



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 603 17 670 T2 2008.10.30**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 515 561 B1**

(51) Int Cl.⁸: **H04N 7/26 (2006.01)**

(21) Deutsches Aktenzeichen: **603 17 670.4**

(96) Europäisches Aktenzeichen: **03 255 624.3**

(96) Europäischer Anmeldetag: **09.09.2003**

(97) Erstveröffentlichung durch das EPA: **16.03.2005**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **21.11.2007**

(47) Veröffentlichungstag im Patentblatt: **30.10.2008**

(73) Patentinhaber:

Mitsubishi Denki K.K., Tokyo, JP

(84) Benannte Vertragsstaaten:

DE, FR, GB

(74) Vertreter:

**Rummler, F., Dipl.-Ing.Univ., Pat.-Anw., 80802
München**

(72) Erfinder:

**Caball, Jordi, Bramley, Surrey GU5 0DS, GB;
Cieplinski, Leszek, Guildford, Surrey, GU1 2UF, GB**

(54) Bezeichnung: **Verfahren und Vorrichtung zur 3D-Teilbandvideokodierung**

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).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren der Verschlüsselung und Entschlüsselung einer Bildsequenz. Die Erfindung betrifft insbesondere 3D-Subband-Codierung, die zeitliche Filterung, gefolgt von räumlicher Filterung, einschließt.

[0002] Die Dokumente „Three-Dimensional Subband Coding with Motion Compensation“ von Jens-Rainer Ohm und „Motion-Compensated 3-D Subband Coding of Video“ von Choi und Woods sind Referenzen des allgemeinen Standes der Technik und beschreiben 3D Subband-Codierung. Kurz gesagt, wird eine Bildsequenz, wie z. B. eine Bildgruppe (Group of Pictures, GOP), in einer Videosequenz durch bewegungskompensierte (motion compensated, MC) zeitliche Analyse gefolgt von einer räumlichen Wavelet-Transformation in räumlich-zeitliche Subbänder aufgeteilt. Die resultierenden Subbänder (Koeffizienten) werden zur Übertragung weiter verschlüsselt.

[0003] Ein Beispiel eines 3D Subband-Codierungssystems ist weiter unten unter Bezug auf [Fig. 1](#) und [Fig. 2](#) ausführlicher beschrieben.

[0004] [Fig. 1](#) ist ein Blockschaltbild eines bewegungskompensierten eingebetteten Nullblock-Codierers (Motion Compensated Embedded Zeroblock Coder, MC-EZBC), der ein Beispiel eines 3D Subband-Codierers ist. Der Codierer beinhaltet ein Modul für zeitliche Bewegungskompensations-Filterung (Motion Compensation Temporal Filtering, MCTF-Modul) **10**. Der Ausgang des MCTF-Moduls **10** ist mit einem Modul für räumliche Transformation **12** und einem Bewegungsabschätzungs-Modul **18** verbunden. Der Ausgang des Moduls für räumliche Transformation ist mit einem eingebetteten Nullblock-Modul (Embedded Zeroblock, (EZBC-Modul) **14** verbunden. Ein Ausgang des Bewegungsabschätzungs-Moduls **18** wird in das MCTF-Modul **10** eingegeben, und ein anderer Ausgang ist mit einem Bewegungsvektorcodierungs-Modul **20** verbunden. Die Ausgänge des EZBC-Moduls **14** und des Bewegungsvektorcodierungs-Moduls **20** sind beide mit einem Paketierungs-Modul **16** verbunden, das komprimierte, zu sendende Videodaten ausgibt.

[0005] Eine GOP (typischerweise 16 Frames) wird in das MCTF-Modul **10** eingegeben, wo MCTF vorgenommen wird. Zusammen mit dem Bewegungsabschätzungs-Modul **18** wird an den Frames eine Bewegungsabschätzung durchgeführt. In diesem Beispiel ist die Bewegungsabschätzung eine Rückwärts-Bewegungsabschätzung mithilfe hierarchischer Blockanpassung mit variabler Größe (Hierarchical Variable Size Block Matching, HVSBM). Bei diesem Schema erfolgt die Bewegungsabschätzung zuerst mit großen Blöcken (64×64 Pixel). Dann wird

jeder Block in vier Subblöcke unterteilt. Bewegungsvektoren für Subblöcke werden durch Verfeinern des Bewegungsvektors des größeren Blocks erzeugt. Dies wird wiederholt, bis eine Mindest-Blockgröße (4×4 Pixel) erreicht ist. Nach Durchführung des Pruning-Prozesses in Rate-Distortion-Optimierung erhält man ein Raster variabler Blockgrößen, und die resultierenden Bewegungsvektoren werden allen Pixeln in den entsprechenden Blöcken zugeordnet.

[0006] Als Nächstes nimmt das MCTF-Modul **10** zeitliche Filterung mithilfe des Bewegungsvektors vor, der vom Bewegungsabschätzungs-Modul **18** eingegeben wird. In diesem Beispiel ist die zeitliche Filterung mithilfe einer zeitlichen Haar-Wavelet-Transformation (Anhebeschema) implementiert, die eine hierarchische zeitliche Dekomposition erzeugt, wie in [Fig. 2](#) gezeigt.

[0007] Insbesondere wird Bezug nehmend auf [Fig. 2](#) jedes aufeinanderfolgende Paar Frames A und B in der GOP unter Verwendung geeigneter Transformationen behandelt, um einen Tiefpassgefilterten Frame und einen Hochpassgefilterten Frame zu erzeugen, um ein Subband erster Ebene zu erzeugen. Für die GOP erzeugt dies 8 Tiefpassgefilterte (t-L) Frames und 8 Hochpassgefilterte Frames (t-H).

[0008] Die Prozedur (Bewegungsabschätzung und Transformation) wird unter Verwendung der Tiefpassgefilterten Frames wiederholt, um vier Tiefpassgefilterte Frames (t-LL) und vier Hochpassgefilterte Frames (t-LH) zu erzeugen. In ähnlicher Weise wird die Prozedur im unteren zeitlichen Subband rekursiv wiederholt, bis nur ein Frame des unteren zeitlichen Subbandes verbleibt. Wie in [Fig. 2](#) gezeigt, resultiert dies in Vier-Ebenen-Dekomposition mit 1 t-LLLL-Frame, 1 t-LLLH-Frame, 2 t-LLH- und t-LLL-Frames, 4 t-LL- und t-LH-Frames und 8 t-H- und t-L-Frames.

[0009] Der zeitlichen Stufe folgt räumliche Dekomposition im Modul für räumliche Transformation **12**. Insbesondere wird jeder Frame eines zeitlichen Subbandes, der in der Dekomposition hervorgehoben ist, die in [Fig. 2](#) gezeigt ist, räumlicher Filterung und Dekomposition mithilfe einer Wavelet-Transformation unterzogen. Das Dokument „Embedded Image Coding using Zerotrees of Wavelets Coefficients“ von J. M. Shapiro, Dezember 1993, beschreibt die Anwendung von Wavelet-Transformationen auf Bildframes. Kurz gesagt, resultiert räumliche Dekomposition in einer hierarchischen Dekomposition ähnlich der in [Fig. 2](#) für die zeitliche Filterung gezeigten Struktur.

[0010] Insgesamt resultiert die räumlich-zeitliche Filterung in einer hierarchischen Anordnung gefilterter Frames, wobei jeder Frame aus Koeffizienten für jeden Pixelort besteht.

[0011] US 6,519,284 veranschaulicht und beschreibt hierarchische Subband-Verschlüsselung.

[0012] Die 3D Subband-Koeffizienten werden im EZBC-Modul **14** räumlich codiert. Alle räumlich-zeitlichen Subbänder werden getrennt codiert.

[0013] Die Bewegungsvektorfelder, die durch das Bewegungsabschätzungs-Modul **18** ausgegeben werden, werden durch das Bewegungsvektorcodierungs-Modul **20** mithilfe verlustfreier Prädiktion und adaptiver arithmetischer Codierung verschlüsselt.

[0014] Das Paketierungs-Modul kombiniert die verschlüsselten Subband-Koeffizienten und die verschlüsselten Bewegungsvektorfelder. In Bitebenen-Abtastung zur Bildung des endgültigen Ausgangsbitstromes werden die räumlichen Subbänder aller Frames eines zeitlichen Subbandes einer GOP verschachtelt. Der Bewegungsvektor-Anteil des Bitstromes wird getrennt in nicht skalierbarer Weise gespeichert.

[0015] In den oben erwähnten Bewegungsabschätzungs-Schritten wird jedem Pixel des prädizierten/aktuellen Frames ein Bewegungsvektor zugeteilt (bei aufeinanderfolgenden Frames A und B und Rückwärts-Bewegungsabschätzung ist Frame B der prädizierte Frame, und Frame A ist der Referenzframe, wie in [Fig. 3](#) gezeigt). Der Bewegungsvektor kann eine gute Übereinstimmung zwischen dem Pixel in Frame B und dem Pixel in Frame A repräsentieren oder auch nicht, basierend auf der Frameverschiebungsdifferenzen (Displaced Frame Difference, DFD), d. h. der Differenz der Pixelwerte für das Pixel in Frame B und das entsprechende Pixel in Frame A, identifiziert durch den Bewegungsvektor. Zeitliche Filterung mithilfe zwei schlecht übereinstimmender Pixel führt zu Problemen, unter anderem DFD mit Video minderwertiger visueller Qualität und geringerer Framerate und mit hoher Energie.

[0016] Das Hauptproblem bei Bewegungskompensation besteht darin, dass die Beziehung zwischen Frames nicht immer perfekt durch ein Bewegungsmodell beschrieben werden kann. Bei blockbasierten Modellen gilt dies immer dann, wenn die Bewegung nicht stückweise translatorisch ist, wie es zum Beispiel im Falle von Abdeckung, Kamerazoom oder Schwenk usw. auftreten kann. In solchen Gebieten können die Pixel als getrennt (nicht erfasst, Bewegungsvektor ungewiss) oder mehrfach verbunden (erfasst, mehrere Pixel in Frame B können Bewegungsvektoren aufweisen, die auf dasselbe Pixel in Frame A zeigen) klassifiziert werden. Es ist nicht zu erwarten, dass Bewegungskompensation in derartigen abgetrennten Regionen effizient ist.

[0017] Um Bewegungsfilterung zwischen schlecht übereinstimmenden Pixeln zu vermeiden, werden die

DFDs zwischen entsprechenden Pixeln in Frame A und B mit einem vorgegebenen Schwellwert verglichen. Bei einem Pixel, bei dem der zugeteilte Bewegungsvektor in einer DFD oberhalb des Schwellwertes resultiert, wird der Bewegungsvektor als schlecht angesehen, und das Pixel wird als getrennt behandelt (Bewegungsvektor ungewiss). Beispielsweise werden Pixel mit größerem mittleren quadratischen DFD-Fehler (Mean Square Error, MSE) als ihre skalierten Varianzen als getrennt klassifiziert.

[0018] Bei mehrfach verbundenen Pixeln (Pixel in Frame A, die mehrere entsprechende Pixel im prädizierten Frame, Frame B aufweisen), wird der absolute DFD-Wert für jeden Bewegungsvektor berechnet, und der Bewegungsvektor und das zugeordnete Pixel in Frame B mit der minimalen zugeordneten DFD wird ausgewählt. Die anderen Pixel in Frame B, die auf dasselbe Pixel in Frame A zeigen, werden als getrennt behandelt.

[0019] In [Fig. 3](#) sind Pixel b8 und a2 getrennt, Pixel a5 ist mehrfach verbunden, und die übrigen Pixel sind verbunden. Angenommen, die DFD für a5/b4 ist größer als die DFD für a5/b5, dann wird der Bewegungsvektor zwischen a5 und b5 behalten, und b4 wird als getrennt behandelt.

[0020] Das Problem mehrfach verbundener und getrennter Pixel wird in den oben erwähnten Dokumenten von Ohm und Choi und Woods sowie in WO02/085026 und US 6,381,276 diskutiert, welches auch erfasste/nicht erfasste Bildgebiete veranschaulicht.

[0021] Einzelheiten zur oben erwähnten zeitlichen Haar-Wavelet-Transformation sind weiter unten dargestellt.

[0022] Für zwei aufeinanderfolgende Frames A und B ist die Hochpassfilterung gegeben durch

$$H(m, n) = (\sqrt{2}/2) \cdot [B(m, n) - A(m - k, n - l)],$$

wobei A (Referenzframe) und B (aktueller Frame) die ursprünglichen Frames sind und H der Hochpassgefilterte Frame ist, m und n die Pixel indizieren und k und l die Bewegungsvektoren sind.

[0023] Die Tiefpassfilterung ist für verbundene Pixel,

$$L(m, n) = H(m + k, n + l) + \sqrt{2} \cdot A(m, n),$$

und für getrennte (erfasste/nicht erfasste) Pixel ausgelegt

$$L(m, n) = \sqrt{2} \cdot A(m, n),$$

[0024] Im Decodierer wird durch Verwendung von L

und H dieselbe Interpolation bei H durchgeführt, um A für verbundene Pixel exakt zu rekonstruieren als

$$A(m, n) = (1/\sqrt{2}) \cdot [L(m, n) - H(m + k, n + l)],$$

und für getrennte (erfasste/nicht erfasste) Pixel als

$$A(m, n) = (1/\sqrt{2}) \cdot L(m, n)$$

[0025] Nachdem A verfügbar ist, kann dieselbe Interpolation bei A durchgeführt werden, um B exakt zu rekonstruieren als

$$B(m, n) = \sqrt{2} \cdot H(m, n) + A(m - k, n - l)$$

[0026] Um getrennte Blöcke zu verschlüsseln, nimmt der Algorithmus nach Stand der Technik eine Vorwärts-Bewegungsabschätzung mit den segmentierten getrennten Blöcken vor. Wenn dies für einen bestimmten Block in kleinerer DFD resultiert, wird vorwärts-bewegungskompensierte Prädiktion (Motion-Compensated Prediction, Vorwärts-MCP) verwendet. Anderenfalls wird die standardmäßige Rückwärts-MCP verwendet. Dies ergibt drei Codierungsmodi:

- DEFAULT (verbundene Blöcke)
- INTRA (getrennte Blöcke unter Verwendung von Rückwärts-MCP)
- REVERSE (getrennte Blöcke unter Verwendung von Vorwärts-MCP)

[0027] Die drei Fälle werden mithilfe eines dreisymboligen Huffman-Codes 0, 10 bzw. 11 repräsentiert, der als Overheadinformation zusammen mit den Bewegungsvektoren übertragen wird. Da für INTRA- und REVERSE-Blöcke MCP statt MCTF verwendet wird, werden die Blöcke im aktuellen Frame, die ursprünglich als INTRA- und REVERSE-Blöcke gewählt wurden, nicht zeitlich Tiefpass-gefiltert.

[0028] Zur Verschlüsselung von Bewegungsvektoren wird Codierung variabler Länge verwendet, da die Bewegungsvektoren nicht gleichförmig verteilt sind. Die Bewegungsvektoren werden durch adaptive arithmetische Codierung der Vektordifferenzen in Bitstrom umgewandelt.

[0029] Die 3D Subband-Koeffizienten werden mithilfe von EZBC codiert. EZBC ist ein eingebetteter Bildcodierungs-Algorithmus, der Nullblockcodierung und Kontextmodellierung von Subband-/Wavelet-Koeffizienten verwendet. Jeder Frame wird separat räumlich codiert, und der resultierende Bitstrom wird an das Paketierungs-Modul weitergegeben.

[0030] Alternative Techniken (MPEG-2, MPEG-4, H.263, H.264) zur Codierung von Bildsequenzen beinhalten Bewegungsabschätzung/-kompensation und räumliche Transformations-Codierung. Einige Frames (I-Frames) sind intracodiert ohne Verwen-

dung von Bewegungsabschätzung. Andere Frames (B- und P-Frames) schließen Block-Bewegungskompensation und räumliche Verschlüsselung des resultierenden Differenzblocks ein. Block-Intracodierung kann auch in B- und P-Frames vorgenommen werden, in denen kein zweckentsprechender Bewegungsvektor in der Bewegungsabschätzung gefunden wurde. Um die Effizienz der Intracodierung zu steigern, kann räumliche Korrektur zwischen benachbarten Blöcken in einem gegebenen Frame ausgenutzt werden. Insbesondere im MPEG-4-AVC/H.264-Codec kann der interessierende Block aus den umgebenden Blöcken nach deren Richtungsinformationen prädiziert werden. Die Differenz zwischen dem tatsächlichen Block und seiner Prädiktion wird dann codiert, und für diesen Block braucht kein Bewegungsvektor codiert und übertragen zu werden.

[0031] Eines der Probleme von MC-EZBC und Interframe-Wavelet-Codierern im Allgemeinen ist die relativ schlechte Leistung bei niedrigeren Bitraten, die hauptsächlich durch Nicht-Skalierbarkeit von Bewegungsvektoren verursacht ist.

[0032] Auch werden beispielsweise im aktuellen Interframe-Wavelet-Videocodierungs-Ansatz schlechte Übereinstimmungen zwischen benachbarten Frames in derselben Weise zeitlich in die zeitlichen Hochpass-Frames gefiltert wie die guten Übereinstimmungen, wobei bewegungskompensierte Prädiktion verwendet wird, was zu schlechter Qualität jener Frames (oder Blöcke innerhalb von Frames) führt, da bei den schlechten Übereinstimmungen, bei denen Bewegungsabschätzung keine Entsprechungen zwischen benachbarten Frames finden konnte, die Blöcke, die als Prädiktion vom Referenzframe aus verwendet wurden, von den zu prädizierenden Blöcken deutlich verschieden sind.

[0033] US2002/0064231 betrifft ein Videoverschlüsselungsverfahren unter Verwendung einer Wavelet-Dekomposition. Nach Wavelet-Dekomposition hängt der Wert eines Pixels immer noch von jenen von umgebenden Pixeln ab, wie in der ursprünglichen Videosequenz.

[0034] Dies wird beim Senden und nachfolgenden Codieren der Koeffizienten im Rahmen universeller Codierung von Bounded-Memory-Tree-Quellen verwendet.

[0035] Aspekte der Erfindung sind in den beiliegenden Ansprüchen dargelegt.

[0036] Die Erfindung stellt ein Verfahren zur Verschlüsselung und Entschlüsselung einer Framesequenz mithilfe von 3D Subband-Dekomposition bereit, das zeitliche Filterung einschließt, dadurch gekennzeichnet, dass unter gewissen Bedingungen bei

mindestens einem Pixel der Wert des Pixels mithilfe des Wertes eines oder mehrerer anderer Pixels in demselben Frame in der zeitlichen Dekomposition approximiert oder prädiziert wird. Die zeitliche Filterung schließt die Herleitung eines Bewegungsvektorfeldes für einen Frame ein und schließt mindestens zwei Frames in die zeitliche Filterung ein. Der approximierte Pixelwert kann verwendet werden, um das Pixel in entweder einem Hochpass- oder einem Tiefpass-Frame in der (zeitlichen) Subband-Dekomposition zu repräsentieren.

[0037] Vorzugsweise wird das Pixel mithilfe eines oder mehrerer angrenzender Pixel in demselben Frame oder einer Kombination davon approximiert. Bei einem gegebenen Pixel kann ein angrenzendes Pixel ein beliebiges der Pixel sein, die das Pixel unmittelbar umgeben. Alternativ kann bei einem Pixel in einem Block von Pixeln, beispielsweise einem 4×4-Block, wie z. B. einem Block, der zur Bestimmung von Bewegungsvektoren beispielsweise bei Anpassung variabler Blockgröße verwendet wird, oder einem Block oder einer Gruppe von Pixeln, in dem/der alle Pixels in der Gruppe oder dem Block denselben Bewegungsvektor aufweisen, oder einem/einer verbundenen oder segmentierten Block oder Gruppe von Pixeln, ein angrenzendes Pixel als ein Pixel angesehen werden, das an den Block oder die Gruppe von Pixeln angrenzt oder diese umgibt.

[0038] Diese „Intra“-Prädiktion kann beispielsweise für getrennte Pixel oder dann verwendet werden, wenn die DFD, die aus der Prädiktion unter Verwendung desselben Frames resultiert, kleiner ist als unter Verwendung von bewegungskompensierter Prädiktion. Ansonsten kann standardmäßige zeitliche Filterung verwendet werden.

[0039] In dieser Spezifikation wird der Begriff „Frame“ verwendet, um eine Bildeinheit, einschließlich nach Filterung, zu beschreiben, der Begriff gilt aber auch für andere ähnliche Terminologie, wie z. B. Bild, Feld, Abbildung oder Untereinheiten oder Regionen eines Bildes, Frames usw. Die Begriffe Pixel und Blöcke oder Gruppen von Pixeln können ggf. synonym verwendet sein.

[0040] Eine Ausführungsform der Erfindung wird unter Bezug auf die zugehörigen Zeichnungen beschrieben, wobei

[0041] [Fig. 1](#) ein Blockschaltbild eines Verschlüsselungssystems ist;

[0042] [Fig. 2](#) eine Grafik ist, die zeitliche Dekomposition einer GOP darstellt;

[0043] [Fig. 3](#) eine Repräsentation eines Framepaares und verbundener und getrennter Pixel ist;

[0044] [Fig. 4](#) ein Diagramm einer Anordnung von Pixeln in einem Frame ist;

[0045] [Fig. 5](#) eine Grafik ist, die Richtungen zur Auswahl von Prädiktionspixeln darstellt.

[0046] Die Erfindung kann mithilfe eines Systems ähnlich dem oben beschriebenen System nach Stand der Technik implementiert werden, das, wie unten beschrieben, Gegenstand von Modifikationen ist.

[0047] Dementsprechend sind die grundlegenden Komponenten des Codierungssystems gemäß der vorliegenden Ausführungsform so, wie in [Fig. 1](#) gezeigt, und arbeiten so, wie oben in Bezug auf den Stand der Technik beschrieben, mit der Ausnahme, dass das MCTF-Modul in Bezug auf die Verarbeitung getrennter Pixel oder Blöcke von Pixeln modifiziert ist.

[0048] Wie oben erwähnt, erfolgt die Bewegungsabschätzung bis hinunter zu einer Blockgröße von 4×4 Pixeln, und jedem Pixel im Block wird derselbe Bewegungsvektor zugewiesen. Ein derartiger 4×4-Pixel-Block ist in [Fig. 4](#) dargestellt, wobei $c(m, n)$ die Pixel des Blocks unter Betrachtung repräsentiert. Angrenzende Pixel sind durch $t(p, q)$ für Pixel in der Zeile oberhalb der obersten Zeile des 4×4-Blocks und durch $1(r, s)$ für Pixel in der Spalte links der linken Spalte des 4×4-Blocks repräsentiert.

[0049] Nach Bestimmung der Bewegungsvektoren werden die Bewegungsvektoren und die zugeordneten DFDs wie beim Stand der Technik verarbeitet, um zu bestimmen, welche Pixel als getrennt (wie oben definiert) behandelt werden müssen.

[0050] Bei jedwedem Pixel im Block, das getrennt ist, wird das Pixel zur weiteren Verarbeitung mithilfe eines nahe gelegenen Pixels in demselben Frame approximiert (oder prädiziert), im Folgenden als Intra-Prädiktion beschrieben. Dies kann in einer Reihe von Weisen implementiert sein.

[0051] In diesem Beispiel gibt es drei Intra-Modi: vertikaler Intra-Prädiktionsmodus, horizontaler Intra-Prädiktionsmodus und gemittelter horizontal-vertikaler Intra-Prädiktionsmodus.

Vertikaler Intra-Prädiktionsmodus:

- $c(0, 0)$, $c(0, 1)$, $c(0, 2)$ und $c(0, 3)$ werden durch $t(0, -1)$ prädiziert
- $c(1, 0)$, $c(1, 1)$, $c(1, 2)$ und $c(1, 3)$ werden durch $t(1, -1)$ prädiziert
- $c(2, 0)$, $c(2, 1)$, $c(2, 2)$ und $c(2, 3)$ werden durch $t(2, -1)$ prädiziert
- $c(3, 0)$, $c(3, 1)$, $c(3, 2)$ und $c(3, 3)$ werden durch $t(3, -1)$ prädiziert

Horizontaler Intra-Prädiktionsmodus:

- $c(0, 0)$, $c(1, 0)$, $c(2, 0)$ und $c(3, 0)$ werden durch $l(-1, 0)$ prädiziert
- $c(0, 1)$, $c(1, 1)$, $c(2, 1)$ und $c(3, 1)$ werden durch $l(-1, 1)$ prädiziert
- $c(0, 2)$, $c(1, 2)$, $c(2, 2)$ und $c(3, 2)$ werden durch $l(-1, 2)$ prädiziert
- $c(0, 3)$, $c(1, 3)$, $c(2, 3)$ und $c(3, 3)$ werden durch $l(-1, 3)$ prädiziert

Gemittelter horizontal-vertikaler Intra-Prädiktionsmodus:

- $c(0, 0)$ wird durch $(t(0, -1) + l(-1, 0))/2$ prädiziert
- $c(1, 0)$ wird durch $(t(1, -1) + l(-1, 0))/2$ prädiziert
- $c(2, 0)$ wird durch $(t(2, -1) + l(-1, 0))/2$ prädiziert
- $c(3, 0)$ wird durch $(t(3, -1) + l(-1, 0))/2$ prädiziert
- $c(0, 1)$ wird durch $(t(0, -1) + l(-1, 1))/2$ prädiziert
- $c(1, 1)$ wird durch $(t(1, -1) + l(-1, 1))/2$ prädiziert
- $c(2, 1)$ wird durch $(t(2, -1) + l(-1, 1))/2$ prädiziert
- $c(3, 1)$ wird durch $(t(3, -1) + l(-1, 1))/2$ prädiziert
- $c(0, 2)$ wird durch $(t(0, -1) + l(-1, 2))/2$ prädiziert
- $c(1, 2)$ wird durch $(t(1, -1) + l(-1, 2))/2$ prädiziert
- $c(2, 2)$ wird durch $(t(2, -1) + l(-1, 2))/2$ prädiziert
- $c(3, 2)$ wird durch $(t(3, -1) + l(-1, 2))/2$ prädiziert
- $c(0, 3)$ wird durch $(t(0, -1) + l(-1, 3))/2$ prädiziert
- $c(1, 3)$ wird durch $(t(1, -1) + l(-1, 3))/2$ prädiziert
- $c(2, 3)$ wird durch $(t(2, -1) + l(-1, 3))/2$ prädiziert
- $c(3, 3)$ wird durch $(t(3, -1) + l(-1, 3))/2$ prädiziert

[0052] Das System kann für den Prädiktionsmodus eingestellt werden, der bevorzugt wird. In jenem Fall kann, wenn der Prädiktionswert für den verwendeten Prädiktionsmodus nicht verfügbar ist (beispielsweise, wenn im vertikalen Intra-Prädiktionsmodus der relevante Wert $t(p, q)$ nicht verfügbar ist) das System dann eingestellt werden, einen alternativen Prädiktionsmodus (beispielsweise den horizontalen Prädiktionsmodus) zu verwenden. Wenn der Prädiktionswert für den anderen Prädiktionsmodus nicht verfügbar ist, wird der Prädiktionswert 128 verwendet.

[0053] Wenn ein Intra-Prädiktionsmodus verwendet wird, wird die bewegungskompensierte zeitliche Filterung (MCTF) in der Anhebe-Implementierung für die entsprechenden Blöcke in zwei aufeinanderfolgenden Frames B (aktuell) und A (Referenz) geringfügig modifiziert. Die Hochpassfilterung H ist für getrennte Pixel ausgelegt durch

$$H(m, n) = (\sqrt{2}/2) \cdot [B(m, n) - B'(m, n)],$$

wobei $B'(m, n)$ die Prädiktionspixel im aktuellen Frame unter Verwendung von Intra-Prädiktion, wie oben erwähnt, sind. Die Hochpassfilterung für verbundene Pixel und die Tiefpassfilterung bleiben dieselben.

[0054] Dann wird der aktuelle Frame für getrennte Pixel rekonstruiert durch

$$B(m, n) = \sqrt{2} \cdot H(m, n) + B'(m, n),$$

wobei $B'(m, n)$ die vorherigen rekonstruierten Pixel sind. Die übrigen Rekonstruktionsgleichungen bleiben unverändert.

[0055] Durch Verwendung dieses Intra-Prädiktions-Ansatzes werden die zeitlichen Hochpass-Frames verbessert, da die verwendeten Prädiktionswerte dichter an den zu prädizierenden Pixelwerten sein können als bei Verwendung des Bewegungsabschätzungs-Prozesses. Um die besten Ergebnisse zu erzielen, werden die Intra-Prädiktionsmodi mit der MC-Prädiktion basierend auf dem Bewegungsabschätzungs-Prozess verglichen, und es wird der Modus (Intra-Prädiktion oder MC-Prädiktion) ausgewählt, der zum niedrigsten durchschnittlichen Verzerrungswert führt.

[0056] Insbesondere werden die Werte der mittleren absoluten Differenz (MAD) für alle Intra-Prädiktionsmodi (vertikal, horizontal und horizontal-vertikal) und für die MC-Prädiktionsmodi (rückwärts und vorwärts) berechnet. Der Modus mit dem niedrigsten MAD-Wert wird als bester Prädiktionsmodus ausgewählt.

[0057] In dieser Ausführungsform werden dieselben drei Intra-Prädiktionsmodi für alle Komponenten verwendet, d. h. für Luminanz (luma, Y) und Chrominanz (chroma, Cb und Cr) plus dem ursprünglichen MCP-Modus, abhängig vom Verzerrungslevel, den jeder Modus mit sich bringt.

[0058] Durch Verwendung eines variablen Gewichtungsparameters kann der Anteil an Pixeln/Blöcken variiert werden, die unterschiedliche Modi verwenden. Beispielsweise kann Setzen der Gewichtungsparameter auf 0 bedeuten, dass alle getrennten Pixel/Blöcke mithilfe von Pixeln aus demselben Frame als Prädiktionswerte prädiziert werden, und Erhöhen des Gewichtungsparameters erhöht den Anteil an Pixeln/Blöcken, die mithilfe von Bewegungskompensations-Prädiktion prädiziert werden.

[0059] Die obigen Intra-Prädiktionsmodi verwenden nur vertikale, horizontale und gemittelte Prädiktion. Jedoch könnten verschiedene andere Intra-Prädiktionsmodi verwendet werden. Beispielsweise können alle Modi verwendet werden, die für den MPEG-4-AVC-Codec definiert sind. Dies sind neun Intra-Prädiktionsmodi, die als vertikale (Modus 0), horizontale (Modus 1), DC-(Modus 2), abwärts-links-diagonale (Modus 3), abwärts-rechts-diagonale (Modus 4), rechts-vertikale (Modus 5), abwärts-horizontale (Modus 6), links-vertikale (Modus 7) bzw. aufwärts-horizontale (Modus 8) Prädiktionsmodi bezeichnet werden. Auch gibt es vier unterschiedliche Intra-Prädiktionsmodi, die für unterschiedliche Blockgrößen und unterschiedliche Farb-

komponenten angewendet werden können.

[0060] Ein vollständiger Satz von Prädiktionsmodus-Richtungen, wie in [Fig. 5](#) gezeigt, ist unten beschrieben,

- Modus 0: Vertikal Aufwärts
- Modus 1: Vertikal Aufwärts Rechts
- Modus 2: Diagonal Aufwärts Rechts
- Modus 3: Horizontal Aufwärts Rechts
- Modus 4: Horizontal Rechts
- Modus 5: Horizontal Abwärts Rechts
- Modus 6: Diagonal Abwärts Rechts
- Modus 7: Vertikal Abwärts Rechts
- Modus 8: Vertikal Abwärts
- Modus 9: Vertikal Abwärts Links
- Modus 10: Diagonal Abwärts Links
- Modus 11: Horizontal Abwärts Links
- Modus 12: Horizontal Links
- Modus 13: Horizontal Aufwärts Links
- Modus 14: Diagonal Aufwärts Links
- Modus 15: Vertikal Aufwärts Links
- Modus 16: DC

[0061] Bei der Bewegungsabschätzung basiert die Auswahl der Blockgröße auf Rate-Distortion-Optimierung. Gegenwärtig erfolgt der Intra-Prädiktionsprozess nach der Bewegungsabschätzung und ist daher nicht Rate-Distortion-optimiert. Die Intra-Prädiktionsmodus-Auswahl könnte in die Rate-Distortion-Optimierung einbezogen werden.

[0062] Ein ähnlicher Intra-Prädiktionsprozess könnte für die Tiefpass-gefilterten Frames eingeführt werden.

[0063] Andere Arten von 3D Dekomposition und Transformationen können verwendet werden. Beispielsweise könnte die Erfindung in einem Dekompositionsschema angewendet werden, in dem zuerst räumliche Filterung und danach zeitliche Filterung vorgenommen wird.

[0064] Die Erfindung kann beispielsweise in einem computerbasierten System oder unter Verwendung geeigneter Hardware und/oder Software implementiert werden. Ein Codierer ist in [Fig. 1](#) gezeigt, und ein entsprechender Decodierer weist entsprechende Komponenten zur Durchführung der inversen Decodierungsoperationen auf.

Patentansprüche

1. Verfahren zur Verschlüsselung einer Framesequenz (A, B) mithilfe von 3D Subband-Dekomposition, wobei das Verfahren Folgendes umfasst: Durchführung von bewegungskompensierter zeitlicher Filterung und räumlicher Filterung der Framesequenz (A, B) zur Erzeugung von 3D Subband-Koeffizienten; und Verschlüsselung der 3D Subband-Koeffizienten für

eine Übertragung,

wobei die zeitliche Filterung Folgendes umfasst:

Bestimmung von Bewegungsvektoren;

Ermittlung getrennter Pixel; und

dadurch gekennzeichnet, dass für mindestens ein getrenntes Pixel $(c(m, n))$ ein Näherungswert basierend auf einem oder mehreren Pixelwerten $(t(p, q), l(r, s))$ in demselben Frame wie das getrennte Pixel berechnet wird.

2. Verfahren nach Anspruch 1, wobei das getrennte Pixel $(c(m, n))$ basierend auf einem oder mehreren angrenzenden Pixelwerten $(t(p, q), r(r, s))$ in demselben Frame wie das getrennte Pixel angenähert wird.

3. Verfahren nach Anspruch 2, wobei vertikal und horizontal angrenzende Pixel in demselben Frame wie das getrennte Pixel als Näherungswerte verwendet werden.

4. Verfahren nach Anspruch 2 oder 3 umfassend: Berechnung mehrerer Näherungswerte für ein getrenntes Pixel, wobei jeder Näherungswert basierend auf einem anderen Satz aus einem oder mehreren Pixeln in demselben Frame wie das getrennte Pixel berechnet wird; Vergleich des Unterschieds zwischen jedem Näherungswert und dem tatsächlichen Wert des getrennten Pixels; und Auswahl des Näherungswerts, der dem tatsächlichen Wert des getrennten Pixels am nächsten kommt.

5. Verfahren nach einem der vorherigen Ansprüche ferner umfassend die Approximation des getrennten Pixelwerts mithilfe von Bewegungskompensation und Werten in anderen Frames als das getrennte Pixel, den Vergleich der Ergebnisse für die Approximationen mithilfe von Pixeln in anderen Frames als das getrennte Pixel, wobei die Approximationen auf Pixeln in demselben Frame wie das getrennte Pixel basieren, und die Auswahl der besten Approximation.

6. Verfahren nach einem der vorherigen Ansprüche, wobei die zeitliche Filterung die Verwendung des angenäherten Pixelwerts in einer Transformationsgleichung einschließt.

7. Verfahren zur Entschlüsselung einer Sequenz von 3D Subband aufgeteilten Frames, die mithilfe eines Verfahrens nach einem der vorherigen Ansprüche verschlüsselt wurden, wobei das Verfahren die Rekonstruktion eines getrennten Pixelwerts in einem Frame basierend auf dem Wert mindestens eines anderen Pixels in demselben Frame wie das getrennte Pixel umfasst.

8. Verschlüsselungs- und/oder Entschlüsselungsvorrichtung, die für die Implementierung des

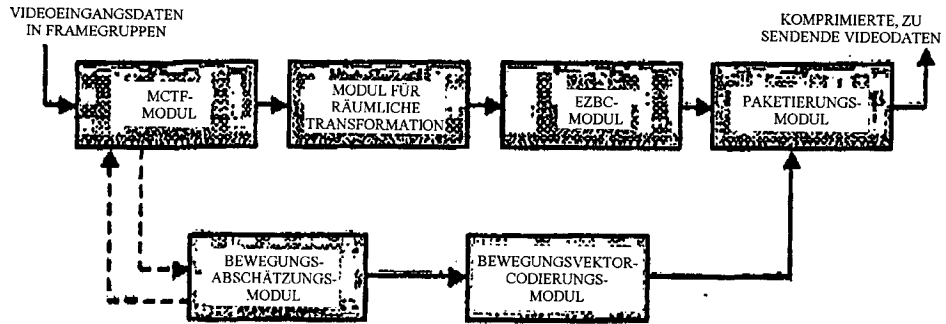
Verfahrens nach einem der vorherigen Ansprüche geeignet ist.

9. Computerprogramm für die Ausführung des Verfahrens nach einem der Ansprüche 1 bis 7.

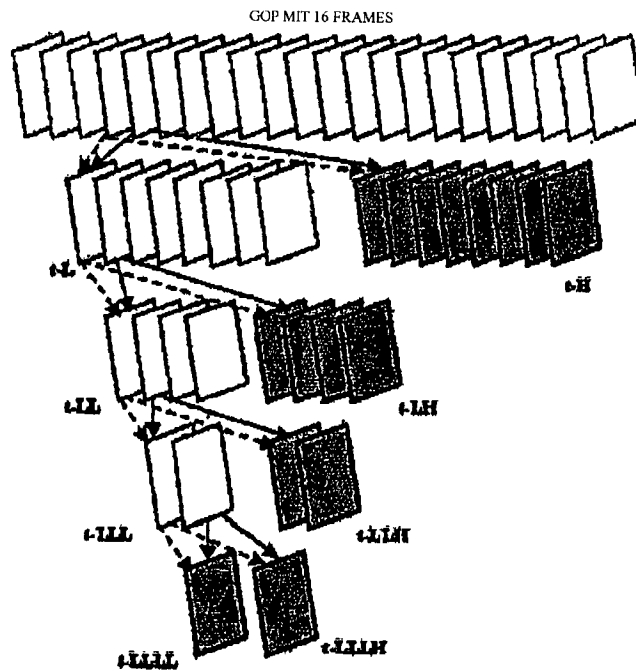
10. Computerprogramm nach Anspruch 9, das auf einem Computer-lesbaren Medium enthalten ist.

Es folgen 3 Blatt Zeichnungen

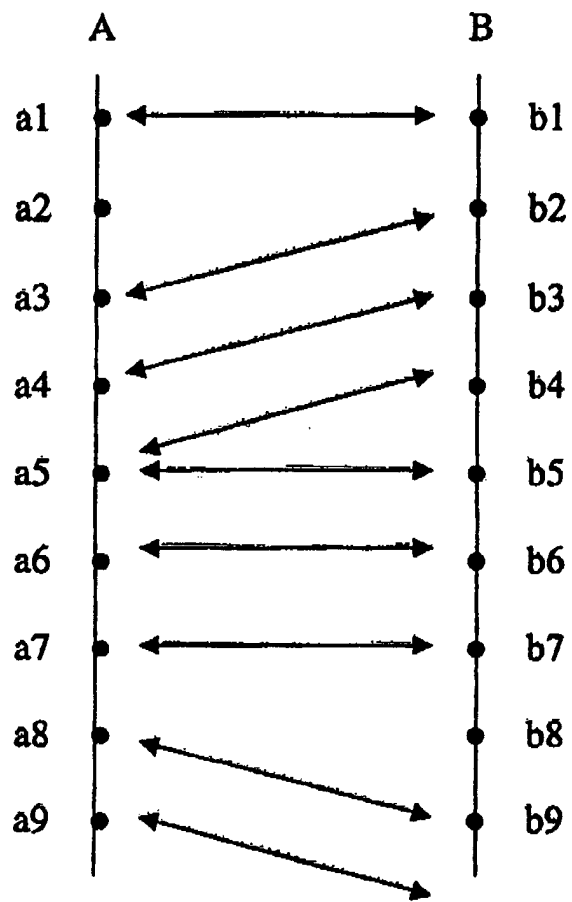
Anhängende Zeichnungen



Figur 1



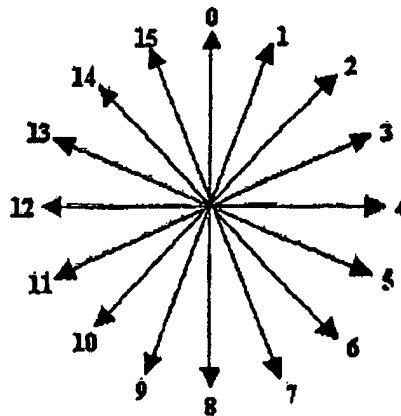
Figur 2



Figur 3

	$\alpha(0,1)$	$\alpha(1,1)$	$\alpha(2,1)$	$\alpha(3,1)$
$\beta(-1,0)$	$\alpha(0,0)$	$\alpha(1,0)$	$\alpha(2,0)$	$\alpha(3,0)$
$\beta(-1,1)$	$\alpha(0,1)$	$\alpha(1,1)$	$\alpha(2,1)$	$\alpha(3,1)$
$\beta(-1,2)$	$\alpha(0,2)$	$\alpha(1,2)$	$\alpha(2,2)$	$\alpha(3,2)$
$\beta(-1,3)$	$\alpha(0,3)$	$\alpha(1,3)$	$\alpha(2,3)$	$\alpha(3,3)$

Figur 4



Figur 5