESELLSCHAFT
ZUR FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. [DE/DE]; Leonrodstrasse 54,
80636 München (DE).
(81) **Bestimmungsstaaten (national): JP, US.**
(84) **Bestimmungsstaaten (regional): europäisches Patent (AT,**
BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE, TR).

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND DEVICE FOR PROCESSING A STEREO AUDIO SIGNAL

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUM VERARBEITEN EINES STEREOAUDIOSIGNALS



12 ... DEVICE FOR ANALYSING

18 ... MEASURE FOR THE BIT QUANTITY

14 ... DEVICE FOR MODIFYING

(57) **Abstract:** The invention relates to a device for processing a stereo audio signal comprising a first channel (L) and a second channel (R). The stereo signal is analysed (12) for obtaining a measure for a bit quantity, whereby said quantity is required by a coder for coding the stereo audio signal using a coding algorithm. The first and the second channel are subsequently modified (14) when the measure for the bit quantity is greater than a predetermined value. Modification is carried out in such a way that the energy of a sum signal of the first and second modified channel (L', R') bears a predetermined ratio in relation to the energy of a sum signal of the first and second channel and that a difference signal of the first and second modified channel is muffled in relation to the difference signal of the first and second channel. The side channel is muffled, especially for audio coders that require a constant output bit rate, when the coding of stereo audio signals cannot observe the output bit rate of the coder. Stereo channel separation is thus abandoned in favour of an increased audio bandwidth or a reduction of quantisation interference.

[Fortsetzung auf der nächsten Seite]

WO 01/43503 A2

**Veröffentlicht:**

— Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

Zur Erklärung der Zweibuchstaben-Codes, und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) Zusammenfassung: Bei einer Vorrichtung zum Verarbeiten eines Stereosignals mit einem ersten Kanal (L) und einem zweiten Kanal (R) wird das Stereosignal zunächst analysiert (12), um ein Maß für eine Menge an Bits zu erhalten, die von einem Codierer benötigt wird, um das Stereoaudiosignal unter Verwendung eines Codieralgorithmus zu codieren. Hierauf werden der erste und der zweite Kanal modifiziert (14), wenn das Maß für die Menge an Bits größer als ein vorbestimmter Wert ist, wobei die Modifikation derart vonstatten geht, daß die Energie eines Summensignals aus dem ersten und zweiten modifizierten Kanal (L', R') in einem vorbestimmten Verhältnis zur Energie eines Summensignals aus dem ersten und zweiten Kanal ist, und daß ein Differenzsignal aus dem ersten und zweiten modifizierten Kanal gegenüber dem Differenzsignal aus dem ersten und zweiten Kanal gedämpft ist. Insbesondere für Audiocodierer, die eine konstante Ausgangsbitrate erfordern, wird im Falle von Stereoaudiosignalen, deren Codierung die Ausgangsbitrate des Codierers nicht einhalten kann, der Seitenkanal gedämpft, wodurch Stereokanaltrennung zugunsten einer erhöhten Audiobandbreite bzw. einer Verringerung von Quantisierungsstörungen aufgegeben wird.

Verfahren und Vorrichtung zum Verarbeiten eines Stereoaudiosignals

Beschreibung

Die vorliegende Erfindung bezieht sich allgemein auf die Codierung von Audiosignalen und insbesondere auf eine Verarbeitung von Stereosignalen.

Ein Stereosignal umfaßt zumindest zwei Kanäle, d. h. einen linken Kanal und einen rechten Kanal. Darüber hinaus können Stereosignale noch einen linken und rechten Surround-Kanal haben. Außerdem besteht die Möglichkeit, daß ein Stereosignal fünf verschiedene Kanäle hat, d. h. einen vorderen linken Kanal, einen vorderen Mitte-Kanal und einen vorderen rechten Kanal sowie einen linken hinteren und einen hinteren rechten Kanal.

Für eine datenreduzierende Codierung von Stereosignalen besteht die Möglichkeit, daß Ähnlichkeiten von zumindest zwei Kanälen ausgenutzt werden, um die Menge an Bits zu reduzieren, die benötigt wird, um ein Stereosignal mit mindestens zwei Kanälen zu codieren.

Ein bekanntes Verfahren zum Verarbeiten von Stereosignalen, um eine effizientere Codierung zu erreichen, wird als Mitte/Seite-Verfahren (M/S-Verfahren) bezeichnet. Beim M/S-Verfahren werden der erste und der zweite Kanal miteinander kombiniert, um einen Mitte-Kanal und einen Seite-Kanal zu erzeugen. Aus Übersichtlichkeitsgründen wird im nachfolgenden nicht mehr von einem ersten und einem zweiten Kanal, sondern von einem linken Kanal (L-Kanal) und von einem rechten Kanal (R-Kanal) gesprochen. Es ist bekannt, daß der Mitte-Kanal gleich der Summe aus linkem Kanal L und rechtem Kanal R, multipliziert mit einem Faktor von 0,5, ist, während der Seite-Kanal die Differenz zwischen dem linken Kanal L und dem rechten Kanal R ist, multipliziert mit einem Faktor

von z. B. 0,5 (andere Faktoren sind ebenfalls möglich). Gleichungsmäßig ausgedrückt bedeutet dies:

$$M = 0,5 \cdot (L + R)$$

$$S = 0,5 \cdot (L - R)$$

Wenn der linke Kanal L und der rechte Kanal R relativ ähnlich zueinander sind, so bringt eine M/S-Verarbeitung eine deutliche Einsparung der zum Codieren benötigten Bitmenge, da der Seite-Kanal relativ weniger Energie als R oder L haben wird. Im Grenzfall, bei dem der linke Kanal L und der rechte Kanal R identisch sind, wird der Mitte-Kanal gleich dem linken Kanal L oder gleich dem rechten Kanal R sein, während der Seite-Kanal 0 ist. Es ist zu sehen, daß somit aufgrund der Tatsache, daß der Seite-Kanal 0 ist, eine theoretisch maximale Bitrateneinsparung beim Codieren von 50% erreicht wird, da nur der Mitte-Kanal codiert werden muß, während für den Seite-Kanal keinerlei Bits aufgewendet werden müssen.

Es besteht somit die allgemeine Regel, daß je ähnlicher sich der rechte und der linke Kanal sind, desto kleiner, d. h. energieärmer, wird der Seite-Kanal sein, und umso weniger Bits werden zur Codierung des Seite-Kanals benötigt.

Ein Hörer wird die Ähnlichkeit von linkem und rechtem Kanal dadurch wahrnehmen, daß, im Falle von identischen Kanälen, ein Sprecher oder ein Orchester genau in der Mitte zwischen den beiden Lautsprechern wahrgenommen wird. Andererseits wird ein Hörer unähnliche Kanäle dadurch wahrnehmen, daß er einen ausgeprägten Stereo-Effekt hat, d. h. daß ein Sprecher, ein Orchester bzw. einzelne Instrumente eines Orchesters genau links und/oder genau rechts lokalisiert werden können. Wird der Fall betrachtet, daß der linke Kanal viel Energie hat, und daß der rechte Kanal nur wenig Energie hat, d. h. der Fall, bei dem z. B. nur ein einziges Instrument ganz links im Aufzeichnungsraum angeordnet ist, und nur im linken Kanal hörbar ist, während auf dem rechten Kanal le-

diglich Rauschen ist, so wird der Mitte-Kanal nach einer M/S-Verarbeitung etwa gleich dem linken Kanal sein. Darüberhinaus wird auch der Seite-Kanal etwa gleich dem linken Kanal sein. In diesem Fall haben somit sowohl der Mitte-Kanal als auch der Seite-Kanal nahezu gleich viel Energie und müssen beide mit einer relativ großen Anzahl von Bits codiert werden. Im Vergleich zum ursprünglichen Fall hat sich bei dieser Signalkonstellation die zur Codierung erforderliche Bitmenge aufgrund der M/S-Codierung nicht verkleinert, sondern im Grenzfall sogar verdoppelt, wenn davon ausgegangen wird, daß der linke Kanal L eine bestimmte Energiemenge umfaßt, während der rechte Kanal R 0 ist. Hier wäre es wesentlich günstiger gewesen, keine M/S-Verarbeitung durchzuführen, sondern lediglich eine L/R-Verarbeitung. Die Auswirkungen auf die Anzahl von Bits, die benötigt werden, um ein Stereosignal zu codieren, reichen somit in einem Extremfall von einer Einsparung von 50% bis zum anderen Extremfall, der eine Verdoppelung der zum Codieren benötigten Bits zur Folge hat. Es muß daher beim Einsatz eines M/S-Verfahrens genau darauf geachtet werden, ob das Stück für eine M/S-Verarbeitung geeignet ist oder nicht. In dem Fall, in dem ein Stereosignal (z. B. ein Testausschnitt von 20 ms, der auch als Frame bezeichnet wird) nicht für eine M/S-Verarbeitung geeignet ist, wird aus Biteffizienzgründen auf die M/S-Verarbeitung verzichtet, und es werden sowohl der linke als auch der rechte Kanal für sich gesehen codiert. Dieser "normale" Fall wird auch als L/R-Verarbeitung bezeichnet.

Übliche Audiocodierverfahren, wie sie beispielsweise verwendet werden, um Audiosignale zu codieren, die gemäß einem der MPEG-Standards decodiert werden, gliedern sich generell in mehrere Schritte. Zunächst wird ein Audiosignal, das beispielsweise in Form von PCM-Abtastwerten vorliegt, wie sie z. B. ein CD-Player ausgibt, mittels einer Zeit-Frequenz-Transformation oder einer Filterbank in eine spektrale Darstellung überführt. Typischerweise wird ein Block mit einer bestimmten Anzahl von Abtastwerten, der auch als "Frame"

bezeichnet wird, verwendet, um einen Block von komplexen Spektralwerten zu erzeugen, die ein Kurzzeitspektrum des Frames von Audioabtastwerten ("Samples") bilden. Die Blockbildung wird unter Verwendung von Transformationsfenstern erreicht, welche beispielsweise 1024 Abtastwerte lang sind. Wenn beispielsweise überlappende Fenster zur Transformation verwendet werden, deren Überlappungsgrad 50% beträgt, werden aus 1024 Abtastwerten 1024 Spektralwerte gebildet. Diese Spektralwerte werden anschließend mittels eines bekannten Iterationsprozesses quantisiert, wonach die quantisierten Spektralwerte einer Entropie-Codierung z. B. unter Verwendung einer Mehrzahl von festgelegten Huffman-Codetabellen unterzogen werden, um schließlich einen Bitstrom zu erhalten, der einerseits die codierten quantisierten Spektralwerte enthält und der andererseits auch Seiteninformationen aufweist, die sich auf die Fenster, auf Skalenfaktoren, die beim Quantisieren berechnet werden sowie auf weitere Informationen beziehen, die zum Decodieren des Bitstroms benötigt werden.

Eine Mitte/Seite-Verarbeitung kann entweder vor der Transformation in den Spektralbereich durchgeführt werden, d. h. unter Verwendung der digitalen zeitdiskreten Abtastwerte. Alternativ kann eine Mitte/Seite-Verarbeitung jedoch auch nach der Transformation, d. h. mit den komplexen Spektralwerten durchgeführt werden. Die letztere Alternative bietet darüberhinaus den Vorteil, daß eine Mitte/Seite-Verarbeitung nicht, wie im Zeitbereich, für das gesamte Spektrum eingesetzt werden kann, sondern auch für bestimmte Frequenzbänder, wenn bestimmte Spektralwerte einer Mitte/Seite-Verarbeitung unterzogen werden, und andere nicht.

Üblicherweise sind Audiocodierer derart ausgelegt, daß sie eine konstante Bitrate, d. h. eine bestimmte Anzahl von Bits pro Sekunde, liefern. Eine andere Randbedingung besteht darin, daß das durch das Quantisieren eingeführte Quantisierungsrauschen wenn möglich derart gewählt wird, daß seine Energie unter der psychoakustischen Maskierungsschwelle oder

Mithörschwelle des Audiosignals ist. Das grundsätzliche Verfahren, um das Quantisierungsrauschen im Frequenzbereich einzustellen, besteht in der "Formung" des Rauschens unter Verwendung der Skalenfaktoren. Zu diesem Zweck wird, wie es bekannt ist, das Spektrum in mehrere Gruppen von Spektralkoeffizienten aufgeteilt, die Skalenfaktorbänder genannte werden, denen ein einziger Skalenfaktor zugeordnet ist. Ein Skalenfaktor stellt einen Multiplikationswert dar, der verwendet wird, um die Amplitude aller Spektralkoeffizienten in diesem Skalenfaktorband zu verändern. Dieser Mechanismus wird verwendet, um die Zuordnung des Quantisierungsrauschens im Spektralbereich, das durch den Quantisierer erzeugt wird, so einzustellen, daß in jedem Skalenfaktorband die Energie des Quantisierungsrauschens unter der psychoakustischen Maskierungsschwelle in diesem Skalenfaktorband ist. Es ist ersichtlich, daß weder das Quantisieren noch das Entropie-Codieren Vorgänge sind, die eine konstante Bitrate begünstigen. Es sei angemerkt, daß - ganz im Gegenteil - beide Verfahren eine variable Bitrate begünstigen. Für Übertragungsanwendungen ist es jedoch oftmals gefordert, daß der Codierer eine konstante Bitrate am Ausgang hat. Um eine konstante Bitrate zu liefern, wird üblicherweise ein sogenanntes Bitreservoir eingesetzt. Wenn das Audiosignal derart beschaffen ist, daß kurzzeitig weniger Bits als durch die äußere Bitrate am Ausgangs des Codierers vorgegeben benötigt werden, so werden Bits dem Bitreservoir zugeordnet, um im Falle eines Audiosignalabschnitts, der mehr Bits zur Codierung benötigt, auch mehr Bits geben zu können, wodurch das Bitreservoir wieder geleert wird.

Es sei darauf hingewiesen, daß eine Randbedingung eines solchen Codierers wie erwähnt die konstante Ausgangsbitrate ist, und daß die andere Randbedingung darin besteht, daß das Quantisierungsrauschen kleiner oder gleich der psychoakustischen Maskierungsschwelle ist, damit es durch das Audiosignal maskiert oder verdeckt wird.

Im nachfolgenden wird auf Möglichkeiten eingegangen, was zu

unternehmen ist, wenn die "innere Bitrate" des Codierers von der äußeren konstanten Ausgangsbitrate abweicht. Ist die innere Bitrate derart niedrig, daß beispielsweise das Bitreservoir auf seinen maximalen Wert aufgefüllt ist, so existiert selbstverständlich kein Problem, da der Quantisierer dann derart gesteuert werden kann, daß er nun noch feiner als nötig quantisiert, wodurch mehr Bits zum Quantisieren benötigt werden. Dies wird so lange durchgeführt, bis die "äußere" konstante Bitrate erreicht ist.

Kritischer ist jedoch der Fall, bei dem die "innere Bitrate" des Codierers höher ist als die ausgangsseitig geforderte konstante Bitrate. Dieser Fall wird auftreten, wenn das Audiosignal schwer zu codieren ist, d. h. wenn der Codierer viele Bits aufwenden muß, um das Audiosignal zu codieren, was auch anschaulich als "hohe Last" des Codierers bezeichnet werden kann. Für die Transformationscodierung besteht der Merksatz, daß sie tonale Stücke relativ effizient codiert werden kann, daß jedoch rauschhafte Signale, die relativ hohe Energien haben, und die darüberhinaus ein relativ kompliziertes Spektrum haben, wie beispielsweise Sprache oder Schlagzeug- bzw. Trommelmusik, relativ wenig komprimiert werden können. Auch Signale, die transient sind, d. h. die ein unregelmäßiges Zeitverhalten haben, können nur relativ aufwendig codiert werden, wenn keine Codierungsartefakte erzeugt werden sollen. Im Fall von transienten Signalen wird bereits bei der Fensterung von langen Fenstern auf kürzere Fenster umgeschaltet, um eine bessere zeitliche Auflösung zu erreichen, bzw. um zu erreichen, daß sich das Quantisierungsrauschen nur über eine kleinere Anzahl von Audioabtastwerten "verschmiert". Im Falle von kurzen Fenstern fallen wesentlich mehr Seiteninformationen an.

Ein Codierer, der feststellt, daß die Ausgangsbitrate nicht reicht, und der auch schon das Bitreservoir "leergefahren" hat, hat nun mehrere Möglichkeiten, um seine innere Bitrate "gewaltsam" zu reduzieren, um das Kriterium der konstanten Ausgangsbitrate zu erfüllen. Eine Möglichkeit besteht darin,

auf die Umschaltung zu kurzen Fenstern zu verzichten. Dies führt jedoch zu hörbaren Codierartefakten.

Eine weitere Möglichkeit besteht darin, die psychoakustische Maskierungsschwelle bei der Quantisierung bewußt zu verletzen, um gröber als eigentlich erforderlich zu quantisieren, um eine niedrigere Bitrate zu erreichen. Auch dies führt zu hörbaren Störungen.

Eine weitere Möglichkeit besteht darin, die Audiobandbreite zu verringern, d. h. nicht mehr die volle Audiobandbreite zu codieren sondern ab einer bestimmten von der Ausgangsbitrate abhängigen Grenzfrequenz die darüberliegenden Spektralwerte zu 0 zu setzen, um somit die Ausgangsbitrate zu reduzieren. Dieses Verfahren führt nicht zu hörbaren Quantisierungsstörungen, führt jedoch zu einem Verlust an Höhen im Audiosignal. Oftmals wird dieser Verlust jedoch weniger stark wahrgenommen als ein hörbares Quantisierungsrauschen.

Ein besonderes Problem bei der Codierung von Stereosignalen besteht in dem als "Stereo Unmasking" bezeichneten Effekt, der nachfolgend kurz dargelegt wird. Wird eine normale L/R-Codierung eingesetzt, so wird sowohl der linke Kanal als auch der rechte Kanal für sich transformiert, quantisiert und codiert, so daß das im linken Kanal und rechten Kanal zur Datenreduzierung eingeführte Quantisierungsrauschen unabhängig vom jeweils anderen Kanal sind. Das heißt, daß das Quantisierungsrauschen im linken Kanal und das Quantisierungsrauschen im rechten Kanal nicht korreliert sind. Wird der Fall betrachtet, daß sich linker und rechter Kanal relativ ähnlich sind, so bedeutet dies, daß ein Hörer nach einer Decodierung dieses Signal so wahrnehmen wird, daß beispielsweise ein Sprecher in der Mitte ist. Der "Stereo-Unmasking"-Effekt besteht nun darin, daß aufgrund der Tatsache, daß das Quantisierungsrauschen in den beiden Kanälen nicht korreliert ist, das Quantisierungsrauschen des linken Kanals links und das Quantisierungsrauschen des rechten Kanals rechts wahrgenommen wird. Eine hohe Verdeckung des Rau-

schens findet aber lediglich in der Mitte statt, wo auch das Nutzsignal ist, jedoch nicht links und rechts.

Die M/S-Codierung hat daher neben ihrer Datenraten-reduzierenden Wirkung bei speziellen Signalen auch den Vorteil, daß das Quantisierungsrauschen sowohl im linken Kanal als auch im rechten Kanal mit dem Quantisierungsrauschen des jeweils anderen Kanals korreliert wird, so daß auch das Quantisierungsrauschen in der Mitte stattfindet und dort von dem Nutzsignal im wesentlichen vollständig bzw. wesentlich besser als im unkorrelierten Fall verdeckt wird. Anders ist der Fall, bei dem der linke und der rechte Kanal relativ unähnlich sind. Wenn hier M/S-Codierung eingesetzt wird, so wird aufgrund des Stereoeffekts das Nutzsignal entweder links oder rechts sein, während aufgrund der M/S-Codierung das Quantisierungsrauschen korreliert ist und eher in der Mitte liegt. Auch hier findet sozusagen ein Stereo-Unmasking statt.

In jüngster Zeit werden immer mehr skalierbare Audiocodierer untersucht. Skalierbare Audiocodierer sind derart angeordnet, daß ihr ausgangsseitiger Bitstrom zumindest eine erste und eine zweite Skalierungsschicht aufweist. Ein Decodierer, der einfach ausgelegt ist, wird aus dem skalierten Bitstrom lediglich die erste Skalierungsschicht entnehmen, die beispielsweise ein codiertes Audiosignal mit reduzierter Bandbreite aufweist bzw. ein mit einem einfachen Codieralgorithmus codiertes Audiosignal ist. Ein anderer Decodierer, der voll ausgelegt ist, wird sowohl die erste Skalierungsschicht als auch die zweite Skalierungsschicht aus dem Bitstrom nehmen, um die erste Skalierungsschicht mit einem ersten Decodierer zu decodieren, und um dann die zweite Skalierungsschicht ebenfalls zu decodieren, die alleine oder zusammen mit der decodierten ersten Skalierungsschicht ein Audiosignal mit voller Bandbreite liefert.

Skalierbare Codierer sind besonders im Bereich der Stereosignale erwünscht, da hier als erste Skalierungsschicht ein

Mono-Signal, d. h. der Mitte-Kanal, verwendet werden kann, während als zweite Skalierungsschicht z. B. der Seite-Kanal genommen werden kann. Ein einfacher Decodierer bzw. ein Decodierer, der auf schnellen Betrieb ausgelegt ist, wird lediglich das Mono-Signal liefern, während ein besserer Decodierer bzw. ein Decodierer, bei dem die Schnelligkeit der Übertragung nicht das entscheidendste Kriterium ist, neben der Mono- bzw. Mitte-Schicht auch die Seite-Schicht nehmen wird, um ein volles Stereosignal am Ausgang des Decodierers zu erzeugen.

Für den Aufbau der Skalierungsschichten existieren verschiedene Möglichkeiten. Die erste Skalierungsschicht kann sich von der zweiten Skalierungsschicht bzw. von einer beliebigen Anzahl weiterer Skalierungsschichten im Audiocodierverfahren selbst, in der Audiobandbreite, in der Audioqualität, bezüglich Mono/Stereo und oder einer Kombination der genannten Qualitätskriterien bzw. weiterer denkbarer Kriterien unterscheiden. Für eine hohe Codiereffizienz wird es angestrebt, daß die zweite Skalierungsschicht eine möglichst kleine Anzahl an Bits aufweist, bzw. daß ein Decodierer, der die zweite Skalierungsschicht decodiert, möglichst umfassend auch die erste Skalierungsschicht verwendet. Wenn ein skalierbarer Codierer für Stereosignale betrachtet wird, der als erste Skalierungsschicht das Mitte-Signal, d. h. das Mono-Signal liefert, und der als zweite Schicht den Seite-Kanal liefert, so ist zu sehen, daß seine Gesamteffizienz umso besser ist, je öfter die M/S-Codierung eingesetzt wird. Diese Forderung steht jedoch bei bestimmten Stereosignalen im Widerspruch zur Biteffizienz, nämlich bei Stereosignalen, die eine hohe Stereokanaltrennung haben. Andererseits liefert die M/S-Verarbeitung eine gewisse "natürliche" Skalierbarkeit und führt zu einer Korrelation des Quantisierungsrauschen im linken Kanal und im rechten Kanal.

Die genannten Probleme bezüglich der M/S-Codierung gelten umso mehr, je mehr ein zu codierendes Audiosignal seine Eigenschaften bezüglich der M/S-Codierung plötzlich ändert.

Hat ein zu codierendes Audiosignal auf einmal nicht mehr die Eigenschaft, daß der linke Kanal ähnlich dem rechten ist, fällt der M/S-Codierungsgewinn weg. Eine Folge wird daher in der Regel eine Zunahme der Quantisierungsstörung womöglich über die psychoakustische Hörschwelle hinaus und/oder eine Reduktion der Audiobandbreite je nach konkreter Implementierung des Codierers sein.

Dieses Problem macht sich besonders stark, aber nicht nur, bei der skalierbaren Audiocodierung bemerkbar, und insbesondere dort, wo die sogenannte Mono-Stereo-Skalierbarkeit benutzt wird, wie es oben ausgeführt worden ist.

Die Aufgabe der vorliegenden Erfindung besteht darin, eine Vorrichtung und ein Verfahren zum Verarbeiten eines Stereoaudiosignals zu schaffen, das zu weniger hörbaren Störungen führt.

Diese Aufgabe wird durch eine Vorrichtung zum Verarbeiten eines Stereoaudiosignals nach Patentanspruch 1 sowie durch ein Verfahren zum Verarbeiten eines Stereoaudiosignals nach Patentanspruch 18 gelöst.

Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, daß es bei Stereoaudiosignalen oftmals günstiger ist, auf eine hohe Stereokanaltrennung zu verzichten, um eine höhere Audiobandbreite und/oder geringere hörbare Störung im Vergleich zu dem Fall zu erreichen, bei dem die Stereokanaltrennung beibehalten wird, während die Audiobandbreite reduziert wird, oder durch das Quantisieren eingeführte Störungen hörbar werden.

Erfahrungsgemäß wird ein Hörer hörbare Quantisierungsstörungen eher als unangenehm empfinden, als eine geringere Stereokanaltrennung. Hörbare Quantisierungsstörungen sind generell ein Fremdkörper in einem Audiosignal, während ein Hörer eines erfindungsgemäß verarbeiteten Stereosignals nicht unbedingt weiß, wie die Stereokanaltrennung des

Ausgangssignals war und somit eine geringere Stereokanaltrennung nicht als Codierartefakt empfinden wird.

Eine Verringerung der Stereokanaltrennung wird somit eingesetzt, um die ausgangsseitige Bitrate des Codierers generell zu verringern, oder auf einen vorbestimmten Wert zu reduzieren.

Eine erfindungsgemäße Vorrichtung zum Verarbeiten eines Stereosignals, das einen ersten Kanal und einen zweiten Kanal aufweist, umfaßt eine Einrichtung zum Analysieren des Stereoaudiosignals, um ein Maß für eine Menge an Bits zu erhalten, die von einem Codierer benötigt wird, um das Stereoaudiosignal unter Verwendung eines Codieralgorithmus zu codieren, und eine Einrichtung zum Modifizieren des ersten und des zweiten Kanals, um einen modifizierten ersten und einen modifizierten zweiten Kanal zu erhalten, wobei die Einrichtung zum Modifizieren auf die Einrichtung zum Analysieren anspricht, um wirksam zu sein, wenn das Maß für die Menge an Bits ein vorbestimmtes Maß überschreitet, und wobei die Einrichtung zum Modifizieren derart ausgestaltet ist, daß ein Summensignal aus dem ersten und zweiten modifizierten Kanal zumindest gemäß einer Charakteristik des Signals, die sich ähnlich zu der Energie des Signals ändert, im wesentlichen gleich der Charakteristik eines Summensignals aus dem ersten und zweiten Kanal ist, und daß ein Differenzsignal aus dem ersten und dem zweiten modifizierten Kanal gegenüber dem Differenzsignal aus dem ersten und zweiten Kanal gedämpft ist.

Es sei darauf hingewiesen, daß die Charakteristik, die ähnlich zur Energie verläuft, die Energie selbst sein kann, aber auch z. B. die Summe von quadrierten Abtastwerten in einer bestimmten Zeitdauer, die Summe von quadrierten Spektralwerten in einem bestimmten Frequenzbereich, die Summe von Abtastwertbeträgen in einer bestimmten Zeitdauer oder die Summe von quadrierten Spektralwerten in einem bestimmten Frequenzbereich ist oder aber auch eine Kombination zwischen

zwei oder mehreren der genannten Charakteristika. Aus Einfachheitsgründen wird im nachfolgenden jedoch beispielhaft von der Energie als Charakteristik, die ähnlich zur Energie verläuft, gesprochen.

Das Modifizieren des Stereoaudiosignals, d. h. das Reduzieren der Kanaltrennung, wird unter der Voraussetzung durchgeführt, daß die Lautstärke des Signals nicht schwankt. Eine reduzierte Kanaltrennung selbst wird nicht zu störenden Artefakten im decodierten Signal führen, eine Schwankung der Lautstärke jedoch schon. Daher werden der erste und der zweite Kanal, z. B. der linke Kanal und der rechte Kanal, derart modifiziert, daß die Lautstärke, d. h. das Summensignal, gegenüber dem unmodifizierten ersten und zweiten Kanal zumindest energiemäßig und bevorzugterweise sogar signalmäßig im wesentlichen gleich bleibt, während das Differenzsignal gedämpft ist.

Die erfindungsgemäße Vorverarbeitung des Stereosignals wird immer dann einsetzen, wenn festgestellt wird, daß die Menge an Bits, die benötigt wird, um das Stereoaudiosignal zu codieren, zu hoch wird. Das Maß für die Menge an Bits, die zum Codieren des Stereoaudiosignals benötigt werden, kann aus dem Stereoaudiosignal durch Analyse desselben auf verschiedene Arten und Weisen abgeleitet werden.

Zunächst kann der Mitte- und der Seite-Kanal des Stereoaudiosignals betrachtet werden, um aufgrund eines Energieverhältnisses bzw. einer Differenz der Logarithmen der Energien derselben festzustellen, wieviel Bits benötigt werden. Ohne die genaue Anzahl von Bits feststellen zu müssen, ist die Schlußfolgerung zulässig, daß im Falle eines kleinen Energieverhältnisses zwischen Mitte- und Seite-Kanal, d. h. im Falle von etwa gleich großen Kanälen, eine hohe Anzahl von Bits nötig sein wird. Je geringer somit das Energieverhältnis zwischen dem Mitte- und dem Seite-Kanal ist, umso mehr Dämpfung des Seite-Kanals wird notwendig sein, um eine bestimmte Ausgangsbitrate zu erreichen. Ein kleines Energie-

verhältnis zwischen dem Mitte- und dem Seite-Kanal liegt vor, wenn das ursprüngliche Audiosignal eine hohe Stereokanaltrennung hat, beispielsweise wenn der linke Kanal viel Energie hat, während der rechte Kanal im wesentlichen Rauschen hat. Ein kleines Energieverhältnis liegt jedoch auch vor, wenn im linken Kanal die Sprache eines Sprechers ist, und wenn im rechten Kanal die Sprache eines anderen Sprechers ist, was dazu führt, daß der linke Kanal und der rechte Kanal unter Umständen gleich viel Energie haben, daß jedoch beide Kanäle unkorreliert sind. Auch in diesem Fall liegt eine hohe Stereosignaltrennung vor, und der Mitte-Kanal und der Seite-Kanal werden eine relativ geringe Differenz der Logarithmen der Energie haben.

Eine weitere Möglichkeit zur Bestimmung des Maßes für eine Menge an Bits besteht jedoch unabhängig von der Beschaffenheit des Mitte-Kanals und des Seite-Kanals darin, den Codierer an sich zu betrachten. Ein Maß für die von einem Codierer benötigte Anzahl an Bits ist die sogenannte Perceptual Entropy (PE), die gleich dem Energieverhältnis zwischen dem Nutzaudiosignal und der für das Nutzaudiosignal berechneten psychoakustischen Mithörschwelle ist. Ist die PE groß, kann gefolgert werden, daß das Audiosignal eine relativ geringe Verdeckungsfähigkeit hat. Ist die PE dagegen klein, d. h. liegt die Energie des Nutzsignals nur knapp über der psychoakustischen Mithörschwelle, so muß das Nutzsignal nur relativ grob quantisiert werden, und das Quantisierungsrauschen ist immer noch unter der psychoakustischen Mithörschwelle "versteckt". Wird festgestellt, daß die Summe aus der, vorzugsweise über einer gewissen Zeit gemittelten, PE des linken Kanals und der, ebenfalls vorzugsweise über einer gewissen Zeit gemittelten, PE für den rechten Kanal über einem vorbestimmten Wert liegt, so wird erfindungsgemäß der Seite-Kanal gedämpft, um die erforderliche Anzahl von Bits zu reduzieren. Dieser alternative Aspekt der vorliegenden Erfindung befaßt sich somit nicht mit dem individuellen Aussehen des Mitte- und des Seite-Kanals, sondern mit dem Stereoaudiosignal selbst, das nicht bezüglich seiner M/S-Co-

dierfähigkeit beurteilt wird, sondern seiner generellen Audiocodierfähigkeit, d. h. der Schwierigkeit, dasselbe zu codieren, um eine bestimmte Zielbitrate zu erreichen.

Eine Verallgemeinerung des zweiten Aspekts besteht darin, irgendeine andere Größe als Maß für die Menge an Bits zu verwenden, welche auf die "Last" des Codierers hinweist. Eine solche Größe kann beispielsweise auch ein Signal sein, das aufgrund transienter Eigenschaften des Audiosignals anzeigt, daß ein Audiocodierer kurze Fenster zum Fenstern verwenden muß, da es Tatsache ist, daß kurze Fenster nicht zuletzt aufgrund der gestiegenen Anzahl von Seiteninformationen ein höhere Bitrate erforderlich machen. Zwecks der vorliegenden Erfindung kann somit die gesamte Palette von Steuergrößen eines Audiocodierers verwendet werden, um ein Maß dafür zu finden, daß bzw. wie stark der Seite-Kanal gedämpft werden muß, um die Ausgangsbitrate des Codierers zu verringern.

Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung führen eine zeitlich zunehmende bzw. zeitlich abnehmende Dämpfung des Seite-Kanals durch, um zu verhindern, daß ein Hörer die abnehmende Stereokanaltrennung unmittelbar wahrnimmt, sondern daß die Verringerung der Stereokanaltrennung nach und nach eintritt bzw. die Vergrößerung der Stereokanaltrennung nach und nach zunimmt, um die codiererseitige Manipulation des Stereoaudiosignals möglichst gut zu verschleiern.

Es sei darauf hingewiesen, daß zum Beibehalten einer nicht-schwankenden Lautstärke aufgrund des Modifizierens das Summensignal des modifizierten linken und rechten Kanals nicht unbedingt zu dem Summensignal des nicht-modifizierten linken und rechten Kanals identisch sein muß, sondern daß es genügt, daß lediglich die Energien der beiden Summensignale im wesentlichen gleich sind bzw. in einem vorbestimmten Verhältnis zueinander liegen. Ein Zuhörer weiß nicht, wie groß die Lautstärke des unmodifizierten Stereoaudiosignals war

und wird es daher nicht als Störung empfinden, wenn durch die Vorverarbeitung eine Lautstärkenveränderung in Richtung höherer Lautstärke bzw. geringerer Lautstärke eingeführt worden ist. Aufgrund der Einfachheit der Implementation wird es jedoch bevorzugt, daß dieses Verhältnis gleich 1 ist.

Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend bezugnehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

Fig. 1 ein Prinzipblockschaltbild der erfindungsgemäßen Vorrichtung zum Verarbeiten eines Stereoaudiosignals;

Fig. 2 eine detailliertere Darstellung einer bevorzugten Ausgestaltung der Einrichtung zum Modifizieren; und

Fig. 3 ein Blockschaltbild einer erfindungsgemäßen Vorrichtung als Vorverarbeitungsstufe für einen skalierbaren Codierer mit Mono/Stereo-Skalierbarkeit.

Fig. 1 zeigt ein Blockschaltbild der erfindungsgemäßen Vorrichtung zum Verarbeiten eines Stereoaudiosignals, das an einem Eingang 10 in die Vorrichtung eingespeist wird und einen ersten Kanal L und einen zweiten Kanal R aufweist. Das Stereoaudiosignal in Form des ersten Kanals L und des zweiten Kanals R wird einerseits in eine Einrichtung 12 zum Analysieren des Stereoaudiosignals eingespeist, und wird andererseits auch in eine Einrichtung 14 zum Modifizieren des ersten und zweiten Kanals eingespeist, um an einem Ausgang 16 einen modifizierten ersten Kanal L' und einen modifizierten zweiten Kanal R' zu erhalten. Generell werden sich der modifizierte erste Kanal L' und der modifizierte zweite Kanal R' am Ausgang 16 vom nicht-modifizierten ersten Kanal L und vom nicht-modifizierten zweiten Kanal R am Eingang 10 darin unterscheiden, daß das am Ausgang 16 anliegende modifizierte Stereoaudiosignal eine geringere Kanaltrennung als

das nicht-modifizierte Stereoaudiosignal am Eingang 10 haben wird.

Die Einrichtung 12 zum Analysieren des Stereoaudiosignals ermittelt ein Maß für eine Menge an Bits, die von einem in Fig. 1 nicht gezeigten Codierer benötigt wird, um das Stereoaudiosignal unter Verwendung eines durch den Codierer vorgegebenen Codieralgorithmus zu codieren. Das Maß für die Bitmenge wird von der Einrichtung 12 zum Analysieren über einen Signalpfad 18 der Einrichtung 14 zum Modifizieren zugeführt. Übersteigt das über den Signalpfad 18 zugeführte Maß für die Bitmenge ein vorbestimmtes Maß, so wird die Einrichtung 14 zum Modifizieren wirksam, um den ersten Kanal L und den zweiten Kanal R zu modifizieren. Erfindungsgemäß wird die Modifikation des ersten und zweiten Kanals derart durchgeführt, daß die Energie der Summe des modifizierten Stereoaudiosignals am Ausgang 16 in einem vorbestimmten Verhältnis und vorzugsweise im wesentlichen gleich der Energie des nicht-modifizierten Stereoaudiosignals am Eingang 10 ist, während jedoch das Differenzsignal, das abgesehen von dem Faktor von z. B. 0,5 dem Seite-Kanal entspricht, im modifizierten Stereoaudiosignal am Ausgang 16 gegenüber dem nicht-modifizierten Stereoaudiosignal am Eingang 10 gedämpft ist.

In Fig. 1 sind zwei Möglichkeiten der Speisung der Einrichtung 12 zum Analysieren dargestellt, die einzeln oder in Kombination verwendet werden können. Die erste Möglichkeit ist durch einen linken Pfeil 15a dargestellt, der gewissermaßen eine Vorwärtskopplung darstellt, d. h. die Einrichtung zum Analysieren des Stereoaudiosignals wird mit dem nicht modifizierten Signal L, R gespeist. Die andere Möglichkeit besteht darin, die Einrichtung 12 zum Analysieren mit dem modifizierten Signal L', R' zu speisen. Insbesondere in Fällen, in denen die Dämpfung des Seite-Signals zeitlich langsam abläuft, ist es unerheblich, ob die Dämpfung abhängig von dem aktuellen nicht modifizierten Signal oder von einem der letzten Verarbeitungsblocks des modifizierten Signals

gewissermaßen rückkopplungsmäßig gesteuert wird. Damit ist es unerheblich, ob das Stereoaudiosignal selbst direkt analysiert wird, oder aber indirekt anhand eines vorausgehenden modifizierten Signals.

Im nachfolgenden wird auf verschiedene Ausgestaltungen der Einrichtung 12 zum Analysieren des nicht-modifizierten Stereoaudiosignals am Eingang 10 eingegangen. Eine Möglichkeit besteht darin, daß die Einrichtung 12 zum Analysieren sowohl den Mitte- als auch den Seite-Kanal des Stereoaudiosignals bildet und dann das Verhältnis der Energien des Mitte- und des Seite-Kanal betrachtet. Das Energieverhältnis zwischen dem Mitte- und dem Seite-Kanal wird vorzugsweise über eine bestimmte Zeit gemittelt, die beispielsweise in der Größenordnung von 10 Audio-Frames liegen kann, was einem Wert von 200 ms entspricht, wenn als Audiocodierer ein MPEG-2-AAC-Codierer eingesetzt wird, der eine Frame-Länge von etwa 20 ms haben kann. Bezüglich des MPEG-2-AAC-Codierers wird auf den Standard ISO/IEC 13818-7 verwiesen, in dem die einzelnen Funktionsblöcke eines Audiocodierers und eines Audio-Decodierers sowie ihr Zusammenwirken detailliert beschrieben sind.

Wird festgestellt, daß das Energieverhältnis bzw. die Differenz der Logarithmen kleiner als ein bestimmter abhängig vom Anwendungsfall empirisch zu bestimmender Wert ist, der beispielsweise zu 6 dB gewählt werden kann, wird die Einrichtung 14 zum Modifizieren aktiviert, um eine Dämpfung des Seite-Kanals zu erreichen, wie es bezugnehmend auf Fig. 2 noch detaillierter ausgeführt werden wird. Gemäß dem im vorstehenden dargelegten ersten Aspekt der vorliegenden Erfindung arbeitet die Einrichtung 12 zum Analysieren des Stereoaudiosignals somit aufgrund einer direkten Untersuchung der MS-Codierfähigkeit des Stereoaudiosignals. Bei einer Implementation dieses ersten Aspekts der vorliegenden Erfindung wird die erfindungsgemäße Vorrichtung zum Verarbeiten des Stereoaudiosignals lediglich dann den Seite-Kanal dämpfen, wenn das Signal eine nicht mehr so gute MS-Codier-

fähigkeit aufweist, weil beispielsweise beide Kanäle entweder energiemäßig und/oder signalmäßig unähnlich zueinander sind. Gemäß diesem Aspekt wird somit eine Stereokanaltrennung immer dann reduziert, wenn das Beibehalten der ursprünglichen Stereokanaltrennung zu einer zu hohen Ausgangsbitrate führen würde, und wenn die Stereokanaltrennung überhaupt hoch war.

Gemäß einem weiteren Aspekt der vorliegenden Erfindung wird die Dämpfung des Seite-Kanals zur Reduktion der ausgangseitigen Codiererbitrate eingesetzt, unabhängig davon, ob das Stereoaudiosignal eine bestimmte MS-Codierfähigkeit hat oder nicht. Dieser zweite erfindungsgemäße Aspekt geht davon aus, daß selbst im Falle einer geringen Stereokanaltrennung immer noch eine weitere Dämpfung des Seite-Kanals erreicht werden kann, um eine vorbestimmte Ausgangsbitrate des Audio-codierers nicht zu überschreiten. Hierzu wird unabhängig von der MS-Codierfähigkeit des Audiosignals die Anzahl an Bits abgeschätzt, die benötigt wird, um das Audiosignal zu codieren.

Wie es in der Technik bekannt ist, verwenden moderne Audio-codierer, und beispielsweise auch ein MPEG-2-AAC-Audiocodierer, ein psychoakustisches Modell, das dazu dient, von einem zu codierenden Audiosignal die frequenzabhängige psychoakustische Maskierungsschwelle zu berechnen. Grob gesagt liefert das psychoakustische Modell als psychoakustische Maskierungsschwelle für jedes Skalenfaktorband einen Energiewert. Liegt das durch den Quantisierer eingeführte Quantisierungsrauschen unter dem Energiewert oder ist das durch die Quantisierungsstörungen eingeführte Rauschen gleich dem Energiewert, so wird entsprechend der Psychoakustiktheorie das eingeführte Quantisierungsrauschen im wesentlichen unhörbar sein.

Das Energieverhältnis bzw. die Differenz der Logarithmen des Audiosignals an sich und seiner psychoakustischen Maskierungsschwelle, die auch als Perceptual Entropy (PE) bezeich-

net wird, liefert somit ein Maß dafür, wieviel Bits zum Codieren des Audiosignals benötigt werden. Ist die PE hoch, so werden viele Bits benötigt, da die Maskierungsfähigkeit des Audiosignals relativ gering ist und somit fein quantisiert werden muß. Ist die PE dagegen klein, so werden relativ wenig Bits benötigt, da das Audiosignal relativ gut maskiert, und somit nur eine relativ grobe Quantisierung erforderlich ist.

Gemäß einem bevorzugten Ausführungsbeispiel wird beim zweiten Aspekt der vorliegenden Erfindung das Maß für die Menge an Bits folgendermaßen bestimmt. Die PE-Werte für die einzelnen Skalenfaktorbänder werden über der Frequenz integriert, d. h. aufsummiert. Dies wird sowohl für den linken als auch für rechten Kanal durchgeführt. Anschließend wird die PE-Summe für den linken Kanal zu der PE-Summe für den rechten Kanal summiert. Dieser Summen-PE-Wert aus linkem und rechtem Kanal stellt den Bitbedarf für einen Frame dar. Dieser Summen-Kanal-PE-Wert wird dann noch vorzugsweise über eine bestimmte Anzahl von Frames, wie z. B. 10, gemittelt, um einen gemittelten PE-Wert für das Stereoaudiosignal zu erhalten. Wenn dieser gemittelte PE-Wert größer oder gleich einem typischerweise empirisch zu bestimmenden vorbestimmten Wert ist, wird die Einrichtung zum Multiplizieren aktiviert, um den Seite-Kanal zu dämpfen.

Verallgemeinert kann somit als Maß für die Menge an Bits, die ein Codierer benötigen wird, irgendeine andere Regelgröße verwendet werden, die ein Maß für die "Last" des Codierers darstellt, wie z. B. ein Steuersignal des Codierers, das den Einsatz von kurzen Fenster beim Fenstern signalisiert. Das Fenstern mit kurzen Fenstern führt per se zu einer höheren Anzahl an Bits, da kürzere Fenster nicht so bit-sparend codiert werden können, wie längere Fenster.

Bezüglich des Dämpfungsbetrags des Seite-Kanals existieren mehrere Möglichkeiten, die sich hier bezüglich ihres Aufwands unterscheiden. Die einfachste Art und Weise besteht

darin, einen vorgegebenen Dämpfungswert als Zielwert zu vereinbaren, der beispielsweise empirisch festgesetzt sein kann. Eine weitere Möglichkeit besteht jedoch auch darin, den Dämpfungswert adaptiv zu bestimmen, d. h. den Seite-Kanal um einen vorbestimmten Inkrementenbetrag zu dämpfen, und dann wieder zu sehen, ob sich die Anzahl von Bits bereits ausreichend verringert hat oder nicht. Es kann dann in eine neue Iterationsschleife mit einem weiteren Inkrementen-Dämpfungsbetrag eingestiegen werden, um wiederum festzustellen, ob die Anzahl von Bits bereits ausreichend gering ist. Dieses Verfahren kann so oft wiederholt werden, bis die vom Codierer benötigte Anzahl von Bits in einem Zielkorridor liegt. Es ist jedoch ersichtlich, daß der Rechenzeit- und Implementierungsaufwand im Falle der adaptiven Dämpfungsanpassung wesentlich höher als im Falle einer vorgegebenen Dämpfung ist. Andererseits liefert eine adaptive Dämpfungsanpassung die besten und genauesten Ergebnisse.

Im nachfolgenden wird auf Fig. 2 eingegangen, in der eine detaillierte Darstellung der Einrichtung 14 zum Modifizieren gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung dargestellt ist. Die Einrichtung 14 zum Modifizieren kann derart aufgefaßt werden, daß sie einen ersten Eingang 20a für den ersten Kanal L und einen zweiten Eingang 20b für den zweiten Kanal R aufweist. Die Einrichtung 14 umfaßt einen ersten Multiplizierer 22a zum Multiplizieren des ersten Kanals L mit einem bestimmten Faktor x, einen zweiten Multiplizierer 22b zum Multiplizieren des ersten Kanals L mit einem Faktor y, einen dritten Multiplizierer zum Multiplizieren des zweiten Kanals R mit dem Faktor x und schließlich einen vierten Multiplizierer 22d zum Multiplizieren des zweiten Kanals R mit dem Faktor y. Darüberhinaus umfaßt die Einrichtung 14 zum Modifizieren einen ersten Summierer 24a zum Summieren des Ausgangssignals des ersten Multiplizierers 22a mit dem Ausgangssignal des vierten Multiplizierers 22d, und einen zweiten Summierer 24b zum Summieren des Ausgangssignals des zweiten Multiplizierers 22b mit dem Ausgangssignal des dritten Multiplizierers 22c. Am Ausgang 26a des er-

sten Summierers 24a liegt schließlich der modifizierte erste Kanal L' an, und am Ausgang 26b des zweiten Summierers 24b liegt der modifizierte zweite Kanal R' an.

Im nachfolgenden wird auf die Bestimmung der beiden Multiplikationsfaktoren x, y eingegangen, um einen gedämpften Seite-Kanal zu erreichen, während der Mitte-Kanal am Ausgang 26a, 26b gleich dem Mitte-Kanal am Eingang 20a, 20b der Einrichtung 14, die in Fig. 2 gezeigt ist, ist. Für die von der Einrichtung 14 zum Modifizieren durchgeführte Signalverarbeitung gilt folgende Matrix:

$$L' = xL + yR \quad (1)$$

$$R' = yL + xR \quad (2)$$

Die Aufgabe besteht nun darin, daß x und y bestimmt werden, damit gilt:

$$L' + R' = L + R = M = M', \quad (3)$$

und daß gilt:

$$L' - R' = S' = \text{Dämpfung} * S = \text{Dämpfung} * (L - R) \quad (4)$$

Es ergibt sich:

$$M = 0,5 (x+y) (L+R) \quad (5)$$

Da M nicht durch die Verarbeitung modifiziert werden soll, gilt ferner folgende Gleichung:

$$x + y = 1 \quad (6)$$

Für den Seite-Kanal ergibt sich:

$$S = 0,5 (x - y) (L - R) \quad (7)$$

Aus Gleichung (7) ergibt sich, daß S um den Faktor x-y reduziert wird, oder, logarithmisch ausgedrückt, durch $10 \cdot \log_{10}(x-y) \text{ dB} = \text{att}$, gedämpft wird. att steht für die Dämpfung, und es gilt, daß att kleiner als 0 dB ist.

Für eine Dämpfung in dB-Schritten ergibt sich:

$$\text{att}(\text{in dB}) = 20 \cdot \log_{10}(x-y) \quad (8)$$

Aus Gleichung (8) ergibt sich dann folgender Ausdruck:

$$\exp(0,05 \text{ att}) = x - y \quad (9)$$

Aus Gleichung (6) und Gleichung (9) ergibt sich für x die Gleichung (10) und für y die Gleichung (11).

$$x = 0,5 * (1 + \exp(0,05 \text{ att})) \quad (10)$$

$$y = 0,5 * (1 - \exp(0,05 \text{ att})) \quad (11)$$

Die Dämpfung "att" (in dB) wird in Abhängigkeit von einer der beschriebenen Regelgrößen bestimmt. Damit ergeben sich mit den Gleichungen (9) und (10) die Faktoren x und y für die durch Fig. 2 dargestellte Dämpfungsmatrix, die sich gleichungsmäßig in den Gleichungen (1) und (2) widerspiegelt. Um Implementierungsaufwand und Rechenaufwand einzusparen, muß keine vollständig adaptive Anpassung der Dämpfung att durchgeführt werden, sondern es kann ein bestimmter Dämpfungswert att, der empirisch festgelegt worden ist, eingesetzt werden, falls das Maß für die Menge an Bits einen vorbestimmten Grenzwert überschreitet.

Erfindungsgemäß wird die Dämpfung nicht schlagartig erhöht, da eine Verringerung der Kanaltrennung, die schlagartig vonstatten geht, unter Umständen zu einer hörbaren Störung bzw. zu einer Verwunderung beim Hörer führen könnte, beispielsweise wenn ein Sprecher zunächst links plaziert war und auf einmal mittig wahrgenommen wird. Daher wird in dem Fall, in dem festgestellt wird, daß der Seite-Kanal zu dämpfen ist,

eine schrittweise Dämpfung des Seite-Kanals beispielsweise unter Verwendung eines vorbestimmten Inkrementwerts, unternommen, derart, daß anschaulich gesprochen der Nachrichtensprecher langsam von der linken Seite in die Mitte "wandert". Wird im entgegengesetzten Fall festgestellt, daß das Maß für die Menge an Bits wieder kleiner als der vorbestimmte Wert ist, wird die Dämpfung nicht schlagartig aufgehoben, sondern wieder langsam auf 0 zurückgeführt, derart, daß um im Beispiel zu bleiben der Sprecher wieder langsam von der Mitte zur Seite "wandert". Diese schrittweise Dämpfung bzw. schrittweise Aufhebung der Dämpfung sollte möglichst langsam stattfinden, damit die Dämpfung des Seite-Kanals praktisch nicht wahrgenommen wird. Die Verringerung der Dämpfung muß jedoch so schnell stattfinden, daß der Codierer aufgrund der hohen Bitrate am Ausgang nicht damit beginnt, die psychoakustische Maskierungsschwelle zu verletzen bzw. Audiobandbreite zu entfernen. Erfindungsgemäß wird daher bei Codierern, die einen Bitreservoir-Mechanismus haben, dieses Bitreservoir ausgenutzt, um die Dämpfung langsam zu erhöhen, bis der Zielwert erreicht ist, bei dem die Dämpfung so hoch ist, daß die vorbestimmte Bitrate am Ausgang des Codierers eingehalten werden kann. Wenn die Dämpfung dann wieder aufgehoben wird, kann das Bitreservoir wieder entleert werden.

Bei der in Fig. 2 dargestellten Implementierung war eine Randbedingung zur Bestimmung von x und y derart beschaffen, daß das Summen-Signal, das dem Mitte-Kanal bis auf den Faktor 0,5 entspricht, nicht verändert wurde. Es sind jedoch Signale denkbar, bei denen der linke und der rechte Kanal zwar ähnlich sind, aber eine Phasenverschiebung im Bereich von 180 Grad zueinander aufweisen. Es sei darauf hingewiesen, daß solche Signale nicht besonders häufig anzutreffen sind, da sie mit Mono-Wiedergabegeräten nicht gut präsentiert werden können. Nichtsdestoweniger sind solche Signale aber denkbar. In diesem Fall würde der Mitte-Kanal M klein und der Seite-Kanal groß werden. Würde dann S so stark gedämpft werden, daß S kleiner als M wird, würde auch die Gesamtlautstärke stark beeinflußt werden. Im Gegensatz zu

einer Verringerung der Stereokanaltrennung ist es jedoch für einen Zuhörer nicht tolerierbar, wenn die Lautstärke stark schwankt, und zwar unabhängig von dem Audiosignal an sich. Eine solche Störung wird ein Hörer als lästig empfinden.

Um diesem Problem aus dem Weg zu gehen, wird es bevorzugt, zusätzlich in der Einrichtung 12 zum Analysieren festzustellen, ob die Phasenverschiebung von L und R in der Nähe von 180 Grad liegt. Wird dies festgestellt, so kann einfach das Vorzeichen von R umgekehrt werden. Zwar geht dann der ursprünglich gewollte räumliche Stereoeffekt verloren, aber der Effekt der reduzierten Lautstärke wird vermieden, was einen Zuhörer weniger stören wird.

Alternativ zur Vorzeichenumkehr könnte der M-Kanal auch in der Einrichtung zum Modifizieren bzw. in einer nachgeschalteten Codierer-Stufe auf einen bestimmten Wert verstärkt werden, derart, daß die Energie des modifizierten M-Kanals in einem vorbestimmten Verhältnis zur Energie des M-Kanals des unmodifizierten Stereoaudiosignals ist. Für das Energieverhältnis wird ein Wert von 1 bevorzugt, wobei jedoch auch durch die Modifiziereinrichtung eine bestimmte Verstärkung oder Dämpfung ausgeführt werden kann, wobei jedoch immer das Verhältnis zum nicht-modifizierten Stereoaudiosignal im wesentlichen beibehalten werden soll, damit ein Zuhörer keine wesentlichen Lautstärkeschwankungen aufgrund der Vorverarbeitung wahrnehmen wird. Selbstverständlich sind kleine Lautstärkeschwankungen nicht so problematisch und manchmal sogar noch nicht wahrnehmbar. Große Lautstärkeschwankungen wird ein Testhörer jedoch als lästig empfinden.

An dieser Stelle sei darauf hingewiesen, daß es unerheblich ist, ob am Eingang 10 der erfindungsgemäßen Vorrichtung zum Verarbeiten eines Stereoaudiosignals zeitdiskrete Abtastwerte anliegen, oder Spektralwerte anliegen. Sämtliche Operationen zur Analyse des Stereoaudiosignals können sowohl mit zeitdiskreten Abtastwerten als auch mit Spektralwerten durchgeführt werden. Darüberhinaus können sämtliche Opera-

tionen in der Einrichtung zum Modifizieren sowohl mit zeitdiskreten Abtastwerten als auch mit Spektralwerten durchgeführt werden. Die erfindungsgemäße Vorrichtung zum Verarbeiten eines Stereoaudiosignals könnte somit auch nach der Zeit-Frequenz-Transformationsstufe eines Zeit/Frequenz-Transformations-basierten Codierers angeordnet sein, wie z. B. einem MPEG-Audio-Codierer. Dieses Konzept ergibt sogar die zusätzliche Möglichkeit, daß die Stereo-Vorverarbeitung frequenzselektiv vorgenommen werden kann, d. h. daß beispielsweise eine unterschiedliche Dämpfung des Signals S in Abhängigkeit der Frequenz durchgeführt werden kann. Dies ist insbesondere sinnvoll, da die Möglichkeit zur Richtungsordnung des menschlichen Gehörs nicht für alle Frequenzen gleich empfindlich ist. Wird somit die erfindungsgemäße Verarbeitung spektralwertweise durchgeführt, so können Spektralwerte des Seite-Kanals um so stärker gedämpft werden, je weniger das menschliche Gehör in diesem Frequenzbereich richtungsabhängig hört, während Spektralwerte nicht oder nur ganz wenig angetastet werden, die in Frequenzbereichen liegen, in denen das menschliche Gehör eine gute Richtungsordnung liefert.

Es sei darauf hingewiesen, daß bei modernen Audiocodierern ohnehin frequenzmäßig unter Verwendung der sogenannten M/S-Maske festgestellt wird, wo eine M/S-Codierung durchgeführt werden soll, und wo eine L/R-Codierung besser ist. In diesem Fall würde die erfindungsgemäße Verarbeitung lediglich auf die Frequenzbereiche angewandt werden, in denen eine MS-Codierung vorhanden ist, d. h. in denen die MS-Maske gesetzt ist. Alternativ könnte auch in mehr Bändern die MS-Maske gesetzt werden, d. h. eine MS-Codierung, durchgeführt werden, wobei in diesen im Vergleich zum bekannten Verfahren zusätzlichen MS-Bändern der Seite-Kanal gedämpft ist, um Bitratenanforderungen einzuhalten.

Im nachfolgenden wird auf Fig. 3 Bezug genommen, in der eine Vorrichtung zum Verarbeiten eines Stereoaudiosignals dargestellt ist, die zusätzlich zu den in Fig. 1 gezeigten Funk-

tionsblöcken ebenfalls einem MS-Codierer 30 sowie einen skalierbaren Codierer 32 umfaßt, der ausgangsseitig einen skalierten Bitstrom BS ausgibt. Der MS-Codierer 30 umfaßt, wie es in der Technik bekannt ist, einen Summierer 30a zum Summieren des modifizierten linken Kanals L' und des modifizierten rechten Kanals R', um nach einer Multiplikation mit einem Multiplizierer 30b, dem ein Faktor von z. B. 0,5 zugeordnet ist, den multiplizierten Mitte-Kanal zu erzeugen. Darüberhinaus umfaßt der MS-Codierer 30 einen Subtrahierer 30c sowie einen weiteren Multiplizierer 30d, um den modifizierten Seite-Kanal S' zu erzeugen, das gegenüber einem Seite-Signal, das aus dem nicht-modifizierten Stereoaudiosignal am Eingang 10 gebildet ist, gedämpft ist. Der Mitte-Kanal M' und der Seite-Kanal S' werden beide in den skalierbaren Codierer 32 eingespeist, der vorzugsweise eine Mono-Stereo-Skalierbarkeit aufweist. Die erste Skalierungsschicht wird das Mono-Signal M' darstellen, und die zweite Skalierungsschicht wird den modifizierten Seite-Kanal S' umfassen. Weitere Skalierungsmöglichkeiten, wie z. B. daß der modifizierte oder nicht-modifizierte Mono-Kanal M' zusätzlich bandbegrenzt ist, und daß in der zweiten Skalierungsschicht neben dem modifizierten Seite-Kanal auch das obere Mono-Band enthalten ist, sind möglich.

Der Effekt der Skalierbarkeit bei dem Mono-Stereo-Codierer 32 ist besonders günstig, wenn eben keine LR-Codierung sondern eine MS-Codierung eingesetzt wird. Die erfindungsgemäße Stereosignalverarbeitung durch die Einrichtungen 12 und 14 ist daher besonders in Verbindung mit dem skalierbaren Codierer 32 besonders vorteilhaft. Um eine Mono-Stereo-Skalierbarkeit zu erhalten, kann nämlich auch eine MS-Codierung eingesetzt werden, wenn sie gegenüber der LR-Codierung eigentlich nicht mehr zu bevorzugen ist. Dies wird eben dadurch erreicht, daß der Seite-Kanal am Eingang des skalierbaren Codierers 32 gegenüber dem unmodifizierten Fall gedämpft ist.

In Fig. 3 ist ferner ein gestrichelter Signalpfad 36 vom

skalierbaren Codierer 32 zur Einrichtung 12 zum Analysieren eingezeichnet. Dieser gestrichelte Signalpfad 36 soll symbolisieren, daß bestimmte Maßnahmen, um ein Maß für die Menge an Bits abzuleiten, die der skalierbare Codierer benötigen wird, um das Stereoaudiosignal am Eingang 10 zu codieren, nicht direkt in der Einrichtung 12 berechnet werden müssen, sondern aus dem skalierbaren Codierer in die Einrichtung 12 ausgegeben werden können, wie z. B. die Perceptual Entropy PE, der Hinweis auf die Verwendung kurzer Fenster usw. Das heißt, daß diese Funktionsblöcke nicht sowohl in der Einrichtung 12 zum Analysieren als auch im skalierbaren Codierer 32 vorhanden sein müssen, sondern daß ihre Implementierung lediglich im skalierbaren Codierer 32 genügt.

In diesem Fall würde die Einrichtung zum Modifizieren 14, um das Maß 18 für die Bitmenge zu bestimmen, zunächst keine Modifikation durchführen. Die in Fig. 3 gezeichnete Vorrichtung wäre dann gewissermaßen in einem "Vorlaufmodus", wo kein Bitstrom geschrieben wird, sondern wo lediglich der erforderliche Dämpfungsgrad für den Seite-Kanal bestimmt wird. Im dann anschließenden Codier-Modus, in dem der Bitstrom BS dann durch den skalierbaren Codierer geschrieben wird, wird die Einrichtung 14 zum Modifizieren mit entsprechend festgelegten Faktoren x, y arbeiten.

Wird die in Fig. 3 gezeigte Vorrichtung mit Spektralwerten für den ersten Kanal L und den zweiten Kanal R betrieben, und ist der skalierbare Codierer ein Zeit/Frequenz-Transformationscodierer, so würde die Stufe des skalierbaren Codierers 32, die die Zeit-Frequenz-Transformation durchführt, dem Eingang 10 vorgeschaltet sein. Die Einrichtungen 12, 14 und 30 wären dann in den skalierbaren Codierer 32 eingebettet.

Die Signalpfade 36a, 36b verdeutlichen, daß auch die modifizierten Kanäle ohne M/S-Codierung zum skalierbaren Codierer geleitet werden können, damit derselbe dann feststellen kann, ob eine M/S- oder L/R-Codierung günstiger ist.

Patentansprüche

1. Vorrichtung zum Verarbeiten eines Stereoaudiosignals, das einen ersten Kanal (L) und einen zweiten Kanal (R) aufweist, mit folgenden Merkmalen:

einer Einrichtung (12) zum Analysieren des Stereoaudiosignals oder eines von dem Stereoaudiosignals abgeleiteten Signals, um ein Maß für eine Menge an Bits zu erhalten, die von einem Codierer (32) benötigt wird, um das Stereoaudiosignal unter Verwendung eines Codieralgorithmus zu codieren; und

einer Einrichtung (14) zum Modifizieren des ersten und zweiten Kanals (L, R), um einen modifizierten ersten und einen modifizierten zweiten Kanal (L', R') zu erhalten,

wobei die Einrichtung (14) zum Modifizieren auf die Einrichtung (12) zum Analysieren anspricht, um wirksam zu werden, wenn das Maß (18) für die Menge an Bits ein vorbestimmtes Maß überschreitet, und

wobei die Einrichtung (12) zum Modifizieren derart ausgestaltet ist, daß eine Charakteristik eines Summensignals aus dem ersten und dem zweiten modifizierten Kanal (L', R'), die ähnlich der Energie des Summensignals verläuft, in einem vorbestimmten Verhältnis zu der Charakteristik eines Summensignals aus dem ersten und dem zweiten Kanal (L, R) ist, und daß ein Differenzsignal aus dem ersten und dem zweiten modifizierten Kanal (L', R') gegenüber einem Differenzsignal aus dem ersten und dem zweiten Kanal (L, R) gedämpft ist.

2. Vorrichtung nach Anspruch 1, bei der die Einrichtung (14) zum Analysieren folgende Merkmale aufweist:

eine Einrichtung zum Bestimmen der Charakteristik der

Summe des ersten und zweiten Kanals über einer vorbestimmten Zeitdauer;

eine Einrichtung zum Bestimmen der Charakteristik der Differenz aus dem ersten und zweiten Kanal über einer vorbestimmten Zeitdauer; und

einer Einrichtung zum Bilden des Verhältnisses der Charakteristik der Summe des ersten und zweiten Kanals und der Charakteristik der Differenz des ersten und zweiten Kanals, wobei das Verhältnis der Charakteristika das Maß (18) für die Menge an Bits ist.

3. Vorrichtung nach Anspruch 1, bei der die Einrichtung (12) zum Analysieren folgende Merkmale aufweist:

eine erste Einrichtung zum Bestimmen eines ersten Charakteristikverhältnisses zwischen dem ersten Kanal und der psychoakustischen Maskierungsschwelle des ersten Kanals über einer vorbestimmten Zeit;

eine zweite Einrichtung zum Bestimmen eines zweiten Charakteristikverhältnisses zwischen dem zweiten Kanal und der psychoakustischen Maskierungsschwelle des zweiten Kanals über einer vorbestimmten Zeit; und

eine Einrichtung zum Summieren des ersten und zweiten Charakteristikverhältnisses, wobei die Summe des ersten und des zweiten Charakteristikverhältnisses auf das Maß (18) für die Menge an Bits hinweist.

4. Vorrichtung nach Anspruch 1, bei der der Codierer (32) angeordnet ist, um ansprechend auf die zeitliche Struktur des Stereoaudiosignals lange oder kurze Fenster für eine Überführung eines zeitlichen Stereoaudiosignals in ein spektrales Stereoaudiosignal zu verwenden, und bei der die Einrichtung (12) zum Analysieren angeordnet ist, um zu erfassen, ob kurze oder lange Fenster im

Codierer (32) eingesetzt werden, wobei das Maß für die Menge an Bits ist, daß kurze Fenster eingesetzt werden.

5. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (14) zum Modifizieren angeordnet ist, um derart wirksam zu werden, daß das Differenzsignal aus dem ersten und zweiten Kanal ausgehend von keiner Dämpfung bis zu einer bestimmten Dämpfung allmählich gedämpft wird, und um derart wirksam zu sein, daß die Dämpfung von der bestimmten Dämpfung zu keiner Dämpfung allmählich reduziert wird.
6. Vorrichtung nach Anspruch 5, bei der die Geschwindigkeit des Dämpfens so langsam als möglich gewählt wird, jedoch noch so schnell, daß ein Bitreservoirmechanismus des Codierers (32) ausgenutzt wird, damit der Codierer (32) weder die Audiobandbreite verringert noch eine psychoakustische Maskierungsschwelle bei einer Quantisierung verletzt.
7. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (14) zum Modifizieren angeordnet ist, um abhängig von dem ermittelten Maß das Differenzsignal adaptiv zu dämpfen.
8. Vorrichtung nach Anspruch 2, bei der die Einrichtung (14) zum Modifizieren angeordnet ist, um abhängig von einem Charakteristikverhältnis, das durch die Einrichtung zum Bilden des Charakteristikverhältnisses erzeugt wird, das Differenzsignal zu dämpfen, so daß die Dämpfung des Differenzsignals hoch ist, wenn das Charakteristikverhältnis klein ist, und daß die Dämpfung des Differenzsignals niedrig ist, wenn das Charakteristikverhältnis hoch ist.
9. Vorrichtung nach Anspruch 7 oder Anspruch 8, bei der die Einrichtung (14) zum Modifizieren derart ausgestaltet ist, daß sie das Differenzsignal derart adaptiv

dämpft, daß das Charakteristikverhältnis des Differenzsignals zum Summensignal zu einem vorbestimmten Wert im wesentlichen gleich ist.

10. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (14) zum Modifizieren folgende Merkmale aufweist:

einen ersten Multiplizierer (22a) zum Multiplizieren des ersten Kanals (L) mit einem ersten Faktor (x);

einen zweiten Multiplizierer (22b) zum Multiplizieren des ersten Kanals (L) mit einem zweiten Faktor (y);

einen dritten Multiplizierer (22c) zum Multiplizieren des zweiten Kanals mit dem ersten Faktor (x);

einen vierten Multiplizierer (22d) zum Multiplizieren des zweiten Kanals (R') mit dem zweiten Faktor (y);

einen ersten Summierer (24a) zum Summieren des Ausgangssignals des ersten Multiplizierers (22a) und des Ausgangssignals des vierten Multiplizierers (22d), um den modifizierten ersten Kanal (L') zu erzeugen; und

einem zweiten Summierer (24b) zum Summieren des Ausgangssignals des dritten Multiplizierers (22c) und des Ausgangssignals des zweiten Multiplizierers (22b), um den modifizierten zweiten Kanal (R') zu erzeugen;

wobei der erste und der zweite Faktor (x, y) derart gewählt sind, daß das Summensignal des ersten und zweiten Kanals und das Summensignal des modifizierten ersten und zweiten Kanals im wesentlichen gleich sind, und das Differenzsignal um einen bestimmten Faktor gedämpft ist.

11. Vorrichtung nach einem der vorhergehenden Ansprüche,

bei der die Einrichtung (12) zum Analysieren ferner folgendes Merkmal aufweist:

eine Einrichtung zum Bestimmen, ob ein Phasenwinkel zwischen dem ersten und zweiten Kanal (L, R) einen Wert in der Nähe von 180° aufweist; und

wobei die Vorrichtung zum Modifizieren (18) ferner folgendes Merkmal aufweist:

eine Einrichtung zum Umkehren des Vorzeichens eines Kanals (L, R), falls der Phasenwinkel in der Nähe von 180° ist.

12. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der der erste und der zweite Kanal (L, R) des Stereosignals durch Spektralwerte gegeben sind, die aus einem zeitlichen Stereosignal durch Überführung in den Spektralbereich erzeugt worden sind, wobei die Einrichtung zum Modifizieren (14) angeordnet ist, um eine frequenzselektive Dämpfung des Differenzsignals durchzuführen.
13. Vorrichtung nach Anspruch 12, bei der die Einrichtung zum Modifizieren angeordnet ist, um in einen Frequenzbereich, in dem die Richtungsordnung des menschlichen Gehörs verringert ist, stärker zu dämpfen als in einem Frequenzbereich, in dem die Richtungsordnung des menschlichen Gehörs nicht verringert ist.
14. Vorrichtung nach einem der vorhergehenden Ansprüche, die ferner folgende Merkmale aufweist:

eine Mitte/Seite-Einrichtung (30) zum Erzeugen eines Mitte-Kanals (M'), der gleich der Hälfte der Summe des modifizierten linken (L') und modifizierten rechten Kanals (R') ist,

eine Seite-Einrichtung (30) zum Erzeugen eines Seite-Kanals, der gleich der Hälfte der Differenz des modifizierten ersten Kanals (L') und des modifizierten zweiten Kanals (R') ist; und

einen skalierbaren Codierer (32), der angeordnet ist, um den Mitte-Kanal (M') zu codieren und in einen Bitstrom (BS) als erste Skalierungsschicht zu schreiben, und der ferner angeordnet ist, um den Seite-Kanal (S') zu codieren und in den Bitstrom (BS) als zweite Skalierungsschicht zu schreiben.

15. Vorrichtung nach Anspruch 14, bei der der skalierbare Codierer (32) angeordnet ist, um für den Fall, daß das Maß für die Menge von Bits einen vorbestimmten Wert überschreitet, eine Bitreservoireinrichtung zu verwenden, damit nicht die Audiobandbreite verringert wird und/oder die psychoakustische Maskierungsschwelle verletzt wird.
16. Verfahren nach einem der vorhergehenden Ansprüche, bei dem die Charakteristik, die ähnlich zur Energie verläuft, die Energie selbst, die Summe von quadrierten Abtastwerten in einer bestimmten Zeitdauer, die Summe von quadrierten Spektralwerten in einem bestimmten Frequenzbereich, die Summe von Abtastwertbeträgen in einer bestimmten Zeitdauer und/oder die Summe von quadrierten Spektralwerten in einem bestimmten Frequenzbereich ist.
17. Verfahren nach einem der vorhergehenden Ansprüche, bei dem das Stereoaudiosignal blockweise verarbeitet wird, und bei dem das beim Analysieren verwendete von dem Stereoaudiosignal abgeleitete Signal das modifizierte Signal eines vorausgehenden Verarbeitungsblocks ist.
18. Verfahren zum Verarbeiten eines Stereoaudiosignals, das einen ersten Kanal (L) und einen zweiten Kanal (R) aufweist, mit folgenden Schritten:

Analysieren (12) des Stereoaudiosignals oder eines von dem Stereoaudiosignal abgeleiteten Signals, um ein Maß für eine Menge an Bits zu erhalten, die von einem Codieralgorithmus benötigt wird, um das Stereoaudiosignal zu codieren; und

Modifizieren (14) des ersten und zweiten Kanals (L, R), um einen modifizierten ersten und einen modifizierten zweiten Kanal (L', R') zu erhalten, falls im Schritt des Analysieren ein Maß (18) für die Menge an Bits bestimmt wird, das ein vorbestimmtes Maß überschreitet, wobei das Modifizieren derart ausgeführt ist, daß eine Charakteristik eines Summensignals aus dem ersten und dem zweiten modifizierten Kanal (L', R'), die ähnlich der Energie des Summensignals verläuft, in einem vorbestimmten Verhältnis zu einer Charakteristik eines Summensignals aus dem ersten und dem zweiten Kanal (L, R) ist, und daß ein Differenzsignal aus dem ersten und dem zweiten modifizierten Kanal (L', R') gegenüber einem Differenzsignal aus dem ersten und dem zweiten Kanal (L, R) gedämpft ist.

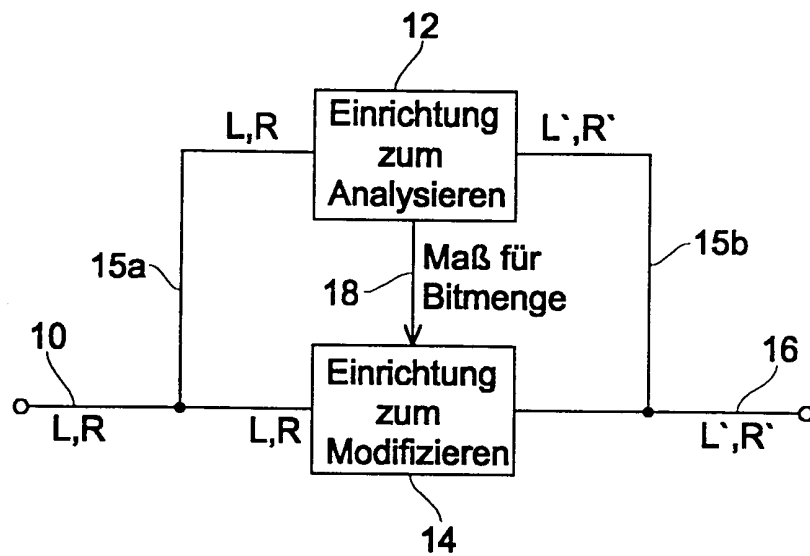


Fig. 1

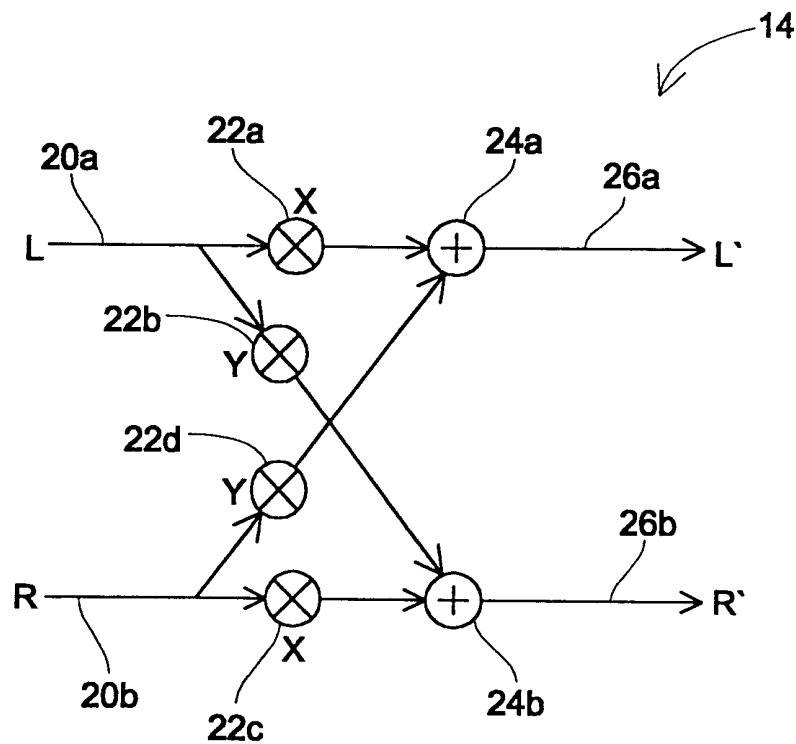


Fig. 2

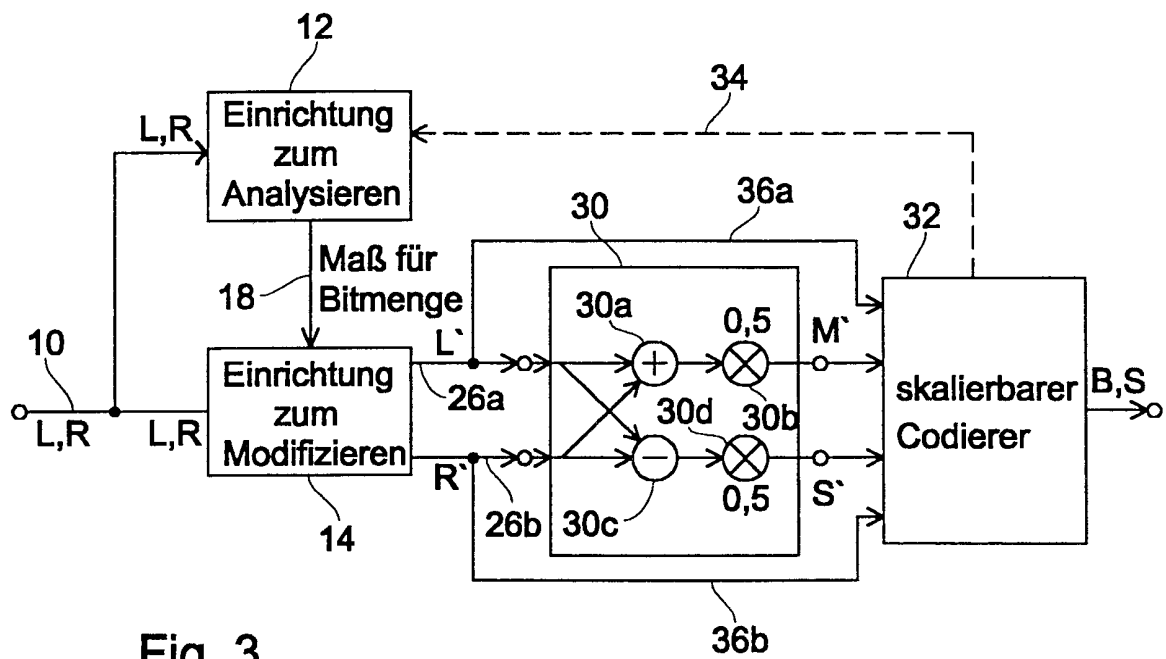


Fig. 3