

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 1 495 464 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
28.09.2005 Patentblatt 2005/39

(21) Anmeldenummer: **02792858.9**

(22) Anmeldetag: **02.12.2002**

(51) Int Cl.7: **G10L 19/02**

(86) Internationale Anmeldenummer:
PCT/EP2002/013623

(87) Internationale Veröffentlichungsnummer:
WO 2003/088212 (23.10.2003 Gazette 2003/43)

(54) **VORRICHTUNG UND VERFAHREN ZUM CODIEREN EINES ZEITDISKRETEN AUDIOSIGNALS UND VORRICHTUNG UND VERFAHREN ZUM DECODIEREN VON CODIERTEN AUDIODATEN**

DEVICE AND METHOD FOR ENCODING A TIME-DISCRETE AUDIO SIGNAL AND DEVICE AND METHOD FOR DECODING CODED AUDIO DATA

DISPOSITIF ET PROCEDE POUR CODER UN SIGNAL AUDIO A TEMPS DISCRET ET DISPOSITIF ET PROCEDE POUR DECODER DES DONNEES AUDIO CODEES

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

(30) Priorität: **18.04.2002 DE 10217297**

(43) Veröffentlichungstag der Anmeldung:
12.01.2005 Patentblatt 2005/02

(73) Patentinhaber: **FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V. 80636 München (DE)**

(72) Erfinder:

- **GEIGER, Ralf**
98704 Wümbach (DE)
- **SPORER, Thomas**
90766 Fürth (DE)
- **BRANDENBURG, Karlheinz**
91054 Erlangen (DE)
- **HERRE, Jürgen**
91054 Erlangen (DE)
- **KOLLER, Jürgen**
98704 Wümbach (DE)
- **DEGUARA, Joachim**
98693 Ilmenau (DE)

(74) Vertreter: **Zinkler, Franz et al**
Patentanwälte Schoppe, Zimmermann, Stöckeler & Zinkler,
Postfach 246
82043 Pullach bei München (DE)

(56) Entgegenhaltungen:

- **MORIYA T ET AL: "A design of lossy and lossless scalable audio coding" 2000 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP). ISTANBUL, TURKEY, JUNE 5-9, 2000, IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING (ICASSP), NEW YORK, NY: IEEE, US, Bd. 2 OF 6, 5. Juni 2000 (2000-06-05), Seiten II889-II892, XP002901946 ISBN: 0-7803-6294-2**
- **GEIGER R ET AL: "AUDIO CODING BASED ON INTEGER TRANSFORMS" PREPRINTS OF PAPERS PRESENTED AT THE 111TH AES CONVENTION, Nr. 5471, 30. November 2001 (2001-11-30), Seiten 1-9, XP008006797**
- **HANS, M.; SCHAFER, R.W.: "Lossless compression of digital audio" SIGNAL PROCESSING MAGAZINE, IEEE, Bd. 18, Nr. 4, Juli 2001 (2001-07), Seiten 21-32, XP001053611**
- **NOLL P AND LIEBCHEN T: "DIGITAL AUDIO: FROM LOSSLESS TO TRANSPARENT CODING" IEEE SIGNAL PROCESSING WORKSHOP, 1999, Seiten 53-60, XP000926389 Poznan**
- **GEIGER R ET AL: "INTMDCT-A LINK BETWEEN PERCEPTUAL AND LOSSLESS AUDIO CODING" 2002 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. PROCEEDINGS. (ICASSP), Bd. 3 OF 4, 13. - 17. Mai 2002, Seiten II-1813-II-1816, XP001097166 Orlando, Florida, USA ISBN: 0-7803-7402-9**

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

EP 1 495 464 B1

- RAAD, M. AND MERTINS, A.: "From lossy to lossless audio coding using SPITH" PROC. OF THE 5TH INT. CONF. ON DIGITAL AUDIO EFFECTS, DAFX-02, [Online] 26. - 28. September 2002, Seiten 245-250, XP002255007 Hamburg, Deutschland Gefunden im Internet:
<URL:www.unibw-hamburg.de/EWEB/ANT/dafx2002/papers/DAFX02_Raad_Mertins_lossy_lossless_coding.pdf> [gefunden am 2003-09-18]

Beschreibung

[0001] Die vorliegende Erfindung bezieht sich auf die Audiocodierung/Audiodecodierung und insbesondere auf skalierbare Codier/Decodier-Algorithmen mit einer psychoakustischen ersten Skalierungsschicht und einer zweiten Skalierungsschicht, die Zusatzaudiosdaten für eine verlustlose Decodierung umfaßt.

[0002] Moderne Audiocodierverfahren, wie z. B. MPEG Layer3 (MP3) oder MPEG AAC verwenden Transformationen wie beispielsweise die sogenannte modifizierte diskrete Kosinustransformation (MDCT), um eine blockweise Frequenzdarstellung eines Audiosignals zu erhalten. Ein solcher Audiocodierer erhält üblicherweise einen Strom von zeitdiskreten Audio-Abtastwerten. Der Strom von Audio-Abtastwerten wird gefenstert, um einen gefensterten Block von beispielsweise 1024 oder 2048 gefensterten Audio-Abtastwerten zu erhalten. Zur Fensterung werden verschiedene Fensterfunktionen eingesetzt, wie z. B. ein Sinus-Fenster, etc.

[0003] Die gefensterten zeitdiskreten Audio-Abtastwerte werden dann mittels einer Filterbank in eine spektrale Darstellung umgesetzt. Prinzipiell kann hierzu eine Fourier-Transformation, oder aus speziellen Gründen eine Abart der Fourier-Transformation, wie z. B. eine FFT oder, wie es ausgeführt worden ist, eine MDCT eingesetzt werden. Der Block von Audio-Spektralwerten am Ausgang der Filterbank kann dann je nach Bedarf weiter verarbeitet werden. Bei den oben bezeichneten Audio-Codierern folgt eine Quantisierung der Audio-Spektralwerte, wobei die Quantisierungsstufen typischerweise so gewählt werden, daß das durch das Quantisieren eingeführt Quantisierungsrauschen unterhalb der psychoakustischen Maskierungsschwelle liegt, d. h. "wegmaskiert" wird. Die Quantisierung ist eine verlustbehaftete Codierung. Um eine weitere Datenmengenreduktion zu erhalten, werden die quantisierten Spektralwerte anschließend beispielsweise mittels einer Huffman-Codierung Entropie-codiert. Durch Hinzufügen von Seiteninformationen, wie z. B. Skalenfaktoren etc. wird aus den Entropie-codierten quantisierten Spektralwerten mittels eines Bitstrom-Multiplexers ein Bitstrom gebildet, der gespeichert oder übertragen werden kann.

[0004] Im Audio-Decodierer wird der Bitstrom mittels eines Bitstrom-Demultiplexers in codierte quantisierte Spektralwerte und Seiteninformationen aufgeteilt. Die Entropie-codierten quantisierten Spektralwerte werden zunächst Entropie-decodiert, um die quantisierten Spektralwerte zu erhalten. Die quantisierten Spektralwerte werden dann invers quantisiert, um decodierte Spektralwerte zu erhalten, die Quantisierungsrauschen aufweisen, das jedoch unterhalb der psychoakustischen Maskierungsschwelle liegt und daher unhörbar sein wird. Diese Spektralwerte werden dann mittels einer Synthese-Filterbank in eine zeitliche Darstellung umgesetzt, um zeitdiskrete decodierte Audio-Abtastwerte zu erhalten. In der Synthese-Filterbank muß ein zum Transformations-Algorithmus inverser Transformations-Algorithmus eingesetzt werden. Außerdem muß nach der Frequenz-Zeit-Rücktransformation das Fenster rückgängig gemacht werden.

[0005] Um eine gute Frequenzselektivität zu erreichen, verwenden moderne Audio-Codierer typischerweise eine Blocküberlappung. Ein solcher Fall ist in Fig. 4a dargestellt. Zunächst werden beispielsweise 2048 zeitdiskrete Audio-Abtastwerte genommen und mittels einer Einrichtung 402 gefenstert. Das Fenster, das die Einrichtung 402 verkörpert, hat eine Fensterlänge von $2N$ Abtastwerten und liefert ausgangsseitig einen Block von $2N$ gefensterten Abtastwerten. Um eine Fensterüberlappung zu erreichen, wird mittels einer Einrichtung 404, die lediglich aus Übersichtlichkeitsgründen in Fig. 4a getrennt von der Einrichtung 402 dargestellt ist, ein zweiter Block von $2N$ gefensterten Abtastwerten gebildet. Die in die Einrichtung 404 eingespeisten 2048 Abtastwerte sind jedoch nicht die an das erste Fenster unmittelbar anschließenden zeitdiskreten Audio-Abtastwerte, sondern beinhalten die zweite Hälfte der durch die Einrichtung 402 gefensterten Abtastwerte und beinhalten zusätzlich lediglich 1024 "neue" Abtastwerte. Die Überlappung ist durch eine Einrichtung 406 in Fig. 4a symbolisch dargestellt, die einen Überlappungsgrad von 50% bewirkt. Sowohl die durch die Einrichtung 402 ausgegebenen $2N$ gefensterten Abtastwerte als auch die durch die Einrichtung 404 ausgegebenen $2N$ gefensterten Abtastwerte werden dann mittels einer Einrichtung 408 bzw. 410 dem MDCT-Algorithmus unterzogen. Die Einrichtung 408 liefert gemäß dem bekannten MDCT-Algorithmus N Spektralwerte für das erste Fenster, während die Einrichtung 410 ebenfalls N Spektralwerte liefert, jedoch für das zweite Fenster, wobei zwischen dem ersten Fenster und dem zweiten Fenster eine Überlappung von 50% besteht.

[0006] Im Decodierer werden die N Spektralwerte des ersten Fensters, wie es in Fig. 4b gezeigt ist, einer Einrichtung 412, die eine inverse modifizierte diskrete Kosinustransformation durchführt, zugeführt. Dasselbe gilt für die N Spektralwerte des zweiten Fensters. Diese werden einer Einrichtung 414 zugeführt, die ebenfalls eine inverse modifizierte diskrete Kosinustransformation durchführt. Sowohl die Einrichtung 412 als auch die Einrichtung 414 liefern jeweils $2N$ Abtastwerte für das erste Fenster bzw. $2N$ Abtastwerte für das zweite Fenster.

[0007] In einer Einrichtung 416, die in Fig. 4b mit TDAC (TDAC = Time Domain Aliasing Cancellation) bezeichnet ist, wird die Tatsache berücksichtigt, daß die beiden Fenster überlappend sind. Insbesondere wird ein Abtastwert y_1 der zweiten Hälfte des ersten Fensters, also mit einem Index $N+k$, mit einem Abtastwert y_2 aus der ersten Hälfte des zweiten Fensters, also mit einem Index k summiert, so daß sich ausgangsseitig, also im Decodierer, N decodierte zeitliche Abtastwerte ergeben.

[0008] Es sei darauf hingewiesen, daß durch die Funktion der Einrichtung 416, die auch als Add-Funktion bezeichnet wird, die in dem durch Fig. 4a schematisch dargestellten Codierer durchgeführte Fensterung gewissermaßen auto-

matisch berücksichtigt wird, so daß in dem durch Fig. 4b dargestellten Decodierer keine explizite "inverse Fensterung" stattzufinden hat.

[0009] Wenn die durch die Einrichtung 402 oder 404 implementierte Fensterfunktion mit $w(k)$ bezeichnet wird, wobei der Index k den Zeitindex darstellt, bezeichnet wird, so muß die Bedingung erfüllt sein, daß das Fenstergewicht $w(k)$ im Quadrat addiert zu dem Fenstergewicht $w(N+k)$ im Quadrat zusammen 1 ergibt, wobei k von 0 bis $N-1$ läuft. Wenn ein Sinus-Fenster verwendet wird, dessen Fenster-Gewichtungen der ersten Halbwelle der Sinus-Funktion folgen, so ist diese Bedingung immer erfüllt, da das Quadrat des Sinus und das Quadrat des Kosinus für jeden Winkel zusammen den Wert 1 ergeben.

[0010] Nachteilig an dem in Fig. 4a beschriebenen Fenster-Verfahren mit anschließender MDCT-Funktion ist die Tatsache, daß die Fensterung durch Multiplikation eines zeitdiskreten Abtastwerts, wenn an ein Sinus-Fenster gedacht wird, mit einer Gleitkommazahl erreicht wird, da der Sinus eines Winkels zwischen 0 und 180 Grad abgesehen von dem Winkel 90 Grad keine Ganzzahl ergibt. Auch wenn ganzzahlige zeitdiskrete Abtastwerte gefenstert werden, entstehen nach dem Fenstern also Gleitkommazahlen.

[0011] Daher ist, auch wenn kein psychoakustischer Codierer verwendet wird, d. h. wenn eine verlustlose Codierung erreicht werden soll, am Ausgang der Einrichtungen 408 bzw. 410 eine Quantisierung notwendig, um eine einigermaßen überschaubare Entropie-Codierung durchführen zu können.

[0012] Wenn also bekannte Transformationen, wie sie anhand von Fig. 4a betrieben worden sind, für ein verlustloses Audiocodieren eingesetzt werden soll, muß entweder eine sehr feine Quantisierung eingesetzt werden, um den resultierenden Fehler aufgrund der Rundung der Gleitkommazahlen vernachlässigen zu können, oder das Fehlersignal muß zusätzlich beispielsweise im Zeitbereich codiert werden.

[0013] Konzepte der ersten Art, also bei denen die Quantisierung so fein eingestellt, daß der resultierende Fehler aufgrund der Rundung der Gleitkommazahlen vernachlässigbar ist, sind beispielsweise in der deutschen Patentschrift DE 197 42 201 C1 offenbart. Hier wird ein Audiosignal in seine spektrale Darstellung überführt und quantisiert, um quantisierte Spektralwerte zu erhalten. Die quantisierten Spektralwerte werden wieder invers quantisiert, in den Zeitbereich überführt und mit dem ursprünglichen Audiosignal verglichen. Liegt der Fehler, also der Fehler zwischen dem ursprünglichen Audiosignal und dem quantisierten/invers quantisierten Audiosignal, oberhalb einer Fehlerschwelle, so wird der Quantisierer rückkopplungsmäßig feiner eingestellt, und der Vergleich wird erneut durchgeführt. Die Iteration ist beendet, wenn die Fehlerschwelle unterschritten wird. Das dann noch möglicherweise vorhandene Restsignal wird mit einem Zeitbereichscodierer codiert und in einen Bitstrom geschrieben, der neben dem Zeitbereichs-codierten Restsignal auch codierte Spektralwerte umfaßt, die gemäß den Quantisierereinstellungen quantisiert worden sind, die zum Zeitpunkt des Abbruchs der Iteration vorhanden waren. Es sei darauf hingewiesen, daß der verwendete Quantisierer nicht von einem psychoakustischen Modell gesteuert werden muß, so daß die codierten Spektralwerte typischerweise genauer quantisiert sind, als dies aufgrund des psychoakustischen Modells sein müßte.

[0014] In der Fachveröffentlichung "A Design of Lossy and Lossless Scalable Audio Coding", T. Moriya u.a., Proc. ICASSP, 2000, ist ein skalierbarer Codierer beschrieben, der als erstes verlustbehaftetes Datenkompressionsmodul z. B. einen MPEG-Codierer umfaßt, der eine blockweise digitale Signalform als Eingangssignal hat und den komprimierten Bitstrom erzeugt. In einem ebenfalls vorhandenen lokalen Decodierer wird die Codierung wieder rückgängig gemacht, und es wird ein codiertes/decodiertes Signal erzeugt. Dieses Signal wird mit dem ursprünglichen Eingangssignal verglichen, indem das codierte/decodierte Signal von dem ursprünglichen Eingangssignal subtrahiert wird. Das Fehlersignal wird dann in ein zweites Modul eingespeist, wo eine verlustlose Bitkonversion verwendet wird. Diese Konversion hat zwei Schritte. Der erste Schritt besteht in einer Konversion von einem Zweierkomplementformat in ein Vorzeichen-Betrag-Format. Der zweite Schritt besteht in der Umwandlung von einer vertikalen Betragsequenz in eine horizontale Bitsequenz in einem Verarbeitungsblock. Die verlustlose Datenumwandlung wird ausgeführt, um die Anzahl von Nullen zu maximieren oder die Anzahl von aufeinanderfolgenden Nullen in einer Sequenz zu maximieren, um eine möglichst gute Komprimierung des zeitlichen Fehlersignals, das als Folge von digitalen Zahlen vorliegt, zu erreichen. Dieses Prinzip basiert auf einem Bit-Slice-Arithmetic-Coding- (BSAC-) Schema, das in der Fachveröffentlichung "Multi-Layer Bit Sliced Bit Rate Scalable Audio Coder", 103. AES-Convention, Preprint Nr. 4520, 1997, dargestellt ist.

[0015] Nachteilig an den vorstehend beschriebenen Konzepten ist die Tatsache, daß die Daten für die verlustlose Erweiterungsschicht, d. h. die Zusatzdaten, die benötigt werden, um eine verlustlose Decodierung des Audiosignals zu erreichen, im Zeitbereich gewonnen werden müssen. Dies bedeutet, daß eine vollständige Decodierung einschließlich einer Frequenz/Zeit-Umsetzung erforderlich ist, um das codierte/decodierte Signal im Zeitbereich zu erhalten, damit mittels einer Abtastwert-weisen Differenzbildung zwischen dem ursprünglichen Audioeingangssignal und dem codierten/decodierten Audiosignal, das aufgrund der psychoakustischen Codierung verlustbehaftet ist, das Fehlersignal berechnet wird. Dieses Konzept ist insbesondere dahingehend nachteilhaft, daß im Codierer, der den Audiodatenstrom erzeugt, sowohl eine komplette Zeit-Frequenz-Umsetzungseinrichtung, wie z. B. eine Filterbank bzw. z. B. ein MDCT-Algorithmus für die Hintransformation benötigt wird, und gleichzeitig, lediglich um das Fehlersignal zu erzeugen, eine komplette inverse Filterbank bzw. ein kompletter Synthesalgorithmus benötigt wird. Der Codierer muß daher zusätzlich zu seinen inhärenten Codierfunktionalitäten auch die komplette Decodierfunktionalität enthalten.

Wenn der Codierer softwaremäßig implementiert ist, so werden hierfür sowohl Speicherkapazitäten als auch Prozessorkapazitäten benötigt, die zu einer Codiererimplementation mit erhöhtem Aufwand führt.

[0016] Die Aufgabe der vorliegenden Erfindung besteht darin, ein weniger aufwendiges Konzept zu schaffen, durch das ein Audiodatenstrom erzeugbar ist, der zumindest nahezu verlustlos decodierbar ist.

[0017] Diese Aufgabe wird durch eine Vorrichtung zum Codieren eines zeitdiskreten Audiosignals nach Patentanspruch 1, durch ein Verfahren zum Codieren eines zeitdiskreten Audiosignals nach Patentanspruch 21, durch eine Vorrichtung zum Decodieren von codierten Audiodaten nach Patentanspruch 22, durch ein Verfahren zum Decodieren von codierten Audiodaten nach Patentanspruch 31 oder durch ein Computer-Programm nach Anspruch 32 oder 33 gelöst.

[0018] Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, daß die Zusatzaudiodaten, die eine verlustlose Decodierung des Audiosignals ermöglichen, dadurch gewonnen werden können, daß ein Block von quantisierten Spektralwerten wie üblich bereitgestellt wird, und dann invers quantisiert wird, um invers quantisierte Spektralwerte zu haben, die aufgrund der Quantisierung mittels eines psychoakustischen Modells verlustbehaftet sind. Diese invers quantisierten Spektralwerte werden dann gerundet, um einen Rundungs-Block von gerundeten invers quantisierten Spektralwerten zu erhalten. Als Referenz zur Differenzbildung wird erfindungsgemäß ein Ganzzahl-Transformationsalgorithmus verwendet, welcher aus einem Block von ganzzahligen zeitdiskreten Abtastwerten einen Ganzzahl-Block von Spektralwerten, der lediglich ganzzahlige Spektralwerte aufweist, erzeugt. Erfindungsgemäß wird nunmehr die Kombination der Spektralwerte im Rundungs-Block und im Ganzzahl-Block spektralwertweise, also im Frequenzbereich durchgeführt, so daß im Codierer selbst kein Synthesalgorithmus, also eine inverse Filterbank oder ein inverser MDCT-Algorithmus etc. benötigt wird. Der Kombinations-Block, der die Differenz-Spektralwerte aufweist, umfaßt aufgrund des Ganzzahl-Transformationsalgorithmus und der gerundeten Quantisierungswerte lediglich ganzzahlige Werte, die auf irgendeine bekannte Art und Weise Entropie-codiert werden können. Es sei darauf hingewiesen, daß zur Entropie-Codierung des Kombinationsblocks beliebige Entropie-Codierer eingesetzt werden können, wie z. B. Huffman-Codierer oder arithmetische Codierer etc.

[0019] Zur Codierung der quantisierten Spektralwerte des Quantisierungsblocks können ebenfalls beliebige Codierer eingesetzt werden, wie z. B. die bekannten für moderne Audiocodierer üblichen Werkzeuge.

[0020] Es sei darauf hingewiesen, daß das erfindungsgemäße Codier/Decodierkonzept kompatibel ist mit modernen Codierwerkzeugen, wie z. B. Fenster-Umschalten, TNS oder Mitte/Seite-Codierung für mehrkanalige Audiosignale.

[0021] Bei einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird zum Liefern eines Quantisierungsblocks von unter Verwendung eines psychoakustischen Modells quantisierten Spektralwerten eine MDCT eingesetzt. Darüber hinaus wird es bevorzugt, als Ganzzahl-Transformationsalgorithmus eine sogenannte IntMDCT einzusetzen.

[0022] Bei einem alternativen Ausführungsbeispiel der vorliegenden Erfindung kann auf die übliche MDCT verzichtet werden, und es kann die IntMDCT als Annäherung für die MDCT verwendet werden, und zwar dahingehend, daß das ganzzahlige Spektrum, das durch den Ganzzahl-Transformationsalgorithmus erhalten wird, einem psychoakustischen Quantisierer zugeführt wird, um quantisierte IntMDCT-Spektralwerte zu erhalten, die dann wieder invers quantisiert und gerundet werden, um mit den ursprünglichen Ganzzahl-Spektralwerten verglichen zu werden. In diesem Fall wird lediglich nur noch eine einzige Transformation benötigt, nämlich die IntMDCT, die aus ganzzahligen zeitdiskreten Abtastwerten ganzzahlige Spektralwerte erzeugt.

[0023] Typischerweise arbeiten Prozessoren mit ganzen Zahlen, bzw. jede Gleitkommazahl ist als eine ganze Zahl darstellbar. Wenn eine Ganzzahl-Arithmetik in einem Prozessor verwendet wird, so kann auf das Runden der invers quantisierten Spektralwerte verzichtet werden, da aufgrund der Arithmetik des Prozessors ohnehin gerundete Werte, nämlich innerhalb der Genauigkeit des LSB, d. h. des niederwertigen Bits, vorliegen. In diesem Fall wird eine vollständig verlustlose Verarbeitung erreicht, d. h. eine Verarbeitung innerhalb der Genauigkeit des verwendeten Prozessorsystems. Alternativ kann jedoch eine Rundung auf eine gröbere Genauigkeit durchgeführt werden, dahingehend, daß das Differenzsignal im Kombinationsblock auf die durch eine Rundungsfunktion festgelegte Genauigkeit gerundet ist. Das Einführen einer Rundung über die inhärente Rundung eines Prozessorsystems hinaus ermöglicht eine Flexibilität dahingehend, den "Grad" der Verlustlosigkeit der Codierung zu beeinflussen, um im Sinne einer Datenkompression einen nahezu verlustlosen Codierer zu schaffen.

[0024] Der erfindungsgemäße Decodierer zeichnet sich dadurch aus, daß aus den Audiodaten sowohl die psychoakustisch codierten Audiodaten als auch die Zusatzaudiodaten extrahiert werden, einer möglicherweise vorhandenen Entropie-Decodierung unterzogen werden und dann wie folgt verarbeitet werden. Zunächst wird der Quantisierungsblock im Decodierer invers quantisiert und unter Verwendung derselben Rundungsfunktion, die auch im Codierer eingesetzt worden ist, gerundet, um dann zu den Entropie-decodierten Zusatzaudiodaten hinzuaddiert zu werden. Im Decodierer liegen dann sowohl eine psychoakustisch komprimierte spektrale Darstellung des Audiosignals als auch eine verlustlose Darstellung des Audiosignals vor, wobei die psychoakustisch komprimierte spektrale Darstellung des Audiosignals in den Zeitbereich umzusetzen ist, um ein verlustbehaftetes codiertes/decodiertes Audiosignal zu erhalten, während die verlustlose Darstellung unter Verwendung eines zum Ganzzahl-Transformationsalgorithmus inversen

Ganzzahl-Transformationsalgorithmus in den Zeitbereich umgesetzt wird, um ein verlustlos oder, wie es ausgeführt worden ist, nahezu verlustlos codiertes/decodiertes Audiosignal zu erhalten.

[0025] Bevorzugte Ausführungsbeispiele der vorliegenden Erfindung werden nachfolgend Bezug nehmend auf die beiliegenden Zeichnungen detailliert erläutert. Es zeigen:

- Fig. 1 ein Blockschaltbild einer bevorzugten Einrichtung zum Verarbeiten von zeitdiskreten Audio-Abtastwerten, um ganzzahlige Werte zu erhalten, aus denen ganzzahlige Spektralwerte ermittelbar sind;
- Fig. 2 eine schematische Darstellung der Zerlegung einer MDCT und einer inversen MDCT in Givens-Rotationen und zwei DCT-IV-Operationen;
- Fig. 3 eine Darstellung zur Veranschaulichung der Zerlegung der MDCT mit 50-Prozent-Überlappung in Rotationen und DCT-IV-Operationen;
- Fig. 4a ein schematisches Blockschaltbild eines bekannten Codierers mit MDCT und 50 Prozent Überlappung;
- Fig. 4b ein Blockschaltbild eines bekannten Decodierers zum Decodieren der durch Fig. 4a erzeugten Werte;
- Fig. 5 ein Prinzipblockschaltbild eines bevorzugten erfindungsgemäßen Codierers;
- Fig. 6 ein Prinzipblockschaltbild eines alternativen erfindungsgemäß bevorzugten Codierers;
- Fig. 7 ein Prinzipblockschaltbild eines erfindungsgemäß bevorzugten Decodierers;
- Fig. 8a eine schematische Darstellung eines Bitstroms mit einer ersten Skalierungsschicht und einer zweiten Skalierungsschicht;
- Fig. 8b eine schematische Darstellung eines Bitstroms mit einer ersten Skalierungsschicht und mehreren weiteren Skalierungsschichten; und
- Fig. 9 eine schematische Darstellung von binär codierten Differenz-Spektralwerten zur Verdeutlichung möglicher Skalierungen hinsichtlich der Genauigkeit (Bits) der Differenz-Spektralwerte und/oder hinsichtlich der Frequenz (Abtastrate) der Differenz-Spektralwerte.

[0026] Im nachfolgenden wird anhand der Fig. 5 bis 7 auf erfindungsgemäße Codiererschaltungen (Fig. 5 und Fig. 6) bzw. eine erfindungsgemäß bevorzugte Decodierschaltung (Fig. 7) eingegangen. Der in Fig. 5 gezeigte erfindungsgemäße Codierer umfaßt einen Eingang 50, in den ein zeitdiskretes Audiosignal einspeisbar ist, sowie einen Ausgang 52, aus dem codierte Audiodaten ausgebar sind. Das am Eingang 50 eingespeiste zeitdiskrete Audiosignal wird in eine Einrichtung 52 zum Liefern eines Quantisierungs-Blocks eingespeist, der ausgangsseitig einen Quantisierungs-Block des zeitdiskreten Audiosignals liefert, der unter Verwendung eines psychoakustischen Modells 54 quantisierte Spektralwerte des zeitdiskreten Audiosignals 50 aufweist. Der erfindungsgemäße Codierer umfaßt ferner eine Einrichtung zum Erzeugen eines Ganzzahl-Blocks unter Verwendung eines Ganzzahl-Transformationsalgorithmus 56, wobei der Ganzzahl-Algorithmus wirksam ist, um aus ganzzahligen zeitdiskreten Abtastwerten ganzzahlige Spektralwerte zu erzeugen.

[0027] Der erfindungsgemäße Codierer umfaßt ferner eine Einrichtung 58 zum inversen Quantisieren des Quantisierungs-Blocks, der von der Einrichtung 52 ausgegeben wird, und, wenn eine andere Genauigkeit als die Prozessor-genauigkeit erforderlich ist, eine Rundungsfunktion. Wenn bis zur Genauigkeit des Prozessorsystems, wie es ausgeführt worden ist, gegangen werden soll, so ist die Rundungsfunktion bereits inhärent bei dem inversen Quantisieren des Quantisierungs-Blocks enthalten, da ein Prozessor, der eine Ganzzahlarithmetik hat, ohnehin nicht in der Lage ist, nicht-ganzzahlige Werte zu liefern. Die Einrichtung 58 liefert somit einen sogenannten Rundungs-Block, der invers quantisierte Spektralwerte umfaßt, die ganzzahlig sind, d. h. inhärent oder explizit gerundet worden sind. Sowohl der Rundungs-Block als auch der Ganzzahl-Block werden einer Kombinationseinrichtung zugeführt, die unter Verwendung einer Differenzbildung einen Differenz-Block mit Differenz-Spektralwerten liefert, wobei der Ausdruck "Differenz-Block" darauf hindeuten soll, daß die Differenz-Spektralwerte Werte sind, die Unterschiede zwischen dem Ganzzahl-Block und dem Rundungs-Block umfassen.

[0028] Sowohl der Quantisierungs-Block, der aus der Einrichtung 52 ausgegeben wird, als auch der Differenz-Block, der aus der Differenzbildungseinrichtung 58 ausgegeben wird, werden einer Verarbeitungseinrichtung 60 zugeführt, die z. B. eine übliche Verarbeitung des Quantisierungs-Blocks durchführt, und die ferner z. B. eine Entropie-Codierung

des Differenz-Blocks bewirkt. Die Einrichtung 60 zum Verarbeiten gibt an dem Ausgang 52 codierte Audiodaten aus, die sowohl Informationen über den Quantisierungs-Block enthalten als auch Informationen über den Differenz-Block umfassen.

[0029] Bei einem ersten bevorzugten Ausführungsbeispiel wird, wie es in Fig. 6 gezeigt ist, das zeitdiskrete Audiosignal mittels einer MDCT in seine spektrale Darstellung umgesetzt und dann quantisiert. Die Einrichtung 52 zum Liefern des Quantisierungsblocks besteht somit aus der MDCT-Einrichtung 52a und einem Quantisierer 52b.

[0030] Darüber hinaus wird es bevorzugt, den Ganzzahl-Block mit einer IntMDCT 56 als ganzzahligem Transformationsalgorithmus zu erzeugen.

[0031] In Fig. 6 ist ferner die in Fig. 5 gezeigte Verarbeitungseinrichtung 60 als Bitstrom-Codiereinrichtung 60a zum Bitstrom-Codieren des Quantisierungs-Blocks, der durch die Einrichtung 52b ausgegeben wird, sowie durch einen Entropie-Codierer 60b zum Entropie-Codieren des Differenz-Blocks dargestellt. Der Bitstrom-Codierer 60a gibt die psychoakustisch codierten Audiodaten aus, während der Entropie-Codierer 60b einen Entropie-codierten Differenz-Block ausgibt. Die beiden Ausgangsdaten der Blöcke 60a und 60b können in geeigneter Weise in einen Bitstrom kombiniert werden, der als erste Skalierungsschicht die psychoakustisch codierten Audiodaten hat, und der als zweite Skalierungsschicht die Zusatzaudiodaten für eine verlustlose Decodierung hat. Der skalierte Bitstrom entspricht dann den in Fig. 5 gezeigten codierten Audiodaten am Ausgang 52 des Codierers.

[0032] Bei einem alternativen bevorzugten Ausführungsbeispiel kann auf den MDCT-Block 52a von Fig. 6 verzichtet werden, wie es in Fig. 5 durch einen gestrichelten Pfeil 62 angedeutet ist. In diesem Fall wird das Ganzzahl-Spektrum, das durch die Ganzzahl-Transformationseinrichtung 56 geliefert wird, sowohl in die Differenz-Bildungseinrichtung 58 eingespeist als auch in den Quantisierer 52b von Fig. 6. Die Spektralwerte, die durch die Ganzzahl-Transformation erzeugt werden, werden hier gewissermaßen als Annäherung für eine übliches MDCT-Spektrum verwendet. Dieses Ausführungsbeispiel hat den Vorteil, daß nur der IntMDCT-Algorithmus im Codierer vorhanden ist, und daß nicht sowohl der IntMDCT-Algorithmus als auch der MDCT-Algorithmus im Codierer vorhanden sein müssen.

[0033] Wieder Bezug nehmend auf Fig. 6 sei darauf hingewiesen, daß die durchgezogenen Blöcke und Linien einen üblichen Audio-codierern nach einem der MPEG-Standards darstellen, während die gestrichelten Blöcke und Linien die Erweiterung eines solchen üblichen MPEG-Codierers darstellen. Es ist also zu sehen, daß keine grundsätzliche Änderung des üblichen MPEG-Codierers erforderlich sind, sondern daß die erfindungsgemäße Gewinnung der Zusatzaudiodaten für eine verlustlose Codierung mittels einer Ganzzahl-Transformation ohne Änderung der Codierer/Decodierer-Grundstruktur hinzugefügt werden können.

[0034] Fig. 7 zeigt ein Prinzipblockschaltbild eines erfindungsgemäßen Decodierers zum Decodieren der an dem Ausgang 52 von Fig. 5 ausgegebenen codierten Audiodaten. Diese werden zunächst in psychoakustisch codierte Audiodaten einerseits und die Zusatzaudiodaten andererseits zerlegt. Die psychoakustisch codierten Audiodaten werden einem üblichen Bitstrom-Decodierer 70 zugeführt, während die Zusatzaudiodaten, wenn sie im Codierer Entropie-codiert worden sind, mittels eines Entropie-Decodierers 72 Entropie-decodiert werden. Am Ausgang des Bitstrom-Decodierers 70 von Fig. 7 liegen quantisierte Spektralwerte vor, die einem inversen Quantisierer 74 zugeführt werden, der prinzipiell identisch zu dem inversen Quantisierer in der Einrichtung von Fig. 6 aufgebaut sein kann. Wird eine Genauigkeit angestrebt, die nicht der Prozessor-Genauigkeit entspricht, so ist im Decodierer ferner eine Rundungseinrichtung 76 vorgesehen, die denselben Rundungsalgorithmus bzw. dieselbe Rundungsfunktion zum Abbilden einer reellen Zahl auf eine Ganzzahl durchführt, wie sie auch in der Einrichtung 58 von Fig. 6 implementiert sein kann. In einem decodiererseitigen Kombinerer 78 werden die gerundeten invers quantisierten Spektralwerte mit den Entropie-codierten Zusatzaudiodaten spektralwertweise vorzugsweise additiv kombiniert, so daß im Decodierer zum einen invers quantisierte Spektralwerte am Ausgang der Einrichtung 74 vorliegen und zum anderen Ganzzahl-Spektralwerte am Ausgang des Kombinerers 78 vorliegen.

[0035] Die ausgangsseitigen Spektralwerte der Einrichtung 74 können dann mittels einer Einrichtung 80 zum Durchführen einer inversen modifizierten diskreten Kosinustransformation in den Zeitbereich umgesetzt werden, um ein verlustbehaftetes psychoakustisch codiertes und wieder decodiertes Audiosignal zu erhalten. Mittels einer Einrichtung 82 zum Durchführen einer inversen Ganzzahl-MDCT (IntMDCT) wird ferner das Ausgangssignal des Kombinerers 78 in seine zeitliche Darstellung umgesetzt, um ein verlustlos codiertes/decodiertes Audiosignal oder ein, wenn eine entsprechende gröbere Rundung eingesetzt worden ist, ein nahezu verlustlos codiertes und wieder decodiertes Audiosignal zu erzeugen.

[0036] Im nachfolgenden wird auf eine besondere bevorzugte Ausführungsform des Entropie-Codierers 60b von Fig. 6 eingegangen. Nachdem in einem üblichen modernen MPEG-Codierer mehrere Codetabellen, die abhängig von einer durchschnittlichen Statistik der quantisierten Spektralwerte ausgewählt werden, vorliegen, wird es bevorzugt, dieselben Codetabellen oder Codebooks auch für die Entropie-Codierung des Differenz-Blocks am Ausgang des Kombinerers 58 zu verwenden. Nachdem der Betrag des Differenz-Blocks, also des Rest-IntMDCT-Spektrums, von der Genauigkeit der Quantisierung abhängt, kann eine Codebuch-Auswahl für den Entropie-Codierer 60b ohne zusätzliche Seiteninformationen durchgeführt werden.

[0037] In einem MPEG-2-AAC-Codierer sind die spektralen Koeffizienten, also die quantisierten Spektralwerte im

Quantisierungs-Block in Skalenfaktorbander gruppiert, wobei die Spektralwerte mit einem Verstärkungsfaktor gewichtet sind, der von einem entsprechenden Skalenfaktor, der einem Skalenfaktorband zugeordnet ist, abgeleitet ist. Da in diesem bekannten Codiererkonzept ein ungleichmäßiger Quantisierer verwendet wird, um die gewichteten Spektralwerte zu quantisieren, hängt die Größe der Restwerte, also der Spektralwerte am Ausgang des Kombinerers 58, nicht nur von den Skalenfaktoren ab, sondern auch von den quantisierten Werten selbst. Nachdem jedoch sowohl die Skalenfaktoren als auch die quantisierten Spektralwerte in dem Bitstrom, der von der Einrichtung 60a von Fig. 6 erzeugt wird, also in den psychoakustisch codierten Audiodaten enthalten sind, wird es bevorzugt, eine Codebuch-Auswahl im Codierer abhängig von der Größe der Differenz-Spektralwerte durchzuführen und ferner im Decodierer die im Codierer verwendete Codetabelle auf der Basis sowohl der im Bitstrom übertragenen Skalenfaktoren als auch der quantisierten Werte zu ermitteln. Nachdem zum Entropie-Codieren der Differenz-Spektralwerte am Ausgang des Kombinerers 58 keine Seiteninformationen übertragen werden müssen, führt die Entropie-Codierung lediglich zu einer Datenratenkompression ohne daß irgendwelche Signalisierungsbits im Datenstrom als Seiteninformationen für den Entropie-Codierer 60b aufgewendet werden müßten.

[0038] In einem Audiocodierer nach dem Standard MPEG-2 AAC wird eine Fensterumschaltung verwendet, um Vorechos in transienten Audiosignalbereichen zu vermeiden. Diese Technik basiert auf der Möglichkeit, Fensterformen individuell in jeder Hälfte des MDCT-Fensters auszuwählen, und erlaubt es, die Blockgröße in aufeinanderfolgenden Blöcken zu variieren. Auf ähnliche Art und Weise ist der Ganzzahl-Transformationsalgorithmus in Form der IntMDCT, auf die Bezug nehmend auf die Fig. 1 bis 3 eingegangen wird, ausgeführt, um ebenfalls unterschiedliche Fensterformen beim Fenstern und beim Zeitbereich-Aliasing-Abschnitt der MDCT-Zerlegung zu verwenden. Es wird daher bevorzugt, sowohl für den Ganzzahl-Transformationsalgorithmus als auch für den Transformationsalgorithmus zum Erzeugen des Quantisierungsblocks dieselben Fensterentscheidungen zu verwenden.

[0039] In einem Codierer nach MPEG-2 AAC existieren ferner mehrere weitere Codierwerkzeuge, von denen lediglich TNS (TNS = Temporal Noise Shaping) und Mitte/Seite- (MS) Stereocodierung erwähnt seien. Bei einer TNS-Codierung wird genauso wie bei einer MS-Codierung eine Modifikation der Spektralwerte vor der Quantisierung durchgeführt. Folglich nimmt die Differenz zwischen den IntMDCT-Werten, also dem Ganzzahl-Block, und den quantisierten MDCT-Werten zu. Erfindungsgemäß ist der Ganzzahl-Transformationsalgorithmus ausgebildet, um sowohl eine TNS-Codierung als auch eine Mitte/Seite-Codierung auch von Ganzzahl-Spektralwerten zuzulassen. Die TNS-Technik basiert auf einer adaptiven Vorwärtsprädiktion der MDCT-Werte über der Frequenz. Dasselbe Prädiktionsfilter, das von einem üblichen TNS-Modul signaladaptiv berechnet wird, wird vorzugsweise ebenfalls dazu verwendet, die Ganzzahl-Spektralwerte zu präzisieren, wobei, falls dadurch nicht-ganzzahlige Werte entstehen, eine nachgeschaltete Rundung eingesetzt werden kann, um wieder ganzzahlige Werte zu erzeugen. Diese Rundung erfolgt vorzugsweise nach jedem Prädiktionsschritt. In dem Decodierer kann das ursprüngliche Spektrum wieder rekonstruiert werden, indem das inverse Filter und dieselbe Rundungsfunktion eingesetzt werden. Auf ähnliche Art und Weise kann die MS-Codierung ebenfalls auf IntMDCT-Spektralwerte angewendet werden, indem gerundete Givens-Rotationen mit einem Winkel von $\pi/4$, basierend auf dem Lifting-Schema, eingesetzt werden. Dadurch können die ursprünglichen IntMDCT-Werte in dem Decodierer wieder rekonstruiert werden.

[0040] Es sei darauf hingewiesen, daß das erfindungsgemäße Konzept in seiner bevorzugten Ausprägung mit der IntMDCT als ganzzahligem Transformationsalgorithmus auf alle MDCT-basierten gehörangepaßten Audiocodierer angewendet werden kann. Lediglich beispielhaft sind solche Codierer Codierer nach MPEG-4 AAC Scalable, MPEG-4 AAC Low Delay, MPEG-4 BSAC, MPEG-4 Twin VQ, Dolby AC-3 etc.

[0041] Es sei besonders darauf hingewiesen, daß das erfindungsgemäße Konzept rückwärts kompatibel ist. Der gehörangepaßte Codierer bzw. Decodierer wird nicht verändert, sondern nur erweitert. Zusatzinformationen für die verlustlosen Komponenten lassen sich rückwärts kompatibel im gehörangepaßt codierten Bitstrom übertragen, beispielsweise bei MPEG-2 AAC im Feld "Ancillary Data". Der Zusatz zum bisherigen gehörangepaßten Decodierer, der in Fig. 7 gestrichelt gezeichnet ist, kann diese Zusatzdaten auswerten und zusammen mit dem quantisierten MDCT-Spektrum vom gehörangepaßten Decodierer das IntMDCT-Spektrum verlustlos rekonstruieren.

[0042] Das erfindungsgemäße Konzept der psychoakustischen Codierung, ergänzt um eine verlustlose bzw. nahezu verlustlose Codierung eignet sich besonders für die Erzeugung, Übertragung und Decodierung von skalierbaren Datenströmen. Es ist bekannt, daß skalierbare Datenströme verschiedene Skalierungsschichten umfassen, von denen zumindest die unterste Skalierungsschicht unabhängig von den höheren Skalierungsschichten übertragen und decodiert werden kann. Weitere Skalierungsschichten oder Enhancement Layers werden bei einer skalierbaren Verarbeitung von Daten der ersten Skalierungsschicht oder Grundlayer hinzugefügt. Ein voll ausgestatteter Codierer kann einen skalierten Datenstrom erzeugen, der eine erste Skalierungsschicht hat, und der prinzipiell beliebig viele weitere Skalierungsschichten aufweist. Ein Vorteil des Skalierungskonzepts besteht darin, daß in dem Fall, in dem ein breitbandiger Übertragungskanal zur Verfügung steht, der vom Codierer erzeugte skalierte Datenstrom komplett, also inklusive aller Skalierungsschichten über den breitbandigen Übertragungskanal übertragen werden kann. Ist dagegen lediglich ein schmalbandiger Übertragungskanal vorhanden, so kann das codierte Signal dennoch über den Übertragungskanal übertragen werden, jedoch lediglich in Form der ersten Skalierungsschicht oder einer bestimmten Anzahl

von weiteren Skalierungsschichten, wobei die bestimmte Anzahl kleiner als die gesamte vom Codierer erzeugte Anzahl von Skalierungsschichten ist. Selbstverständlich kann der Codierer bereits, angepaßt auf den Kanal, mit dem er verbunden ist, die Grundskalierungsschicht oder erste Skalierungsschicht und eine vom Kanal abhängige Anzahl weiterer Skalierungsschichten erzeugen.

[0043] Auf Decodierseite hat das skalierbare Konzept ferner den Vorteil, daß es rückwärts kompatibel ist. Dies bedeutet, daß ein Decodierer, der nur in der Lage ist, die erste Skalierungsschicht zu verarbeiten, einfach die zweite und weitere Skalierungsschichten im Datenstrom ignoriert und ein nützliches Ausgangssignal erzeugen kann. Ist der Decodierer dagegen ein typischerweise modernerer Decodierer, der mehrere Skalierungsschichten aus dem skalierten Datenstrom verarbeiten kann, so kann dieser Codierer mit demselben Datenstrom wie ein Basis-Decodierer angesprochen werden.

[0044] Bei der vorliegenden Erfindung besteht die grundsätzliche Skalierbarkeit darin, daß der Quantisierungsblock, also die Ausgabe des Bitstrom-Codierers 60a in eine erste Skalierungsschicht 81 von Fig. 8 geschrieben wird, die, wenn Fig. 6 betrachtet wird, psychoakustisch codierte Daten z.B. für einen Frame umfaßt. Die vorzugsweise Entropie-codierten Differenz-Spektralwerte, die durch die Kombinationseinrichtung 58 erzeugt werden, werden bei einfacher Skalierbarkeit in die zweite Skalierungsschicht geschrieben, die in Fig. 8a mit 82 bezeichnet ist und somit die Zusatzaudiodaten für einen Frame umfaßt.

[0045] Ist der Übertragungskanal vom Codierer zum Decodierer ein breitbandiger Übertragungskanal, so können beide Skalierungsschichten 81 und 82 zum Decodierer übertragen werden. Ist jedoch der Übertragungskanal ein schmalbandiger Übertragungskanal, in den nur die erste Skalierungsschicht "paßt", so kann einfach die zweite Skalierungsschicht vor der Übertragung vom Datenstrom entfernt werden, so daß ein Decodierer nur mit der ersten Skalierungsschicht angesprochen wird.

[0046] Auf Decodierseite kann ein "Basis-Decodierer", der nur die psychoakustisch codierten Daten verarbeiten kann, die zweite Skalierungsschicht 82 einfach weglassen, sofern er diese über einen breitbandigen Übertragungskanal empfangen hat. Ist der Decodierer jedoch ein voll ausgestatteter Decodierer, der sowohl einen psychoakustischen Decodieralgorithmus als auch einen Ganzzahl-Decodieralgorithmus umfaßt, so kann dieser voll ausgestattete Decodierer sowohl die erste Skalierungsschicht als auch die zweite Skalierungsschicht zum Decodieren nehmen, um ein verlustlos codiertes und wieder decodiertes Ausgangssignal zu erzeugen.

[0047] Bei einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung werden, wie es in Fig. 8a schematisch dargestellt ist, wieder in einer ersten Skalierungsschicht die psychoakustisch codierten Daten für einen Frame sein. Die zweite-Skalierungsschicht von Fig. 8a wird nun jedoch feiner skaliert, so daß aus dieser zweiten Skalierungsschicht in Fig. 8a mehrere Skalierungsschichten entstehen, wie z.B. eine (kleinere) zweite Skalierungsschicht, eine dritte Skalierungsschicht, eine vierte Skalierungsschicht, etc.

[0048] Die von dem Addierer 58 ausgegebenen Differenz-Spektralwerte eignen sich besonders gut für eine weitere Unter-Skalierung, wie es anhand von Fig. 9 dargestellt wird. Fig. 9 stellt schematisch binär codierte Spektralwerte dar. Jede Zeile 90 in Fig. 9 stellt einen binär codierten Differenz-Spektralwert dar. In Fig. 9 sind die Differenz-Spektralwerte nach der Frequenz geordnet, wie es durch einen Pfeil 91 angedeutet ist. So hat ein Differenz-Spektralwert 92 eine höhere Frequenz als der Differenz-Spektralwert 90. Die erste Spalte des Tableaus in Fig. 9 stellt das höchstwertige Bit eines Differenz-Spektralwerts vor. Die zweite Stelle stellt das Bit mit der Wertigkeit MSB-1 dar. Die dritte Spalte stellt ein Bit mit der Wertigkeit MSB-2 dar. Die drittletzte Spalte stellt ein Bit mit der Wertigkeit LSB+2 dar. Die vorletzte Spalte stellt ein Bit mit der Wertigkeit LSB+1 dar. Schließlich stellt die letzte Spalte ein Bit mit der Wertigkeit LSB, also das niederstwertige Bit eines Differenz-Spektralwerts, dar.

[0049] Bei einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung wird eine Genauigkeits-Skalierung dahingehend vorgenommen, daß als zweite Skalierungsschicht die z.B. 16 höchstwertigsten Bits eines Differenz-Spektralwerts genommen werden, um dann, falls gewünscht, durch den Entropie-Codierer 60b entropie-codiert zu werden. Ein Decodierer, der die zweite Skalierungsschicht verwendet, erhält ausgangsseitig Differenz-Spektralwerte mit einer Genauigkeit von 16 Bits, so daß die zweite Skalierungsschicht zusammen mit der ersten Skalierungsschicht ein verlustlos decodiertes Audiosignal in CD-Qualität liefert. Es ist bekannt, daß Audioabtastwerte in CD-Qualität mit einer Breite von 16 Bits vorliegen.

Wird in den Codierer dagegen ein Audiosignal in Studioqualität eingespeist, also ein Audiosignal mit Abtastwerten, wobei jeder Abtastwert 24 Bits umfaßt, so kann der Codierer ferner eine dritte Skalierungsschicht erzeugen, die die letzten acht Bit eines Differenz-Spektralwerts umfaßt und ebenfalls je nach Bedarf entropie-codiert wird (Einrichtung 60 von Fig. 6).

[0050] Ein voll ausgestatteter Decodierer, der den Datenstrom mit der ersten Skalierungsschicht, der zweiten Skalierungsschicht (16 höchstwertige Bits der Differenz-Spektralwerte) und der dritten Skalierungsschicht (8 niederwertige Bits eines Differenz-Spektralwerts) erhält, so kann der Decodierer unter Verwendung aller drei Skalierungsschichten ein verlustlos codiertes/decodiertes Audiosignal in Studioqualität, also mit einer am Ausgang des Decodierers vorliegenden Wortbreite eines Abtastwerts von 24 Bits liefern.

[0051] Es sei darauf hingewiesen, daß im Studiobereich höhere Wortlängen der Abtastwerte üblich sind als im Con-

sumer-Bereich. Im Consumer-Bereich beträgt die Wortbreite 16 Bits bei einer Audio-CD, während im Studiobereich 24 Bit oder 20 Bit eingesetzt werden.

[0052] Basierend auf dem Konzept der Skalierung im IntMDCT-Bereich können somit, wie es dargelegt worden ist, alle drei Genauigkeiten (16 Bit, 20 Bit oder 24 Bit) oder beliebige um minimal 1 Bit skalierte Genauigkeiten skalierbar codiert werden.

[0053] Hierbei wird das mit 24-Bit-Genauigkeit dargestellte Audiosignal mit Hilfe der inversen IntMDCT im ganzzahligen Spektralbereich dargestellt und mit einem Gehör angepaßt MDCT-basierten Audiocoder-Ausgangssignal skalierbar kombiniert.

[0054] Die ganzzahligen Differenzwerte, die zur verlustlosen Darstellung vorhanden sind, werden nun nicht vollständig in einer Skalierungsschicht codiert, sondern zunächst mit geringerer Genauigkeit. Erst in einer weiteren Skalierungsschicht werden die zur exakten Darstellung nötigen Restwerte übertragen. Alternativ könnte jedoch auch in einer weiteren Skalierungsschicht ein Differenz-Spektralwert insgesamt, also mit 24 Bits beispielsweise, dargestellt sein, so daß zum Decodieren dieser weiteren Skalierungsschicht die darunterliegende Skalierungsschicht nicht benötigt wird. Dieses Szenario führt jedoch insgesamt zu einer höheren Bitstromgröße, kann jedoch dann, wenn die Bandbreite des Übertragungskanal unproblematisch ist, zu einer Vereinfachung im Decodierer beitragen, da im Decodierer dann nicht mehr Skalierungsschichten kombiniert werden müssen, sondern immer eine Skalierungsschicht allein zum Decodieren ausreicht.

[0055] Werden beispielsweise die unteren acht LSB, wie es in Fig. 9 dargestellt ist, zunächst nicht übertragen, so erreicht man eine Skalierbarkeit zwischen 24 Bit und 16 Bit.

[0056] Zur Rücktransformation der mit geringerer Genauigkeit übertragenen Werte in den Zeitbereich werden die übertragenen Werte bevorzugt auf den ursprünglichen Bereich, beispielsweise 24 Bit, zurückskaliert, indem dieselben beispielsweise mit 2^8 multipliziert werden. Auf die entsprechend zurückskalierten Werte wird dann eine inverse IntMDCT angewandt.

[0057] Bei der erfindungsgemäßen Genauigkeitsskalierung im Frequenzbereich wird es ferner bevorzugt, auch die Redundanz in den LSBs zu nützen. Hat ein Audiosignal beispielsweise im oberen Frequenzbereich sehr wenig Energie, so äußert sich dies auch beim IntMDCT-Spektrum in sehr kleinen Werten, die beispielsweise deutlich kleiner als die z. B. bei 8 Bit möglichen Werten (-128,...,127) sind. Dies äußert sich in einer Komprimierbarkeit der LSB-Werte des IntMDCT-Spektrums. Ferner wird darauf hingewiesen, daß bei sehr kleinen Differenz-Spektralwerten typischerweise eine Anzahl von Bits von MSB bis MSB-n gleich Null sind, und daß dann erst bei einem Bit mit einer Wertigkeit MSB-n-1 die erste, führende 1 in einem binär codierten Differenz-Spektralwert auftritt. In einem solchen Fall, wenn ein Differenz-Spektralwert in der zweiten Skalierungsschicht nur Nullen umfaßt, eignet sich eine Entropie-Codierung zur weiteren Datenkomprimierung besonders gut.

[0058] Gemäß einem weiteren Ausführungsbeispiel der vorliegenden Erfindung wird für die zweite Skalierungsschicht 82 von Fig. 8a eine Abtastraten-Skalierbarkeit bevorzugt. Eine Abtastraten-Skalierbarkeit wird dadurch erreicht, daß in der zweiten Skalierungsschicht, wie es in Fig. 9 rechts dargestellt ist, die Differenz-Spektralwerte bis zu einer ersten Grenzfrequenz enthalten sind, während in einer dritten Skalierungsschicht die Differenz-Spektralwerte mit einer Frequenz zwischen der ersten Grenzfrequenz und der maximalen Frequenz enthalten sind. Natürlich kann auch eine weitere Skalierung durchgeführt werden, so daß aus dem gesamten Frequenzbereich mehrere Skalierungsschichten gemacht werden.

[0059] Bei einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung umfaßt die zweite Skalierungsschicht in Fig. 9 Differenz-Spektralwerte bis zu einer Frequenz von 24 kHz, was einer Abtastrate von 48 kHz entspricht. Die dritte Skalierungsschicht enthält dann die Differenz-Spektralwerte von 24 kHz bis 48 kHz, was einer Abtastrate von 96 kHz entspricht.

[0060] Es sei ferner darauf hingewiesen, daß in der zweiten Skalierungsschicht und der dritten Skalierungsschicht nicht notwendigerweise sämtliche Bits eines Differenz-Spektralwertes codiert sein müssen. So könnten in einer weiteren Form der kombinierten Skalierbarkeit die zweite Skalierungsschicht, die Bits MSB bis MSB-x der Differenz-Spektralwerte bis zu einer bestimmten Grenzfrequenz umfassen. Eine dritte Skalierungsschicht könnte dann die Bits MSB bis MSB-x der Differenz-Spektralwerte von der ersten Grenzfrequenz bis zur maximalen Frequenz umfassen. Eine vierte Skalierungsschicht könnte dann die restlichen Bits für die Differenz-Spektralwerte bis zur Grenzfrequenz umfassen. Die letzte Skalierungsschicht könnte dann die restlichen Bits der Differenz-Spektralwerte für die oberen Frequenzen umfassen. Dieses Konzept wird zu einer Aufteilung des Tableaus in Fig. 9 in vier Quadranten führen, wobei jeder Quadrant eine Skalierungsschicht darstellt.

[0061] Bei der frequenzmäßigen Skalierbarkeit wird bei einem bevorzugtem Ausführungsbeispiel der vorliegenden Erfindung eine Skalierbarkeit zwischen 48 kHz und 96 kHz Abtastrate beschrieben. Das 96 kHz-Abtastsignal wird im IntMDCT-Bereich in der verlustlosen Erweiterungsschicht zunächst nur bis zu Hälfte codiert und übertragen. Falls der obere Teil nicht zusätzlich übertragen wird, wird er im Decodierer zu Null angenommen. Bei der inversen IntMDCT (gleiche Länge wie im Encoder) entsteht dann ein 96 kHz-Signal, das im oberen Frequenzbereich keine Energie enthält und daher ohne Qualitätsverluste auf 48 kHz unter-abgetastet werden kann.

[0062] Die vorstehende Skalierung der Differenz-Spektralwerte in Quadranten von Fig. 9 mit festen Grenzen ist hinsichtlich der Größe der Skalierungsschichten günstig, da in einer Skalierungsschicht tatsächlich lediglich z. B. 16 Bits bzw. 8 Bits oder die Spektralwerte bis zu der Grenzfrequenz bzw. über der Grenzfrequenz enthalten sein müssen.

[0063] Eine alternative Skalierung besteht darin, die Quadrantengrenzen in Fig. 9 gewissermaßen "aufzuweichen".

Am Beispiel der Frequenzskalierbarkeit würde das bedeuten, nicht einen sogenannten "Brickwall-Tiefpaß" anzuwenden, dahingehend, daß die Differenz-Spektralwerte vor einer Grenzfrequenz unverändert sind und nach der Grenzfrequenz gleich Null sind. Stattdessen könnten die Differenz-Spektralwerte auch mit irgendeinem beliebigen Tiefpaß gefiltert sein, der die Spektralwerte unterhalb der Grenzfrequenz schon etwas beeinträchtigt, aber oberhalb der Grenzfrequenz dazu führt, daß hier ebenfalls noch Energie vorhanden ist, obgleich die Differenz-Spektralwerte energiemäßig abnehmend sind. In einer solchermaßen erzeugten Skalierungsschicht sind dann auch Spektralwerte oberhalb der Grenzfrequenz enthalten. Da diese Spektralwerte jedoch relativ klein sind, sind sie durch eine Entropie-Codierung effizient codierbar. Die höchste Skalierungsschicht hätte in diesem Fall die Differenz zwischen den kompletten Differenz-Spektralwerten und den in der zweiten Skalierungsschicht enthaltenen Spektralwerten.

[0064] Ähnlich kann auch die Genauigkeitsskalierung gewissermaßen aufgeweicht werden. So kann die erste Skalierungsschicht auch Spektralwerte mit z. B. mehr als 16 Bit haben, wobei nächste Skalierungsschicht dann noch die Differenz hat. Allgemein gesagt hat die zweite Skalierungsschicht somit die Differenz-Spektralwerte mit geringerer Genauigkeit, während in der nächsten Skalierungsschicht der Rest, also die Differenz zwischen den kompletten Spektralwerten und den in der zweiten Skalierungsschicht enthaltenen Spektralwerten übertragen wird. Damit wird eine variable Genauigkeitsreduzierung erreicht.

[0065] Das erfindungsgemäße Verfahren zum Codieren bzw. Decodieren ist vorzugsweise auf einem digitalen Speichermedium, wie z. B. einer Diskette, mit elektronisch auslesbaren Steuersignalen gespeichert, wobei die Steuersignale so mit einem programmierbaren Computersystem zusammenarbeiten können, dass das Codier- und/oder Decodierverfahren ausgeführt werden kann/können. In anderen Worten ausgedrückt liegt also ein Computer-Programm-Produkt mit auf einem maschinenlesbaren Träger gespeichertem Programmcode zur Durchführen des Codierverfahrens und/oder des Decodierverfahrens vor, wenn das Programmprodukt auf einem Rechner abläuft. Die erfindungsgemäßen Verfahren können also in einem Computer-Programm mit einem Programmcode zur Durchführen der erfindungsgemäßen Verfahren, wenn das Programm auf einem Computer abläuft, realisiert sein.

[0066] Im nachfolgenden wird als Beispiel für einen ganzzahligen Transformationsalgorithmus auf den IntMDCT-Transformationsalgorithmus eingegangen, der in "Audio Coding Based on Integer Transforms" 111-te AES-Versammlung, New York, 2001, beschrieben ist. Die IntMDCT ist besonders günstig, da sie die attraktiven Eigenschaften der MDCT hat, wie z. B. eine gute spektrale Darstellung des Audiosignals, eine kritische Abtastung und eine Blocküberlappung aufweist. Die gute Approximation der MDCT durch eine IntMDCT erlaubt es ferner, bei dem in Fig. 5 gezeigten Codierer nur einen Transformationsalgorithmus zu verwenden, wie es durch einen Pfeil 62 in Fig. 5 dargestellt ist. Anhand der Fig. 1 bis 4 werden die wesentlichen Eigenschaften dieser speziellen Form eines ganzzahligen Transformationsalgorithmus erklärt.

[0067] Fig. 1 zeigt ein Übersichtsdiagramm für die erfindungsgemäß bevorzugte Vorrichtung zum Verarbeiten von zeitdiskreten Abtastwerten, die ein Audiosignal darstellen, um ganzzahlige Werte zu erhalten, auf denen aufbauend der Int-MDCT-Ganzzahl-Transformationsalgorithmus arbeitet. Die zeitdiskreten Abtastwerte werden durch die in Fig. 1 gezeigte Vorrichtung gefenstert und optional in eine spektrale Darstellung umgesetzt. Die zeitdiskreten Abtastwerte, die an einem Eingang 10 in die Vorrichtung eingespeist werden, werden mit einem Fenster w mit einer Länge, die $2N$ zeitdiskreten Abtastwerten entspricht, gefenstert, um an einem Ausgang 12 ganzzahlige gefensterte Abtastwerte zu erreichen, welche dazu geeignet sind, um mittels einer Transformation und insbesondere der Einrichtung 14 zum Ausführen einer ganzzahligen DCT in eine spektrale Darstellung umgesetzt zu werden. Die ganzzahlige DCT ist ausgebildet, um aus N Eingangswerten N Ausgangswerte zu erzeugen, was im Gegensatz zu der MDCT-Funktion von Fig. 4a steht, die aus $2N$ gefensterten Abtastwerten aufgrund der MDCT-Gleichung lediglich N Spektralwerte erzeugt.

[0068] Zum Fenstern der zeitdiskreten Abtastwerte werden zunächst in einer Einrichtung 16 zwei zeitdiskrete Abtastwerte ausgewählt, die zusammen einen Vektor von zeitdiskreten Abtastwerten darstellen. Ein zeitdiskreter Abtastwert, der durch die Einrichtung 16 ausgewählt wird, liegt im ersten Viertel des Fensters. Der andere zeitdiskrete Abtastwert liegt in dem zweiten Viertel des Fensters, wie es anhand von Fig. 3 noch detaillierter ausgeführt wird. Der durch die Einrichtung 16 erzeugte Vektor wird nunmehr mit einer Drehmatrix der Dimension 2×2 beaufschlagt, wobei diese Operation nicht unmittelbar durchgeführt wird, sondern mittels mehrerer sogenannten Lifting-Matrizen.

[0069] Eine Lifting-Matrix hat die Eigenschaft, daß sie nur ein Element aufweist, das von dem Fenster w abhängt und ungleich "1" oder "0" ist.

[0070] Die Faktorisierung von Wavelet-Transformationen in Lifting-Schritte ist in der Fachveröffentlichung "Factoring Wavelet Transforms Into Lifting Steps", Ingrid Daubechies und Wim Sweldens, Preprint, Bell Laboratories, Lucent Technologies, 1996, dargestellt. Allgemein ist ein Lifting-Schema eine einfache Beziehung zwischen perfekt rekonstruierenden Filterpaaren, die dasselbe Tiefpaß- oder Hochpaßfilter haben. Jedes Paar komplementärer Filter kann in Lifting-

Schritte faktorisiert werden. Insbesondere gilt dies für Givens Rotationen. Es sei der Fall betrachtet, bei dem die Polyphasenmatrix eine Givens-Rotation ist. Es gilt dann:

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \sin \alpha & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{\cos \alpha - 1}{\sin \alpha} \\ 0 & 1 \end{pmatrix} \quad (1)$$

[0071] Jede der drei rechts des Gleichheitszeichens stehenden Lifting-Matrizen hat als Hauptdiagonalelemente den Wert "1". Ferner ist in jeder Lifting-Matrix ein Nebendiagonalelement gleich 0, und ein Nebendiagonalelement vom Drehwinkel α abhängig.

[0072] Der Vektor wird nunmehr mit der dritten Lifting-Matrix, d. h. der Liftingmatrix ganz rechts in obiger Gleichung, multipliziert, um einen ersten Ergebnisvektor zu erhalten. Dies ist in Fig. 1 durch eine Einrichtung 18 dargestellt. Es wird nunmehr der erste Ergebnisvektor mit einer beliebigen Rundungsfunktion, die die Menge der reellen Zahlen in die Menge der ganzen Zahlen abbildet, gerundet, wie es in Fig. 1 durch eine Einrichtung 20 dargestellt ist. Am Ausgang der Einrichtung 20 wird ein gerundeter erster Ergebnisvektor erhalten. Der gerundete erste Ergebnisvektor wird nunmehr in eine Einrichtung 22 zum Multiplizieren desselben mit der mittleren, d. h. zweiten, Lifting-Matrix eingespeist, um einen zweiten Ergebnisvektor zu erhalten, der in einer Einrichtung 24 wiederum gerundet wird, um einen gerundeten zweiten Ergebnisvektor zu erhalten. Der gerundete zweite Ergebnisvektor wird nunmehr in eine Einrichtung 26 eingespeist, und zwar zum Multiplizieren desselben mit der links in der obigen Gleichung aufgeführten, d. h. ersten, Lifting-Matrix, um einen dritten Ergebnisvektor zu erhalten, der schließlich noch mittels einer Einrichtung 28 gerundet wird, um schließlich an dem Ausgang 12 ganzzahlige gefensterte Abtastwerte zu erhalten, die nun, wenn eine spektrale Darstellung derselben gewünscht wird, durch die Einrichtung 14 verarbeitet werden müssen, um an einem Spektralausgang 30 ganzzahlige Spektralwerte zu erhalten.

[0073] Vorzugsweise ist die Einrichtung 14 als Ganzzahl-DCT oder Integer-DCT ausgeführt.

[0074] Die diskrete Kosinus-Transformation gemäß Typ 4 (DCT-IV) mit einer Länge N ist durch folgende Gleichung gegeben:

$$X_i(m) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} x(k) \cos\left(\frac{\pi}{4N} (2k+1)(2m+1)\right) \quad (2)$$

[0075] Die Koeffizienten der DCT-IV bilden eine orthonormale N x N Matrix. Jede orthogonale N x N-Matrix kann in N(N-1)/2 Givens-Rotationen zerlegt werden, wie es in der Fachveröffentlichung P. P. Vaidyanathan, "Multirate Systems And Filter Banks", Prentice Hall, Englewood Cliffs, 1993, ausgeführt ist. Es sei darauf hingewiesen, daß auch weitere Zerlegungen existieren.

[0076] Bezüglich der Klassifikationen der verschiedenen DCT-Algorithmen sei auf H. S. Malvar, "Signal Processing With Lapped Transforms", Artech House, 1992, verwiesen. Allgemein unterscheiden sich die DCT-Algorithmen durch die Art ihrer Basisfunktionen. Während die DCT-IV, die hier bevorzugt wird, nicht-symmetrische Basisfunktionen umfaßt, d. h. eine Kosinus-Viertelwelle, eine Cosinus-3/4-Welle, eine Cosinus-5/4-Welle, eine Cosinus-7/4-Welle, etc., hat die diskrete Kosinustransformation z. B. vom Typ II (DCT-II), achsensymmetrische und punktsymmetrische Basisfunktionen. Die 0-te Basisfunktion hat einen Gleichanteil, die erste Basisfunktion ist eine halbe Kosinuswelle, die zweite Basisfunktion ist eine ganze Kosinuswelle, usw. Aufgrund der Tatsache, daß die DCT-II den Gleichanteil besonders berücksichtigt, wird sie bei der Videocodierung verwendet, nicht aber bei der Audiocodierung, da bei der Audiocodierung im Gegensatz zur Videocodierung der Gleichanteil nicht von Relevanz ist.

[0077] Im nachfolgend wird darauf eingegangen, wie der Drehwinkel α der Givens-Rotation von der Fensterfunktion abhängt.

[0078] Eine MDCT mit einer Fensterlänge von 2N kann in eine diskrete Kosinustransformation vom Typ IV mit einer Länge N reduziert werden. Dies wird dadurch erreicht, daß die TDAC-Operation explizit im Zeitbereich durchgeführt wird, und daß dann die DCT-IV angewandt wird. Bei einer 50%igen Überlappung überlappt die linke Hälfte des Fensters für einen Block t mit der rechten Hälfte des vorausgehenden Blocks, d. h. des Blocks t-1. Der überlappende Teil zwei aufeinanderfolgender Blöcke t-1 und t wird im Zeitbereich, d. h. vor der Transformation, folgendermaßen vorverarbeitet, d. h. zwischen dem Eingang 10 und dem Ausgang 12 von Fig. 1 verarbeitet:

$$\begin{pmatrix} \tilde{x}_i(k) \\ \tilde{x}_{i-1}(N-1-k) \end{pmatrix} = \begin{pmatrix} w(\frac{N}{2}+k) & -w(\frac{N}{2}-1-k) \\ w(\frac{N}{2}-1-k) & w(\frac{N}{2}+k) \end{pmatrix} \begin{pmatrix} x_i(\frac{N}{2}+k) \\ x_i(\frac{N}{2}-1-k) \end{pmatrix} \quad (3)$$

[0079] Die mit der Tilde bezeichneten Werte sind die Werte am Ausgang 12 von Fig. 1, während die ohne Tilde in der obigen Gleichung bezeichnete x Werte die Werte am Eingang 10 bzw. hinter der Einrichtung 16 zum Auswählen sind. Der Laufindex k läuft von 0 bis N/2-1, während w die Fensterfunktion darstellt.

[0080] Aus der TDAC-Bedingung für die Fensterfunktion w gilt folgender Zusammenhang:

$$w\left(\frac{N}{2}+k\right)^2 + w\left(\frac{N}{2}-1-k\right)^2 = 1 \quad (4)$$

[0081] Für bestimmte Winkel α_k , $k = 0, \dots, N/2-1$ kann diese Vorverarbeitung im Zeitbereich als Givens-Rotation geschrieben werden, wie es ausgeführt worden ist.

[0082] Der Winkel α der Givens-Rotation hängt folgendermaßen von der Fensterfunktion w ab:

$$\alpha = \arctan[w(N/2-1-k) / w(N/2 + k)] \quad (5)$$

[0083] Es sei darauf hingewiesen, daß beliebige Fensterfunktionen w eingesetzt werden können, solange sie diese TDAC-Bedingung erfüllen.

[0084] Im nachfolgenden wird anhand von Fig. 2 ein kaskadierter Codierer und Decodierer beschrieben. Die zeitdiskreten Abtastwerte $x(0)$ bis $x(2N-1)$, die durch ein Fenster gemeinsam "gefenstert" werden, werden zunächst derart durch die Einrichtung 16 von Fig. 1 ausgewählt, daß der Abtastwert $x(0)$ und der Abtastwert $x(N-1)$, d. h. ein Abtastwert aus dem ersten Viertel des Fensters und ein Abtastwert aus dem zweiten Viertel des Fensters, ausgewählt werden, um den Vektor am Ausgang der Einrichtung 16 zu bilden. Die sich überkreuzenden Pfeile stellen schematisch die Lifting-Multiplikationen und anschließenden Rundungen der Einrichtungen 18, 20 bzw. 22, 24 bzw. 26, 28 dar, um am Eingang der DCT-IV-Blöcke die ganzzahligen gefensterten Abtastwerte zu erhalten.

[0085] Wenn der erste Vektor wie oben beschrieben verarbeitet ist, wird ferner ein zweiter Vektor aus den Abtastwerten $x(N/2-1)$ und $x(N/2)$, d. h. wieder ein Abtastwert aus dem ersten Viertel des Fensters und ein Abtastwert aus dem zweiten Viertel des Fensters, ausgewählt und wiederum durch den in Fig. 1 beschriebenen Algorithmus verarbeitet. Analog dazu werden sämtliche anderen Abtastwertpaare aus dem ersten und zweiten Viertel des Fensters bearbeitet. Die selbe Verarbeitung wird für das dritte und vierte Viertel des ersten Fensters durchgeführt. Nunmehr liegen am Ausgang 12 $2N$ gefensterte ganzzahlige Abtastwerte vor, die nunmehr so, wie es in Fig. 2 dargestellt ist, in eine DCT-IV-Transformation eingespeist werden. Insbesondere werden die ganzzahligen gefensterten Abtastwerte des zweiten und dritten Viertels in eine DCT eingespeist. Die gefensterten ganzzahligen Abtastwerte des ersten Viertels des Fensters werden in eine vorausgehende DCT-IV zusammen mit den gefensterten ganzzahligen Abtastwerten des vierten Viertels des vorausgehenden Fensters verarbeitet. Analog dazu wird das vierte Viertel der gefensterten ganzzahligen Abtastwerte in Fig. 2 mit dem ersten Viertel des nächsten Fensters zusammen in eine DCT-IV-Transformation eingespeist. Die mittlere in Fig. 2 gezeigte ganzzahlige DCT-IV-Transformation 32 liefert nunmehr N ganzzahlige Spektralwerte $y(0)$ bis $y(N-1)$. Diese ganzzahligen Spektralwerte können nunmehr beispielsweise einfach Entropie-codiert werden, ohne daß eine dazwischenliegende Quantisierung erforderlich ist, da die Fensterung und Transformation ganzzahlige Ausgangswerte liefert.

[0086] In der rechten Hälfte von Fig. 2 ist ein Decodierer dargestellt. Der Decodierer bestehend aus Rücktransformation und "inverser Fensterung" arbeitet invers zum Codierer. Es ist bekannt, daß zur Rücktransformation einer DCT-IV eine inverse DCT-IV verwendet werden kann, wie es in Fig. 2 dargestellt ist. Die Ausgangswerte der Decodierer-DCT-IV 34 werden nunmehr, wie es in Fig. 2 dargestellt ist, mit den entsprechenden Werten der vorausgehenden Transformation bzw. der nachfolgenden Transformation invers verarbeitet, um aus den ganzzahligen gefensterten Abtastwerten am Ausgang der Einrichtung 34 bzw. der vorausgehenden und nachfolgenden Transformation wieder zeitdiskrete Audio-Abtastwerte $x(0)$ bis $x(2N-1)$ zu erzeugen.

[0087] Die ausgangsseitige Operation geschieht durch eine inverse Givens-Rotation, d. h. derart, daß die Blöcke

26, 28 bzw. 22, 24 bzw. 18, 20 in der entgegengesetzten Richtung durchlaufen werden. Dies sei anhand der zweiten Lifting-Matrix von Gleichung 1 näher dargestellt. Wenn (im Codierer) der zweite Ergebnisvektor durch Multiplikation des gerundeten ersten Ergebnisvektors mit der zweiten Liftingmatrix (Einrichtung 22) gebildet wird, so ergibt sich folgender Ausdruck:

$$(x, y) \mapsto (x, y + x \sin \alpha) \quad (6)$$

[0088] Die Werte x , y auf der rechten Seite von Gleichung 6 sind Ganzzahlen. Dies trifft jedoch für den Wert $x \sin \alpha$ nicht zu. Hier muß die Rundungsfunktion r eingeführt werden, wie es in der nachfolgenden Gleichung

$$(x, y) \mapsto (x, y + r(x \sin \alpha)) \quad (7)$$

dargestellt ist. Diese Operation führt die Einrichtung 24 aus.

[0089] Die inverse Abbildung (im Decodierer) ist folgendermaßen definiert:

$$(x', y') \mapsto (x', y' - r(x' \sin \alpha)) \quad (8)$$

[0090] Aufgrund dem Minuszeichens vor der Rundungsoperation wird ersichtlich, daß die ganzzahlige Approximierung des Lifting-Schritts umgekehrt werden kann, ohne daß ein Fehler eingeführt wird. Die Anwendung dieser Approximation auf jeden der drei Lifting-Schritte führt zu einer ganzzahligen Approximation der Givens-Rotation. Die gerundete Rotation (im Codierer) kann umgekehrt werden (im Decodierer), ohne daß ein Fehler eingeführt wird, und zwar indem die inversen gerundeten Lifting-Schritte in umgekehrter Reihenfolge durchlaufen werden, d. h. wenn beim Decodieren der Algorithmus von Fig. 1 von unten nach oben durchgeführt wird.

[0091] Wenn die Rundungsfunktion r punktsymmetrisch ist, ist die inverse gerundete Rotation identisch zu der gerundeten Rotation mit dem Winkel $-\alpha$ und lautet folgendermaßen:

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \quad (9)$$

[0092] Die Lifting-Matrizen für den Decodierer, d. h. für die inverse Givens-Rotation, ergibt sich in diesem Fall unmittelbar aus Gleichung (1), indem lediglich der Ausdruck "sin" durch den Ausdruck "-sin α " ersetzt wird.

[0093] Im nachfolgenden wird anhand von Fig. 3 noch einmal die Zerlegung einer üblichen MDCT mit überlappenden Fenstern 40 bis 46 dargelegt. Die Fenster 40 bis 46 überlappen jeweils zu 50%. Pro Fenster werden zunächst Givens-Rotationen innerhalb des ersten und zweiten Viertels eines Fensters bzw. innerhalb des dritten und vierten Viertels eines Fensters ausgeführt, wie es durch die Pfeile 48 schematisch dargestellt ist. Dann werden die rotierten Werte, d. h. die gefensterten ganzzahligen Abtastwerte derart in eine N-zu-N-DCT eingespeist, daß immer das zweite und dritte Viertel eines Fensters bzw. das vierte und erste Viertel eines darauffolgenden Fensters gemeinsam mittels eines DCT-IV-Algorithmus in eine spektrale Darstellung umgesetzt wird.

[0094] Es werden daher die üblichen Givens-Rotation in Lifting-Matrizen zerlegt, die sequentiell ausgeführt werden, wobei nach jeder Lifting-Matrix-Multiplikation ein Rundungsschritt eingeführt wird, derart, daß die Gleitkomma-Zahlen unmittelbar nach ihrer Entstehung gerundet werden, derart, daß vor jeder Multiplikation eines Ergebnisvektors mit einer Lifting-Matrix der Ergebnisvektor lediglich Ganzzahlen hat.

[0095] Die Ausgangswerte bleiben also immer ganzzahlig, wobei es bevorzugt wird, auch ganzzahlige Eingangswerte zu verwenden. Dies stellt keine Einschränkung dar, da jegliche beispielsweise PCM-Abtastwerte, wie sie auf einer CD abgespeichert sind, ganzzahlige Zahlenwerte sind, deren Wertebereich je nach Bitbreite variiert, d. h. abhängig davon, ob die zeitdiskreten digitalen Eingangswerte 16-Bit-Werte oder 24-Bit-Werte sind. Dennoch ist, wie es ausgeführt worden ist, der gesamte Prozeß invertierbar, indem die inversen Rotationen in umgekehrter Reihenfolge ausgeführt werden. Es existiert somit eine ganzzahlige Approximation der MDCT mit perfekter Rekonstruktion, also eine verlustlose Transformation.

[0096] Die gezeigte Transformation liefert ganzzahlige Ausgangswerte statt Gleitkommawerte. Sie liefert eine perfekte Rekonstruktion, so daß kein Fehler eingeführt wird, wenn eine Vorwärts- und dann eine Rückwärtstransformation ausgeführt werden. Die Transformation ist gemäß einem bevorzugten Ausführungsbeispiel der vorliegenden Erfindung

ein Ersatz für die modifizierte diskrete Cosinustransformation. Auch andere Transformationsverfahren können jedoch ganzzahlig ausgeführt werden, so lange eine Zerlegung in Rotationen und eine Zerlegung der Rotationen in Lifting-Schritte möglich ist.

[0097] Die ganzzahlige MDCT hat die meisten günstigen Eigenschaften der MDCT. Sie hat eine überlappende Struktur, wodurch eine bessere Frequenzselektivität als bei nichtüberlappenden Blocktransformationen erhalten wird. Aufgrund der TDAC-Funktion, die bereits beim Fenstern vor der Transformation berücksichtigt wird, wird eine kritische Abtastung beibehalten, so daß die Gesamtanzahl von Spektralwerten, die ein Audiosignal darstellen, gleich der Gesamtanzahl von Eingangs-Abtastwerten ist.

[0098] Verglichen mit einer normalen MDCT, die Gleitkomma-Abtastwerte liefert, zeigt sich bei der beschriebenen bevorzugten ganzzahligen Transformation, daß lediglich in dem Spektralbereich, in dem wenig Signalpegel ist, das Rauschen im Vergleich zur normalen MDCT erhöht ist, während sich diese Rauscherhöhung bei signifikanten Signalpegeln nicht bemerkbar macht. Dafür bietet sich die ganzzahlige Verarbeitung für eine effiziente Hardware-Implementation an, da lediglich Multiplikationsschritte verwendet werden, die ohne weiteres in Verschieben-Addieren-Schritte (Shift/Add-Schritte) zerlegt werden können, welche einfach und schnell hardwaremäßig implementiert werden können. Selbstverständlich ist auch eine Software-Implementation möglich.

[0099] Die ganzzahlige Transformation liefert eine gute spektrale Darstellung des Audiosignals und bleibt dennoch im Bereich der ganzen Zahlen. Wenn sie auf tonale Teile eines Audiosignals angewandt wird, resultiert dies in einer guten E-nergiekonzentrierung. Damit kann ein effizientes verlustloses Codierschema aufgebaut werden, indem einfach die in Fig. 1 dargestellte Fensterung/Transformation mit einem Entropiecodierer kaskadiert wird. Insbesondere ein gestapeltes Codieren (Stacked Coding) unter Verwendung von Escape-Werten, wie es in MPEG AAC eingesetzt wird, ist günstig. Es wird bevorzugt, alle Werte um eine bestimmte Potenz von zwei herunterzuskalieren, bis sie in eine erwünschte Codetabelle passen, und dann die weggelassenen niederwertigen Bits zusätzlich zu codieren. Im Vergleich zu der Alternative der Verwendung von größeren Codetabellen ist die beschriebene Alternative hinsichtlich des Speicherverbrauchs zum Speichern der Codetabellen günstiger. Ein nahezu verlustloser Codierer könnte auch dadurch erhalten werden, daß einfach bestimmte der niederwertigen Bits weggelassen werden.

[0100] Insbesondere für tonale Signale ermöglicht eine Entropie-Codierung der ganzzahligen Spektralwerte einen hohen Codiergewinn. Für transiente Teile des Signals ist der Coergewinn niedrig, und zwar aufgrund des flachen Spektrums transienter Signale, d. h. aufgrund einer geringen Anzahl von Spektralwerten, die gleich oder nahezu 0 sind. Wie es in J. Herre, J. D. Johnston: "Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS)" 101. AES Convention, Los Angeles, 1996, Preprint 4384, beschrieben ist, kann diese Flachheit jedoch verwendet werden, indem eine lineare Prädiktion im Frequenzbereich verwendet wird. Eine Alternative ist eine Prädiktion mit offener Schleife. Eine andere Alternative ist der Prädiktor mit geschlossener Schleife. Die erste Alternative, d. h. der Prädiktor mit offener Schleife, wird TNS genannt. Die Quantisierung nach der Prädiktion führt zu einer Adaption des resultierenden Quantisierungsrauschens an die zeitliche Struktur des Audiosignals und verhindert daher Vorechos in psychoakustischen Audiocodierern. Für ein verlustloses Audiocodieren ist die zweite Alternative, d. h. mit einem Prädiktor mit geschlossener Schleife, geeigneter, da die Prädiktion mit geschlossener Schleife eine genaue Rekonstruktion des Eingangssignals erlaubt. Wenn diese Technik auf ein erzeugtes Spektrum angewendet wird, muß ein Rundungsschritt nach jedem Schritt des Prädiktionsfilters durchgeführt werden, um im Bereich der Ganzzahlen zu bleiben. Durch Verwenden des inversen Filters und derselben Rundungsfunktion kann das ursprüngliche Spektrum genau wieder hergestellt werden.

[0101] Um die Redundanz zwischen zwei Kanälen zur Datenreduktion auszunutzen, kann auch eine Mitte-Seite-Codierung verlustlos eingesetzt werden, wenn eine gerundete Rotation mit einem Winkel $\pi/4$ verwendet wird. Im Vergleich zur Alternative des Berechnens der Summe und Differenz des linken und rechten Kanals eines Stereosignals hat die gerundete Rotation den Vorteil der Energieerhaltung. Die Verwendung sogenannter Joint-Stereo-Codieretechniken kann für jedes Band ein- oder ausgeschaltet werden, wie es auch im Standard MPEG AAC durchgeführt wird. Weitere Drehwinkel können ebenfalls berücksichtigt werden, um eine Redundanz zwischen zwei Kanälen flexibler reduzieren zu können.

Patentansprüche

1. Vorrichtung zum Codieren eines zeitdiskreten Audiosignals, um codierte Audiodaten zu erhalten, mit folgenden Merkmalen:

einer Einrichtung (52) zum Liefern eines Quantisierungs-Blocks von unter Verwendung eines psychoakustischen Modells (54) quantisierten Spektralwerten des zeitdiskreten Audiosignals;

einer Einrichtung (58) zum inversen Quantisieren des Quantisierungsblocks und zum Runden der invers quan-

tisierten Spektralwerte, um einen Rundungs-Block von gerundeten invers quantisierten Spektralwerten zu erhalten;

einer Einrichtung (56) zum Erzeugen eines Ganzzahl-Blocks von ganzzahligen Spektralwerten unter Verwendung eines Ganzzahl-Transformationsalgorithmus, der ausgebildet ist, um aus einem Block von ganzzahligen zeitdiskreten Abtastwerten den Ganzzahl-Block von Spektralwerten zu erzeugen;

eine Kombinationseinrichtung (58) zum Bilden eines Differenz-Blocks, der von einer spektralwertweisen Differenz zwischen dem Rundungs-Block und dem Ganzzahl-Block abhängt, um einen Differenz-Block mit Differenz-Spektralwerten zu erhalten; und

einer Einrichtung (60) zum Verarbeiten des Quantisierungs-Blocks und des Differenz-Blocks, um codierte Audiodaten zu erzeugen, die Informationen über den Quantisierungs-Block und Informationen über den Differenz-Block umfassen.

2. Vorrichtung nach Anspruch 1, bei der die Einrichtung (52) zum Liefern ausgebildet ist, um aus einem Zeit-Block von zeitlichen Audiosignalwerten mittels einer MDCT einen MDCT-Block von MDCT-Spektralwerten zu erzeugen, und um den MDCT-Block unter Verwendung eines psychoakustischen Modells zu quantisieren, um den Quantisierungs-Block zu erzeugen, der quantisierte MDCT-Spektralwerte aufweist.
3. Vorrichtung nach Anspruch 2, bei der die Einrichtung (56) zum Erzeugen des Ganzzahl-Blocks ausgebildet ist, um eine IntMDCT auf den Zeit-Block auszuführen, um den Ganzzahl-Block zu erzeugen, der IntMDCT-Spektralwerte aufweist.
4. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (52) zum Liefern ausgebildet ist, um unter Verwendung eines Gleitkomma-Transformationsalgorithmus den Quantisierungs-Block zu berechnen.
5. Vorrichtung nach einem der Ansprüche 1 bis 3, bei der die Einrichtung (52) zum Liefern ausgebildet ist, um den Quantisierungs-Block unter Verwendung des von der Einrichtung (56) zum Erzeugen erzeugten Ganzzahl-Blocks zu berechnen.
6. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (60) zum Verarbeiten ausgebildet ist, um den Quantisierungs-Block einer Entropie-Codierung zu unterziehen (60a), um einen Entropie-codierten Quantisierungs-Block zu erhalten, um den Rundungs-Block einer Entropie-Codierung zu unterziehen (60b), um einen Entropie-codierten Rundungs-Block zu erhalten, und um den Entropie-codierten Quantisierungs-Block in eine erste Skalierungsschicht eines skalierten Datenstroms, der die codierten Audiodaten darstellt, zu überführen, und um den Entropie-codierten Rundungs-Block in eine zweite Skalierungsschicht des skalierten Datenstroms zu überführen.
7. Vorrichtung nach Anspruch 6, bei der die Einrichtung (60) zum Verarbeiten ferner ausgebildet ist, um für die Entropie-Codierung des Quantisierungs-Blocks eine aus einer Mehrzahl von Codetabellen abhängig von den quantisierten Spektralwerten zu verwenden, und bei der die Einrichtung (60) zum Verarbeiten ferner ausgebildet ist, um für die Entropie-Codierung des Differenz-Blocks eine aus einer Mehrzahl von Codetabellen abhängig von einer Eigenschaft eines Quantisierers auszuwählen, der bei einer Quantisierung zum Erzeugen des Quantisierungs-Blocks verwendbar ist.
8. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (52) zum Liefern ausgebildet ist, um abhängig von einer Beschaffenheit des Audiosignals eines einer Mehrzahl von Fenstern zum Fenstern eines zeitlichen Blocks von Audiosignalwerten zu verwenden, und bei der die Einrichtung (56) zum Erzeugen ausgebildet ist, um für den ganzzahligen Transformationsalgorithmus dieselbe Fensterauswahl zu treffen.
9. Vorrichtung nach einem der Ansprüche 1 bis 8,

bei der die Einrichtung zum Erzeugen ausgebildet ist, um einen ganzzahligen Transformationsalgorithmus zu verwenden, der folgende Schritte aufweist:

Fenstern der zeitdiskreten Abtastwerte mit einem Fenster (w) mit einer Länge, die $2N$ zeitdiskreten Abtastwerten entspricht, um gefensterte zeitdiskrete Abtastwerte für eine Umsetzung der zeitdiskreten Abtastwerte in eine spektrale Darstellung mittels einer Transformation zu liefern, die aus N Eingangswerten N Ausgangswerte erzeugen kann, wobei das Fenster folgende Teilschritte aufweist:

Auswählen (16) eines zeitdiskreten Abtastwerts aus einem Viertel des Fensters und eines zeitdiskreten Abtastwerts aus einem anderen Viertel des Fensters, um einen Vektor von zeitdiskreten Abtastwerten zu erhalten;

Beaufschlagen des Vektors mit einer quadratischen Drehmatrix, deren Dimension mit der Dimension des Vektors übereinstimmt, wobei die Drehmatrix durch eine Mehrzahl von Lifting-Matrizen darstellbar ist, wobei eine Lifting-Matrix nur ein Element aufweist, das von dem Fenster (w) abhängt und ungleich 1 oder 0 ist, wobei der Teilschritt des Beaufschlagens folgende Unterschritte aufweist:

Multiplizieren (18) des Vektors mit einer Lifting-Matrix, um einen ersten Ergebnisvektor zu erhalten;

Runden (20) einer Komponente des ersten Ergebnisvektors mit einer Rundungsfunktion (r), die eine reelle Zahl auf eine ganze Zahl abbildet, um einen gerundeten ersten Ergebnisvektor zu erhalten; und

sequentielles Durchführen der Schritte des Multiplizierens (22) und Rundens (24) mit einer anderen Lifting-Matrix, bis alle Lifting-Matrizen abgearbeitet sind, um einen gedrehten Vektor zu erhalten, der einen ganzzahligen gefensterten Abtastwert aus dem Viertel des Fensters und einen ganzzahligen gefensterten Abtastwert aus dem anderen Viertel des Fensters aufweist, und

Durchführen des Schritts des Fensters für alle zeitdiskreten Abtastwerte der verbleibenden Viertel des Fensters, um $2N$ gefilterte ganzzahlige Abtastwerte zu erhalten; und

Umsetzen (14) von N gefensterten ganzzahligen Abtastwerten in eine spektrale Darstellung durch eine Ganzzahl-DCT für Werte mit den gefilterten ganzzahligen Abtastwerten des zweiten Viertels und des dritten Viertels des Fensters, um N ganzzahlige Spektralwerte zu erhalten.

10. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (52) zum Liefern des Quantisierungs-Blocks ausgebildet ist, um vor einem Quantisierungsschritt (52b) eine Prädiktion von Spektralwerten über der Frequenz unter Verwendung eines Prädiktionsfilters durchzuführen, um Prädiktionsrestspektralwerte zu erhalten, die nach einer Quantisierung den Quantisierungs-Block darstellen; bei der ferner eine Prädiktionseinrichtung vorgesehen ist, die ausgebildet ist, um eine Prädiktion über der Frequenz der ganzzahligen Spektralwerte des Ganzzahl-Blocks durchzuführen, wobei ferner eine Rundungseinrichtung vorgesehen ist, um Prädiktionsrestspektralwerte aufgrund der Ganzzahl-Spektralwerte zu runden, die den Rundungs-Block darstellen.

11. Verfahren nach einem der vorhergehenden Ansprüche, bei der das zeitdiskrete Audiosignal zumindest zwei Kanäle aufweist, bei der die Einrichtung (52) zum Liefern ausgebildet ist, um eine Mitte/Seite-Codierung mit Spektralwerten des zeitdiskreten Audiosignals durchzuführen, um nach einer Quantisierung von Mitte/Seite-Spektralwerten den Quantisierungs-Block zu erhalten, und bei der die Einrichtung (56) zum Erzeugen des Ganzzahl-Blocks ausgebildet ist, um ebenfalls eine Mitte/Seite-Codierung durchzuführen, die der Mitte/Seite-Codierung der Einrichtung (52) zum Liefern entspricht.

12. Vorrichtung nach einem der vorhergehenden Ansprüche, bei der die Einrichtung (60) zum Verarbeiten ausgebildet ist, um einen MPEG-2-AAC-Datenstrom zu erzeugen, wobei in einem Feld Ancillary Data Zusatzinformationen für den Ganzzahl-Transformationsalgorithmus eingebracht sind.

13. Vorrichtung nach einem der vorhergehenden Ansprüche,

bei der die Einrichtung zum Verarbeiten (60) ausgebildet ist, um die codierten Audiodaten als Datenstrom mit einer Mehrzahl von Skalierungsschichten auszugeben.

14. Vorrichtung nach Anspruch 13,

bei der die Einrichtung zum Verarbeiten (60) ausgebildet ist, um in eine erste Skalierungsschicht (81) Informationen über den Quantisierungsblock einzufügen, und um in eine zweite Skalierungsschicht (82) Informationen über den Differenzblock einzufügen.

15. Vorrichtung nach Anspruch 13,

bei der die Einrichtung zum Verarbeiten (60) ausgebildet ist, um in eine erste Skalierungsschicht Informationen über den Quantisierungsblock einzufügen, und um die Informationen über den Differenzblock in zumindest eine zweite und eine dritte Skalierungsschicht einzufügen.

16. Vorrichtung nach Anspruch 15,

bei der in der zweiten Skalierungsschicht Differenz-Spektralwerte mit reduzierter Genauigkeit enthalten sind, und in einer oder mehreren höheren Skalierungsschichten ein restlicher Teil der Differenz-Spektralwerte enthalten ist.

17. Vorrichtung nach Anspruch 15 oder 16,

bei der die Informationen über den Differenzblock binär codierte Differenz-Spektralwerte umfassen, bei der die zweite Skalierungsschicht für Differenz-Spektralwerte eine Anzahl von Bits von einem höchstwertigen Bit (MSB) zu einem niederwertigen Bit (MSB-x) für einen Differenz-Spektralwert umfaßt, und bei der die dritte Skalierungsschicht eine Anzahl von Bits ausgehend von einem niederwertigen Bit (MSB-x-1) bis zu einem niederstwertigen Bit (LSB) umfaßt.

18. Vorrichtung nach Anspruch 17,

bei der das zeitdiskrete Audiosignal in einer Form von Abtastwerten mit einer Breite von 24 Bits vorliegt, und bei der die Einrichtung zum Verarbeiten (60) ausgebildet ist, um höherwertige 16 Bits der Differenz-Spektralwerte in die zweite Skalierungsschicht einzufügen, und um restliche 8 Bits eines Differenz-Spektralwerts in die dritte Skalierungsschicht einzufügen, so daß ein Decodierer unter Verwendung der zweiten Skalierungsschicht eine CD-Qualität erreicht, wobei ein Decodierer unter Verwendung auch der dritten Skalierungsschicht eine Studio-Qualität erreicht.

19. Vorrichtung nach Anspruch 15,

bei der die Einrichtung (60) zum Verarbeiten ausgebildet ist, um in eine zweite Skalierungsschicht zumindest einen Teil von Differenz-Spektralwerten zur Darstellung eines tiefpaß-gefilterten Signals einzufügen und in zumindest eine weitere Skalierungsschicht eine Differenz zwischen den Differenz-Spektralwerten in der zweiten Skalierungsschicht und ursprünglichen Differenz-Spektralwerten einzufügen.

20. Vorrichtung nach Anspruch 15 oder 19,

bei der die Einrichtung (60) zum Verarbeiten ausgebildet ist, um in eine zweite Skalierungsschicht zumindest einen Teil von Differenz-Spektralwerten bis zu einer bestimmten Grenzfrequenz einzufügen, und um in einer dritten Skalierungsschicht zumindest einen Teil von Differenz-Spektralwerten ab der bestimmten Grenzfrequenz zu einer höheren Frequenz einzufügen.

21. Verfahren zum Codieren eines zeitdiskreten Audiosignals, um codierte Audiodaten zu erhalten, mit folgenden Schritten:

Liefern (52) eines Quantisierungs-Blocks von unter Verwendung eines psychoakustischen Modells (54) quantisierten Spektralwerten des zeitdiskreten Audiosignals;

inverses Quantisieren (58) des Quantisierungsblocks und zum Runden der invers quantisierten Spektralwerte, um einen Rundungs-Block von gerundeten invers quantisierten Spektralwerten zu erhalten;

Erzeugen (56) eines Ganzzahl-Blocks von ganzzahligen Spektralwerten unter Verwendung eines Ganzzahl-Transformationsalgorithmus, der ausgebildet ist, um aus einem Block von ganzzahligen zeitdiskreten Abtastwerten den Ganzzahl-Block von Spektralwerten zu erzeugen;

Bilden (58) eines Differenz-Blocks, der von einer spektralwertweisen Differenz zwischen dem Rundungs-Block

und dem Ganzzahl-Block abhängt, um einen Differenz-Block mit Differenz-Spektralwerten zu erhalten; und

Verarbeiten (60) des Quantisierungs-Blocks und des Differenz-Blocks, um codierte Audiodaten zu erzeugen, die Informationen über den Quantisierungs-Block und Informationen über den Differenz-Block umfassen.

22. Vorrichtung zum Decodieren von codierten Audiodaten, die aus einem zeitdiskreten Audiosignal durch Liefern (52) eines Quantisierungs-Blocks von unter Verwendung eines psychoakustischen Modells (54) quantisierten Spektralwerten des zeitdiskreten Audiosignals, durch inverses Quantisieren. (58) des Quantisierungsblocks und Runden der invers quantisierten Spektralwerte, um einen Rundungs-Block von gerundeten invers quantisierten Spektralwerten zu erhalten, durch Erzeugen (56) eines Ganzzahl-Blocks von ganzzahligen Spektralwerten unter Verwendung eines Ganzzahl-Transformationsalgorithmus, der ausgebildet ist, um aus einem Block von ganzzahligen zeitdiskreten Abtastwerten den Ganzzahl-Block von Spektralwerten zu erzeugen, und durch Bilden (58) eines Differenz-Blocks, der von einer spektralwertweisen Differenz zwischen dem Rundungs-Block und dem Ganzzahl-Block abhängt, um einen Differenz-Block mit Differenz-Spektralwerten zu erhalten, erzeugt worden sind, mit folgenden Merkmalen:

einer Einrichtung (70) zum Verarbeiten der codierten Audiodaten, um einen Quantisierungs-Block und einen Differenz-Block zu erhalten;

einer Einrichtung (74) zum inversen Quantisieren und Runden des Quantisierungs-Blocks, um einen ganzzahligen invers quantisierten Quantisierungs-Block zu erhalten;

einer Einrichtung (78) zum spektralwertweisen Kombinieren des ganzzahligen Quantisierungs-Blocks und des Differenz-Blocks, um einen Kombinations-Block zu erhalten; und

einer Einrichtung (82) zum Erzeugen einer zeitlichen Darstellung des zeitdiskreten Audiosignals unter Verwendung des Kombinations-Blocks und unter Verwendung eines zu dem Ganzzahl-Transformationsalgorithmus inversen Ganzzahl-Transformationsalgorithmus.

23. Vorrichtung zum Decodieren nach Anspruch 22, bei der die codierten Audiodaten skaliert sind und eine Mehrzahl von Skalierungsschichten umfassen, bei der die Einrichtung (70) zum Verarbeiten der codierten Audiodaten ausgebildet ist, um als erste Skalierungsschicht den Quantisierungsblock aus den codierten Audiodaten zu ermitteln, und um als zweite Skalierungsschicht den Differenzblock aus den codierten Audiodaten zu ermitteln.

24. Vorrichtung nach Anspruch 22, bei der die Informationen über den Differenzblock binär codierte Differenz-Spektralwerte umfassen, bei der die codierten Audiodaten skaliert sind und eine Mehrzahl von Skalierungsschichten umfassen, bei der die Einrichtung (70) zum Verarbeiten der codierten Audiodaten ausgebildet ist, um als erste Skalierungsschicht den Quantisierungsblock aus den codierten Audiodaten zu ermitteln, und um als zweite Skalierungsschicht eine Darstellung der Differenz-Spektralwerte mit reduzierter Genauigkeit zu extrahieren.

25. Vorrichtung nach Anspruch 24, bei der die Einrichtung (70) zum Verarbeiten der codierten Audiodaten ausgebildet ist, um als zweite Skalierungsschicht eine Anzahl von Bits ausgehend von einem höchstwertigen Bit bis zu einem niederwertigen Bit, das höherwertig als ein niederwertiges Bit eines Differenz-Spektralwerts ist, zu extrahieren, und bei der die Einrichtung (82) zum Erzeugen einer zeitlichen Darstellung des zeitdiskreten Audiosignals ausgebildet ist, um vor einer Verwendung des Ganzzahl-Transformationsalgorithmus fehlende Bits für einen Differenz-Spektralwert synthetisch zu erzeugen.

26. Vorrichtung nach Anspruch 25, bei der die Einrichtung (82) ausgebildet ist, um zur synthetischen Erzeugung eine Heraufskalierung der zweiten Skalierungsschicht durchzuführen, wobei bei der Heraufskalierung ein Skalierungsfaktor verwendet wird, der gleich 2^n ist, wobei n gleich der Anzahl von niederwertigen Bits ist, die in der zweiten Skalierungsschicht nicht enthalten sind, oder um zur synthetischen Erzeugung einen Dithering-Algorithmus einzusetzen.

27. Vorrichtung nach Anspruch 22, bei der die codierten Audiodaten skaliert sind und eine Mehrzahl von Skalierungsschichten umfassen, und

bei der die Einrichtung (70) zum Verarbeiten der codierten Audiodaten ausgebildet ist, um als erste Skalierungsschicht den Quantisierungsblock aus den codierten Audiodaten zu ermitteln, und um als zweite Skalierungsschicht tiefpassgefilterte Differenz-Spektralwerte zu ermitteln.

28. Vorrichtung nach Anspruch 22 oder 27, bei der die codierten Audiodaten skaliert sind und eine Mehrzahl von Skalierungsschichten umfassen, und bei der die Einrichtung (70) zum Verarbeiten der codierten Audiodaten ausgebildet ist, um als erste Skalierungsschicht den Quantisierungsblock aus den codierten Audiodaten zu ermitteln, und um als zweite Skalierungsschicht Differenz-Spektralwerte bis zu einer ersten Grenzfrequenz zu ermitteln, wobei die erste Grenzfrequenz kleiner als eine maximale Frequenz eines Differenz-Spektralwerts ist, der in einem Codierer erzeugbar ist.

29. Vorrichtung nach Anspruch 28, bei der die Einrichtung (82) zum Erzeugen einer zeitlichen Darstellung ausgebildet ist, um Eingangswerte in einen Ganzzahl-Transformationsalgorithmus voller Länge, die oberhalb der Grenzfrequenz der zweiten Skalierungsschicht sind, auf einen vorbestimmten Wert zu setzen, und um die zeitliche Darstellung des zeitdiskreten Audiosignals nach dem Verwenden des inversen Ganzzahl-Transformationsalgorithmus mit einem Faktor herunterabzutasten, der entsprechend einem Verhältnis einer maximalen Frequenz eines Differenz-Spektralwerts, der von einem Codierer erzeugbar ist, und der Grenzfrequenz gewählt wird.

30. Vorrichtung nach Anspruch 29, bei der der vorbestimmte Wert für alle Eingangswerte oberhalb der Grenzfrequenz gleich Null ist.

31. Verfahren zum Decodieren von codierten Audiodaten, die aus einem zeitdiskreten Audiosignal durch Liefern, inverses Quantisieren, Erzeugen, Bilden und Verarbeiten erzeugt worden sind, mit folgenden Schritten:

Verarbeiten (70) der codierten Audiodaten, um einen Quantisierungs-Block und einen Differenz-Block zu erhalten;

inverses Quantisieren (74) des Quantisierungs-Blocks und Runden, um einen ganzzahligen invers quantisierten Quantisierungs-Block zu erhalten;

spektralwertweises Kombinieren (78) des ganzzahligen Quantisierungs-Blocks und des Differenz-Blocks, um einen Kombinations-Block zu erhalten; und

Erzeugen (82) einer zeitlichen Darstellung des zeitdiskreten Audiosignals unter Verwendung des Kombinations-Blocks und unter Verwendung eines zu dem Ganzzahl-Transformationsalgorithmus inversen Ganzzahl-Transformationsalgorithmus.

32. Computer-Programm mit einem Programmcode zur Durchführung des Verfahrens zum Codieren nach Anspruch 21, wenn das Programm auf einem Computer abläuft.

33. Computer-Programm mit einem Programmcode zur Durchführung des Verfahrens zum Decodieren nach Anspruch 31, wenn das Programm auf einem Computer abläuft.

Claims

1. Apparatus for coding a time-discrete audio signal to obtain coded audio data, comprising:

means (52) for providing a quantization block of spectral values of the time-discrete audio signal quantized using a psychoacoustic model (54);

means (58) for inversely quantizing the quantization block and for rounding the inversely quantized spectral values to obtain a rounding block of rounded inversely quantized spectral values;

means (56) for generating an integer block of integer spectral values using an integer transform algorithm formed to generate the integer block of spectral values from a block of integer time-discrete samples;

combination means (58) for forming a difference block depending on a spectral value-wise difference between the rounding block and the integer block, to obtain a difference block with difference spectral values; and

means (60) for processing the quantization block and the difference block to generate coded audio data including information on the quantization block and information on the difference block.

2. Apparatus of claim 1, wherein means (52) for providing is formed to generate a MDCT block of MDCT spectral values from a time block of temporal audio signal values by means of an MDCT, and quantize the MDCT block using a psychoacoustic model to generate the quantization block comprising quantized MDCT spectral values.

3. Apparatus of claim 2, wherein the means (56) for generating the integer block is formed to execute an IntMDCT on the time block to generate the integer block comprising IntMDCT spectral values.

4. Apparatus of one of the preceding claims, wherein the means (52) for providing is formed to calculate the quantization block using a floating-point transform algorithm.

5. Apparatus of one of claims 1 to 3, wherein the means (52) for providing is formed to calculate the quantization block using the integer block generated by means (56) for generating.

6. Apparatus of one of the preceding claims, wherein the means (60) for processing is formed to subject (60a) the quantization block to entropy coding, to obtain an entropy-coded quantization block, to subject (60b) the rounding block to entropy coding, to obtain an entropy-coded rounding block, and to convert the entropy-coded quantization block to a first scaling layer of a scaled data stream representing the coded audio data, and to convert the entropy-coded rounding block to a second scaling layer of the scaled data stream.

7. Apparatus of claim 6, wherein the means (60) for processing is further formed to use one of the plurality of code tables depending on the quantized spectral values for the entropy coding of the quantization block, and wherein the means (60) for processing is further formed to select one of a plurality of code tables depending on a property of a quantizer usable in a quantization for generating the quantization block for the entropy coding of the difference block.

8. Apparatus of one of the preceding claims, wherein the means (52) for providing is formed to use one of a plurality of windows for windowing a temporal block of audio signal values depending on a property of the audio signal, and wherein the means (56) for generating is formed to make the same window selection for the integer transform algorithm.

9. Apparatus of one of claims 1 to 8, wherein the means for generating is formed to use an integer transform algorithm, comprising:

windowing the time-discrete samples with a window (w) with a length corresponding to $2N$ time-discrete samples, to provide windowed time-discrete samples for a conversion of the time-discrete samples to a spectral representation by means of a transform capable of generating N output values from N input values, wherein the windowing comprises the following substeps:

selecting (16) a time-discrete sample from a quarter of the window and a time-discrete sample from another quarter of the window to obtain a vector of time-discrete samples;

applying a square rotation matrix the dimension of which matches the dimension of the vector to the vector, wherein the rotation matrix is representable by a plurality of lifting matrices, wherein a lifting matrix only

comprises one element dependent on the window (w) and being unequal to 1 or 0, wherein the substep of applying comprises the following substeps:

5 multiplying (18) the vector by a lifting matrix to obtain a first result vector;

rounding (20) a component of the first result vector with a rounding function (r) mapping a real number to an integer to obtain a rounded first result vector; and

10 sequentially performing the steps of multiplying (22) and rounding (24) with another lifting matrix, until all lifting matrices are processed, to obtain a rotated vector comprising an integer windowed sample from the quarter of the window and an integer windowed sample from the other quarter of the window, and

performing the step of windowing for all time-discrete samples of the remaining quarters of the window to obtain 2N filtered integer values; and

15 converting (14) N windowed integer samples to a spectral representation by an integer DCT for values with the filtered integer samples of the second quarter and the third quarter of the window, to obtain N integer spectral values.

20 **10.** Apparatus of one of the preceding claims, wherein the means (52) for providing the quantization block is formed to perform a prediction of spectral values over the frequency using a prediction filter prior to a quantization step (52b), to obtain prediction residual spectral values representing the quantization block after a quantization; wherein also a prediction means is provided, which is formed to perform a prediction over the frequency of the integer spectral values of the integer block, wherein also rounding means is provided to round prediction residual spectral values due to the integer spectral values representing the rounding block.

30 **11.** Apparatus of one of the preceding claims, wherein the time-discrete audio signal comprises at least two channels, wherein the means (52) for providing is formed to perform center/side coding with spectral values of the time-discrete audio signal to obtain the quantization block after quantization of center/side spectral values, and wherein the means (56) for generating the integer block is formed to also perform center/side coding corresponding to the center/side coding of the means (52) for providing.

35 **12.** Apparatus of one of the preceding claims, wherein the means (60) for processing is formed to generate a MPEG-2 AAC data stream, wherein in a field Ancillary Data ancillary information for the integer transform algorithm is introduced.

40 **13.** Apparatus of one of the preceding claims, wherein the means for processing (60) is formed to output the coded audio data as data stream with a plurality of scaling layers.

45 **14.** Apparatus of claim 13, wherein the means for processing (60) is formed to insert information on the quantization block into a first scaling layer (81), and to insert information on the difference block into a second scaling layer (82).

50 **15.** Apparatus of claim 13, wherein the means for processing (60) is formed to insert information on the quantization block into a first scaling layer, and to insert the information on the difference block into at least a second and a third scaling layer.

16. Apparatus of claim 15, wherein in the second scaling layer difference spectral values with reduced accuracy are contained, and in one or more higher scaling layers a residual part of the difference spectral values is contained.

55 **17.** Apparatus of claim 15 or 16, wherein the information on the difference block includes binarily coded difference spectral values, wherein the second scaling layer for difference spectral values includes a number of bits from a most significant bit (MSB) to a less significant bit (MSB-x) for a difference spectral value, and

wherein the third scaling layer includes a number of bits starting from a less significant bit (MSB-x-1) to a least significant bit (LSB).

18. Apparatus of claim 17,

wherein the time-discrete audio signal is present in form of samples with a width of 24 bits, and wherein the means for processing (60) is formed to insert more significant 16 bits of difference spectral values into the second scaling layer, and to insert residual 8 bits of a difference spectral value into the third scaling layer, so that a decoder reaches CD quality using the second scaling layer, wherein a decoder reaches studio quality using also the third scaling layer.

19. Apparatus of claim 15,

wherein the means (60) for processing is formed to insert at least part of difference spectral values for representation of a low-pass filtered signal into a second scaling layer, and to insert a difference between the difference spectral values in the second scaling layer and original difference spectral values into at least one further scaling layer.

20. Apparatus of claim 15 or 19,

wherein the means (60) for processing is formed to insert at least part of different spectral values up to a certain cut-off frequency into a second scaling layer, and to insert at least part of difference spectral values from the certain cut-off frequency to a higher frequency into a third scaling layer.

21. Method of coding a time-discrete audio signal to obtain coded audio data, comprising:

providing (52) a quantization block of spectral values of a time-discrete audio signal quantized using a psychoacoustic model (54);

inversely quantizing (58) the quantization block and rounding the inversely quantized spectral values to obtain a rounding block of rounded inversely quantized spectral values;

generating (56) an integer block of integer spectral values using an integer transform algorithm formed to generate the integer block of spectral values from a block of integer time-discrete samples;

forming (58) a difference block depending on a spectral value-wise difference between the rounding block and the integer block, to obtain a difference block with difference spectral values; and

processing (60) the quantization block and the difference block to generate coded audio data including information on the quantization block and information on the difference block.

22. Apparatus for decoding coded audio data having been generated from a time-discrete audio signal by providing (52) a quantization block of spectral values of the time-discrete audio signal quantized using a psychoacoustic model (54), by inversely quantizing (58) the quantization block and rounding the inversely quantized spectral values to obtain a rounding block of rounded inversely quantized spectral values, by generating (56) of an integer block of integer spectral values using an integer transform algorithm formed to generate the integer block of spectral values from a block of integer time-discrete samples, and by forming (58) a difference block depending on a spectral value-wise difference between the rounding block and the integer block, to obtain a difference block with difference spectral values, comprising:

means (70) for processing the coded audio data to obtain a quantization block and a difference block;

means (74) for inversely quantizing and rounding the quantization block to obtain an integer inversely quantized quantization block;

means (78) for spectral value-wise combining the integer quantization block and the difference block to obtain a combination block; and

means (82) for generating a temporal representation of the time-discrete audio signal using the combination block and using an integer transform algorithm inverse to the integer transform algorithm.

23. Apparatus for decoding of claim 22,
 wherein the coded audio data is scaled and includes a plurality of scaling layers,
 wherein the means (70) for processing the coded audio data is formed to ascertain the quantization block from
 the coded audio data as first scaling layer, and to ascertain the difference block from the coded audio data as
 second scaling layer.

24. Apparatus of claim 22,
 wherein the information on the difference block includes binarily coded difference spectral values,
 wherein the coded audio data is scaled and includes a plurality of scaling layers,
 wherein the means (70) for processing the coded audio data is formed to ascertain the quantization block from
 the coded audio data as first scaling layer, and to extract a representation of the difference spectral values with
 reduced accuracy as second scaling layer.

25. Apparatus of claim 24,
 wherein the means (70) for processing the coded audio data is formed to extract a number of bits starting from a
 most significant bit to a less significant bit, which is more significant than a least significant bit of a difference
 spectral value, as second scaling layer, and
 wherein the means (82) for generating a temporal representation of the time-discrete audio signal is formed to
 synthetically generate missing bits for a difference spectral value before using the integer transform algorithm.

26. Apparatus of claim 25,
 wherein the means (82) is formed to perform an upscaling of the second scaling layer for the synthetical generation,
 wherein in the upscaling a scale factor is used, which equals 2^n , wherein n is the number of less significant bits
 not contained in the second scaling layer, or to employ a dithering algorithm for the synthetical generation.

27. Apparatus of claim 22,
 wherein the coded audio data is scaled and includes a plurality of scaling layers, and
 wherein the means (70) for processing coded audio data is formed to ascertain the quantization block from the
 coded audio data as first scaling layer, and to ascertain low-pass filtered difference spectral values as second
 scaling layer.

28. Apparatus of claim 22 or 27,
 wherein the coded audio data is scaled and includes a plurality of scaling layers, and
 wherein the means (70) for processing the coded audio data is formed to ascertain the quantization block of the
 coded data as first scaling layer, and to ascertain difference spectral values up to a first cut-off frequency as second
 scaling layer, wherein the first cut-off frequency is smaller than the maximum frequency of a difference spectral
 value, which may be generated in a coder.

29. Apparatus of claim 28,
 wherein the means (82) for generating a temporal representation is formed to set input values in an integer trans-
 form algorithm of full length, which are above the cut-off frequency of the second scaling layer, to a predetermined
 value, and to downsample the temporal representation of the time-discrete audio signal after using the inverse
 integer transform algorithm by a factor chosen corresponding to a ratio of a maximum frequency of a difference
 spectral value, which may be generated by a coder, and the cut-off frequency.

30. Apparatus of claim 29,
 wherein the predetermined value for all input values above the cut-off frequency is zero.

31. Method of decoding coded audio data having been generated from a time-discrete audio signal by providing,
 inversely quantizing, generating, forming, and processing, comprising:

processing (70) the coded audio data to obtain a quantization block and a difference block;

inversely quantizing (74) the quantization block and rounding to obtain an integer inversely quantized quan-
 tization block;

spectral value-wise combining (78) the integer quantization block and the difference block to obtain a combi-
 nation block; and

generating (82) a temporal representation of the time-discrete audio signal using a combination block and using an integer transform algorithm inverse to the integer transformation algorithm.

32. Computer program with a program code for performing the method of coding of claim 21, when the program is executed on a computer.

33. Computer program with a program code for performing the method of decoding of claim 31, when the program is executed on a computer.

Revendications

1. Dispositif pour coder un signal audio discret dans le temps, pour obtenir des données audio codées, aux caractéristiques suivantes :

un dispositif (52) destiné à fournir un bloc de quantification de valeurs spectrales du signal audio discret dans le temps quantifiées à l'aide d'un modèle psycho-acoustique (54) ;

un dispositif (58) destiné à quantifier inversement le bloc de quantification et à arrondir les valeurs spectrales quantifiées inversement, pour obtenir un bloc d'arrondi de valeurs spectrales quantifiées inversement arrondies ;

un dispositif (56) destiné à générer un bloc de nombre entier de valeurs spectrales de nombre entier à l'aide d'un algorithme de transformation de nombre entier qui est réalisé de manière à générer, à partir d'un bloc de valeurs de balayage discrètes dans le temps de nombre entier, le bloc de nombre entier de valeurs spectrales ;

un dispositif de combinaison (58) destiné à former un bloc de différence qui dépend d'une différence en valeurs spectrales entre le bloc d'arrondi et le bloc de nombre entier, pour obtenir un bloc de différence avec des valeurs spectrales de différence ;

un dispositif (60) destiné à traiter le bloc de quantification et le bloc de différence, pour générer des données audio codées comportant des informations sur le bloc de quantification et des informations sur le bloc de différence.

2. Dispositif selon la revendication 1, dans lequel le dispositif (52) destiné à fournir est réalisé pour générer, à partir d'un bloc de temps de valeurs de signal audio dans le temps, au moyen d'une MDCT, un bloc MDCT de valeurs spectrales MDCT, et pour quantifier le bloc MDCT à l'aide d'un modèle psycho-acoustique, pour générer le bloc de quantification qui présente des valeurs spectrales MDCT quantifiées.

3. Dispositif selon la revendication 2, dans lequel le dispositif (56) destiné à générer le bloc de nombre entier est réalisé de manière d'effectuer une IntMDCT sur le bloc de temps, pour générer le bloc de nombre entier présentant des valeurs spectrales IntMDCT.

4. Dispositif selon l'une des revendications précédentes, dans lequel le dispositif (52) destiné à fournir est réalisé de manière à calculer, à l'aide d'un algorithme de transformation à virgule flottante, le bloc de quantification.

5. Dispositif selon l'une des revendications 1 à 3, dans lequel le dispositif (52) destiné à fournir est réalisé de manière à calculer le bloc de quantification à l'aide du dispositif (56) destiné à générer le bloc de nombre entier généré.

6. Dispositif selon l'une des revendications précédentes, dans lequel le dispositif (60) destiné à traiter est réalisé de manière à soumettre le bloc de quantification à un codage entropique (60a), pour obtenir un bloc de quantification codé de manière entropique, à soumettre le bloc d'arrondi à un codage entropique (60b), pour obtenir un bloc d'arrondi codé de manière entropique, et

à transmettre le bloc de quantification codé de manière entropique dans une première couche de modulation d'un flux de données modulé qui représente les données audio codées, et à transmettre le bloc d'arrondi codé de manière entropique dans une deuxième couche de modulation du flux de données modulé.

7. Dispositif selon la revendication 6,

dans lequel le dispositif (60) destiné à traiter est réalisé, par ailleurs, de manière à utiliser, pour le codage entropique du bloc de quantification, l'un parmi une pluralité de tableaux de code en fonction des valeurs spectrales quantifiées, et

dans lequel le dispositif (60) destiné à traiter est réalisé, en outre, de manière à sélectionner, pour le codage entropique du bloc de différence, l'un parmi une pluralité de tableaux de code en fonction d'une propriété d'un dispositif de quantification pouvant être utilisé lors d'une quantification pour générer le bloc de quantification.

8. Dispositif selon l'une des revendications précédentes,

dans lequel le dispositif (52) destiné à fournir est réalisé de manière à utiliser, en fonction d'une qualité du signal audio, l'une parmi une pluralité de fenêtres, pour la division en fenêtres d'un bloc dans le temps des valeurs de signal audio, et

dans lequel le dispositif (56) destiné à générer est réalisé de manière à effectuer, pour l'algorithme de transformation de nombre entier, la même sélection de fenêtre.

9. Dispositif selon l'une des revendications 1 à 8,

dans lequel le dispositif destiné à générer est réalisé de manière à utiliser un algorithme de transformation de nombre entier, présentant les étapes suivantes consistant à :

diviser en fenêtres les valeurs de balayage discrètes dans le temps avec une fenêtre (w) d'une longueur correspondant à 2N valeurs de balayage discrètes dans le temps, pour fournir des valeurs de balayage discrètes dans le temps divisées en fenêtres pour une conversion des valeurs de balayage discrètes dans le temps en une représentation spectrale au moyen d'une transformation pouvant générer, à partir de N valeurs d'entrée, N valeurs de sortie, la division en fenêtres présentant les étapes partielles suivantes :

sélectionner (16) une valeur de balayage discrète dans le temps d'un quart de la fenêtre et une valeur de balayage discrète dans le temps d'un autre quart de la fenêtre, pour obtenir un vecteur de valeurs de balayage discrètes dans le temps ;

soumettre le vecteur à une matrice de rotation carrée dont la dimension coïncide avec la dimension du vecteur, la matrice de rotation pouvant être représentée par une pluralité de matrices d'élévation, une matrice d'élévation ne présentant qu'un seul élément qui est fonction de la fenêtre (w) et est différent de 1 ou 0, l'étape partielle de la soumission présentant les sous-étapes suivantes consistant à :

multiplier (18) le vecteur par une matrice d'élévation, pour obtenir un premier vecteur de résultat ; arrondir (20) une composante du premier vecteur de résultat par une fonction d'arrondi (r) qui reproduit un nombre réel sur un nombre entier, pour obtenir un premier vecteur de résultat arrondi ; et réaliser de manière séquentielle les étapes de multiplication (22) et d'arrondi (24) par une autre matrice d'élévation, jusqu'à ce que toutes les matrices d'élévation soient traitées, pour obtenir un vecteur tourné qui présente une valeur de balayage divisée en fenêtres de nombre entier du quart de la fenêtre et une valeur de balayage divisée en fenêtres de nombre entier de l'autre quart de la fenêtre, et réaliser l'étape de division en fenêtres pour toutes les valeurs de balayage discrètes dans le temps des quarts restants de la fenêtre, pour obtenir 1N valeurs de balayage de nombre entier filtrées ; et convertir (14) N valeurs de balayage de nombre entier divisées en fenêtres en une représentation spectrale par une DCT de nombre entier pour des valeurs avec les valeurs de balayage de nombre entier filtrées du deuxième quart et du troisième quart de la fenêtre, pour obtenir N valeurs spectrales de nombre entier.

10. Dispositif selon l'une des revendications précédentes,

dans lequel le dispositif (52) destiné à fournir le bloc de quantification est réalisé de manière à réaliser, avant une étape de quantification (52b), une prédiction de valeurs spectrales sur la fréquence, à l'aide d'un filtre de prédiction, pour obtenir des valeurs spectrales résiduelles de prédiction qui représentent, après une quantification, le bloc de quantification ;

dans lequel est, par ailleurs, prévu un dispositif de prédiction qui est réalisé de manière à effectuer une prédiction sur la fréquence des valeurs spectrales de nombre entier du bloc de nombre entier, par ailleurs étant prévu un dispositif d'arrondi, pour arrondir les valeurs spectrales résiduelles de prédiction sur base des valeurs spectrales de nombre entier représentant le bloc d'arrondi.

11. Procédé selon l'une des revendications précédentes,

dans lequel le signal audio discret dans le temps présente au moins deux canaux,
 dans lequel le dispositif (52) destiné à fournir est réalisé de manière à effectuer un codage de centre/côté
 avec des valeurs spectrales du signal audio discret dans le temps, pour obtenir, après une quantification de valeurs
 spectrales de centre/côté, le bloc de quantification, et

dans lequel le dispositif (56) destiné à générer le bloc de nombre entier est réalisé de manière à effectuer
 également un codage centre/côté correspondant au codage de centre/côté du dispositif (52) destiné à fournir.

12. Dispositif selon l'une des revendications précédentes,

dans lequel le dispositif (60) destiné à traiter est réalisé de manière à générer un flux de données MPEG-
 2-AAC, dans un champ Ancillary Data étant introduits des informations auxiliaires pour l'algorithme de transfor-
 mation de nombre entier.

13. Dispositif selon l'une des revendications précédentes,

dans lequel le dispositif destiné à traiter (60) est réalisé de manière à sortir les données audio codées comme
 flux de données avec une pluralité de couches de modulation.

14. Dispositif selon la revendication 13,

dans lequel le dispositif destiné à traiter (60) est réalisé de manière à introduire dans une première couche
 de modulation (81) des informations sur le bloc de quantification et à introduire dans une deuxième couche de
 modulation (82) des informations sur le bloc de différence.

15. Dispositif selon la revendication 13,

dans lequel le dispositif destiné à traiter (60) est réalisé de manière à introduire dans une première couche
 de modulation des informations sur le bloc de quantification et à introduire les informations sur le bloc de différence
 dans au moins une deuxième et une troisième couche de modulation.

16. Dispositif selon la revendication 15,

dans lequel sont contenues, dans la deuxième couche de modulation, des valeurs spectrales de différence
 à précision réduite et, dans une ou plusieurs couches de modulation supérieures, une partie restante des valeurs
 spectrales de différence.

17. Dispositif selon la revendication 15 ou 16,

dans lequel les informations sur le bloc de différence comportent des valeurs spectrales de différence codées
 de manière binaire,

dans lequel la deuxième couche de modulation comporte, pour les valeurs spectrales de différence, un cer-
 tain de bits d'un bit à la valeur la plus élevée (MSB) à un bit de valeur inférieure (MSB-x) pour une valeur spectrales
 de différence, et

dans lequel la troisième couche de modulation comporte un nombre de bits partant d'un bit à la valeur la
 plus basse (MSB-x-1) à un bit à la valeur la plus élevée (LSB).

18. Dispositif selon la revendication 17,

dans lequel le signal audio discret dans le temps est présent sous forme de valeurs de balayage avec une
 largeur de 24 bits, et

dans lequel le dispositif destiné à traiter (60) est réalisé de manière à introduire 16 bits de valeur supérieure
 des valeurs spectrales de différence dans la deuxième couche de modulation, et à introduire les 8 bits restants
 d'une valeur spectrale de différence dans la troisième couche de modulation, de sorte qu'un décodeur atteigne,
 à l'aide de la deuxième couche de modulation, une qualité CD, un décodeur atteignant, à l'aide également de la
 troisième couche de modulation, une qualité de studio.

19. Dispositif selon la revendication 15,

dans lequel le dispositif (60) destiné à traiter est réalisé de manière à introduire dans une deuxième couche
 de modulation au moins une partie des valeurs spectrales de différence pour la représentation d'un signal filtré
 passe-bas et à introduire dans au moins une autre couche de modulation une différence entre les valeurs spectrales
 de différence dans la deuxième couche de modulation et les valeurs spectrales de différence originales.

20. Dispositif selon la revendication 15 ou 19,

dans lequel le dispositif (60) destiné à traiter est réalisé de manière à introduire dans une deuxième couche
 de modulation au moins une partie des valeurs spectrales de différence jusqu'à une fréquence limite déterminée,

et à introduire dans une troisième couche de modulation au moins une partie des valeurs spectrales de différence de la fréquence limite déterminée à une fréquence supérieure.

21. Procédé pour coder un signal audio discret dans le temps, pour obtenir des données audio codées, aux étapes suivantes consistant à :

fournir (52) un bloc de quantification de valeurs spectrales du signal discret dans le temps quantifiées à l'aide d'un modèle psycho-acoustique (54) ;
quantifier inversement (58) le bloc de quantification et arrondir les valeurs spectrales quantifiées inversement, pour obtenir un bloc d'arrondi de valeurs spectrales quantifiées inversement arrondies ;
générer (56) un bloc de nombre entier de valeurs spectrales de nombre entier à l'aide d'un algorithme de transformation de nombre entier réalisé de manière à générer, à partir d'un bloc de valeurs de balayage discrètes dans le temps de nombre entier, le bloc de nombre entier de valeurs spectrales ;
former (58) un bloc de différence qui est fonction d'une différence en valeurs spectrales entre le bloc d'arrondi et le bloc de nombre entier, pour obtenir un bloc de différence avec des valeurs spectrales de différence ; et
traiter (60) le bloc de quantification et le bloc de différence, pour générer des données audio codées comportant les informations sur le bloc de quantification et des informations sur le bloc de différence.

22. Dispositif pour décoder des données audio codées qui ont été générées à partir d'un signal audio discret dans le temps en fournissant (52) un bloc de quantification de valeurs spectrales du signal audio discret dans le temps quantifiées à l'aide d'un modèle psycho-acoustique (54), en quantifiant inversement (58) le bloc de quantification et en arrondissant les valeurs spectrales quantifiées inversement, pour obtenir un bloc d'arrondi de valeurs spectrales quantifiées inversement arrondies, en générant (56) un bloc de nombre entier de valeurs spectrales de nombre entier à l'aide d'un algorithme de transformation de nombre entier réalisé de manière à générer, à partir d'un bloc de valeurs de balayage discrètes dans le temps de nombre entier, le bloc de nombre entier de valeurs spectrales, et en formant (58) un bloc de différence qui est fonction d'une différence en valeurs spectrales entre le bloc d'arrondi et le bloc de nombre entier, pour obtenir un bloc de différence à valeurs spectrales de différence, aux caractéristiques suivantes :

un dispositif (70) destiné à traiter les données audio codées, pour obtenir un bloc de quantification et un bloc de différence ;
un dispositif (74) destiné à quantifier inversement et à arrondir le bloc de quantification, pour obtenir un bloc de quantification quantifié inversement de nombre entier ;
un dispositif (78) destiné à combiner en valeurs spectrales le bloc de quantification de nombre entier et le bloc de différence, pour obtenir un bloc de combinaison ; et
un dispositif (82) destiné à générer une représentation dans le temps du signal audio discret dans le temps à l'aide du bloc de combinaison et à l'aide d'un algorithme de transformation inverse de nombre entier inverse à l'algorithme de transformation de nombre entier.

23. Dispositif pour décoder selon la condition 22,
dans lequel les données audio codées sont modulées et comportent une pluralité de couches de modulation, dans lequel le dispositif (70) destiné à traiter les données audio codées est réalisé de manière à déterminer, comme première couche de modulation, le bloc de quantification à partir des données audio codées, et à déterminer, comme deuxième couche de modulation, le bloc de différence à partir des données audio codées.

24. Dispositif selon la revendication 22,
dans lequel les informations sur le bloc de différence comportent des couvertures valeurs spectrales de différence codées de manière binaires,
dans lequel les données audio codées sont modulées et comportent une pluralité de couches de modulation, dans lequel le dispositif (70) destiné à traiter les données audio codées est réalisé de manière à déterminer, comme première couche de mesure, le bloc de quantification à partir des données audio codées et à extraire, comme deuxième couche de modulation, une représentation des valeurs spectrales de différence à précision réduite.

25. Dispositif selon la revendication 24,
dans lequel le dispositif (70) destiné à traiter les données audio codées est réalisé de manière à extraire, comme deuxième couche de modulation, un nombre de bits partant d'un bit à la valeur la plus élevée jusqu'à un bit de valeur inférieure qui est d'une valeur supérieure à un bit à la valeur la plus basse d'une valeur spectrale de

différence, et

dans lequel le dispositif (82) destiné à générer une représentation dans le temps du signal audio discret dans le temps est réalisé de manière à générer synthétiquement, avant une utilisation de l'algorithme de transformation de nombre entier, des bits manquants pour une valeur spectrale de différence.

26. Dispositif selon la revendication 25,

dans lequel le dispositif (82) est réalisé de manière à effectuer, pour la génération synthétique, une modulation vers le haut de la deuxième couche de modulation, lors de la modulation vers le haut étant utilisé un facteur de modulation qui est égal à 2^n , n étant égal au nombre de bits de valeur basse qui ne sont pas contenus dans la deuxième couche de modulation, ou à utiliser, pour la génération synthétique, un algorithme de Dithering.

27. Dispositif selon la revendication 22,

dans lequel les données audio codées sont modulées et comportent une pluralité de couches de modulation, et

dans lequel le dispositif (70) destiné à traiter les données audio codées est réalisé de manière à déterminer, comme première couche de mesure, le bloc de quantification à partir des données audio codées et à déterminer, comme deuxième couche de modulation, des valeurs spectrales de différence filtrées passe-bas.

28. Dispositif selon la revendication 22 ou 27,

dans lequel les données audio codées sont modulées et comportent une pluralité de couches de modulation, et

dans lequel le dispositif (70) destiné à traiter les données audio codées est réalisé de manière à déterminer, comme première couche de modulation, le bloc de quantification à partir des données audio codées, et à déterminer, comme deuxième couche de modulation, des valeurs spectrales de différence jusqu'à une première fréquence limite, la première fréquence limite étant inférieure à une fréquence maximale d'une valeur spectrale de différence pouvant être générée dans un codeur.

29. Dispositif selon la revendication 28,

dans lequel le dispositif (82) destiné à générer une représentation dans le temps est réalisé de manière à régler des valeurs d'entrée dans un algorithme de transformation de nombre de pleine longueur qui se situent au-dessus de la fréquence limite de la deuxième couche de modulation à une valeur prédéterminée, et à balayer vers le bas la représentation dans le temps du signal audio discret dans le temps, après l'utilisation de l'algorithme de transformation de nombre entier inverse, d'un facteur qui est choisi en fonction d'un rapport entre une fréquence maximale d'une valeur spectrale de différence pouvant être générée par un codeur et la fréquence limite.

30. Dispositif selon la revendication 29,

dans lequel la valeur prédéterminée est, pour toutes les valeurs d'entrée au-dessus de la fréquence limite, égale à zéro.

31. Procédé pour décoder des données audio codées qui ont été générées à partir d'un signal audio discret dans le temps par fourniture, quantification inverse, génération, formation et traitement, aux étapes suivantes consistant à :

traiter (70) les données audio codées, pour obtenir un bloc de quantification et un bloc de différence ;
quantifier inversement (74) le bloc de quantification et arrondir, pour obtenir un bloc de quantification quantifié inversement de nombre entier ;
combiner en valeurs spectrales (78) le bloc de quantification de nombre entier et le bloc de différence, pour obtenir un bloc de combinaison ;
générer (82) une représentation dans le temps du signal audio discret dans le temps à l'aide du bloc de combinaison et à l'aide d'un d'algorithme de transformation de nombre entier inverse à l'algorithme de transformation de nombre entier.

32. Programme d'ordinateur avec un code de programme pour la mise en oeuvre du procédé pour coder selon la revendication 21, lorsque le programme est exécuté sur un ordinateur.

33. Programme d'ordinateur avec un code de programme pour la mise en oeuvre du procédé pour décoder selon la revendication 31, lorsque le programme est exécuté sur un ordinateur.

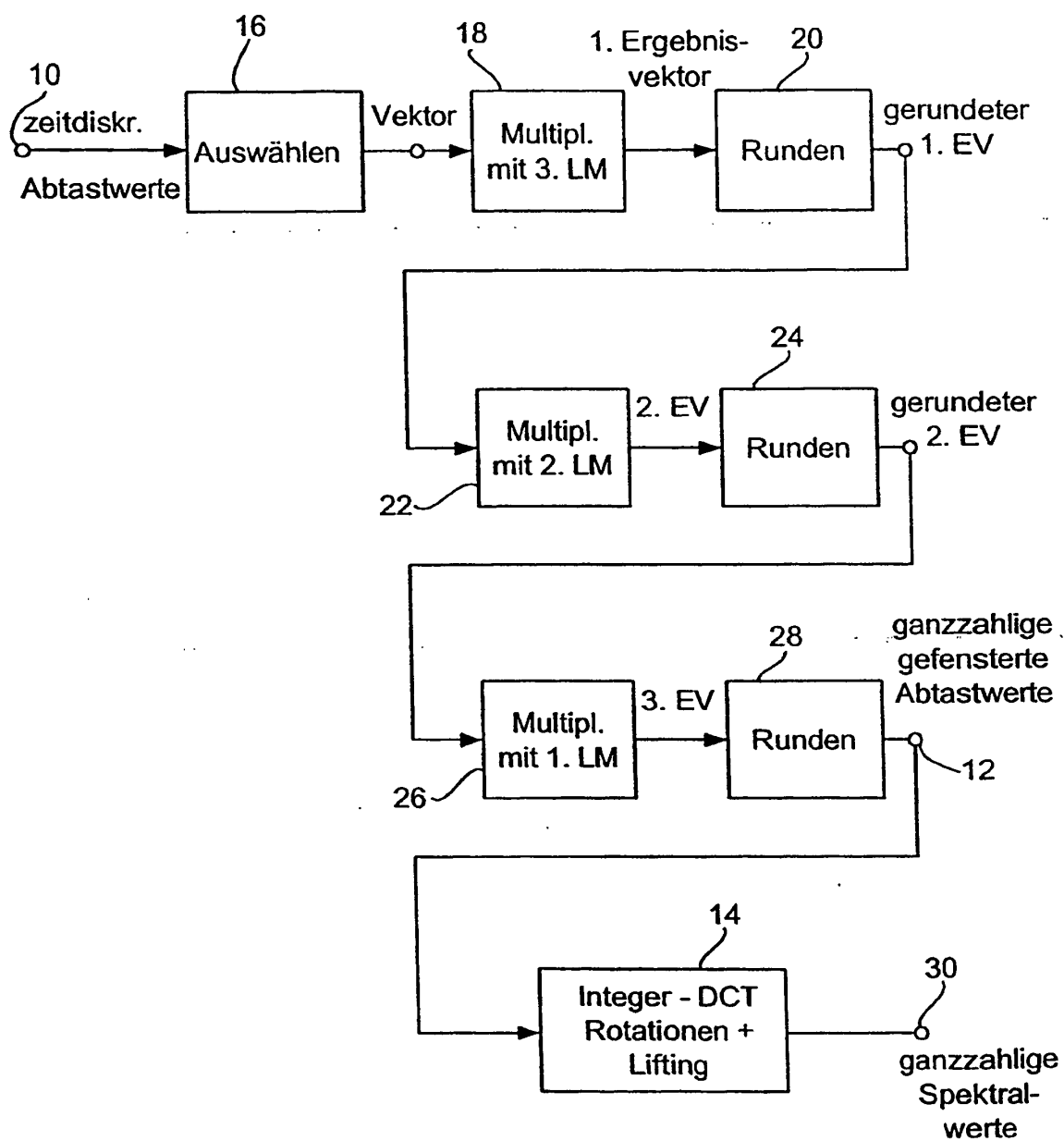


FIG. 1

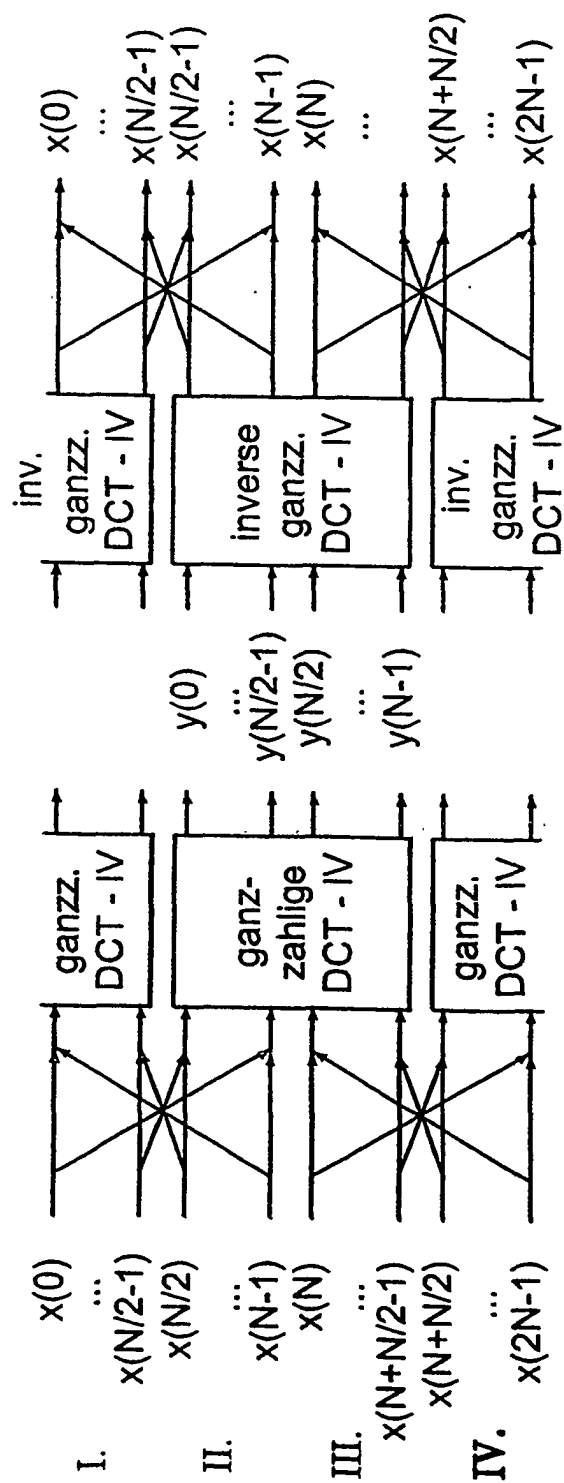


FIG. 2

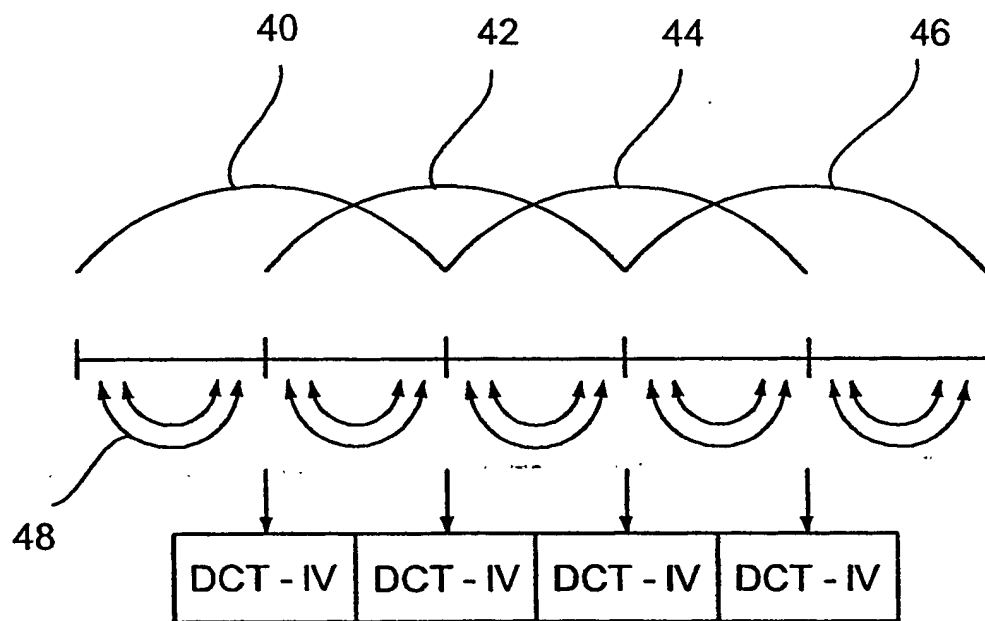


FIG. 3

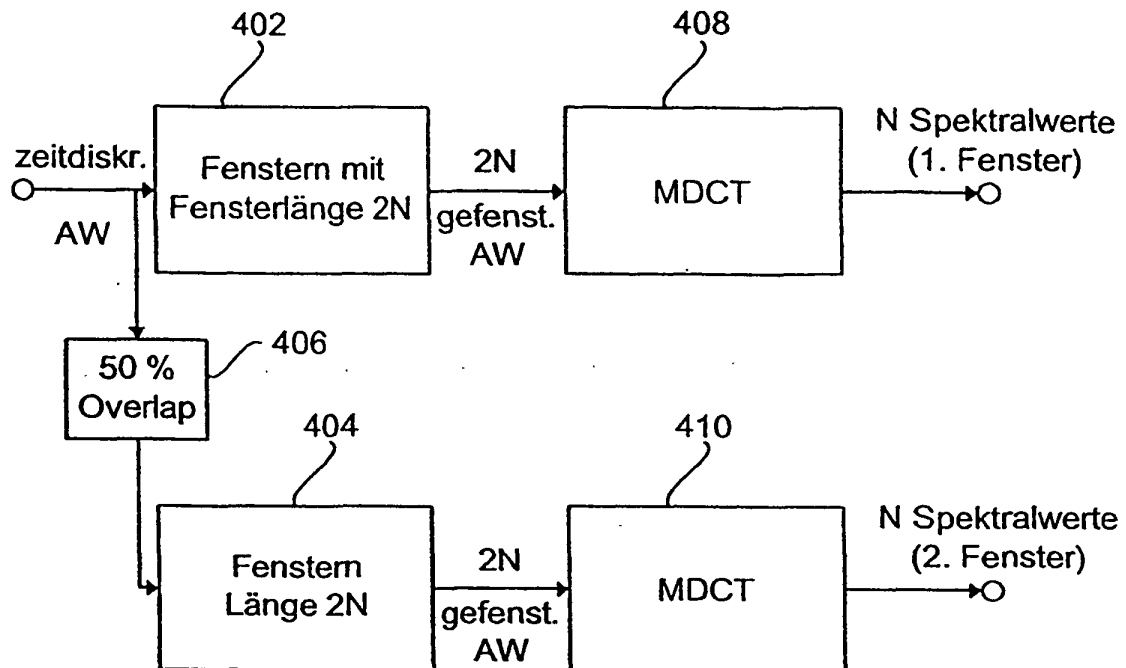


FIG. 4a

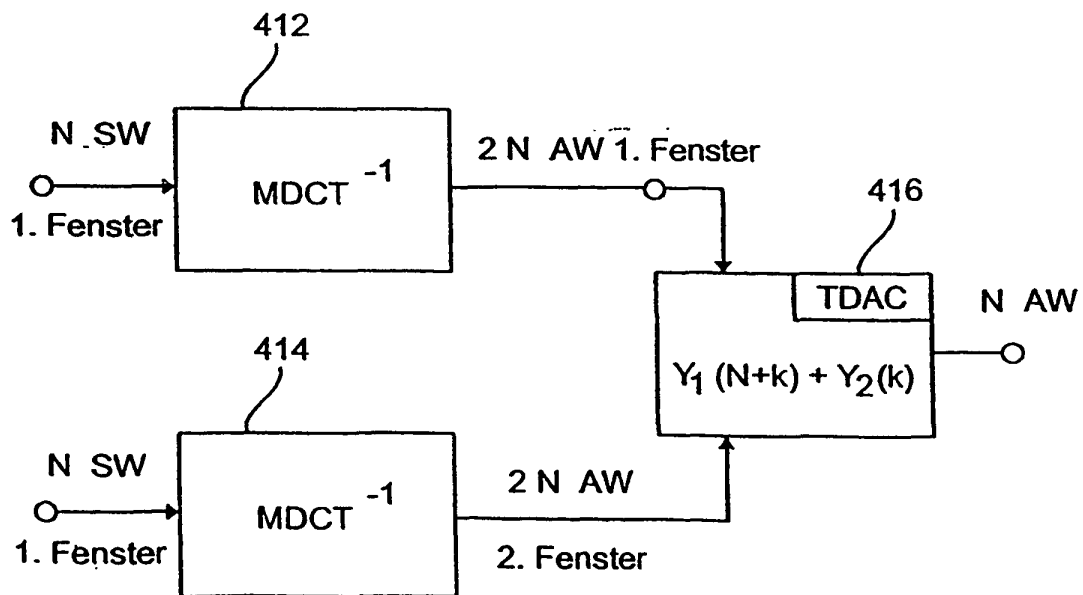


FIG. 4b

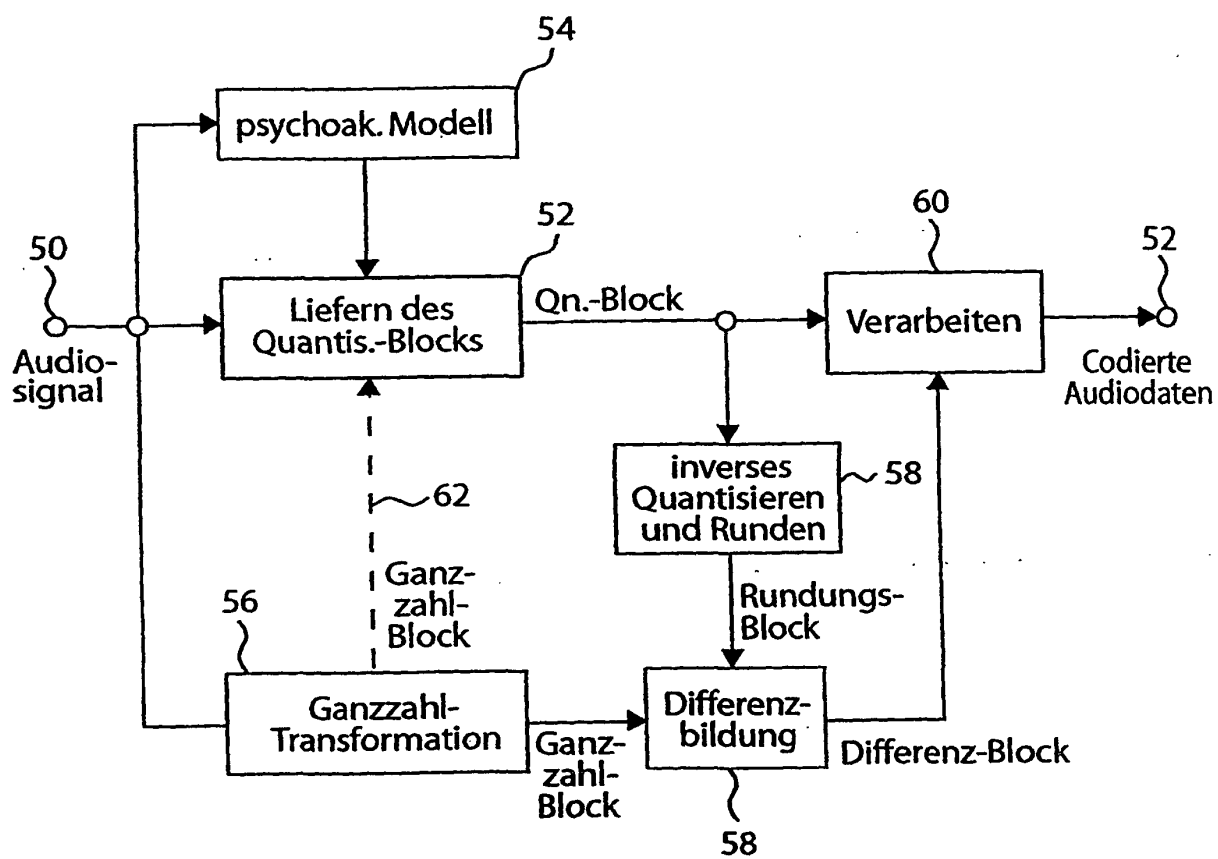


FIG. 5

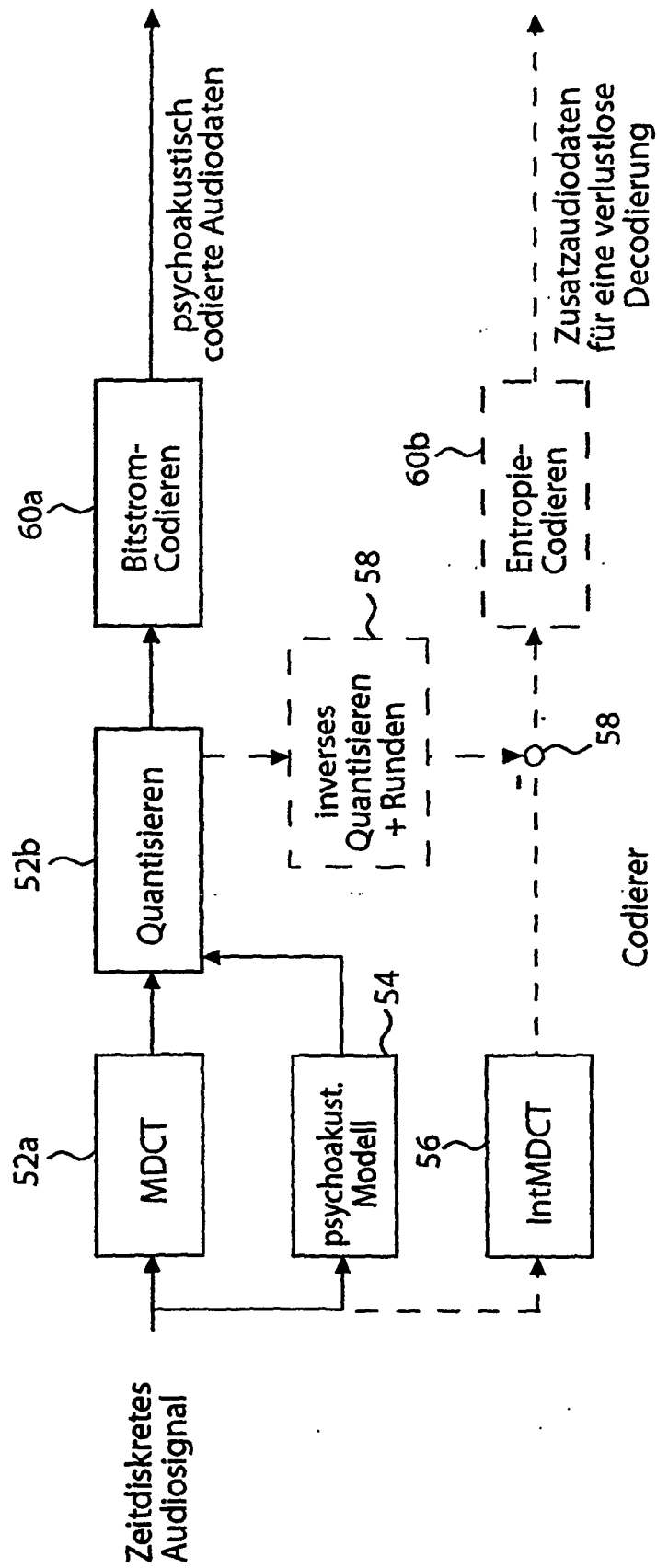


FIG. 6

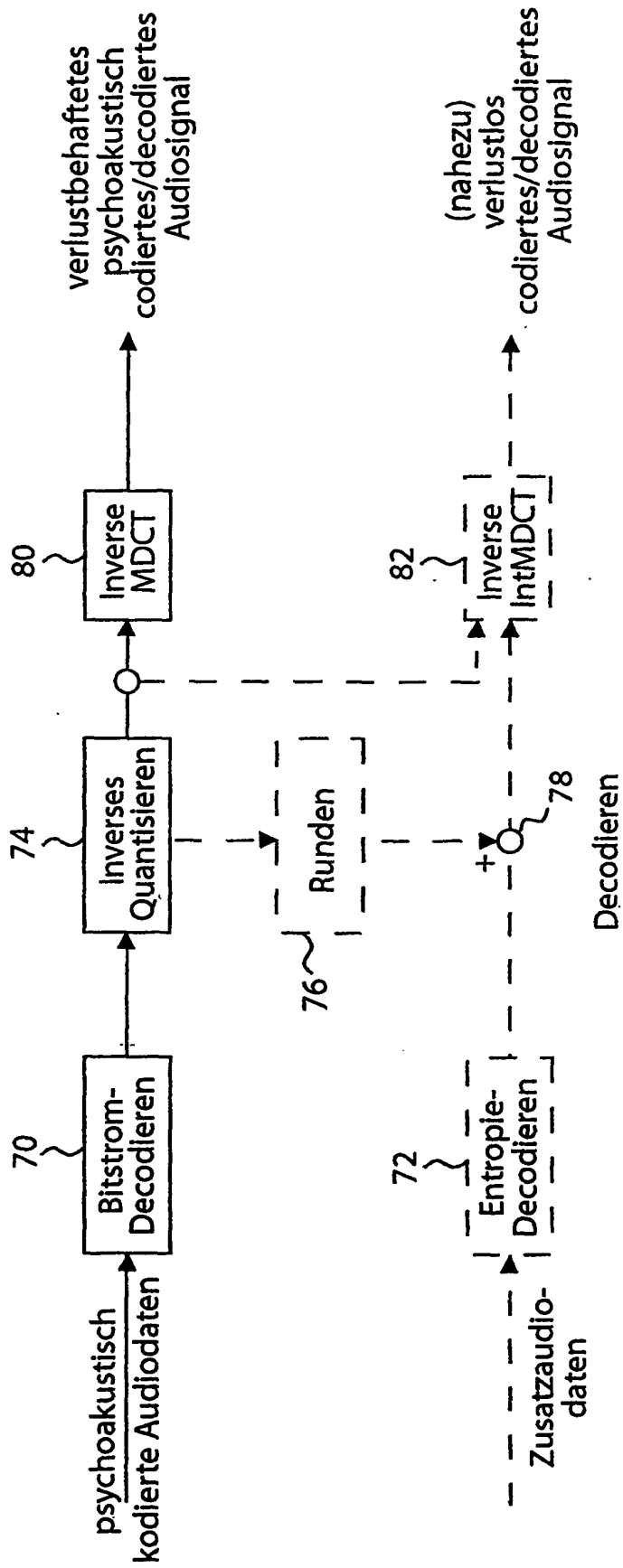


FIG. 7

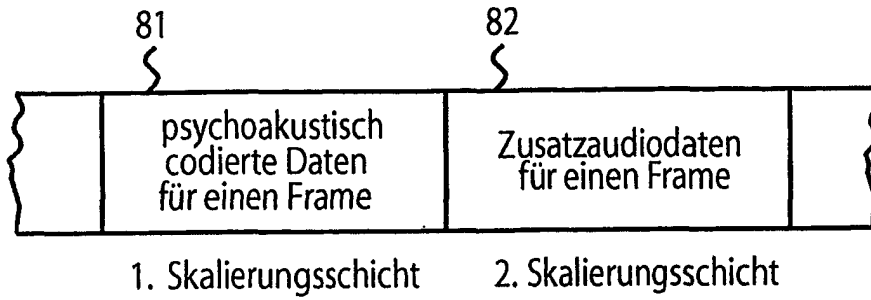


FIG. 8a

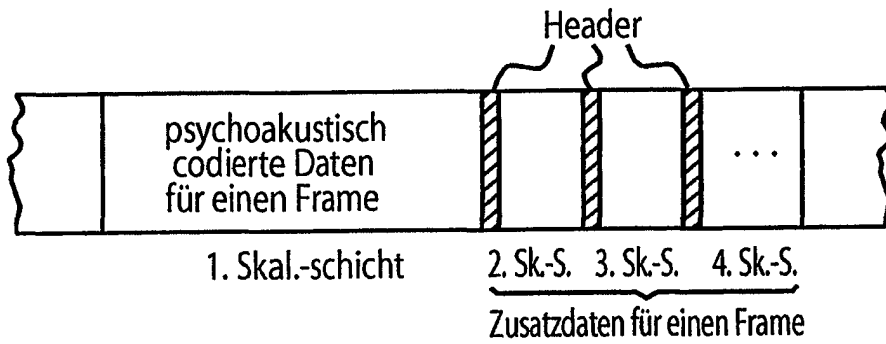


FIG. 8b

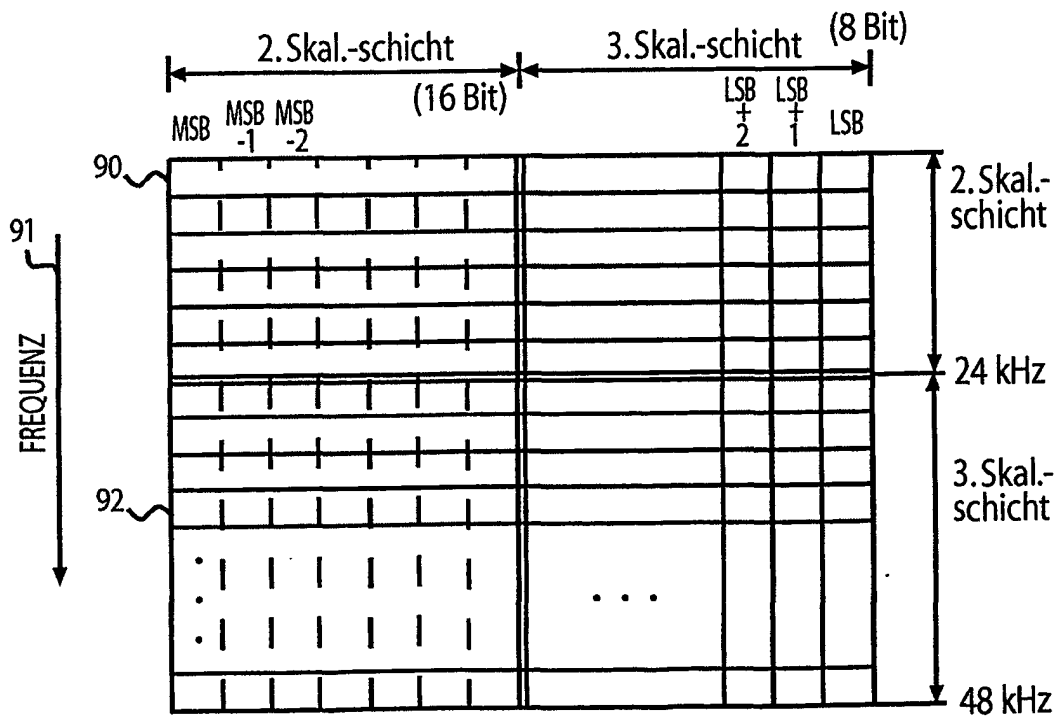


FIG. 9 (Differenzspektalwerte)