



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 698 36 473 T2** 2007.09.20

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 0 940 041 B1**

(21) Deutsches Aktenzeichen: **698 36 473.2**

(86) PCT-Aktenzeichen: **PCT/IB98/01241**

(96) Europäisches Aktenzeichen: **98 935 253.9**

(87) PCT-Veröffentlichungs-Nr.: **WO 1999/016251**

(86) PCT-Anmeldetag: **12.08.1998**

(87) Veröffentlichungstag

der PCT-Anmeldung: **01.04.1999**

(97) Erstveröffentlichung durch das EPA: **08.09.1999**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **22.11.2006**

(47) Veröffentlichungstag im Patentblatt: **20.09.2007**

(51) Int Cl.<sup>8</sup>: **H04N 7/36** (2006.01)

**H04N 7/26** (2006.01)

**H04N 5/14** (2006.01)

(30) Unionspriorität:

**97202912 23.09.1997 EP**

(84) Benannte Vertragsstaaten:

**DE, FR, GB, IT**

(73) Patentinhaber:

**Koninklijke Philips Electronics N.V., Eindhoven,  
NL**

(72) Erfinder:

**DE HAAN, Gerard, NL-5656 AA Eindhoven, NL;  
SCHUTTEN, Jan, Robert, NL-5656 AA Eindhoven,  
NL; PELAGOTTI, Anna, NL-5656 AA Eindhoven,  
NL**

(74) Vertreter:

**Volmer, G., Dipl.-Ing., Pat.-Anw., 52066 Aachen**

(54) Bezeichnung: **BEWEGUNGSSCHÄTZUNG UND BEWEGUNGSKOMPENSIERTE INTERPOLATION**

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 vorliegende Erfindung bezieht sich auf ein Verfahren und eine Anordnung zur Bewegungsschätzung, und auf eine Videowiedergabeanordnung mit einer bewegungskompensierten Interpolationsanordnung.

**[0002]** Der Artikel: "Spatio-temporal segmentation of Video data" von John Wang und Edward H. Adelson, in "Proceedings of SPIE", Heft 2182, "Image and Video Processing II", 1994, Seiten 120–131, [37] beschreibt, dass Bildsegmentierung eine kräftige semantische Beschreibung von Videobildern, wesentlich bei dem Verstehen von Bildern und effizienter Manipulation von Bilddaten. Insbesondere definiert Segmentierung auf Basis von Bildbewegung Gebiete, die eine gleiche Bewegung erfahren wobei ein Bildcodierungssystem erlaubt ist um Videosequenzen effizienter darzustellen. Der Artikel beschreibt ein allgemeines iteratives Netzwerk zur Segmentierung von Videodaten. Aufgabe unserer räumlichzeitlichen Segmentierung ist, eine geschichtete Bilddarstellung des Videos für Bildcodierungsapplikationen, wobei Videodaten auf einfache Weise als ein Satz sich verlagernder Schichten beschrieben werden.

**[0003]** Der Artikel: "A Unified Mixture Framework for Motion Segmentation: Incorporating Spatial Coherence and Estimating the Number of Models" von Yair Weiss und Edward H. Adelson, in "1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'96)" Seiten 321–326, beschreibt, dass Beschreibung einer Videosequenz in Termen einer geringen Anzahl kohärent sich verlagernder Segmente nützlich ist für Aufgaben, die von Videokompression bis Ereignisperzeption reichen. Eine hoffnungsvolle Annäherung ist, das Bewegungssegmentierungsproblem in einem Mischungsschätzungsrahmen zu sehen. Aber bestehende Formulierungen benutzen im Allgemeinen nur die Bewegungsdaten und benutzen nicht die statischen Stichwörter beim Segmentieren der Sequenz. Weiterhin ist die Anzahl Modelle entweder im Voraus spezifiziert oder außerhalb des Mischungsmodellrahmens geschätzt. Der Artikel zeigt, wie räumliche Beschränkungen den Mischformulieren hinzugefügt werden können und bieten eine Variante des EM Algorithmus, der die Form sowie die Bewegungsbeschränkungen benutzt.

**[0004]** Der Artikel: "Representing Moving Images with Layers" von Wang, J.Y.A. und Adelson, E.H., in "IEEE Transactions on Image Processing", "Special Issue: Image Sequence Compression" Heft 3, Nr. 5, September 1994, Seiten 625–638, beschreibt ein System zum Darstellen sich verlagernder Bilder mit Sätzen sich überlappender Schichten. Jede Schicht enthält eine Intensitätsmappe, die die additiven Werte jedes Pixels definiert, zusammen mit einer Alphamappe, die wie eine Maske wirksam ist, welche die Transparenz angibt. Die Schichten sind in der Tiefe geordnet und sie verdecken einander entsprechend den Regeln der Zusammensetzung. Geschwindigkeitsmappen definieren, wie die Schichten in der Zeit gefaltet werden müssen. Die geschichtete Darstellung ist flexibler als Standardbildtransformationen und kann viele wichtige Eigenschaften von natürlichen Bildsequenzen erfassen. Der Artikel beschreibt einige Verfahren zum Zerlegen von Bildsequenzen in Schichten unter Anwendung von Bewegungsanalyse und beschreibt, wie die Darstellung für Bildcodierungs- und andere Applikationen verwendet werden kann.

**[0005]** Bewegungsvektoren werden in einem Bereich von Applikationen, wie Codierung, Rauschunterdrückung und Abtastratenumwandlung verwendet. Einige dieser Applikationen, insbesondere die Frameratenumwandlung, erfordern, dass die wirkliche Bewegung von Objekten geschätzt wird [10, 11]. Andere Applikationen, beispielsweise verschachtelt-zu-sequentielle Abtastumwandlung, erfordern eine hohe Genauigkeit der Bewegungsvektoren zum Erreichen einer niedrigen Amplitude restlicher Alias [12, 13]. Zum Schluss gibt es eine Kategorie von Applikationen, beispielsweise Konsumentenapplikationen von Bewegungsschätzung, wobei die Kosten des Bewegungsschätzers von wesentlicher Bedeutung sind [14, 15]. Es wurden bereits viele Algorithmen vorgeschlagen zum Erzielen einer wirklichen Bewegungsschätzung [3, 10, 11, 15–17]. Es wurden ebenfalls Algorithmen vorgeschlagen zum Verwirklichen einer Bewegungsschätzung mit einem niedrigen Komplexitätspegel, beispielsweise [3, 14, 15, 18–20] und zusätzlich zu den pixelrekursiven Algorithmen, die meistens eine Subpixelgenauigkeit ermöglichen, siehe beispielsweise [21, 22] wobei eine Anzahl Blockübereinstimmungsalgorithmen genannt wurden, die äußerst genaue Bewegungsvektoren ergeben [10, 23, 24].

**[0006]** Voreinigigen Jahren wurde eine rekursiver Suchblockanpassungsanordnung vorgeschlagen, die wirkliche Bewegungsschätzung, wie diese für Frameratenumwandlung erforderlich ist, mit der niedrigen Komplexitätsbeschränkung kombiniert wird, die für Konsumentenapplikationen erforderlich ist [3]. Dieser Entwurf wurde in einer IC von Philips (MELZONIC, SAA4991) kommerzialisiert [6,25], wobei Bewegungsschätzungs- und Kompensationstechniken angewandt werden, um die Bewegungsdarstellung von Filmmaterial, wenn im Fernsehen vorgeführt, zu verbessern, und um Unschärfe von Bilddetails zu eliminieren in dem Fall, dies auftritt, wenn Sequenzen mit einer Bilderneuerungsrate wiedergegeben werden, die von der Übertragungsrate abweicht. Die meist herausfordernde Aufgabe einer derartigen Verarbeitung ist die Schätzung von Bewegungs-

vektoren, die angeben, ob an einer bestimmten Stelle des Schirms, Objekte sich ggf. verlagern, und sollte dies der Fall sein, wie schnell und in welcher Richtung. Bei der bekannten IC wird diese Aufgabe durch eine sog. Blockanpassungsschaltung durchgeführt, die das Bild in Blöcke aufteilt und einen Bewegungsvektor für jeden Block mit Pixeln berechnet durch Minimierung eines Anpassungskriteriums. Die Gefahr einer derartigen Verarbeitung ist, dass das bewegungskompensierte Bild, interpoliert aus benachbarten Bildern die Bewegungsvektoren verwendend, Blockstörungen aufweisen kann, wenn das Bewegungsvektorfeld an unerwünschten Inhomogenitäten leidet. Zum Reduzieren dieser Gefahr auf einen akzeptierbaren Pegel benutzt die IC in [6] eine Blockanpassungsschaltung mit einer verbesserten Konsistenz auf Basis einer räumlichen und zeitlichen Prädiktion von Kandidatvektoren [3]. Ein vorteilhafter Nebeneffekt dieser Annäherung der Bewegungsschätzung ist die sehr signifikante Reduktion der Verarbeitungsleistung, erforderlich für die Funktion, was insbesondere dem sehr beschränkten Kandidatvektorzählwert zuzuschreiben ist.

**[0007]** Der Artikel: "Layered representation for motion analysis" von J.Y.A. Wang und E.A. Adelson in "Proceedings of the 1993 IEEE Computer Society conference on Computer vision and pattern recognition", Seiten 361–366 [29] beschreibt einen Satz von Techniken zur Segmentierung von Bildern in kohärent sich verlagernde Gebiete, wobei ein affine Bewegungsanalysen- und Clustertechniken angewandt werden. Ein Bild wird in einen Satz von Schichten zerlegt, zusammen mit Information über Verstopfung und Tiefenordnung. Eine Szene wird in vier Schichten analysiert und danach wird eine Sequenz mit einem einzigen Bild jeder Schicht dargestellt, zusammen mit assoziierten Bewegungsparametern.

**[0008]** Es ist u. a. eine Aufgabe der vorliegenden Erfindung, einen Bewegungsschätzer mit einer weiter reduzierten Komplexität zu schaffen, Dazu schafft ein erste Aspekt der vorliegenden Erfindung ein Verfahren und eine Anordnung, wie in den Ansprüchen 1 und 9 definiert. Ein zweiter Aspekt der vorliegenden Erfindung schafft ein Verfahren und eine Anordnung, wie in den Ansprüchen 7 und 10 definiert. Ein dritter Aspekt der vorliegenden Erfindung schafft eine Videowiedergabeanordnung, wie in Anspruch 11 definiert. Vorteilhafte Ausführungsformen sind in den Unteransprüchen definiert.

**[0009]** In einem Verfahren zum Schätzen von Bewegung nach einem Hauptaspekt der vorliegenden Erfindung werden wenigstens zwei Bewegungsparametersätze aus den Eingangsvideodaten erzeugt, wobei ein Bewegungsparametersatz ein Satz mit Parametern ist, die Bewegung in einem Bild beschreiben, wobei mit Hilfe dieses Bewegungsparametersatzes Bewegungsvektoren berechnet werden können. Ein Bewegungsparametersatz gibt eine Nullgeschwindigkeit für alle Bildteile in einem Bild und jeder Bewegungsparametersatz hat entsprechende örtliche Übereinstimmungsfehler, wie Übereinstimmungsfehler, ermittelt je Pixelblock. Ausgangsbewegungsdaten werden aus den Eingangsvideodaten ermittelt, und zwar in Abhängigkeit von wenigstens zwei Bewegungsparametersätzen, wobei die Bedeutung jedes Bewegungsparametersatzes (ermittelt durch Gewichtungsfaktoren  $W$ , siehe die Gleichungen 17, 18 und zwischen den Gleichungen 20, 21) beim Berechnen der Ausgangsbewegungsdaten von den örtlichen Anpassungsfehlern der Bewegungsparametersätze abhängig ist. Örtliche Anpassungsfehlern sollen im Gegensatz zu globalen Anpassungsfehlern, wie Anpassungsfehler, berechnet für das ganze Bild verstanden werden.

**[0010]** Bei einem Verfahren zur Bewegungskompensation von Videodaten nach einem anderen Aspekt der vorliegenden Erfindung werden wenigstens zwei Bewegungsparametersätze aus den Eingangsvideodaten erzeugt, wobei ein Bewegungsparametersatz eine Nullgeschwindigkeit angibt, und jeder Bewegungsparametersatz entsprechende Anpassungsfehler hat, und Ausgangsvideodaten aus den Eingangsvideodaten in Abhängigkeit von den wenigstens zwei Bewegungsparametersätzen interpoliert werden, wobei die Bedeutung jedes Bewegungsparametersatzes beim Berechnen der Ausgangsvideodaten von den Anpassungsfehlern der Bewegungsparametersätze abhängig ist.

**[0011]** Bei dem Verfahren der Bewegungskompensation von Videodaten nach einem anderen Aspekt der vorliegenden Erfindung werden aus Eingangsvideodaten wenigstens zwei Bewegungsparametersätze erzeugt, wobei der eine Bewegungsparametersatz eine Nullgeschwindigkeit angibt, und wobei jeder Bewegungsparametersatz entsprechende Übereinstimmungsfehler hat, und Ausgangsdaten von den Eingangsdaten interpoliert werden, und zwar in Abhängigkeit von den wenigstens zwei Bewegungsparametersätzen, wobei die Bedeutung jedes Bewegungsparametersatzes beim Berechnen der Ausgangsvideodaten von den Übereinstimmungsfehlern der Bewegungsparametersätze abhängig ist.

**[0012]** In einer Ausführungsform ist die Reduktion derart signifikant, dass die Verarbeitung an einer völlig programmierbaren Anordnung durchgeführt werden kann, insbesondere dem Philips TriMedia Prozessor.

**[0013]** Ausführungsbeispiele der vorliegenden Erfindung sind in der Zeichnung dargestellt und werden im Fol-

genden näher beschrieben. Es zeigen:

[0014] [Fig. 1](#) zwei Möglichkeiten zur Bewegungskompensation nach der vorliegenden Erfindung,

[0015] [Fig. 2](#) eine erste Ausführungsform eines bewegungskompensierten Interpolators nach der vorliegenden Erfindung,

[0016] [Fig. 3](#) eine zweite Ausführungsform eines bewegungskompensierten Interpolators nach der vorliegenden Erfindung,

[0017] [Fig. 4](#) ein Blockschaltbild eines Bewegungsparameterschätzers nach der vorliegenden Erfindung und

[0018] [Fig. 5](#) einen bevorzugten Parameterschätzer nach der vorliegenden Erfindung.

[0019] In [5] wurde ein Verfahren beschrieben, wie eine globale Bewegung von einer Bildsequenz geschätzt werden soll. Es wird vorausgesetzt, dass Bewegung in dem Bild mit einer zweidimensionalen linearen Gleichung erster Ordnung beschrieben werden kann, und zwar unter Anwendung von  $\bar{D}(\bar{x}, n)$  für den Verlagerungsvektor an der Stelle  $\bar{x}$  in dem Bild mit dem Index  $n$ :

$$\bar{D}(\bar{x}, n) = \begin{Bmatrix} p_1(n) + p_3(n)x \\ p_2(n) + p_3(n)y \end{Bmatrix} \quad (1)$$

[0020] Es ist erkannt, dass wenn wir nur globale Bewegungsvektoren schätzen möchten, die Eingabe zu den Parameterberechnungsmitteln einfacher sein kann als was in [5] beschrieben worden ist.

[0021] Wenn nur derartige globale Bewegungsvektoren verfügbar sind, wird das Problem der Aufwärtsmischung der Teil der Verarbeitung mit der höchsten Herausforderung. [1, 4] beschreibt ein Verfahren zur robusten bewegungskompensierten zeitlichen Interpolation von Bilddaten. Der Grundgedanke bestand aus einem Mediafilter mit drei Abgriffen, das ein Ausgangspixel erzeugt, selektiert um entweder das bewegungskompensierte Pixel  $mcl(left)$  aus dem vorhergehenden Teilbild  $n-1$  zu sein, das bewegungskompensierte Pixel  $mcr(right)$  aus dem nächsten Teilbild  $n$ , oder der nicht-bewegungskompensierte Mittelwert  $av$  von den beiden benachbarten Teilbildern  $n-1, n$ :

$$F_i(\bar{x}, n - \frac{1}{2}) = med(mcl, av, mcr) \quad (2)$$

Mit

$$\begin{aligned} mcl &= F(\bar{x} - \alpha \bar{D}(\bar{x}, n), n-1) \\ av &= \frac{1}{2} ( F(\bar{x}, n) + F(\bar{x}, n-1) ) \\ mcr &= F(\bar{x} + (1-\alpha) \bar{D}(\bar{x}, n), n) \\ med(a, b, c) &= \begin{cases} a, & (b \leq a \leq c \vee c \leq a \leq b) \\ b, & (a \leq b \leq c \vee c \leq b \leq a) \\ c, & (otherwise) \end{cases} \end{aligned} \quad (3)$$

[0022] Die bei der Bewegungskompensation verwendeten Pixel sind in [Fig. 1](#) schematisch dargestellt. Obwohl ziemlich robust, könnte ein noch robusterer Algorithmus betrachtet werden für unseren sehr begrenzten Bewegungsschätzervorschlag, der ein Mediafilter mit drei Abgriffen umfasst, das ein Ausgangspixel erzeugt, durch Selektion entweder des entsprechenden Pixels  $l(left)$  in dem vorhergehenden Teilbild  $n-1$ , des entsprechenden Pixels  $r(right)$  in dem nächsten Teilbild  $n$ , oder der bewegungskompensierte Mittelwert  $mav$  von den beiden benachbarten Teilbildern  $n-1, n$ :

$$F_i(\bar{x}, n - \frac{1}{2}) = \text{med}(l, mcav, r) \quad (7)$$

Mit

$$l = F(\bar{x}, n-1) \quad (8)$$

$$mcav = \frac{1}{2} ( F(\bar{x} - \alpha \bar{D}(\bar{x}, n), n-1) + F(\bar{x} + (1-\alpha) \bar{D}(\bar{x}, n), n) ) \quad (9)$$

$$r = F(\bar{x}, n) \quad (10)$$

**[0023]** Dieser Aufwärtsmischer aber, der tatsächlich sehr robust ist, begrenzt den Vorteil von Bewegungskompensation wesentlich (die Bewegungskompensation ist nur auf die unteren Frequenzen begrenzt). Deswegen wird nach der vorliegenden Erfindung der Aufwärtsmischer zwischen die erste und die zweite Möglichkeit angepasst, und zwar abhängig von der erwarteten Qualität der Bewegungsvektoren. Ein vorteilhaftes Merkmal der vorgeschlagenen Interpolatoren ist, dass Umschaltung zwischen den zwei robusten Möglichkeiten nicht sehr kritisch ist. Dies bedeutet, dass eine ziemlich grobe Entscheidung akzeptierbar ist, die mit geringer Verarbeitungsleistung an einer Version der Eingangssequenz mit einer (räumlich) reduzierten Größe durchgeführt werden kann. Diese Eingangssequenz reduzierter Größe wird verwendet zum berechnen von Anpassungsfehlern, erhalten mit (wenigstens) zwei Bewegungsvektoren je Stelle, entweder erzeugt aus einem Parametermodell oder dem Nullvektor.

**[0024]** Das Ergebnis einer Segmentierung, die das Bild in Schichten aufteilt, wobei das Nullvektormodell oder das berechnete Parametermodell geeigneter ist. Die Segmentierungsmaske SM wird nun als extra Eingabe des Aufwärtsmischer UC verwendet, der die Maske SM zum Umschalten/Blenden zwischen den beiden oben beschriebenen Aufwärtsmischern verwendet (siehe [Fig. 2](#)). Im Falle eines gültigen Parametermodells neigt der Aufwärtsmischer zu der Interpolation der Gleichung 2, sonst zu der Interpolation der Gleichung 7.

**[0025]** In [Fig. 2](#) werden die Werte l und r (siehe [Fig. 1](#)) einer ersten Mittelwertbestimmungsschaltung AV1 zugeführt um den Wert av zu erzeugen. Ein erstes Mittelwertfilter MED1 bestimmt den Mittelwert der Werte av, mcl und mcr. Die Werte mcl und mcr werden einer zweiten Mittelwertbestimmungsschaltung AV2 zugeführt zum Erzeugen des Wertes mcav. Ein zweites Mittelwertfilter MED2 bestimmt den Mittelwert der Werte mcav, l und r. Der Aufwärtsmischer UC1 liefert den interpolierten Wert aus den Ausgangssignalen der Mittelwertfilter MED1, MWD2 in Abhängigkeit von der Segmentierungsmaske SM. Das Ausgangssignal des Aufwärtsmischers UC1 wird einer Wiedergabeeinheit (D) zur Wiedergabe der Ausgangsvideodaten (n-1/2) zwischen den Eingangsvideodaten (n, n-1) zugeführt.

**[0026]** Von dieser Stelle an können Erweiterungen in Richtung vieler Schichten betrachtet werden wobei viele Parameterschätzer PE1...Pen (siehe [Fig. 3](#), die einen geschichteten auf Parametern basierten Schätzer und Aufwärtsmischer zeigt) parallel laufen, die je Parametermodelle für verschiedene, nicht unbedingt feste Teile von Bildern erzeugen. Diese Parameterschätzer PEi sind wieder der Eingang einer Segmentierungsschaltung SC, die die Teile des Bildes findet, für die jedes Modell gültig ist, oder mit anderen Worten, bestimmt eine Segmentierungsmaske SM, die das beste Interpolationsverfahren (Parametersatz) für jeden Teil des Bildes angibt. Der Aufwärtsmischer UC2 sollte wieder das bestmögliche Interpolationsverfahren für jede einzelne Schicht innerhalb des Bildes in Abhängigkeit von der Segmentierungsmaske SM wählen.

**[0027]** In [Fig. 3](#) werden Daten des aktuellen Bildes von dem Eingangsteilbild n und vorhergehende Bilddaten von dem Eingangsteilbild n-1 den Parameterschätzern PE2...Pen zugeführt, und zwar zum Ermitteln von Bewegungsparametern p21-p2m...pn1-pnm. Ein erster Parameterschätzer PE1 liefert Nullparameter. Die Eingangsteilbilder n und n-1 werden ebenfalls der Segmentierungsschaltung SC zugeführt, und zwar über Abwärtssampler D1, D2. Der Aufwärtsmischer UC2 berechnet Bewegungsvektoren auf die Art und Weise, wie durch die Gleichung 1 aus dem Parametersatz angegeben, durch die Segmentierungsmaske SM angegeben, zum Interpolieren des Ausgangsteilbildes n-1/2 von den Eingangsteilbildern n und n-1. Die Gewichtungsfaktoren W werden nachstehend anhand der Gleichungen 17 und 18 erläutert, und zwischen den Gleichungen 20 und 21. Jeder Parameterschätzer PE2...PEn umfasst eine Fehlerberechnung zum Einstellen der Bewegungsparameter. Die Berechnung beschränkt sich vorzugsweise auf diejenigen Bildteile, die mit dem Parametersatz übereinstimmen, behandelt durch den betreffenden Parameterschätzer PE. Dazu wird ein Gewicht W zugeordnet, wobei dieses Gewicht dem Betrag der Übereinstimmung entspricht (Art von Fuzzy-Logik). Schlussend-

lich bei der Interpolation wird für jedes Pixel derjenige Parametersatz verwendet, der den geringsten Schätzungsfehler für jedes Pixel ergibt.

**[0028]** In dem nachfolgenden Teil dieser Beschreibung, werden bevorzugte Parameterschätzer beschrieben.

**[0029]** In [2] werden Verfahren zum Schätzen globaler Bewegungsparameter aus einer Bildsequenz beschrieben. Das Buch richtet sich auf mehrere Möglichkeiten zum Lösen des mehrdimensionalen Optimierungsproblems, wie auf Gradienten basierte Verfahren, simuliertes Glühen, usw. Nach einem weiteren Aspekt der vorliegenden Erfindung werden diese Bewegungsparameter mit einer wesentlich reduzierten Vorgangszählung geschätzt um entweder die Kosten des Siliziums zu reduzieren oder sogar die Verarbeitung an einer programmierbaren Architektur zu ermöglichen (insbesondere den TriMedia Prozessor von Philips).

**[0030]** In [5] wurde ein Verfahren beschrieben zum Schätzen von globalen Bewegungsparametern aus einer Bildsequenz. Es wird dabei vorausgesetzt, dass Bewegung in dem Bild mit einer zweidimensionalen linearen Gleichung erster Ordnung beschrieben werden kann. Komplexere parametrische Bewegungsmodelle wurden bereits vorgeschlagen [2] und können tatsächlich in Kombination mit der vorliegenden Erfindung angewandt werden, werden aber in dieser Beschreibung nicht näher erläutert. In [5] wurde das Parametermodell zum Erzeugen interessanter Kandidatvektoren für einen auf Blöcke basierten Bewegungsschätzer verwendet. Das Eingangssignal für den Parameterschätzer war das vorhergehende Ausgangsvektorfeld, erhalten aus diesem auf Blöcke basierten Schätzer. Es dürfte einleuchten, dass wenn wir nur globale Bewegungsvektoren schätzen wollen, der Eingang zu den Parameterberechnungsmitteln einfacher sein kann.

**[0031]** Wenn wir und der Deutlichkeit halber auf die vier Parametermodell der Gleichung 1 beschränken, definieren wir zunächst den Parametervektor  $\bar{P}$ :

$$\bar{P}(n) = \begin{pmatrix} p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \end{pmatrix} \quad (11)$$

und definieren unsere Aufgabe als Selektion von  $\bar{C}_p(n)$  aus einer Anzahl Kandidatparametervektoren  $\bar{P}(n)$  als derjenige, der den minimalen Wert eines Anpassungskriteriums hat, berechnet entsprechend der nachfolgenden Gleichung:

$$\epsilon(\bar{C}_p, n) = \sum_x |F(\bar{x}, n) - F(\bar{x} - \bar{D}(\bar{x}, n), n-1)| \quad (12)$$

**[0032]** Die Berechnung dieser Fehlerfunktion kann weitgehend vereinfacht werden, und zwar durch Anwendung einer starken Unterabtastung. Versuche geben an, dass mit einem Anpassungskriterium, berechnen an nur 300 Pixeln je Teilbild, gute Ergebnisse erzielt werden können, d.h. mit einem Abtastfaktor der Größenordnung 1000. Weitaus effektiver ist aber eine angehäuften Unterabtastung, d.h. die selektierten Pixel bilden Gruppen, die spärlich über das Feld verteilt sind.

**[0033]** Der Vorschlag zum Durchführen der Minimierung erfordert, dass ein Prädiktionsvektor genommen wird (nun wenigstens dreidimensional, in unserem Beispiel vierdimensional), dass wenigstens ein Aktualisierungsvektor hinzugefügt wird, und dass entsprechend der Gleichung 13 der beste selektiert wird. Gute Ergebnisse könnten versuchsweise erzielt werden, wenn ein Kandidatvektorsatz  $CS_p(n)$  erzeugt wird, der drei Kandidatparametervektoren  $\bar{C}_p(n)$  enthält, und zwar entsprechend:

$$CS_p(n) = \left\{ \bar{C}_p(n) \mid \bar{C}_p(n) = \bar{P}(n-1) + m\bar{U}_p(n), \bar{U}_p(n) \in US_p(n), m = -1, 0, 1 \right\} \quad (13)$$

wobei  $US_p(n)$  entsprechend der nachfolgenden Gleichung selektiert worden ist:

$$US_p(n) = \left\{ \begin{bmatrix} i \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ i \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \end{bmatrix} \right\}, (i=1,2,4,8,16) \quad (14)$$

**[0034]** Strafen können zu dem Anpassungsfehler einzelner Kandidatvektoren (Parametersätze) zum Erhalten beispielsweise zeitlicher Geschmeidigkeit, hinzugefügt werden. Auch zeitliche Filterung der Parametervektoren, entweder innerhalb oder außerhalb der Prädiktionsschleife, wird derart betrachtet, dass dies eine plötzliche Änderung der Bewegungsvektoren von dem einen Bild zum anderen vermeidet.

**[0035]** Obschon in der bisherigen Beschreibung bereits vorgeschlagen wurde, dass das parametrische Bewegungsmodell die globale Bewegung des ganzen Bildes beschreibt, lassen sich Alternativen bedenken, wobei das Bild in einige, beispielsweise 9 große Blöcke aufgeteilt wird, und mögliche Prädiktionen sind nicht nur die zeitliche Prädiktion, sondern auch eine oder mehrere räumliche Prädiktionen. Eine weitere Alternative umfasst Segmentierung und eine feste Anzahl Parameterschätzer laufen parallel, je sich auf ein Segment des Bildes richtend, angegeben durch den Segmentierungsalgorithmus, der über ein vorhergehendes Bild läuft.

**[0036]** Der Vorgangszählwert ist unglaublich niedrig. Berechnung des Fehlerkriterium beträgt etwa 1000 Vorgänge je Kandidatvektor je Wiederholung. Für die beschriebene Implementierung führt dies zu:  $\frac{3 \cdot 16 \cdot 1000}{720 \cdot 288} \approx \frac{48}{207} \approx 0.23$  Vorgängen je Pixel. Dies ist eine Reduktion um eine andere Größenordnung von zwei Größenordnungen, im Vergleich zu dem Schätzer von [6].

**[0037]** [Fig. 4](#) zeigt ein Blockschaltbild eines Bewegungsparameterschätzers nach der vorliegenden Erfindung. Ein erster und ein zweiter Kandidatparametersatz Cp1, Cp2 werden einem Multiplexer MUX und einem Parameter-zu-Vektorwandler PVC zugeführt, zum Erhalten zweier Kandidatbewegungsvektoren Cv1, Cv2. Der erste Kandidatparametersatz Cp1 ist der vorhergehende Ausgangsparametersatz P(n) des Multiplexer MUX. Der zweite Kandidatparametersatz Cp2 wird dadurch erhalten, dass ein Aktualisierungsparametersatz Up zu dem ersten Kandidatparametersatz Cp1 hinzuaddiert wird (Addierer AD). Der Aktualisierungsparametersatz Up wird dadurch erhalten, dass das Ergebnis eines mod(n)-Zählers CNT einer Nachschlagtabelle LUT zugeführt wird. Die Kandidatbewegungsvektoren Cv1, Cv2 werden einem Fehlerberechner EC zugeführt, dem das aktuelle und das vorhergehende Teilbild n, n-1 auch zugeführt werden, und zwar zum Erhalten zweier Fehler E1, E2. Eine Minimumschaltung MIN bestimmt, welcher Fehler der kleinere ist zum Erhalten eines Selektionssignals s für den Multiplexer MUX zum Erhalten des Ausgangsparametersatzes P(n).

**[0038]** Der nachfolgende Teil dieser Beschreibung befasst sich mit einem bevorzugten Verfahren zum Schätzen von Bewegungsparametern aus Videodaten. Bewegungsschätzung wird beim Codieren und bei der Abstratenumwandlung von Videodaten angewandt. Obschon meistens die Bildrate dieser Videodaten am Eingang des Bewegungsschätzers fest ist, kann die Bildrate der Videoquelle, von der diese Daten herrühren, von der der verarbeiteten Daten abweichen. Insbesondere tritt dies auf, wenn Filmmaterial zu Video umgesetzt wird, oder wenn Videomaterial von der einen Norm in eine andere Norm umgesetzt wird, und zwar irgendwo in der Videokette vor dem Bewegungsschätzer.

**[0039]** Eine übliche Art und Weise mit den erforderlichen Bildratenumwandlungen fertig zu werden, ist die Verwendung des jüngsten Bildes bis ein neues verfügbar wird. Wenn von einer niedrigen Bildrate zu einer höheren Bildrate umgesetzt wird, bedeutet dies die Wiederholung von Quellenbildern in dem neuen Format, während eine Umwandlung von einer hohen zu einer niedrigeren Rate zu einer gelegentlichen Überspringung von Bildern des Quellenmaterials. In beiden Fällen zeigt das resultierende Video ein unregelmäßiges Bewegungsmuster ("judder"), was die übliche Annahme bei Bewegungsschätzern, dass Bewegung eine starke zeitliche Konsistenz hat, verletzt. Bei Bewegungsschätzern, die versuchen, Vorteil aus dieser Voraussetzung zu ziehen, durch Verwendung zeitlicher Prädiktionsvektoren, führt das Problem dazu, dass das unregelmäßige Bewegungsverhalten den Nutzen dieser zeitlichen Prädiktionsvektoren eliminiert. Dadurch kann eine wesentliche Verringerung der geschätzten Bewegungsvektoren das Ergebnis sein.

**[0040]** In [9] wurde eine Lösung dieses Problems beschrieben, für Filmmaterial, das in eine 50 Hz Fernsehnorm umgewandelt wird. Der Gedanke dabei ist, den Vektorprädiktionsspeicher umlaufen zu lassen, wenn



ein wiederholtes Bild auftritt. In [8] wurde ein Verfahren beschrieben, wobei der Bildspeicher das "vorhergehende" Bild speichert, das in Umlauf gebracht wird, bis ein nicht wiederholtes Bild auftrat. Eine Charakteristik, die von den beiden bekannten Verfahren geteilt wird, ist, dass das Muster bekannt sein soll um die Speichersteuerung zu ändern.

**[0041]** Es ist nun u. a. eine Aufgabe dieses Aspektes der vorliegenden Erfindung, ein sehr robustes Bewegungsschätzungsverfahren zu schaffen, das überhaupt keine Vorkenntnisse des Wiederholungsmusters braucht um auf zuverlässige Art und Weise Bewegung zu schätzen. Dazu nimmt der Bewegungsschätzer zeitliche Prädiktionsvektoren aus mehr als nur einem vorhergehenden Bilderpaar (ebensoviel wie die maximale Länge des Wiederholungsmusters) und selektiert die besten daraus als Basis für den Schätzungsprozess, oder benutzt sie alle als Kandidat in einem Anpassungsprozess.

**[0042]** Diese Lösung ist wirtschaftlich berechtigt, insbesondere bei objektbasierten Bewegungsschätzern, wobei die Anzahl zu speichernder Bewegungsvektoren sehr gering ist. Eine Software-Version des Algorithmus ist dargestellt um in Echtzeit in dem Philips TM 1000 (TriMedia) Prozessor zu laufen.

**[0043]** [Fig. 5](#) zeigt einen bevorzugten Parameterschätzer nach der vorliegenden Erfindung. Aktuelle Bilddaten aus dem aktuellen Teilbild  $n$  und vorhergehende Bilddaten aus dem vorhergehenden Teilbild  $n-1$  werden einer Bewegungsparameterschätzungseinheit MPE zugeführt zum Erhalten von Bewegungsparametern  $P(n)$ . Bildverzögerungen  $D1, D2, \dots, Dn$  liefern verzögerte Versionen  $TP1, TP2, \dots, TPn$  der Bewegungsparameter  $P(n)$  zu der Bewegungsparameterschätzungseinheit MPE.

**[0044]** Der nachfolgende Teil dieser Beschreibung bezieht sich auf eine geschichtete Bewegungsschätzung, d.h. das Bild wird in eine Anzahl Schichten segmentiert.

**[0045]** Es wurden bereits Region-basierte Bewegungsschätzer eingeführt als Alternative für Block-basierte Bewegungsschätzer. Block-basierte Bewegungsschätzer sind in die internationalen Normen für Videokompression aufgenommen, wie H.261/263 (Videokonferenz über ISDN-Leitungen), MPEG-1 (Multimedia) und MPEG-2 (ganz-digitale Fernsehapplikationen). Obschon diese Normen nicht ein bestimmtes Bewegungsschätzungsverfahren spezifizieren, wird Block-basierte Bewegungsschätzung eine natürliche Wahl.

**[0046]** Die Verwendung aber von Blöcken als Einheiten für Bewegungsschätzung kann zu sperrenden Artefakten führen, weil die Begrenzungen von Objekten im Allgemeinen nicht mit den Blockgrenzen übereinstimmen, und Nachbarblöcken können wesentlich andere Bewegungsvektoren zugeordnet werden, wenn es keine räumlich-zeitliche Konsistenzbeschränkung gibt.

**[0047]** Eine viel versprechende Annäherung zum Lösen des Problems von Blockartefakten und zum Schaffen einer genaueren Prädiktion über sich verlagernder Ränder ist, das Bewegungsfeld zu segmentieren. Bewegungsinformation und Musterinformation (Intensität, Konturstruktur) werden verwendet um eine Gebiets-basierte (beliebig geformte) Bewegungsschätzung zu erzielen, wobei das nächste Ziel ist, Objekte zu behandeln und möglicherweise MPEG-4 "Audio-Visuelle Objekte".

**[0048]** Es wurden bereits viele Verfahren vorgeschlagen [2] zum Segmentieren von Bildern und zum Schätzen von Bewegungsparametern für diese Segmente aus eine Bildsequenz. Je nach der Strategie beim Durchführen der Segmentierung können diese Verfahren in Verfahren mit Aufwärtsstruktur und Verfahren mit Abwärtsstruktur, sowie mit einer geschichteten Darstellung aufgeteilt werden. Wir werden die Kennzeichen der einzelnen Kategorien kurz zusammenfassen.

#### Verfahren mit einer Aufwärtsstruktur.

**[0049]** Die Verarbeitung startet mit einer Intraframe-Segmentierung des Bildes auf Basis von Musterinformation oder auf Basis eines vorher berechneten dichten Bewegungsvektorfeldes. Die Segmentierung führt im Allgemeinen zu einer Anzahl kleiner Gebiete. Diese Gebiete werden danach zusammengefügt, wobei im Allgemeinen Information über ihre Bewegung verwendet wird, d.h. Gebiete mit ähnlicher Bewegung werden zu einem einzigen Gebiet zusammengefügt und Bewegungsparameter werden danach neu berechnet. Diese Prozedur erweist sich als sehr beliebt, wenn das Ziel Objekt-orientierte Codierung ist. Beispiele in [26–29].

#### Verfahren mit einer Abwärtsstruktur.

**[0050]** Die Verarbeitung startet mit einer Ausgangsbildsegmentierung in große Gebiete. Diese werden unter-



teilt, wobei dem berechneten Bewegungsmodell Genauigkeit fehlt, und es werden Bewegungsparameter neu berechnet. Die Ausgangssegmentierung ist im Allgemeinen auf einer "geändert/nicht geändert" Regel basiert, d.h. das aktuelle Bild wird mit dem vorhergehenden Bild verglichen; wenn in derselben Lage der Leuchtdichtewert in dem einen Frame wesentlich anders ist als der Leuchtdichtewert in dem anderen Frame, wird dieses Pixel als "geändert" gemerkt, sonst als "nicht geändert". Daraufhin kann der als "geändert" klassifizierte Teil bewegungskompensiert werden, und zwar entsprechend dem Bewegungsfeld, berechnet für dieses Gebiet, und die oben beschriebene Prozedur wird wiederholt um die verschiedenen Bewegungsgebiete zu identifizieren. Beispiele in [11, 30, 31, 36].

**[0051]** Die zwei Techniken können auch kombiniert werden, so kann beispielsweise der Start der jeweiligen Ausgangssegmentierungen beliebig sein, oder basiert auf einer vorhergehenden Schätzung [11,32], und in beiden Richtungen gibt es aufeinander folgende Verfeinerungen. Die Schätzung und die Segmentierung können auch simultan durchgeführt werden, und zwar unter Anwendung einer statistischen Annäherung der Analyse von Bildsequenzen, beispielsweise mit einem Verfahren der maximalen Wahrscheinlichkeitsschätzung [33].

#### Verfahren mit einer geschichteten Darstellung.

**[0052]** Die ideale Szenensegmentierung führt zu separaten Objekten und betrifft dreidimensionale Information, aber dies lässt sich schwer erzielen und ist relativ intensiv. Deswegen werden die Videodaten segmentiert und als einen Satz sich verlagernder Schichten beschrieben, d.h. die Bildteile, die eine ähnliche Bewegung erfahren, sogar wenn nicht verbunden. Danach wie die Ordnung (Tiefe) der Schichten ermittelt. Beispiele in [29, 34, 35]. Ein Modell, d viel weniger kompliziert ist als das voll-dreidimensionale Modell und weniger kompliziert als ein Modell, das mit allen Objekten in der Sequenz klar kommt, wurde bereits vorgeschlagen. Da dies das in einer bevorzugten Ausführungsform angenommene Modell ist, wird dies in den nachfolgenden Abschnitten näher beschrieben.

**[0053]** Entsprechend diesen Verfahren der geschichteten Darstellung werden die Videodaten segmentiert und als einen Satz sich verlagernder Schichten beschrieben, d.h. als Satz von Gebieten, die eine gleiche Bewegung erfahren, sogar wenn nicht verbunden. Die Tiefenordnung der Schichten kann danach ermittelt werden. Eine geschichtete Darstellung einer Videosequenz ist für verschiedene Applikationen interessant, wie Abtastratenumwandlung, Objektverfolgung, Videokompression, Codierung, Videonotierung und Indexierung. Eine Anzahl Algorithmen wurde bereits präsentiert für eine geschichtete Bewegungsdarstellung [29.34–37].

**[0054]** Einer der wesentlichen Punkte in diesen Algorithmen ist die Art und Weise, wie das Bewegungsschätzung/Segmentierungsproblem gelöst wird. Es wurden bereits zwei Hauptannäherungen vorgeschlagen.

#### Sequentielle Annäherung

**[0055]** Die sequentielle Annäherung zerlegt viele Schichten durch sequentielle Schätzung einer dominierenden Bewegung, entsprechend dem, was bei dem Verfahren mit Abwärtsstruktur gemacht wird. Der Hauptnachteil einer derartigen Annäherung ist der, dass, da die schlussendliche Segmentierung noch nicht bekannt ist, während die eine Schicht behandelt wird, ein Teil des Bildes mit einer anderen Bewegung in die Schätzung der Bewegungsparameter eingeschlossen sein kann, was die Ergebnisse beeinflusst.

#### Die simultane Annäherung

**[0056]** Die simultane Annäherung versucht alle Schichten in dem Bild simultan zu schätzen. Dies kann durch Verwendung eines vorher berechneten dichten Bewegungsvektorfeldes gemacht werden. Der Ausgangssatz mit: Bewegungsmodellen kann durch Verwendung eines angehäuften Algorithmus über das bestimmte Bewegungsvektorfeld [29] hergeleitet werden. Beim Berechnen des Bewegungsvektorfeldes werden im Allgemeinen einige Gleichmäßigkeitsannahmen gemacht. Dies kann zu einem Bewegungsvektorfeld führen, in dem die Grenzen nicht mit Grenzen von Objekten/Schichten übereinstimmen, so dass eine einwandfreie Segmentierung nicht möglich ist. Auf alternative Weise kann das Problem als ein stochastisches Problem formuliert werden und eine maximale Wahrscheinlichkeitsschätzung der vielen Modelle und deren Trägerschichten können unter Anwendung eines Erwartungsmaximierungsalgorithmus [36] durchgeführt werden. Der Hauptnachteil der zwei letzten Verfahren ist ihre Komplexität.

**[0057]** Ein anderer wesentlicher Punkt ist die Art und Weise, wie die Bewegungsparameter geschätzt werden. Abhängig davon, ob die Schätzung von Bewegungsparametern an dem Leuchtdichtesignal selber oder nicht

durchgeführt wird, kann sie als direkt oder als indirekt klassifiziert werden. Die direkten Verfahren werden im Allgemeinen als robuster betrachtet. In [2] werden viele Verfahren zum Schätzen globaler Bewegungsparameter aus einer Bildsequenz beschrieben. Es wurden bereits mehrere Möglichkeiten zum Lösen des mehrdimensionalen Optimierungsproblems, wie Gradienten-basierte Verfahren, simuliertes Glühen, usw. vorgeschlagen. Es ist nun das Ziel des vorgeschlagenen Algorithmus, diese Bewegungsparameter mit einem wesentlich reduzierten Vorgangszählwert zu schätzen um eine Bewegungsschätzung an einer programmierbaren Architektur zu ermöglichen.

**[0058]** Es ist nun der Gegenstand des vorliegenden Teils dieser Beschreibung, Bildsequenzen zu schätzen/Segmentieren mit einem wesentlich reduzierten Vorgangszählwert um entweder die Kosten des zugeordneten Siliziums zu reduzieren oder sogar eine Verarbeitung an einer programmierbaren Architektur (insbesondere der Philips TriMedia Prozessor) zu ermöglichen.

**[0059]** Der aktuelle Aspekt der vorliegenden Erfindung befasst sich mit Bewegungsschätzung/Segmentierung befasst, abzielend auf eine geschichtete Darstellung. Um die Kosten der Implementierung möglichst niedrig zu halten richten wir uns auf eine Implementierung als ein direktes Verfahren, obschon eine indirekte Version denkbar scheint. Es schafft eine elegante Lösung für das Problem "was war zuerst, das Huhn oder das Ei" einer kombinierten Bewegungsschätzung/Segmentierung. Die Lösung besteht aus einem Gewichtungsprozess, der die Verunreinigung der Optimierungskriterien eines Parameterschätzers für eine bestimmte Schicht durch Information besorgt von den anderen Parameterschätzern, die parallel laufen, begrenzt. Das Entwerfen eines Bewegungsschätzers um in Echtzeit in bestehenden programmierbaren Architekturen zu laufen setzt strenge Grenzen an das Problem der Bewegungsschätzung, da die Komplexität des Algorithmus drastisch reduziert werden soll. Zu diesem Zweck ist ein geschichteter Bewegungsschätzer gewählt worden, da man glaubt, dass es im Grunde einfacher ist eine programmierbare Architektur zu implementieren als beispielsweise eine Block-basierte Bewegungsschätzung, wenn man besieht, dass es weniger Schichten als Blöcke gibt.

**[0060]** In einer geschichteten Darstellung wird ein Bild in eine Anzahl Schichten aufgeteilt, d.h. Teile des Bildes erfahren eine kohärente Bewegung, sogar wenn sie nicht verbunden sind. Es wird vorausgesetzt, dass die sichtbare Bewegung (optischer Fluss) in einer Sequenz dann mit parametrischen Modellen beschrieben werden kann, d.h. es ist nur auf eine Kombination von Kamerabewegung und starre Bewegung undurchsichtiger Objekte zurückzuführen. Folglich kann ein einziger Satz mit Bewegungsparametern für jede Schicht statt für das Bewegungsfeld selber geschätzt werden.

**[0061]** Segmentierung einer Sequenz von Bildern in Gebiete, die gleiche Bewegungen erfahren und simultan ihre Bewegung schätzen ist an sich ein Problem, da die zwei Zuordnungen voneinander abhängig sind. Um die Bewegung in einem einzigen Gebiet einwandfrei zu schätzen soll das Gebiet bekannt sein. Um die Gebiete des Bildes, die kohärent bewegen, zu ermitteln, soll ihre Bewegung bekannt sein. Ein neues Verfahren für eine quasi-simultane Bewegungsschätzung und Segmentierung bis zu einer festen Anzahl Schichten wird präsentiert. Wir richten uns auf das Problem der Schätzung der Bewegungsparameter für jede Schicht und segmentieren gleichzeitig das Bild durch Einführung einer Hierarchie, d.h. dadurch, dass ein andere Bewertung gegeben wird. Die zwei Ziele, für die diese Hierarchie bestimmt ist, sind:

- Das Vermeiden, dass eine bestimmte Schicht in Teilen des Bildes geschätzt wird, die durch Schichten bedeckt sind, die in der Hierarchie eine höhere Bewertung haben.
- Das Vermeiden, dass eine bestimmte Schicht durch Teile des Bildes verschmutzt werden, die besser durch Schichten bedeckt sind, die in der Hierarchie eine geringere Bewertung haben.

**[0062]** Die Parametervektoren werden danach parallel geschätzt, und zwar unter Anwendung einer rekursiven Annäherung, d.h. der früh geschätzte Parametervektor für jede Schicht wird als eine Prädiktion verwendet, zu der Aktualisierungsvektoren hinzuaddiert werden. Der selektierte Parametervektor ist derjenige, der zu dem kleinsten Anpassungsfehler führt. Danach werden die Parametervektoren aller Schichten zusammen bei der Segmentierung des Bildes in die gewünschten verschiedenen Schichten verwendet.

**[0063]** Die Bewegung jeder Schicht 1 wird durch ein einfaches Bewegungsmodell beschrieben. Es wird vorausgesetzt, dass die Bewegung innerhalb einer Schicht mit einem zweidimensionalen linearen Modell erster Ordnung beschrieben werden kann.

$$\bar{D}(\bar{x}, l, n) = \begin{Bmatrix} s_x(l, n) + d_x(l, n) \cdot x \\ s_y(l, n) + d_y(l, n) \cdot y \end{Bmatrix} \quad (15)$$

wobei  $\bar{D}(\bar{x}, l, n)$  für den Verlagerungsvektor der Schicht 1 an der Stelle  $\bar{x} = (x, y)^T$  in dem Bild mit dem Index  $n$  verwendet wird. Mit diesem vier-Parametermodell können horizontale und vertikale geradlinige Bewegungen ("Pan" und "Tilt") sowie "Zoom" beschrieben werden. Es wurden bereits komplexere parametrische Bewegungsmodelle vorgeschlagen [2] und können tatsächlich in Kombination mit dem vorgeschlagenen Algorithmus angewandt werden, werden aber an dieser Stelle nicht näher beschrieben. In den Versuchen wurde dieses Bewegungsmodell mit verschiedenen Freiheitsgraden verwendet:

- Alle vier Parameter frei.
- Die Parameter  $s_x$  und  $s_y$  frei,  $d_x$  und  $d_y$  gekoppelt mit einem festen Verhältnis entsprechend dem Seitenverhältnis des Bildes (drei-Parametermodell).
- Die Parameter  $s_x$  und  $s_y$  frei,  $d_x$  und  $d_y$  fest auf Null (Zwei-Parameter, Translationsmodell).
- Die Parameter  $s_r$  frei,  $s_y$ ,  $d_x$  und  $d_y$  fest auf Null (Ein-Parameter, "Panning"-Modell).

**[0064]** In einer Ausführungsform hat eine erste Schicht 4 oder 8 freie Parameter, während jede nachfolgende Schicht weniger freie Parameter hat als die vorhergehende Schicht, und zwar zum Reduzieren der rechnerische Belastung.

**[0065]** Der vorliegenden Erfindung liegt die Erkenntnis zugrunde, dass der Nullvektor (keine Bewegung) sehr üblich ist und in Videosequenzen wichtig, und insbesondere wichtig für die beabsichtigte Anwendung in der Abstratenumwandlung. Deswegen startet der vorgeschlagene Algorithmus mit einer Schicht 0, wobei die Bewegung durch den Nullparametervektor beschrieben wird (der offenbar nicht geschätzt wird). Die Parametervektoren zusätzlicher Schichten 1,  $1 > 0$  werden einzeln geschätzt, und zwar durch ihre betreffenden Parameterschätzer  $PE_1$ .

**[0066]** Jeder  $PE_1$  hat dasselbe Basisprinzip wie die dreidimensionale rekursive Suchblockanpassungsschaltung von [3]. Ein vorher geschätzter Parametervektor wird entsprechend einem pseudobeliebigen Rauschvektor aktualisiert, wonach der am besten passende Parametervektor gewählt wird.

**[0067]** In Anbetracht des Parametermodells der Gleichung (15) werden die Parameter der Schicht 1,  $1 > 0$  betrachtet als ein Parametervektor  $\bar{P}_1$ :

$$\bar{P}_1(n) = \begin{pmatrix} s_x(l, n) \\ s_y(l, n) \\ d_x(l, n) \\ d_y(l, n) \end{pmatrix} \quad (16)$$

und wir definieren unsere Aufgabe als das Wählen von  $\bar{P}_1(n)$  aus einer Anzahl Kandidatparametervektoren  $\bar{CP}_1(n)$  als derjenige, der den minimalen Wert eines Anpassungskriteriums hat. Die Fehlerfunktion wird entsprechend der nachfolgenden Gleichung berechnet:

$$\epsilon'(\bar{CP}_1(n)) = \epsilon(\bar{CP}_1(n)) - \sum_{\bar{x} \in X_1} W_{\bar{x}}(\bar{x}) \cdot \Pi(\bar{CP}_1(n)) \quad (17)$$

wobei Strafen  $\Pi(\bar{CP}_1(n))$  zu dem Anpassungsfehler einzelner Kandidatvektoren (Parametersätze) hinzugefügt werden um beispielsweise eine räumliche Glätte zu erhalten und wobei  $\epsilon$  Folgendes ist:

$$\epsilon(\bar{CP}_1(n)) = \sum_{\bar{x} \in X_1} W_{\bar{x}}(\bar{x}) \cdot |F_s(\bar{x}, n) - F_r(\bar{x}) - \bar{D}(\bar{x}, l, n, n - 1)| \quad (18)$$

wobei  $W_{\bar{x}}(\bar{x})$  ein Gewichtungsfaktor ist, der von der Lage abhängig ist und wobei  $\bar{x} \in X_1$  der Leuchtdichtewert an der Stelle  $\bar{x}$  in dem unterabgetasteten Bild mit dem Index  $n$  ist, und wobei  $X_1$  ein Satz der Stellen  $\bar{x}$  ist, wobei die Bewegung der Schicht 1 geschätzt werden soll (die Mode der Selektion der Stellen  $\bar{x}$  wird nachstehend noch näher erläutert).

**[0068]** Die Bilder werden mit einem Faktor 4 horizontal und 2 vertikal auf Teilbildbasis unterabgetastet, wobei ein unterabgetastetes Bild  $F_s(n)$  aus jedem ursprünglichen Teilbild  $F(n)$  erzeugt wird. Dies liefert einen starken Beitrag zu der gewünschten Reduktion des Vorgängezahlwertes. Die Unterabtastung ist erlaubt, weil die Objekte, für die Bewegung geschätzt wird, groß genug sind. Um eine Pixelgenauigkeit oder sogar eine Subpixelgenauigkeit in dem ursprünglichen Pixelgitter von  $F$  zu erzielen ist Interpolation in dem Unterabtastungsgitter [7] erforderlich.

**[0069]** Die vorgeschlagene Minimierung zeigt eine gewisse Übereinstimmung mit der Strategie, die in [3, 7] angewandt worden ist, d.h. einen Prädiktionsvektor (in diesem Fall vierdimensional) nehmen, wenigstens einen Aktualisierungsvektor hinzufügen und den hend der Gleichung 18 selektieren. Gute Ergebnisse könnten versuchsweise erzielt werden, wenn ein Kandidatparametersatz  $\bar{CP}_i(n)$  erzeugt wird, der drei Kandidaten  $S_{\bar{CP}_i(n)}$  enthält, und zwar entsprechend:

$$S_{\bar{CP}_i(n)} = \left\{ \bar{CP}_i(n) \mid \bar{CP}_i(n) = \bar{P}_i(n-1) + m \bar{UP}_i(n), \right. \\ \left. \bar{UP}_i(n) \in S_{\bar{UP}_i(n)}, m = -1, 0, 1 \right\} \quad (19)$$

wobei Aktualisierungsparameter  $\bar{UP}_i(n)$  aus dem Aktualisierungsparametersatz  $S_{\bar{UP}_i(n)}$  selektiert wird:

$$S_{\bar{UP}_i(n)} = \left\{ \begin{bmatrix} i \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ i \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ i \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ i \end{bmatrix} \right\}. \quad (20)$$

$(i = 1, 2, 4, 8, 16)$

**[0070]** Zeitliche Filterung der Parametervektoren, innerhalb sowie außerhalb der Prädiktionsschleife, wird angewandt um eine plötzliche Änderung von Bewegungsvektoren von dem einen Bild zu dem anderen zu vermeiden.

**[0071]** Der bisher beschriebene Algorithmus führt eine einzige Wiederholung an einem Paar Eingangsbilder durch. Eine schnellere Konvergenz des Algorithmus wird mit vielen Wiederholungen der Parameterschätzer an demselben Paar Eingangsbilder erreicht, in dem vorliegenden Fall wird  $\bar{P}_i(n-1)$  in der Gleichung (19) durch den Ausgang der vorhergehenden Wiederholung  $\bar{P}_i(n)$  nach der Ausgangswiederholung an einem Bilderpaar ersetzt.

**[0072]** Es wird eine hierarchische Struktur der Schichten vorgeschlagen. Dies wird wie folgt erzielt:

- Selektion von Stellen  $\bar{x}$  in  $X_1$ , wobei Bildteile ausgeschlossen werden, die durch Schichten mit einer höheren Bewertung bedeckt sind.
- Innerhalb von  $X_1$  Reduktion des Effektes von Bildteilen, die möglicherweise besser bedeckt sind durch Schichten mit einer niedrigeren Bewertung in der Hierarchie: Zuordnung von höheren Gewichtungen  $w_i(\bar{x})$  zu den Pixeln, die der Schicht 1 in der vorhergehenden Segmentierung zugeordnet wurden.

**[0073]** Jeder Schätzer, mit Ausnahme des höchsten in der Hierarchie (der Null-Schätzer), minimiert einen Anpassungsfehler, der in Gebieten berechnet worden ist, in denen alle Schätzer eines höheren Pegels nicht erfolgreich waren in dem vorhergehenden Bild. Der Satz mit Stellen  $X_1$  wird mit den Stellen  $\bar{x}$  gefüllt, wobei der Anpassungsfehler aller höher bewerteten Schichten den mittleren Blockanpassungsfehler um einen festen Faktor übersteigt.

**[0074]** Versuche geben an, dass gute Ergebnisse erreicht werden, wenn die Anzahl Stellen in  $X_1$  auf einige 2–5% aller Pixel in dem Bild begrenzt wird. Am effektivsten ist eine angehäuften Unterabtastung innerhalb des Bereichs, d.h. die selektierten Pixel bilden Gruppen spärlich verteilt über das ganze Bild. In der aktuellen Applikation werden maximal 50 Cluster von 16 Pixeln gewählt (3% aller Pixel in  $F_s$ ).

**[0075]** Eine einwandfreie Selektion von  $X_1$  ist notwendig um zu vermeiden, dass der aktuelle Schätzer Bewegung schätzt, die bereits durch vorhergehende Schichten bedeckt ist.

**[0076]** Der stellenabhängige Gewichtungsfaktor  $w_i(\bar{x})$  wird durch die Segmentierungsmaske  $SM(n-1)$  bestimmt, gefunden in dem vorhergehenden Bild. Stellen  $\bar{x}$ , die zu der aktuellen Schicht 1 gehören, werden entsprechend der Segmentierungsmaske einen Gewichtungsfaktor haben, der größer ist als Eins, wobei Stellen, die zu einer anderen Schicht gehören, einen Gewichtungsfaktor gleich Eins haben. Eine einwandfreie Selektion von  $w_i(\bar{x})$  ist notwendig um zu vermeiden, dass der aktuelle Schätzer Bewegung schätzt, die durch nachfolgende Schichten in der Hierarchie bedeckt werden können.

**[0077]** Der Segmentierungsschritt ist der meist kritische Schritt in dem Algorithmus. Die Aufgabe ist es, eine der Schichten, d.h. ein Bewegungsmodell, in dem Bild jeder Gruppe von Pixeln zuzuordnen. Dies wird im Grun-

de dadurch erreicht, dass das am besten assende Modell jeder Gruppe von Pixeln zugeordnet wird (einen Block  $\bar{B}$ , der typischerweise 8×8 Pixel auf Framebasis groß ist).

**[0078]** Für jede Schicht wird ein Anpassungsfehler berechnet, und zwar entsprechend:

$$\sum_{\bar{B}} \epsilon(\bar{B}, l, n) = |F_l(\bar{x} + (1 - \alpha)\bar{D}(\bar{x}, l, n), n) - F_l(\bar{x} - \alpha\bar{D}(\bar{x}, l, n), n - 1)| \quad (21)$$

**[0079]** Die Segmentierungsmaske  $\bar{B}$  ordnet die Schicht 1 mit der niedrigsten  $\epsilon$  dem Block  $SM(\bar{B}, n)$  zu. Die zeitliche Lage der Segmentierung wird durch  $\alpha$  definiert, wobei dieser Wert in unseren Versuchen auf  $\frac{1}{2}$  gesetzt wurde.

**[0080]** Um Verarbeitungsleistung zu sparen braucht die Segmentierungsmaske SM nicht für jeden Block  $\bar{B}$  berechnet zu werden. Stattdessen können die berechneten Blöcke in einem Fünfermuster unterabgetastet werden, wonach die fehlenden Stellen in der Segmentierungsmaske interpoliert werden (beispielsweise durch Wahl der am meisten auftretenden Schichtnummer aus einer Umgebung) [7].

**[0081]** Segmentierung ist schwieriger je nachdem es mehr Schichten gibt, da die Segmentierungsaufgabe mehr und mehr derjenigen einer Vollsuchblockanpassungsschaltung entspricht. Um zu vermeiden, dass ein Ausgang des Bewegungsschätzers Inkonsistenz ähnlich der einer Vollsuchblockanpassungsschaltung hat, sind zusätzliche (Glättungs) Beschränkungen zu dem Algorithmus hinzugefügt worden. Aktuelle Glättungsbeschränkungen bestehen aus:

- Räumlicher Glättung: wobei ein größeres Fenster bei der Berechnung der  $\epsilon$  genommen wird als die Größe des Blocks  $\bar{B}$  dem die Schicht zugeordnet ist.
- Zeitliche Glättung: wobei die berechnete  $\epsilon$  einer Schicht um einen Bonuswert reduziert wird, wenn diese Schicht bei der Segmentierung des vorhergehenden Bildes gewählt wird.
- Räumliche Glättung: durch Verwendung eines Majoritätsfilters zum Entfernen bemerkenswerter Punkte bei der Segmentierung.

**[0082]** Als Ergebnis von Versuchen wurde bei der ersten Implementierung von TriMedia eine dreischichtige Struktur gewählt. Die Schicht 0 wird nicht geschätzt, entsprechend keiner Bewegung, d.h. alle Parameter gleich 0. Schicht 1 hat zwei freie Parameter und die Schicht 2 hat nur einen freien Parameter. Der Parameterschätzer der Schicht 1 wiederholt 5 mal und der Schätzer der Schicht 2 wiederholt 3 mal bei jedem Eingangsbilderpaar.

**[0083]** Eine einfache Vorfilterung der Unterabtastung wird durch die Mittelwertbestimmung von Pixelwerten in einem Block von 4×2 Pixeln erreicht. Dies dauert etwa 10 Vorgänge je unterabgetasteter Ausgangspixel oder  $\frac{180 \cdot 144 \cdot 10}{720 \cdot 288} \approx 1,25$  Vorgänge je Pixel des Eingangsgitters (CCIR 601/625 Zeilen/2:1).

**[0084]** Berechnung des Fehlerkriteriums in einem Parameterschätzer dauert etwa 1000 Vorgänge je Kandidatenvektor je Wiederholung. Für die beschriebene Implementierung führt dies zu  $\frac{3 \cdot (5 + 3) \cdot 1000}{720 \cdot 288} \approx 0,12$  Vorgängen je Pixel (dies deckt nicht alle Funktionen der Parameterschätzung). Die Berechnung des Fehlerkriteriums bei der Segmentierung dauert etwa 10 Voran je Schicht je Block folglich  $\frac{3 \cdot (72 \cdot 90/2) \cdot 10}{720 \cdot 288} \approx 0,47$  Vorgänge je Pixel (dies deckt nicht alle Funktionen der Segmentierung). Dies ist eine Reduktion einer anderen Größenordnung im Vergleich zu dem Schätzer von MELZONIC (SAA4991) [3]. Messungen in teilweise optimierten Code für TriMedia gegen einen erreichten Vorgangszahlwert von etwa 1,25 für die Unterabtastung an. 1.0 für den Parameterschätzer und 6.1 Vorgänge je Pixel für die Segmentierung.

**[0085]** Der vorgeschlagene geschichtete Bewegungsschätzer wurde simuliert, einschließlich der Verwendung des resultierenden Verlagerungsvektors für Bilddatenumwandlung von 25 Hz Film in 50 Hz Wiedergabe.

**[0086]** Das Vektorfeld, herrührend aus dem Bewegungsschätzer, hat sich als hoch konsistent gezeigt und als durchaus geeignet zur Abtastratenumwandlung. Die erhaltene Qualität wird als interessant betrachtet und für die meisten Szenen vergleichbar mit der Qualität, die mit MELZONIC (SAA4991) erreicht wird.

**[0087]** Der vorgeschlagene Bewegungsschätzalgorithmus hat keine Vektorbereichsbegrenzung auf Grund der Implementierung, was ein Vorteil gegenüber MELZONIC (SAA4991) ist.

**[0088]** Es ist ein Verfahren zum Erweitern globaler Bewegungsschätzalgorithmen zu der Schätzung von Bewegungsparametern in einer geschichteten Darstellung präsentiert worden. Eine feste Anzahl Parameterschätzer läuft parallel, die je Parameter für eine einzige Bildschicht berechnen. Eine Segmentierung ordnet je-

der Teil des Bildes der richtigen Schicht zu.

**[0089]** Obschon die Schätzer parallel arbeiten, gibt es dennoch eine gewisse Hierarchie. Jeder Schätzer, ausgenommen der höchste in der Hierarchie, arbeitet an Bildteilen, wo Schätzer mit einer höheren Bewertung in der Hierarchie in dem vorhergehenden Bild nicht erfolgreich waren. Zweitens wird vermieden, dass jeder Schätzer verschmutzt wird durch Teile des Bildes, die durch Schätzer mit einer niedrigeren Bewertung in der Hierarchie, besser bedeckt werden.

**[0090]** Versuche zeigen, dass das vorliegende Ergebnis nicht weit entfernt ist von demjenigen, was mit einem speziellen Entwurf erreicht wird: Natürliche Bewegung mit dem MELZONIC (SAA4991). Der Algorithmus ist aber viel geeigneter zur Implementierung in Software bei einem Prozessor wie TriMedia.

**[0091]** Zum Schluss können die Algorithmen nach der vorliegenden Erfindung interessant sein für andere Anwendungsbereiche von Bewegungsschätzung, wie Videokompression und Codierung, Videokommentar und Indexierung, Objektverfolgung und Rauschunterdrückung.

**[0092]** Ein erster Aspekt der Erfindung kann wie folgt zusammengefasst werden. Es wird ein neues Verfahren zur globalen bewegungskompensierten Aufwärtsmischung beschrieben und es werden Möglichkeiten angeboten um den Vorschlag zur Anwendung in einer geschichteten Videodarstellung zu erweitern. Im Wesentlichen werden Parameter geschätzt, die die globale Bewegung beschreiben, vorzugsweise unter Anwendung einer rekursiven Annäherung. Die mit diesen Parametern erzeugten örtlichen Bewegungsvektoren werden verwendet zum Erzeugen eines bewegungskompensierten Bildes. Gleichzeitig wird eine Segmentierungsmaske über ein Bild reduzierter Größe berechnet, wobei der Ausgang zum Schalten zwischen verschiedenen Parametersätzen oder Implementierungsverfahren benutzt wird. Eine interessante preisgünstige Version wird detailliert beschrieben, die geeignet ist zum Implementieren in aktuell verfügbaren völlig programmierbaren Anordnungen (natürliche Bewegung in einem TriMedia).

**[0093]** Die nachfolgenden hervorspringenden Merkmale bevorzugter Ausführungsformen sind bemerkenswert. Ein Verfahren und eine Anordnung, das dieses Verfahren verwirklicht, zur Bewegungskompensation von Videodaten, das die nachfolgenden Elemente umfasst:

Wenigstens zwei Mittel zum Berechnen globaler Bewegungsparameter aus den Eingangs-Videodaten, Interpolationsmittel zum Berechnen von Ausgangs-Videodaten aus einem oder mehreren Eingangsteilbildern, in Abhängigkeit von den wenigstens zwei Sätzen globaler Bewegungsparameter, wobei eines der wenigstens zwei Mittel zum Berechnen globaler Bewegungsparameter Parameter schafft, die eine Nullgeschwindigkeit für das ganze Bild angeben, ungeachtet des Bildinhaltes.

**[0094]** Vorzugsweise ist das Interpolationsmittel ein Filter statistischer Ordnung, beispielsweise ein Mittelwertfilter mit drei Abgriffen, das ein Ausgangspixel erzeugt, entweder aus:

- dem entsprechenden Pixel in dem vorhergehenden Teilbild, dem entsprechenden Pixel in dem nächsten Teilbild und dem bewegungskompensierten Mittelwert der beiden Nachbarbildern (erste Möglichkeit), oder
- dem bewegungskompensierten Pixel aus dem vorhergehenden Teilbild, dem bewegungskompensierten Pixel aus dem nächsten Teilbild, und dem nicht bewegungskompensierten Mittelwert der beiden Nachbarbildern (zweite Möglichkeit).

**[0095]** Vorzugsweise aktiviert ein Segmentierungssignal die erste Entscheidung, in dem Fall, dass der örtliche Bewegungsvektor, der aus dem zweiten Parametersatz berechnet worden ist, die beste Übereinstimmung mit dem Eingangsbild reduzierter Größe ergibt.

**[0096]** Vorzugsweise wird das Segmentierungssignal von einer Version des Eingangssignals mit reduzierter Größe hergeleitet.

**[0097]** Ein Verfahren und eine Anordnung zur Verwirklichung dieses Verfahrens zur Bewegungskompensation von Videodaten, das die nachfolgenden Elemente umfasst:

- wenigstens zwei Mittel zum Berechnen globaler Bewegungsparameter aus den Eingangs-Videodaten,
- Interpolationsmittel zum Berechnen von Ausgangs-Videodaten aus einem oder mehreren Eingangsteilbildern, in Abhängigkeit von den wenigstens zwei Sätzen globaler Bewegungsparameter und einem Segmentierungssignal, hergeleitet von einer Version des Eingangssignals reduzierter Größe.

**[0098]** Vorzugsweise schafft eines der globalen Bewegungsparameterberechnungsmittel Parameter, die eine

Nullgeschwindigkeit für das ganze Bild angeben, ungeachtet des Bildinhaltes.

**[0099]** Ein zweiter Aspekt der vorliegenden Erfindung kann wie folgt zusammengefasst werden. Ein neues Verfahren zu globalen Bewegungsparameterschätzung wird beschrieben. Im Wesentlichen werden Parameter, die die globale Bewegung in dem Bild beschreiben, geschätzt, und zwar unter Anwendung einer rekursiven Annäherung, d.h. es wird eine frühere n-dimensionale (n ist die Anzahl Parameter in dem Bewegungsmodell) Schätzung angewandt als eine Prädiktion, zu der (n-dimensionale) Aktualisierungsvektoren hinzu addiert werden. Der Ausgangsparametervektor ist derjenige, der zudem kleinsten Übereinstimmungsfehler führt. Die extrem geringe Komplexität des Algorithmus und die hohe Qualität machen es sehr interessant für künftigen Gebrauch in Fernseh- und Multimedia-Applikationen, die möglicherweise in völlig programmierbaren Anordnungen, wie TriMedia laufen.

**[0100]** Die nächsten hervorspringenden Merkmale einer bevorzugten Ausführungsform sind bemerkenswert. Ein Verfahren und eine Anordnung zum Verwirklichen dieses Verfahrens zum Schätzen von Bewegungsparametern (des Parametervektors) einer Bildsequenz umfasst die nachfolgenden Elemente:

- Mittel zum Liefern eines Prädiktionsparametervektors, d.h. eine vorher berechnete Bewegungsparameterschätzung,
- Mittel zum Selektieren wenigstens eines Aktualisierungsparametervektors aus einem Aktualisierungssatz,
- Mittel zum Addieren des genannten Prädiktionsvektors zu dem genannten wenigstens einen Aktualisierungsvektor,
- Mittel zum berechnen der Qualität (Kostenfunktion) der resultierenden, wenigstens zwei Parametervektoren, und zwar unter Anwendung von Daten aus wenigstens zwei Teilbildern,
- Mittel zum Selektieren des besten Vektors aus den oben genannten wenigstens zwei Parametervektoren auf Basis deren Qualität,
- Mittel zum Ausliefern des selektierten Parametervektors als Bewegungsparameterschätzung.

**[0101]** Vorzugsweise werden Strafen, eine zeitliche Filterung und eine zeitliche und/oder räumliche Prädiktion angewandt.

**[0102]** Ein dritter Aspekt der vorliegenden Erfindung kann wie folgt zusammengefasst werden. Es wird ein Verfahren zum Schätzen von Bewegungsparametern aus Videodaten beschrieben. Die vorliegende Erfindung ermöglicht eine zeitliche prädiktive Bewegungsschätzung an Videodaten, und zwar auf Grund einfacher Bildratenumwandlungstechniken (Wiederholung des jüngsten Bildes), zeigt eine unregelmäßige Bewegung. Die Lösung besteht aus der Verwendung vieler zeitlicher Prädiktionsvektoren, genommen aus mehreren vorhergehenden Bilderpaaren. Diese Lösung ist wirtschaftlich berechtigt, insbesondere bei objektbasierten Bewegungsschätzern, wobei die Anzahl zu speichernder Bewegungsvektoren sehr gering ist. Eine Software-Version des Algorithmus ist dargestellt um in Echtzeit in dem Philips TM1000 (TriMedia) Prozessor zu laufen.

**[0103]** Die nachfolgenden hervorragenden Merkmale einer bevorzugten Ausführungsform sind bemerkenswert. Ein Verfahren, und eine Anordnung zum Verwirklichen dieses Verfahrens, zum Schätzen von Bewegungsparametervektoren aus Videodaten, was für wenigstens einige Bildteile wenigstens zwei (zeitliche) Prädiktionsvektoren ergibt, geschätzt aus Daten verschiedener vorhergehender Bilderpaare. Vorzugsweise sind die oben genannten wenigstens zwei Prädiktionsvektoren Kandidaten in einem Vektorselektionsprozess, wobei der Ausgangsvektor für ein Bild (einen Bildteil) ermittelt wird. Auf vorteilhafte Weise wird entsprechend einer Kriteriumfunktion der bessere der oben genannten wenigstens zwei Prädiktionsvektoren als Basis zum Berechnen von Kandidatvektoren verwendet (beispielsweise ein Aktualisierungsprozess), die Eingang eines Vektorselektionsprozesses sind, wobei der Ausgangsvektor für ein Bild (einen Bildteil) ermittelt wird. Vorzugsweise wird die Entscheidungsinformation (welcher der wenigstens zwei Prädiktionsvektoren der bessere ist, entsprechend einer Kriteriumfunktion) über eine Anzahl aufeinander folgender Bilder (Bildteile), zum Detektieren von Bildwiederholungsmustern verwendet (beispielsweise 3-2 Transport und 2-2 Transport von Filmmaterial, aber auch andere Muster auf Grund von Quelle-Ziel-Bildfrequenzfehlanspassungen).

**[0104]** Ein vierter Aspekt der vorliegenden Erfindung bezieht sich auch eine Verbindung von Bewegungs-schätzung und Segmentierung von Videodaten und kann wie folgt zusammengefasst werden. Ein Verfahren zum Segmentieren eines Bildes in eine feste Anzahl Schichten und Schätzung von Bewegungsparametern für einzelne Schichten wird beschrieben. Die vorliegende Erfindung schafft eine Lösung für das "Huhn-Ei-Problem" einer kombinierten Bewegungsschätzung und Segmentierung. Die Lösung besteht aus einem Gewichtsprozess, der die Verunreinigung des Optimierungskriteriums eines Parameterschätzers für eine bestimmte Schicht durch Information, erledigt von den anderen Parameterschätzern, die parallel laufen, begrenzt. Die



äußerst geringe Komplexität des Algorithmus und die hohe Qualität machen es sehr interessant für künftige Anwendung in Fernseh- und Multimedia-Applikationen. Eine Software-Version des Algorithmus Informationssignal dargestellt um in Echtzeit in dem Philips TM1000 (TriMedia) Prozessor zu laufen.

**[0105]** In einer bevorzugten Ausführungsform wird ein geschichteter Bewegungsschätzungsalgorithmus vorgeschlagen, der eine quasi-simultane Bewegungsschätzung/Segmentierung bis zu einer maximalen Anzahl Schichten ermöglicht. Die Schätzung führt zu einem einzigen Bewegungsparametersatz je Schicht, und zu einer Segmentierungsabbildung, die diese Sätze verschiedenen Teilen des Bildes (Bewegungsschichten) zuordnet. Bewegung in einer Schicht wird mit einem Maximum von vier Parametern modelliert, die imstande sind, eine Pan-Bewegung, Tilt-Bewegung oder Zoom-Bewegung zu beschreiben. Das Konzept zeigt eine gewisse Hierarchie, d.h. eine Rangordnung der Bewegungsschichten. Auf diese Weise schließt die Bewegungsparameterschätzung, die sich auf eine einzige Schicht bezieht, Teile des Bildes aus, die durch eine Schicht beschrieben worden sind, die eine höhere Bewertung in der Hierarchie hat und nicht durch Teile des Bildes verschmutzt wird, die durch Schichten mit einer niedrigeren Bewertung in der Hierarchie besser beschrieben worden sind. Der Konzept führt zu einer, sehr niedrigen Vorgangszahlwert. Er hat sich als sehr gut erwiesen, sogar in kritischen Umwandlungsapplikationen, insbesondere in Bildratenaufwärtsmischung. Es ist eine Variante mit drei Schichten geplant um in Echtzeit in einem Philips Trimedia Prozessor zu laufen.

**[0106]** Die nachfolgenden hervorragenden Merkmale bevorzugter Ausführungsformen sind bemerkenswert. Ein Verfahren und eine Anordnung zum Verwirklichen dieses Verfahrens zum Segmentieren eines Bildes in einen Satz von Schichten mit einer gewissen Rangordnung und zum Schätzen von Bewegungsparametern für jede Schicht, weist die nachfolgenden Verfahrensschritte auf:

- einen Parameterschätzungsprozess (PE) für jede Schicht in dem aktuellen Bild, wobei eine Kriteriumfunktion optimiert wird, und zwar auf Basis von (Gruppen von) Pixeln von wenigstens zwei Bildern,
- einen Segmentierungsprozess (SP), wobei Bildteilen Bewegungsparametersätze zugeordnet werden,
- einen Gewichtungsprozess (WP) zum Definieren des einzelnen Effektes von Information aus verschiedenen Teilen auf die Kriteriumfunktion eines Bewegungsparameterschätzers, in dem der WP
- den Effekt von Information von denjenigen Bildteilen reduziert oder eliminiert, die ein erstes Kriterium erfüllen, und
- den Effekt von Information von denjenigen Bildteilen steigert, die ein zweites Kriterium erfüllen.

**[0107]** Vorzugsweise wird das erste Kriterium erfüllt, wenn bei einer vorhergehenden Wiederholung des Algorithmus auf dasselbe oder auf ein anderes Bilderpaar, die Bildteile in Gebiete fallen, die auf entsprechende Weise von beliebigen Bewegungsparametersätzen beschrieben wurden, geschätzt von PEen, die bei Schichten mit einer höheren Bewertung aktiv sind. "Auf entsprechende Weise" bedeutet in diesem Zusammenhang, dass eine Fehlerfunktion, die die Parametersätze der PEen verwendet, die bei Schichten mit einer höheren Bewertung in der Hierarchie aktiv sind, unter einer Schwelle bleiben (entweder fest oder angepasst, beispielsweise an einen Mittelwertfehler).

**[0108]** Vorzugsweise wird das zweite Kriterium erfüllt, wenn in einer vorhergehenden Wiederholung des Algorithmus an demselben oder einem anderen Bilderpaar, die Bildteile in Gebiete fielen, die durch die Bewegungsparametersätze beschrieben wurden, die von diesem bestimmten PE Bewegungsparametersätze geschätzt wurden. "Am besten" bedeutet in diesem Zusammenhang, dass eine Fehlerfunktion, wobei die Parametersätze des bestimmten PEs, niedriger ist als die eines der anderen PEen.

**[0109]** Vorzugsweise ist diese Fehlerfunktion auf der bewegungskompensierten Differenz zwischen den Pixeln in dem aktuellen Teilbild und den entsprechenden Pixeln in dem vorhergehenden Teilbild basiert, und zwar unter Verwendung der Parametersätze, die bewertet werden müssen (direktes Verfahren).

**[0110]** Vorzugsweise basiert diese Fehlerfunktion auf der Differenz zwischen Bewegungsvektoren, berechnet mit irgendeinem Verfahren, und Bewegungsvektoren, herrührend aus dem Bewegungsparametersatz, der bewertet werden soll (indirektes Verfahren).

**[0111]** Vorzugsweise werden Bildteile, die dem ersten Kriterium entsprechen, in der Fehlerfunktion eines bestimmten PEs eliminiert und wird dieses erste Kriterium derart angepasst, dass das Bildgebiet, in dem die Kriteriumfunktion berechnet wird, in einem bestimmten Bereich bleibt (Steuerschleife um die maximal verfügbare Verarbeitungsleistung auf effiziente Weise zu benutzen).

**[0112]** Vorzugsweise arbeitet der PE und/oder der SP und/oder der WPP an herunter gemischten und/oder unterabgetasteten Videodaten.

**[0113]** Ein Verfahren und eine Anordnung zum Verwirklichen dieses Verfahrens zum Segmentieren eines Bildes in einen Rangordnungssatz von Schichten und zum Schätzen von Bewegungsparametern für jede Schicht, der die nachfolgenden Verfahrensschritte umfasst:

- einen iterativen Parameterschätzungsprozess für jede Schicht in dem aktuellen Bild zum Optimieren einer Kriteriumfunktion auf Basis selektierter (Gruppen von) Pixel(n) von wenigstens zwei Bildern, einen Segmentierungsprozess, wobei jedem Teil des Bildes einer der Bewegungsparametersätze zugeordnet wird,
- einen Selektionsprozess um zu definieren, auf welche (Gruppen von) Pixel(n) von den wenigstens zwei Bildern der (die) Bewegungsparameterschätzer ihre Kriteriumfunktion optimieren soll(en),
- wobei der Parameterschätzungsprozess dieses Prozess an den Daten öfter als die anderen Prozesse wiederholt.

**[0114]** Vorzugsweise selektiert der Selektionsprozess für eine bestimmte Schicht diejenigen (Gruppen von) Pixel(n), für die die Parametersätze von Schichten mit einer höheren Bewegung in der Hierarchie, in einem vorhergehenden Bild keine befriedigenden Ergebnisse entsprechend einer Regel gebracht hat. Vorzugsweise betrifft diese Regel den Vergleich einer Fehler(summe) von (Gruppen von) Pixeln mit einer festen oder adaptiven Schwelle.

**[0115]** Vorzugsweise wird die Schwelle derart angepasst, dass die Anzahl Pixel, für die die Kriteriumfunktion berechnet wird, innerhalb eines bestimmten Bereichs bleibt.

**[0116]** Vorzugsweise ist die Kriteriumsumme ein summierter Fehler, berechnet zwischen selektierten (Gruppen von) Pixeln von dem vorhergehenden Bild und entsprechenden Pixeln von dem aktuellen Bild, kompensiert auf Bewegung entsprechend den Kandidatbewegungsparametern.

**[0117]** Vorzugsweise wird der Beitrag der selektierten Pixel zu der Kriteriumfunktion gewichtet, und zwar abhängig davon, welcher Schicht sie zugeordnet wurde (in dem vorhergehenden Bild).

**[0118]** Vorzugsweise wird der Beitrag der selektierten Pixel zu der Kriteriumfunktion gesteigert, wenn sie vorher derselben Schicht zugeordnet wurden.

**[0119]** Ein Verfahren und eine Anordnung zum Verwirklichen dieses Verfahrens zum Segmentieren eines Bildes in einen bewerteten Satz von Schichten und zum Schätzen von Bewegungsparametern für jede Schicht, wobei dieses Verfahren die nachfolgenden Verfahrensschritte umfasst:

- einen Parameterschätzungsprozess (PE) für jede Schicht in dem aktuellen Bild zum Optimieren einer Kriteriumfunktion auf Basis von (Gruppen von) Pixeln von wenigstens zwei Bildern,
- einen Segmentierungsprozess (SP) zum Zuordnen eines der Bewegungsparametersätze zu jedem Teil des Bildes,
- einen Selektionsprozess um zu definieren, welche (Gruppen von) Pixel(n) von den wenigstens zwei Bildern der (die) Bewegungsparameterschätzer ihre Kriteriumfunktion optimieren soll(en), wobei der Selektionsprozess es ermöglicht, dass nur ein kleiner Teil der Pixel einen Beitrag zu der von den PEs optimierten Kriteriumfunktion liefert, ungeachtet der Größe der Schicht, der diese Parameter von dem Segmentierungsprozess zugeordnet werden.

**[0120]** Obschon ursprünglich entworfen um als Applikation in dem Philips TriMedia Prozessor zu laufen, sind mehr Applikationen möglich. Insbesondere kann der Konzept für VGA-Controller der nächsten Generation entworfen werden. Da dies speziellen Silizium ist, sind die Kosten vernachlässigbar. Ein derartiger VGA-Controller kann eine verbesserte Leistung haben, und zwar im Vergleich zu der TriMedia-Lösung, weil in speziellem Silizium viel mehr Verarbeitungsleistung verfügbar ist. Weiterhin wird erwartet, dass wenn mehr als zwei parallele Parameterschätzer angewandt werden, die Leistung auf einen Pegel gebracht werden kann, der ggf. bessert ist als der der aktuellen "High-End"-Lösungen zu möglicherweise geringeren Kosten.

**[0121]** Es sei bemerkt, dass die oben genannten Ausführungsformen die vorliegende Erfindung erläutern statt begrenzen, und dass der Fachmann imstande sein wird, im Rahmen der beiliegenden Patentansprüche viele alternative Ausführungsformen zu entwerfen. In den Ansprüchen sollen eingeklammerte Bezugszeichen nicht als den Anspruch begrenzend betrachtet werden. Die vorliegende Erfindung kann mit Hilfe von Hardware mit verschiedenen einzelnen Elementen, und mit Hilfe eines auf geeignete Art und Weise programmierten Computers implementiert werden. In dem Anspruchsanspruch können verschiedene dieser Mittel von ein und demselben Hardware-Item verkörpert werden. In den Patentansprüchen schließt der Ausdruck "enthalten" das Vorhandensein anderer Elemente oder Verfahrensschritte als diejenigen, die hier in einem Anspruch genannt

werden, nicht aus.

# Bezugsmaterial:

- [1] US-A-5,534,946 (Attorneys' docket PHN 14,066).
- [2] A. M. Tekalp, "Digital Video Processing", Prentice Hall Signal Processing Series, ISBN 0-13190075-7, pp. 200–203.
- [3] G. de Haan, P.W.A.C. Biezen, H. Huijgen and O.A. Ojo, "True Motion Estimation with 3-D Recursive Search Block-Matching", IEEE Transactions on Circuits and Systems for Video Technology, Vol.3, October 1993, pp. 368–388.
- [4] G. de Haan, P.W.A.C. Biezen, H. Huijgen, and O.A. Ojo, "Graceful Degradation in Motion Compensated Field-Rate Conversion", in: Signal Processing of HDTV, V, L. Stenger, L. Chiariglione and M. Akgun (Eds.), Elsevier 1994, pp. 249–256.
- [5] WO-A-97/046022 (PCT/IB97/00548), published on 04.12.1997 and prior art under Art. 54(3) EPC (Attorneys' docket PHN 16,112).
- [6] G. de Haan, J. Kettenis, and B. Deloore, 'IC for Motion Compensated 100 Hz TV, with a Smooth Motion Movie-Mode', International Conference on Consumer Electronics, ICCE 95, June 1995, Chicago.
- [7] G. de Haan, P.W.A.C. Biezen, "Sub-pixel motion estimation with 3-D recursive search block-matching", Signal Processing: Image Communication 6 (1994), pp. 229–239.
- [8] WO-A-98/011721 (PCT/IB97/00884), published on 19.03.1998 and prior art under Art. 54(3) EPC (Attorneys' docket PHN 15,943).
- [9] US-A-5,495,300 (Attorneys' docket PHN 14,079).
- [10] G. Thomas, "Television motion measurement for DATV and other applications," BBC Research Report, no. BBC RD 1987/11, 1987.
- [11] R. Thoma and M. Bierling, "Motion compensating interpolation considering covered and uncovered background," Signal Processing: Image Communications 1, pp. 191–212, 1989.
- [12] F. Wang, D. Anastassiou, and A. Netravali, "Time-recursive deinterlacing for IDTV and pyramid coding," Signal Processing: Image Communications 2, pp. 365–374, 1990.
- [13] Kwon, Seo, Kim, and Kim, "A motion adaptive deinterlacing method," IEEE Transactions on Consumer Electronics, vol. 38, pp. 145–150, August 1992.
- [14] G. de Haan and H. Huijgen, "New algorithm for motion estimation," in Chiariglione [38], pp. 109–116.
- [15] G. de Haan and H. Huijgen, "Motion estimation for TV picture enhancement," in Signal Processing of HDTV III (H. Yasuda and L. Chiariglione, eds.), pp. 241–248, Elseviers Science Publishers B.V., 1992.
- [16] T. Reuter, "A modified block-matching algorithm with vector reliability checking and adaptive smoothing," in Third International Conference on Image Processing and its Applications, (England), University of Warwick, July 1989.
- [17] J. Konrad and E. Dubois, "A comparison of stochastic and deterministic solution methods in bayesian estimation of 2-d motion," Image and Vision Computing, vol. 8, pp. 304–317, November 1990.
- [18] J. Jain and A. Jain, "Displacement measurement and its application in interframe image coding," IEEE Transactions on Communications, COM-29, no. 12, 1981.
- [19] T. Koga, K. Linuma, A. Hirano, Y. Iijima, and T. Ishiguro, "Motion-compensated interframe coding for video conferencing," in IEEE, Proceedings of the NTC 81, G5.3.1., (New Orleans LA), 1981.
- [20] R. Srinivasan and K. Rao, "Predictive coding based on efficient motion estimation," IEEE Transactions on Communication, no. 8, pp. 888–896, 1985.
- [21] H. Musmann, P. Pirsch, and J. Grallert, "Advances in picture coding," Proceedings of the IEEE, vol. 73, pp. 523–548, April 1985.
- [22] A. Netravali and J. Robbins, "Motion compensated television coding," Bell Systems Technical Journal, no. 3, pp. 629–668, 1979.
- [23] M. Ziegler, "Hierarchical motion estimation using the phase correlation method in 140 Mbit/s HDTV-coding," in Chiariglione [38], pp. 131–137.
- [24] DE-C 40 23 449.
- [25] G. de Haan, J. Kettenis, and B. Deloore, "IC for motion compensated 100 Hz TV, with a smooth motion movie-mode," IEEE Transactions on Consumer Electronics, vol. 42, pp. 165–174, May 1996.
- [26] J. G. Choi and S.-D. Kim, "Multi-stage segmentation of optical flow field," Signal Processing, vol. 54, pp. 109–118, 1996.
- [27] D. Bagni, R. Lancini, P. Vicari, and S. Tubaro, "Motion estimation method using region-based segmentation methods," in Proc. International Workshop on HDTV '96, (Los Angeles), p. Sess. A2, October 1996.
- [28] D. LeQuang, Z. Zaccarin, and S. Caron, "Object-oriented coding using successive motion field segmentation and estimation," in Proc. International Conference on Image Processing (ICIP'95), (Washington D.C.), pp. 207–210, October 1995.

- [29] J. Y. A. Wang and E. H. Adelson, "Layered representation for motion analysis," in Proceedings of the IEEE Computer Vision and Pattern Recognition Conference, pp. 361–366, 1993.
- [30] P. Csillag and L. Boroczky, "Frame rate conversion based on acceleration and motion-based segmentation," in SPIE, vol. 2952, pp. 438–448, 1996.
- [31] F. Dufaux and F. Moscheni, "Motion estimation techniques for digital tv: a review and a new contribution," in Proceeding of the IEEE, vol. 83 n.6, pp. 858–876 1995.
- [32] S. Jeannin, "On the combination of a polynomial motion estimation with a hierarchical segmentation based video coding scheme," in Proc. International Conference on Image Processing (ICIP'96), (Lausanne, Switzerland), pp. 489–492, September 1996.
- [33] K. E. Matthews and N. M. Namazi, "Simultaneous motion parameter estimation and image segmentation using the EM algorithm," in Proc. International Conference on Image Processing (ICIP'95), (Washington D.C.), pp. 542–545, October 1995.
- [34] T. Darrel and D. Fleet, "Second-order method for occlusion relationships in motion layers," Tech. Rep. 314, MIT Media Laboratory Vision and Modelling Group, 1995.
- [35] H. S. Sawhney, S. Ayer, and M. Gorkani, "Model-based 2D-3D dominant motion estimation for mosaicing video representation:" On the net, 1995. A shorter version appeared in the IEEE Intl. Conf. on Computer Vision, Cambridge, MA, USA, June 1995.
- [36] H. S. Sawhney and S. Ayer, "Layered representation of motion video using robust maximum-likelihood estimation of mixture models and MDL encoding." On the net, 1995. A shorter version appeared in the IEEE Intl. Conf. on Computer Vision, Cambridge, MA, USA, June 1995.
- [37] J. Y. A. Wang and E. H. Adelson, "Spatio-temporal segmentation of video data," in Proceeding of the SPIE: Image and Video Processing II, vol. 2182, (San Jose), pp. 120–131, 1994.
- [38] L. Chiariglione, ed., Signal Processing of HDTV II, Elseviers Science Publishers B.V., 1990. This collection work includes references [14] and [23] and belongs to the general state of the art.

### Patentansprüche

1. Verfahren zum Schätzen von Bewegung in Videodaten, wobei das Verfahren die nachfolgenden Verfahrensschritte umfasst:

- das Liefern (PE1–PE<sub>n</sub>) von wenigstens zwei Bewegungsparametersätzen von Eingangs-Videodaten ( $n$ ,  $n-1$ ), wobei der eine Bewegungsparametersatz eine Null-Geschwindigkeit für alle Teile in einem Bild angibt, und jeder Bewegungsparametersatz entsprechende örtliche Übereinstimmungsfehler für alle Bildteile in einem Bild hat, wobei ein Bewegungsparametersatz ein Satz mit Parametern ist, die eine Bewegung in einem Bild beschreiben, wobei mit Hilfe dieses Bewegungsparametersatzes Bewegungsvektoren berechnet werden können,
- das Ermitteln eines eingestellten Bewegungsparametersatzes in Abhängigkeit von globalen Übereinstimmungsfehlern für jeden Bewegungsparametersatz anders als der Null-Geschwindigkeitssatz, und
- das Ermitteln (MPE) von Ausgangsbewegungsdaten ( $P(n)$ ) aus den genannten Videodaten ( $n$ ,  $n-1$ ) in Abhängigkeit von dem genannten einen Bewegungsparametersatz und dem genannten eingestellten Bewegungsparametersatz oder den genannten eingestellten Bewegungsparametersätzen, wobei die Bedeutung Bewegungsparametersatzes bei der Berechnung der genannten Ausgangsbewegungsdaten ( $P(n)$ ) von den örtlichen Übereinstimmungsfehlern der Bewegungsparametersätze abhängig ist.

2. Verfahren nach Anspruch 1, wobei der Lieferungsschritt der genannten Bewegungsparametersätze die nachfolgenden Schritte umfasst:

- das Liefern wenigstens eines vorher berechneten Bewegungsparametersatzes als wenigstens einen ersten Prädiktionsparametervektor (Cp1);
- das Hinzufügen (AD) wenigstens eines Aktualisierungsparametervektors (UP) zu dem wenigstens einen ersten Prädiktionsparametervektor zum Erhalten weiterer Prädiktionsparametervektoren (Cp2); und
- das Selektieren (PVC, EC, MIN, MUX) eines Vektors aus den genannten ersten und weiteren Prädiktionsparametervektoren (CP1, CP2).

3. Verfahren nach Anspruch 2, wobei der genannte Lieferungsschritt den Schritt der Lieferung von Bewegungsparametersätzen (Tp1–TP<sub>n</sub>) umfasst, berechnet für wenigstens zwei vorhergehende Teilbilder.

4. Verfahren nach Anspruch 1, wobei die genannten globalen Übereinstimmungsfehler für Bildteile entsprechend den Bildteilen zugeordneten Gewichtungen berechnet werden.

5. Verfahren nach Anspruch 4, wobei für eine bestimmte eingestellte Bewegungsparametersatzermittlung (PEi) die genannten Gewichtungen für diejenigen Teile erhöht werden, für die der bestimmte Bewegungspara-

metersatz den niedrigsten örtlichen Übereinstimmungsfehler im Vergleich mit den örtlichen Übereinstimmungsfehlern anderer eingestellter Bewegungsparametersatzermittlungen (PE1–PE<sub>n</sub>) hat.

6. Verfahren nach Anspruch 4, wobei für eine bestimmte eingestellte Bewegungsparametersatzermittlung (PE<sub>i</sub>) die genannten Gewichtungen für diejenigen Teile verringert werden, für die der örtliche Übereinstimmungsfehler einer anderen eingestellten Bewegungsparametersatzermittlung (PE1–PE<sub>n</sub>) einen bestimmten Schwellenwert unterschreitet.

7. Verfahren zur Bewegungskompensation von Videodaten, wobei das Verfahren die nachfolgenden Verfahrensschritte umfasst:

- das Liefern (PE1–PE<sub>n</sub>) von wenigstens zwei Bewegungsparametersätzen von Eingangs-Videodaten ( $n$ ,  $n-1$ ), wobei der eine Bewegungsparametersatz eine Null-Geschwindigkeit für alle Bildteile in einem Bild angibt, und jeder Bewegungsparametersatz entsprechende örtliche Übereinstimmungsfehler für alle Bildteile in einem Bild hat, wobei ein Bewegungsparametersatz ein Satz mit Parametern ist, die eine Bewegung in einem Bild beschreiben, wobei mit Hilfe dieses Bewegungsparametersatzes Bewegungsvektoren berechnet werden können,
- das Ermitteln eines eingestellten Bewegungsparametersatzes in Abhängigkeit von globalen Übereinstimmungsfehlern für jeden Bewegungsparametersatz anders als der Null-Geschwindigkeitssatz, und
- das Interpolieren (UC2) von Ausgangsvideodaten ( $n-1/2$ ) von den genannten Eingangsvideodaten ( $n$ ,  $n-1$ ) in Abhängigkeit von dem genannten einen Bewegungsparametersatz und dem genannten eingestellten Bewegungsparametersatz oder den eingestellten Bewegungsparametersätzen, wobei die Bedeutung jedes Bewegungsparametersatzes bei der Berechnung der genannten Ausgangsbewegungsdaten ( $n-1/2$ ) von den örtlichen Übereinstimmungsfehlern der Bewegungsparametersätze abhängig ist.

8. Verfahren nach Anspruch 7, wobei der genannte Interpolationsschritt Folgendes liefert:

- einen Medianwert (MED2) eines entsprechenden Pixels ( $l$ ) in einem vorhergehenden Teilbild ( $n-1$ ), eines entsprechenden Pixels ( $r$ ) in einem nächsten Teilbild ( $n$ ), und einen bewegungskompensierten Mittelwert ( $mcav$ ) aus den beiden vorhergehenden und nächsten Teilbildern ( $n-1$  bzw.  $n$ ), wenn der Übereinstimmungsfehler des Bewegungsvektors ( $D$ ), verwendet zur Berechnung des genannten bewegungskompensierten Mittelwertes ( $mcav$ ) den Übereinstimmungsfehler des Null-Bewegungsvektors übersteigt, oder
- einen Medianwert (MED1) eines bewegungskompensierten Pixels ( $mcl$ ) aus dem vorhergehenden Teilbild ( $n-1$ ), eines bewegungskompensierten Pixels ( $mcr$ ) aus dem nächsten Teilbild ( $n$ ), und einen nicht-bewegungskompensierten Mittelwert ( $av$ ) aus den beiden vorhergehenden und nächsten Teilbildern ( $n-1$  bzw.  $n$ ), wenn der Übereinstimmungsfehler des Bewegungsvektors ( $D$ ), verwendet zum Liefern der genannten bewegungskompensierten Pixel, den Übereinstimmungsfehler des Null-Bewegungsvektors unterschreitet.

9. Anordnung zur Bewegungsschätzung in Videodaten, wobei die Anordnung die nachfolgenden Elemente umfasst:

- Mittel zum Liefern (PE1–PE<sub>n</sub>) von wenigstens zwei Bewegungsparametersätzen von Eingangs-Videodaten ( $n$ ,  $n-1$ ), wobei der eine Bewegungsparametersatz eine Null-Geschwindigkeit für alle Bildteile in einem Bild angibt, und jeder Bewegungsparametersatz entsprechende örtliche Übereinstimmungsfehler für alle Bildteile in einem Bild hat, wobei ein Bewegungsparametersatz ein Satz mit Parametern ist, die eine Bewegung in einem Bild beschreiben, wobei mit Hilfe dieses Bewegungsparametersatzes Bewegungsvektoren berechnet werden können,
- Mittel zum Ermitteln eines eingestellten Bewegungsparametersatzes in Abhängigkeit von globalen Übereinstimmungsfehlern für jeden Bewegungsparametersatz anders als der Null-Geschwindigkeitssatz, und
- Mittel zum Ermitteln (MPE) von Ausgangsbewegungsdaten ( $P(n)$ ) aus den genannten Videodaten ( $n$ ,  $n-1$ ) in Abhängigkeit von dem genannten einen Bewegungsparametersatz und dem genannten eingestellten Bewegungsparametersatz oder den genannten eingestellten Bewegungsparametersätzen, wobei die Bedeutung Bewegungsparametersatzes bei der Berechnung der genannten Ausgangsbewegungsdaten ( $P(n)$ ) von den örtlichen Übereinstimmungsfehlern der Bewegungsparametersätze abhängig ist.

10. Anordnung zur Bewegungskompensation von Videodaten, wobei diese Anordnung die nachfolgenden Elemente umfasst:

- Mittel zum Liefern (PE1–PE<sub>n</sub>) von wenigstens zwei Bewegungsparametersätzen von Eingangs-Videodaten ( $n$ ,  $n-1$ ), wobei der eine Bewegungsparametersatz eine Null-Geschwindigkeit für alle Bildteile in einem Bild angibt, und jeder Bewegungsparametersatz entsprechende örtliche Übereinstimmungsfehler für alle Bildteile in einem Bild hat, wobei ein Bewegungsparametersatz ein Satz mit Parametern ist, die eine Bewegung in einem Bild beschreiben, wobei mit Hilfe dieses Bewegungsparametersatzes Bewegungsvektoren berechnet werden können,

- Mittel zum Ermitteln eines eingestellten Bewegungsparametersatzes in Abhängigkeit von globalen Übereinstimmungsfehlern für jeden Bewegungsparametersatz anders als der Null-Geschwindigkeitssatz, und
- Mittel zum Interpolieren (UC2) von Ausgangsvideodaten ( $n-1/2$ ) von den genannten Eingangsvideodaten ( $n$ ,  $n-1$ ) in Abhängigkeit von dem genannten einen Bewegungsparametersatz und dem genannten eingestellten Bewegungsparametersatz oder den eingestellten Bewegungsparametersätzen, wobei die Bedeutung jedes Bewegungsparametersatzes bei der Berechnung der genannten Ausgangsbewegungsdaten ( $n-1/2$ ) von den örtlichen Übereinstimmungsfehlern der Bewegungsparametersätze abhängig ist.

11. Videowiedergabeanordnung, die Folgendes umfasst:

- eine Anordnung zur Bewegungskompensation von Videodaten nach Anspruch 10,
- eine Wiedergabeeinheit (D) zur Wiedergabe der genannten Eingangs-Videodaten ( $n$ ,  $n-1$ ) und der genannten Ausgangs-Videodaten ( $n-1/2$ ).

Es folgen 2 Blatt Zeichnungen

Anhängende Zeichnungen

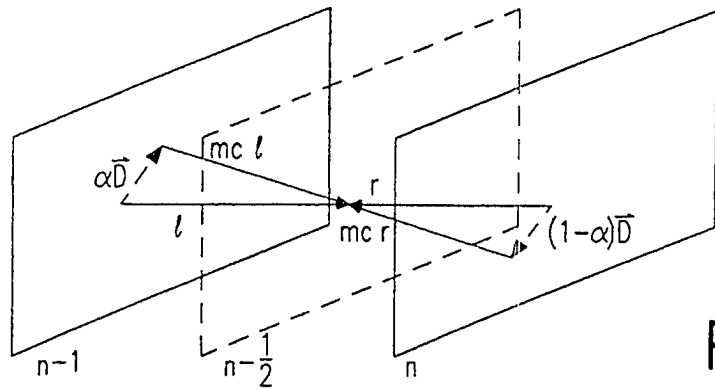


FIG. 1

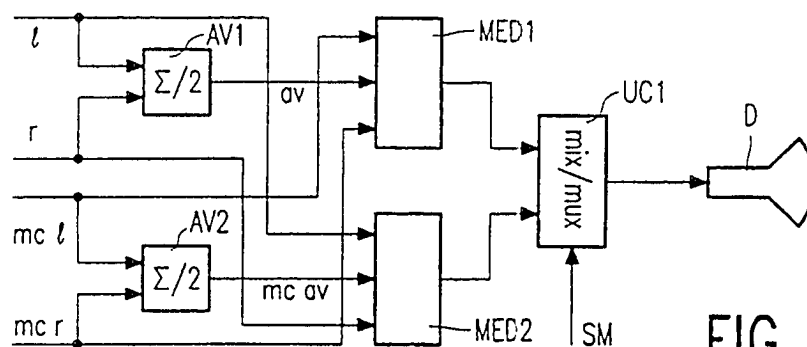


FIG. 2

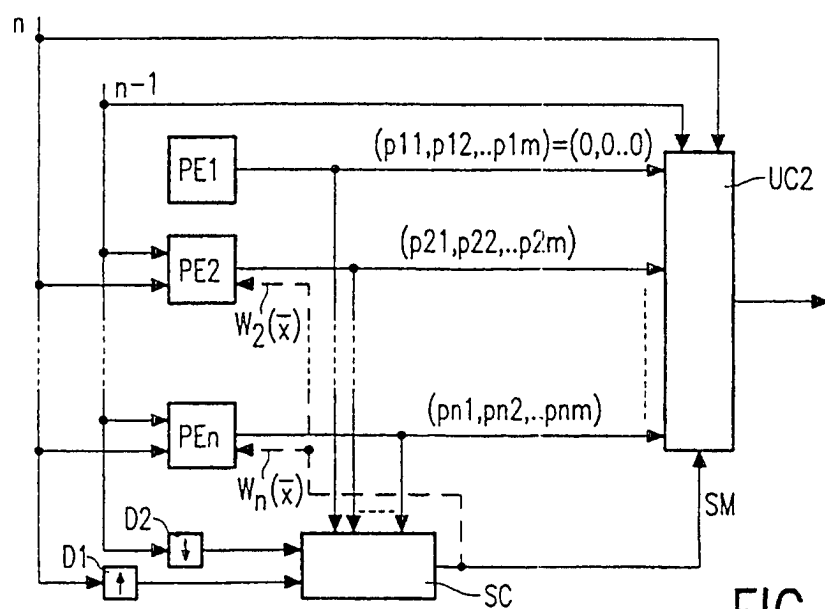


FIG. 3



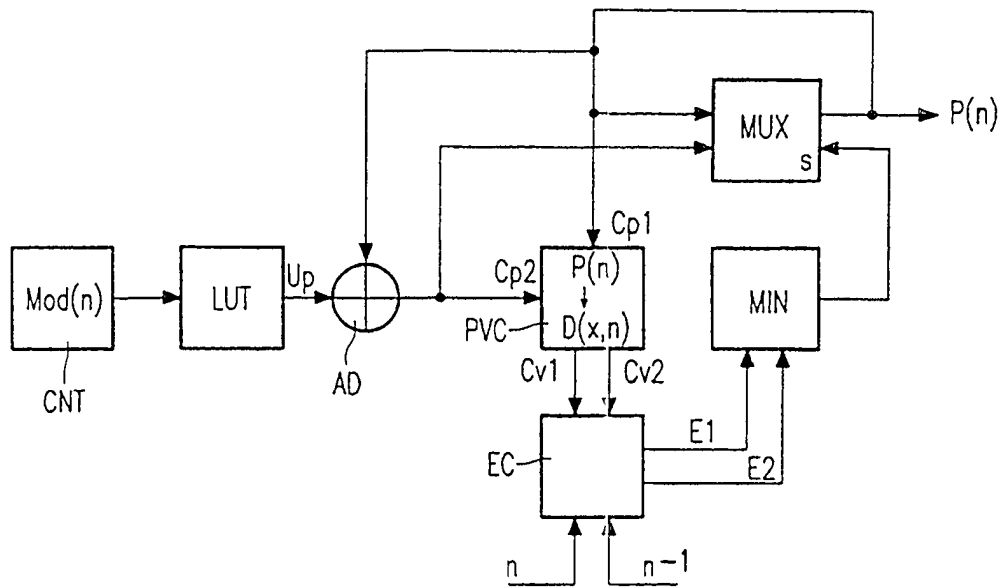


FIG. 4

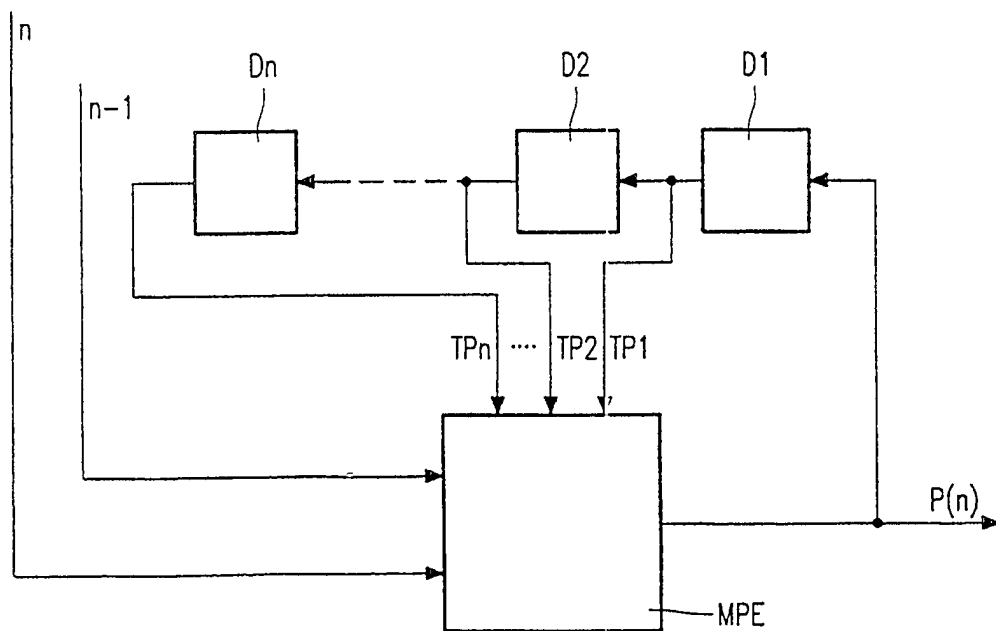


FIG. 5