

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

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
12. Juni 2003 (12.06.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/049452 A2

(51) Internationale Patentklassifikation⁷: **H04N 7/34, 7/26**

(21) Internationales Aktenzeichen: PCT/DE02/03390

(22) Internationales Anmeldedatum:
12. September 2002 (12.09.2002)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
101 58 658.2 30. November 2001 (30.11.2001) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von
US): **ROBERT BOSCH GMBH** [DE/DE]; Postfach 30 02
20, 70442 Stuttgart (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **DAHLHOFF, Achim**
[DE/DE]; Bremenberg 24, 52072 Aachen (DE). **WIEN,**
Mathias [DE/DE]; Mauerstrasse 7, 52064 Aachen (DE).

(81) Bestimmungsstaaten (national): CN, JP, KR, US.

(84) Bestimmungsstaaten (regional): europäisches Patent (AT,
BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR,
IE, IT, LU, MC, NL, PT, SE, SK, TR).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu
veröffentlichen nach Erhalt des Berichts

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

(54) Title: METHOD FOR CONDUCTING A DIRECTED PREDICTION OF AN IMAGE BLOCK

(54) Bezeichnung: VERFAHREN ZUR GERICHTETEN PRÄDIKTION EINES BILDBLOCKES

I	A	B	C	D
E	a	b	c	d
F	e	f	g	h
G	i	j	k	l
H	m	n	o	p

(57) Abstract: Edge picture elements (A, B, C, D; E, F, G, H; I) of already predicted neighboring image blocks (A', B') are used in order to conduct a directed prediction of an image block (C'). According to the direction of prediction, the picture elements used for the prediction are pre-filtered along a block edge of the image block (C') to be currently predicted in order to suppress prediction artifacts.

(57) Zusammenfassung: Zur gerichteten Prädiktion eines Bildblockes (C') werden Randbildpunkte (A, B, C, D; E, F, G, H; I) bereits prädizierter Nachbarbildblöcke (A', B') herangezogen. In Abhängigkeit der Prädiktionsrichtung werden die zur Prädiktion verwendeten Bildpunkte entlang eines Blockrandes des aktuell zu prädizierenden Bildblockes (C') vorgefiltert, um Prädiktionsartefakte zu

unterdrücken.

WO 03/049452 A2

5

10 Verfahren zur gerichteten Prädiktion eines Bildblockes

Stand der Technik

15 Die Erfindung geht aus von einem Verfahren zur gerichteten
Prädiktion eines aus Bildpunkten bestehenden Bildblockes,
insbesondere für die Intra-Prädiktion von Bildblöcken mit
einer Größe von mehr als 4 x 4 Bildpunkten.

20 Aus [1], [http://standard.pictel.com/ftp/video-
site/0109_San/](http://standard.pictel.com/ftp/video-site/0109_San/), insbesondere Seiten 9 bis 11, ist es
bekannt, für den Videocodierungs-Standard H.26L eine Intra-
Prädiktion auf 4 x 4 Pixel (Bildpunkt)-Blöcken oder auf 16 x
16 Pixel-Blöcken durchzuführen. Intra-Prädiktion bedeutet,
25 daß zur Generierung eines Prädiktionssignals nur bereits
übertragene Informationen eines aktuellen Bildes
herangezogen werden (z.B. wird das erste Bild einer Video-
Sequenz immer "Intra"codiert). Zur 4 x 4-Prädiktion stehen
bei [1] sechs Modi zur Verfügung, die sich in der Richtung,
aus der die Prädiktion erfolgt, unterscheiden. Dabei werden
30 die Randpixel bereits prädizierter (codierter)
Nachbarbildblöcke zur Prädiktion herangezogen.

35 Vorteile der Erfindung

Mit den Maßnahmen des Anspruchs 1 werden zur Prädiktion eines aktuellen Bildblockes ebenfalls Randbildpunkte mindestens eines bereits prädizierten Nachbarbildblockes herangezogen. Jedoch werden zusätzlich jene zur Prädiktion verwendeten Bildpunkte in Abhängigkeit der Prädiktionsrichtung entlang eines Blockrandes des aktuell zu prädizierenden Bildblockes vorgefiltert. Diese Vorfilterung führt zu verbesserten Prädiktionseigenschaften, insbesondere für Blöcke mit großer Kantenlänge, d.h. einer Kantenlänge von mehr als vier Bildpunkten.

Die Erfindung geht aus von der Erkenntnis, daß wenn eine gerichtete Prädiktion für Blockgrößen mit mehr als 4×4 Bildpunkten, z.B. 4×8 , 8×4 , 8×8 , 16×8 , ..., 16×16 eingesetzt werden soll, im prädizierten Bildblock sehr prägnante Strukturen entstehen, die zu starken Artefakten im rekonstruierten Bild führen. Durch die erfindungsgemäße Vorfilterung werden diese Artefakte durch Glättung des Prädiktionssignals verhindert, was zu einer subjektiven und objektiven Verbesserung des Codierverfahrens für die Übertragung der derart prädizierten Bildblöcke führt.

Mit den Maßnahmen des Anspruchs 2 werden die zur Prädiktion verwandten Bildpunkte eines Nachbarbildblockes zu einem Prädiktionsvektor zusammengefaßt, und es wird eine Tiefpaßfilterung entlang der Richtung des Prädiktionsvektors vorgenommen. Diese Maßnahme vereinfacht den Verarbeitungsaufwand gegenüber einer möglichen Einzelfilterung.

Durch die stärkste Gewichtung der Tiefpaßfilterung an der Nullposition gemäß Anspruch 3, d.h. an der Position eines gerade zu prädizierenden Bildpunktes, bleibt der Verlauf des Prädiktionsvektors erhalten, die Filterung führt aber zu

einer Glättung des Vektors. Der Einfluß der benachbarten Prädiktionswerte auf die aktuelle Position wird geringer, je weiter sie von der aktuellen Position entfernt sind. Diese Vorgehensweise dämpft die Artefaktbildung besonders wirkungsvoll bei ansteigenden Blockgrößen.

Die Maßnahmen des Anspruchs 4 tragen ebenfalls vorteilhaft zur Artefaktunterdrückung bei.

Die Ansprüche 5 bis 7 zeigen vorteilhafte Maßnahmen der Blockrandbehandlung auf. Insbesondere durch die Fortsetzung des Prädiktionsvektors mit seinen Randwerten vereinfacht sich der Verarbeitungs- und Speicheraufwand, was zur Steigerung der Codiereffizienz führt.

Zeichnungen

Anhand der Zeichnungen werden Ausführungsbeispiele der Erfindung erläutert. Es zeigen

Figur 1 einen zu prädizierenden Bildblock mit seinen Nachbarbildblöcken,

Figur 2 einen aktuell zu prädizierenden Bildblock mit seinen Einzelbildpunkten und den Bildpunkten der Nachbarbildblöcke an seinen Kanten,

Figur 3 die Prädiktionsrichtungen der verschiedenen Prädiktionsmodi,

Figur 4 einen Bildblock der Größe 4 x 4 Bildpunkte mit prädizierten Bildpunkten unter Verwendung von bereits prädizierten Bildpunkten eines Nachbarblockes,

Figur 5 einen Bildblock der Größe 8 x 8 Bildpunkte mit ungefilterter Prädiktion,

Figur 6 den Bildblock aus Figur 5, jedoch mit erfindungsgemäßer Vorfilterung.

Beschreibung von Ausführungsbeispielen

Bevor auf die eigentliche Erfindung eingegangen wird, wird zum besseren Verständnis die aus [1] bekannte richtungsabhängige Intra-Prädiktion näher erläutert. Die Intra-Prädiktion gemäß [1] wird auf 4 x 4-Pixel-Blöcken oder auf 16 x 16-Pixel-Blöcken durchgeführt. Zur 4 x 4-Prädiktion stehen sechs Modi zur Verfügung, die sich in der Richtung, aus der die Prädiktion erfolgt, unterscheiden. Dabei werden die Randpixel (Randbildpunkte) bereits codierter Nachbarbildblöcke zur Prädiktion herangezogen. Mögliche Nachbarblöcke A' und B' zu dem aktuell zu prädicierenden Bildblock C' sind in Figur 1 dargestellt. Figur 2 zeigt den aktuellen Bildblock C' mit seinen Bildpunkten a, b, c, ..., p mit den Randpixeln der Nachbarbildblöcke und zwar A, B, C, D des nach oben anschließenden Bildblockes A' bzw. E, F, G, H für den nach links anschließenden Nachbarbildblock B' sowie I des am linken oberen Eck anschließenden Nachbarbildblockes (in Figur 1 gestrichelt).

Neben der DC (Gleichanteile)-Prädiktion, in der der Mittelwert aus den Positionen A bis H als Prädiktion für alle Pixel des Bildblocks C' verwendet wird, stehen fünf Modi mit verschiedenen Prädiktionsrichtungen zur Verfügung. Neben horizontaler und vertikaler Prädiktion sind auch drei schräge Prädiktionsrichtungen möglich. Die Prädiktionsrichtungen sind in Figur 3 dargestellt. In Figur 4 ist ein Beispiel für die Prädiktion mit dem Modus 1 gegeben. In diesem Modus werden nur die Pixel A bis D zur Prädiktion verwendet. In Modus 1 erfolgt die Prädiktion wie folgt:

Präd. von Pixel	durch
a	$(A+B) / 2$
e	B
b,i	$(B+C) / 2$
f,m	C
c,j	$(C+D) / 2$
d,g,h,k,l,n,o,p	D

Die Prädiktion auf der Basis von 16 x 16 Pixel-Blöcken wird in Bildbereichen mit wenig Aktivität eingesetzt. Hier stehen die DC-Prädiktion, horizontal und vertikale Prädiktion sowie eine etwas aufwendigere "plane"- Prädiktion zur Verfügung. Für eine vollständige Beschreibung der Intra-Prädiktionsmethoden wird auf das H.26L-Test-Modell (TML) [1] verwiesen.

Die für die 4 x 4- Prädiktion vorgestellten Verfahren lassen sich direkt auf andere Blockgrößen erweitern. Hierzu werden die Prädiktionen entsprechend fortgesetzt.

Soll eine gerichtete Prädiktion der Art, wie sie für 4 x 4-Bildblöcke zuvor definiert wurde, für Bildblöcke mit mehr als 4 x 4 Bildpunkten, z.B. 4 x 8, 8 x 4, 8 x 8, 16 x 8, ..., 16 x 16, eingesetzt werden, so können im prädizierten Bildblock sehr prägnante Strukturen entstehen, die zu starken Artefakten im rekonstruierten Bild führen. Die Maßnahmen nach der Erfindung verhindern diese Artefakte und führen dadurch zu einer subjektiven wie objektiven Verbesserung des Codierungsverfahrens.

Die Beschreibung geht von den im TML [1] für die Intra-Codierung von 4 x 4-Bildblöcken verwendeten gerichteten Prädiktions-Modi aus. Es werden Bildblöcke der Größen 4 x 8, 8 x 4, 8 x 8, 8 x 16, 16 x 8 und 16 x 16 betrachtet. Die Filterung wird an den Kanten mit Länge > 4 Pixel, d.h. für

Kanten der Länge 8 oder 16 Pixel eingesetzt. Das Verfahren kann ebenso für andere Blockgrößen und Prädiktionsrichtungen verwendet werden.

- 5 Im folgenden werden die Pixel des aktuellen Bildblocks der Größe M Zeilen und N Spalten mit

$$P(i,j), \quad i = 0, \dots, N-1, j = 0, \dots, M-1$$

- 10 bezeichnet. Die zur Prädiktion verwendeten Pixel der Nachbarbildblöcke befinden sich an den Positionen $P(-1,j)$ und $P(i,-1)$. Sie werden im folgenden im Prädiktionsvektor $V(n)$ der Länge N_f zusammengefaßt. Im Fall der zuvor vorgestellten Modus-1- Prädiktion für 4 x 4 Blöcke ist also
15 mit $N_f = 4$:

$$V(n) = P(n,-1) = [A,B,C,D], \quad n = 0, \dots, 3.$$

- 20 Zur Prädiktion eines Bildblockes wird nicht $V(n)$, sondern eine gefilterte Version $V_f(n)$ verwendet. Zur Filterung wird ein (Tiefpaß-)Filter geeigneter Länge L herangezogen, z.B.:

Kanten- länge	Filter $h(n)$	Filterlänge L
8	7/32, 18/32, 7/32	3
16	3/32, 7/32, 12/32, 7/32, 3/32	5

- 25 Die Vorfilterung erfolgt in Abhängigkeit der Prädiktionsrichtung. Im vorgestellten Modus 1 entlang des oberen Blockrandes des aktuell zu prädizierenden Bildblockes C' . Die Vorfilterung wird entlang der Richtung des Prädiktionsvektors $V(n)$ vorgenommen.

- 7 -

Durch die Tiefpaßfilterung wird eine Glättung des Prädiktionssignals erreicht. Es bestehen mehrere Möglichkeiten zur Randbehandlung.

- Konstante Fortsetzung mit $V(0)$ und $V(N_f-1)$ an den Rändern
- 5 - Für den rechten und unteren Rand Verwendung von bereits codierten Pixeln, die über den Blockrandbereich hinausgehen, z.B. $V(n) = P(n, -1)$ mit $n = N, \dots, N_f-1$
- Für den linken und oberen Rand: Fortsetzung des Prädiktionsvektors entlang der Kante, z.B.:
- 10 $V(n) = P(n, -1)$, für $n = 0, \dots, N_f-1$
- $V(n) = P(-1, -(n+2))$, für $n = -\text{int}(L/2), \dots, -1$
- Hier bedeutet $\text{int}(x)$ "rounding by truncation".

Die Wirkung der Vorfilterung wird in Figur 5 beispielhaft für einen 8 x 8-Bildblock bei einer Prädiktion mit Modus 1 gezeigt. Figur 5 zeigt die ungefilterte Prädiktion, Figur 6 zeigt den prädizierten Bildblock bei Vorfilterung mit dem oben für die Kantenlänge 8 Pixel gegebenen Filter der Länge 3.

Nachfolgend wird ein mögliches Ausführungsbeispiel der Vorfilterung angegeben.

Gegeben ist ein 8 x 8-Block $P(i, j)$, $i, j = 0, \dots, 7$, mit einem Prädiktionsvektor $V(n)$, $n = 0, \dots, 7$.

Es werde mit einem 3-Tap-Filter für Bildblöcke der Kantenlänge 8 gefiltert, mit den Filterkoeffizienten $h(n) = (7 \ 18 \ 7)/32$.

Die Vorfilterung mit konstanter Fortsetzung an den Rändern ergibt:

$$V_f(0) = (7 \cdot V(0) + 18 \cdot V(0) + 7 \cdot V(1)) / 32$$

$$V_f(1) = (7 \cdot V(0) + 18 \cdot V(2) + 7 \cdot V(3)) / 32$$

$$V_f(2) = (7 \cdot V(1) + 18 \cdot V(3) + 7 \cdot V(4)) / 32$$

...

$$V_f(7) = (7*V(6) + 18*V(7) + 7*V(7))/32$$

Der Vektor $V_f(n)$, $n = 0, \dots, 7$ wird dann zur Prädiktion des
5 aktuellen Bildblocks C' eingesetzt.

Wie Fig. 5 zeigt sind bei grober Quantisierung ohne die
Verwendung des erfindungsgemässen Verfahrens starke
Artefakte der Prädiktion sichtbar. Mit dem
erfindungsgemässen Verfahren zeigt sich statt dessen ein
10 weicher Verlauf (bei Filterung mit einem Tiefpass). In
Bildblöcken in denen alle AC (Wechselanteil)-Koeffizienten
zu Null quantisiert wurden, lässt sich bei Kenntnis des
Prädiktionsmodus und des verwendeten Filters der Bildblock
aus den Randpixeln der Nachbarbildblöcke direkt bestimmen.

15 Literatur:

[1] G. Bjontegaard und Thomas Wiegand (eds.), 'H.26L Test
Model Long-Term (TML) 8', Doc. VCEG-N10, ITU-T Q6/SG16 VCEG,
20 14. Meeting, Santa Barbara, CA, September 2001

5

10 Ansprüche

1. Verfahren zur gerichteten Prädiktion eines aus
Bildpunkten (a, b, ..., p) Bildblockes (C'), insbesondere
für die Intra- Prädiktion von Bildblöcken mit einer Größe
15 von mehr als 4 x 4 Bildpunkten, wobei Randbildpunkte (A, B,
C, D; E, F, G, H) mindestens eines bereits prädizierten
Nachbarbildblockes (A', B') zur Prädiktion eines aktuellen
Bildblockes (C') herangezogen werden und wobei jene
Bildpunkte, die in Abhängigkeit der Prädiktionsrichtung zur
20 Prädiktion verwendet werden, entlang eines Blockrandes des
aktuellen zu prädizierenden Bildblockes (C') vorgefiltert
werden.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß
25 die zur Prädiktion des aktuellen Bildblockes verwendeten
Bildpunkte des mindestens einen Nachbarbildblockes zu einem
Prädiktionsvektor zusammengefaßt werden und daß eine
Tiefpaßfilterung der für die Prädiktion verwendeten
Bildpunkte entlang der Richtung des Prädiktionsvektors
30 vorgenommen wird.

3. Verfahren nach Anspruch 2, dadurch gekennzeichnet, daß
bei der Tiefpaßfilterung die Nullposition am stärksten
gewichtet wird.

35

4 Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß für die Tiefpaßfilterung auch Bildpunkte verwendet werden, die über den Blockrandbereich des aktuellen Bildblockes hinausgehen.

5

5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, daß die über den Blockrandbereich hinausgehenden Bildpunkte für die Prädiktion durch Fortsetzung des Prädiktionsvektors an den Blockrändern gewonnen werden.

10

6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß zur Fortsetzung des Prädiktionsvektors die Vektorkomponentenwerte am Anfang bzw. Ende des Prädiktionsvektors verwendet werden.

15

7. Verfahren nach Anspruch 5, dadurch gekennzeichnet, daß zur Fortsetzung des Prädiktionsvektors bereits prädizierte Bildpunkte von Nachbarbildblöcken herangezogen werden, die nicht unmittelbar einen Blockrand mit dem aktuellen zu prädizierenden Bildblock gemeinsam haben.

20

I	A	B	C	D
E	a	b	c	d
F	e	f	g	h
G	i	j	k	l
H	m	n	o	p

Fig. 2

	A'
B'	C'

Fig. 1

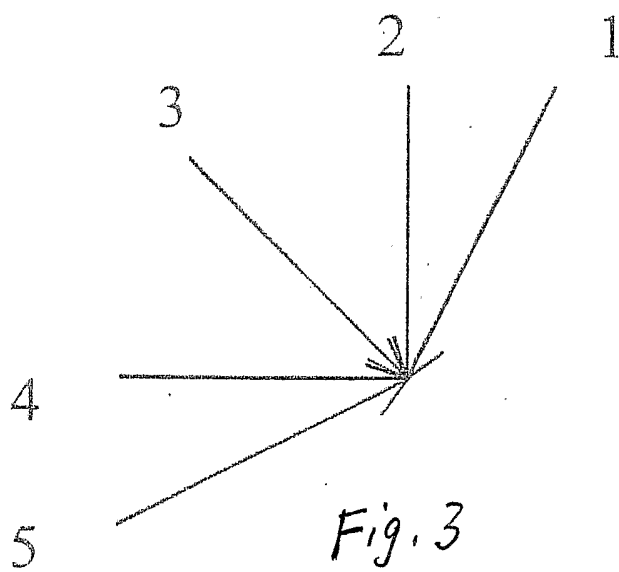


Fig. 3

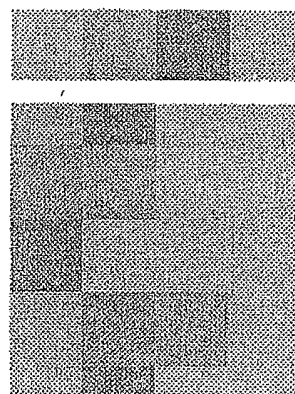


Fig. 4

Fig. 6

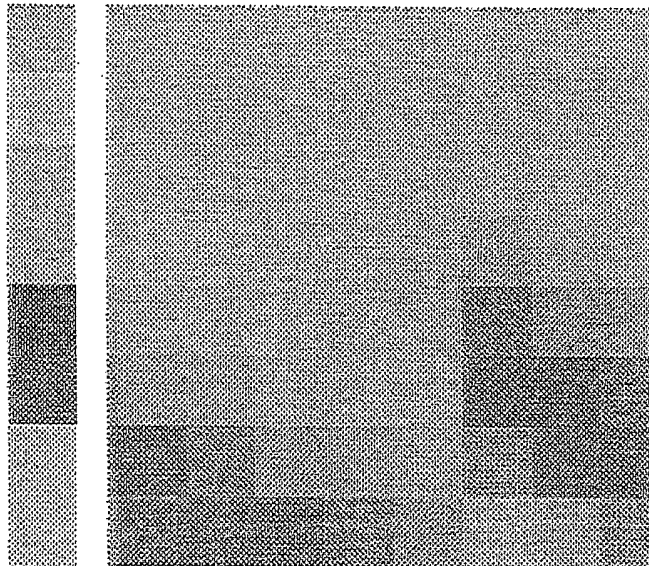


Fig. 5

