



Προχωρημένες μέθοδοι κωδικοποίησης βίντεο

MPEG-4 AVC (Advanced Video Coding) HEVC

T. Wiegand, G. Sullivan, G. Bjontegaard and A. Luthra,
Overview of the H.264/AVC video coding standard,
IEEE Trans. on Circuits and Systems for Video Technology, Vol. 13, pp. 560-576, July 2003.

G. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand,
Overview of the High Efficiency Video Coding (HEVC) Standard,
IEEE Trans. on Circuits and Systems for Video Technology, Vol. 22, pp. 1649-1668, Dec. 2012.



MPEG-4 AVC

Βελτιωμένες επιδόσεις έναντι MPEG-2 και απλό MPEG-4 (ASP)

Πρώτη έκδοση 2003 (ISO, ITU)

Κλιμακωτή κωδικοποίηση (2007)

Κωδικοποίηση πολλών όψεων / τρισδιάστατη (2009)

Δορυφορική τηλεόραση 1.5 Mbits/s

Επίγεια και καλωδιακή ψηφιακή τηλεόραση

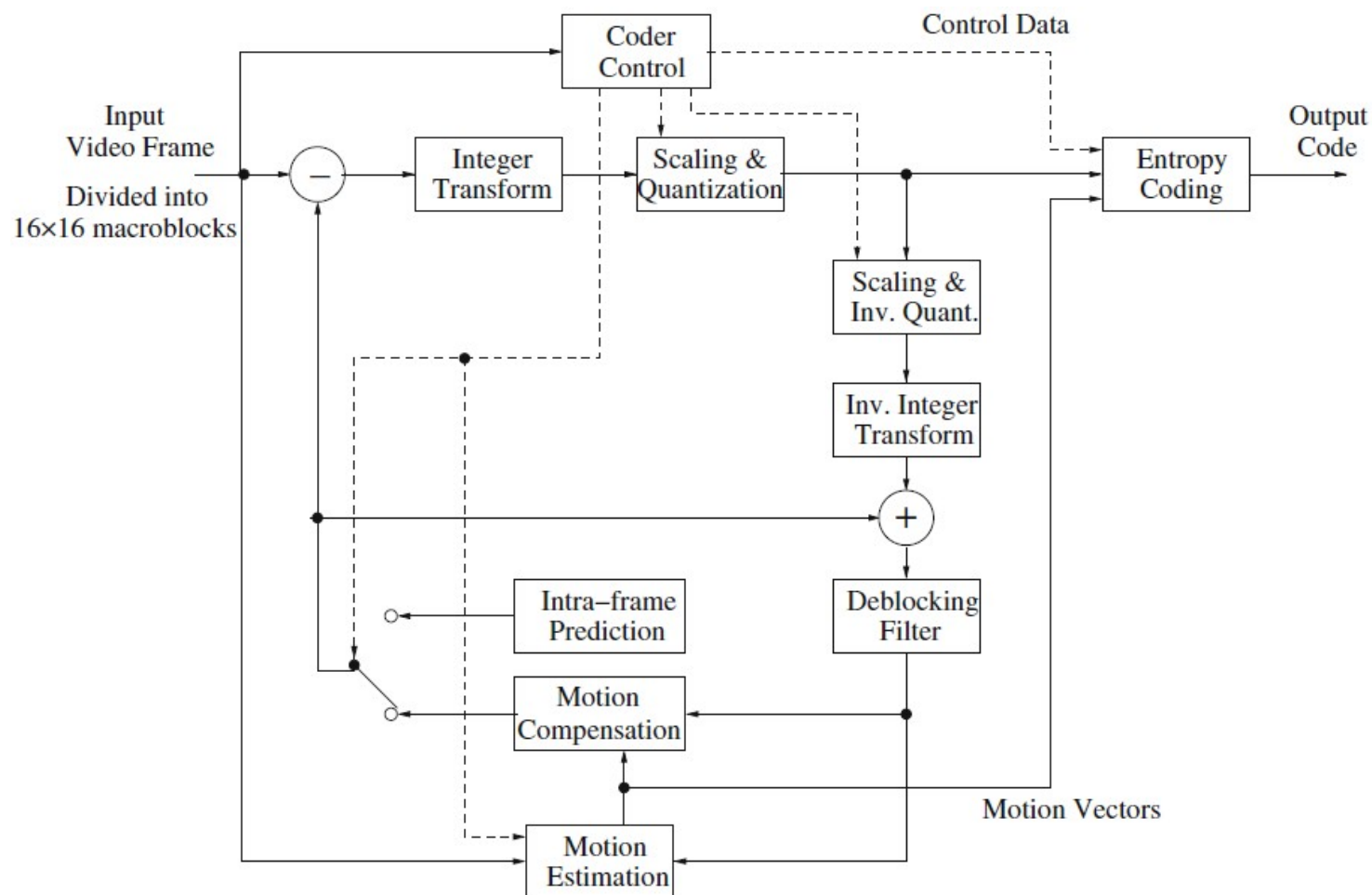
Blu-ray Disc, HD DVD

Επικοινωνίες σε Δίκτυα (Ethernet, DSL, ασύρματα, κινητά, ...)

Flash, YouTube

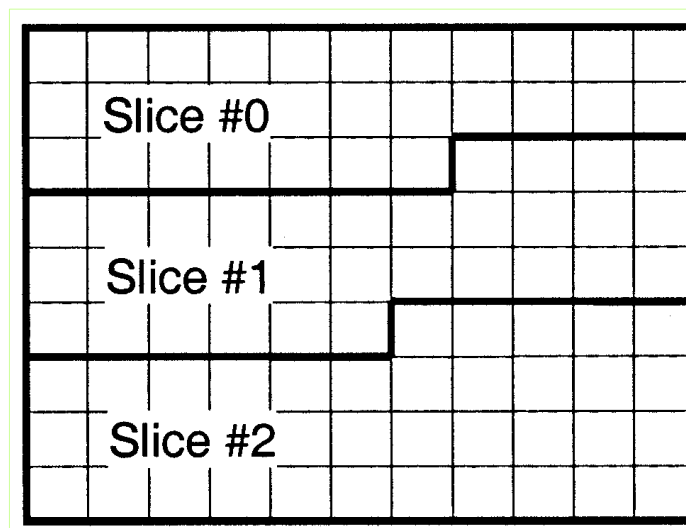


Υβριδικός κωδικοποιητής

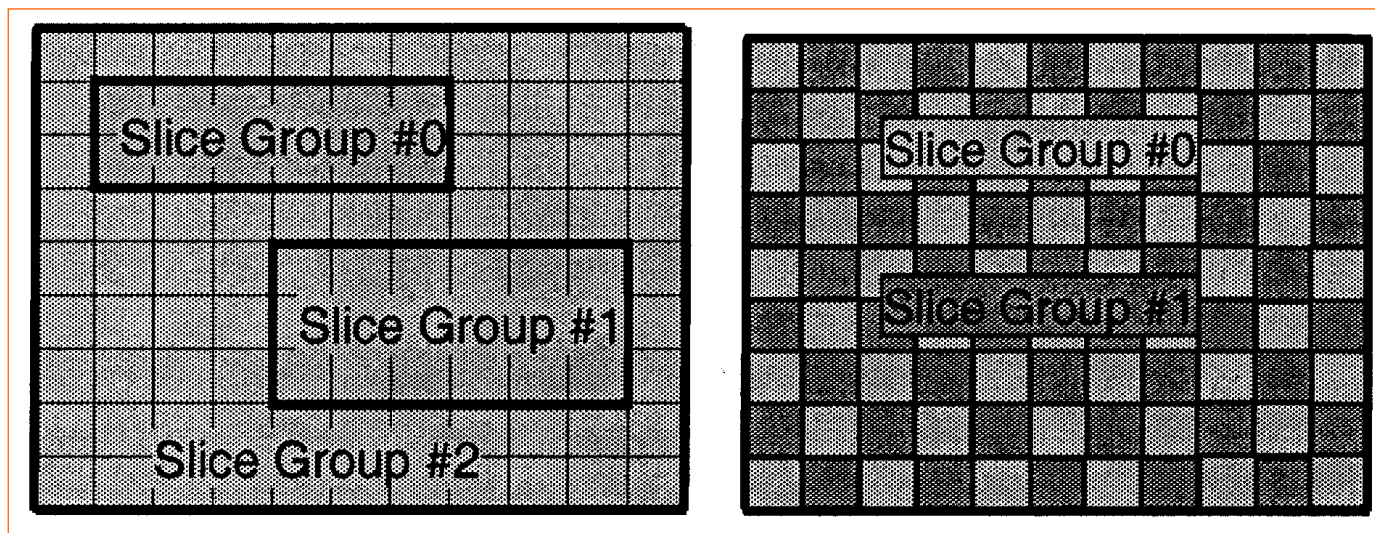


Οργάνωση μάκρο-μπλοκ

Συστηματικός
τεμαχισμός



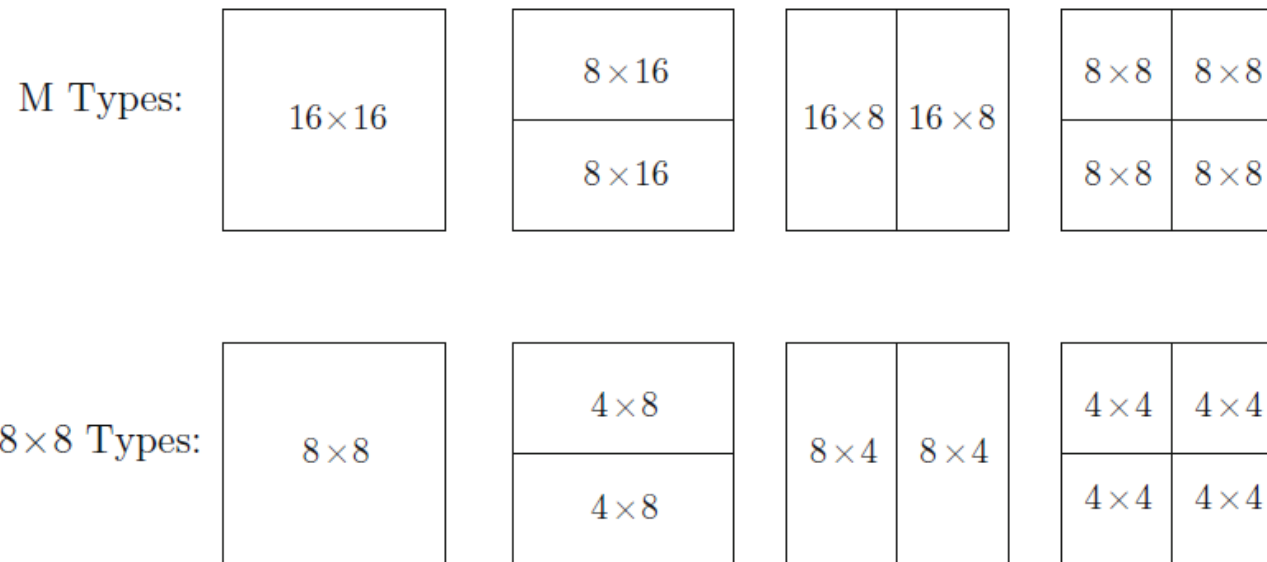
Ευέλικτη διάταξη μάκρο-μπλοκ



Περιοχές ενδιαφέροντος



Αντιστάθμιση κίνησης



Μεταβλητό μέγεθος μπλοκ

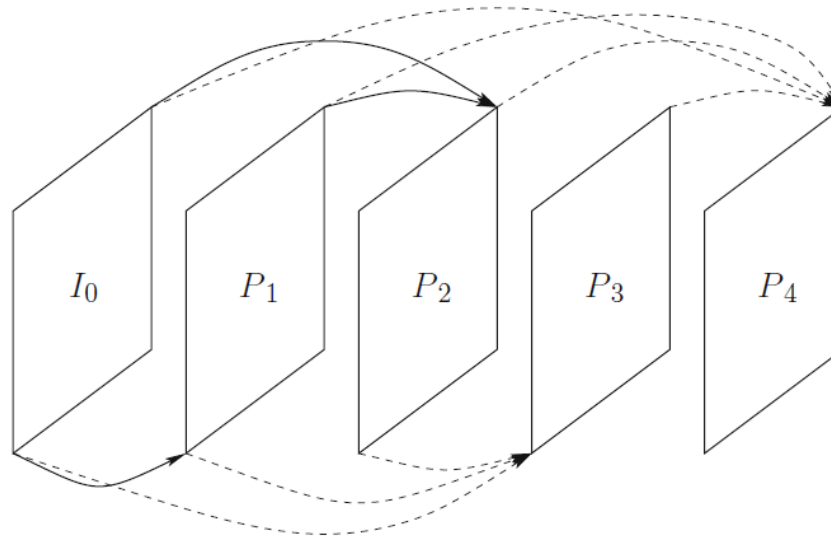
Ακρίβεια $1/4$ pixel

Γραμμική παρεμβολή



Ομάδα εικόνων

- No B-frames
- Πολλαπλές αναφορές



- Ιεραρχική δομή πρόβλεψης

$I_0 B_1 B_2 B_3 B_4 B_5 B_6 B_7 B_8 B_9 B_{10} B_{11} I_{12}$



Μετασχηματισμός

Αριθμητική ακεραίων

$$\mathbf{F} = \mathbf{T} \times \mathbf{f} \times \mathbf{T}^T$$

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \quad \mathbf{H}_{\text{inv}} = \begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

Κανονικοποίηση (ορθοκανονικότητα) μέσω κβαντιστή



Κβαντισμός

$$\mathbf{F} = \mathbf{T} \times \mathbf{f} \times \mathbf{T}^T$$

$$\hat{\mathbf{F}} = \text{round} \left[(\mathbf{H} \times \mathbf{f} \times \mathbf{H}^T) \cdot \mathbf{M}_{\mathbf{f}} / 2^{15} \right]$$

$$\tilde{\mathbf{f}} = \text{round} \left[(\mathbf{H}_{\text{inv}} \times (\hat{\mathbf{F}} \cdot \mathbf{V}_i) \times \mathbf{H}_{\text{inv}}^T) / 2^6 \right]$$



Κβαντισμός

$$\mathbf{M}_f = \begin{bmatrix} \mathbf{m}(QP, 0) & \mathbf{m}(QP, 2) & \mathbf{m}(QP, 0) & \mathbf{m}(QP, 2) \\ \mathbf{m}(QP, 2) & \mathbf{m}(QP, 1) & \mathbf{m}(QP, 2) & \mathbf{m}(QP, 1) \\ \mathbf{m}(QP, 0) & \mathbf{m}(QP, 2) & \mathbf{m}(QP, 0) & \mathbf{m}(QP, 2) \\ \mathbf{m}(QP, 2) & \mathbf{m}(QP, 1) & \mathbf{m}(QP, 2) & \mathbf{m}(QP, 1) \end{bmatrix}$$

$$\mathbf{V}_i = \begin{bmatrix} \mathbf{v}(QP, 0) & \mathbf{v}(QP, 2) & \mathbf{v}(QP, 0) & \mathbf{v}(QP, 2) \\ \mathbf{v}(QP, 2) & \mathbf{v}(QP, 1) & \mathbf{v}(QP, 2) & \mathbf{v}(QP, 1) \\ \mathbf{v}(QP, 0) & \mathbf{v}(QP, 2) & \mathbf{v}(QP, 0) & \mathbf{v}(QP, 2) \\ \mathbf{v}(QP, 2) & \mathbf{v}(QP, 1) & \mathbf{v}(QP, 2) & \mathbf{v}(QP, 1) \end{bmatrix}$$

QP	Positions in $\mathbf{M}_f$ (0, 0), (0, 2) (2, 0), (2, 2)	Positions in $\mathbf{M}_f$ (1, 1), (1, 3) (3, 1), (3, 3)	Remaining $\mathbf{M}_f$ positions
0	13107	5243	8066
1	11916	4660	7490
2	10082	4194	6554
3	9362	3647	5825
4	8192	3355	5243
5	7282	2893	4559

QP	Positions in $\mathbf{V}_i$ (0, 0), (0, 2) (2, 0), (2, 2)	Positions in $\mathbf{V}_i$ (1, 1), (1, 3) (3, 1), (3, 3)	Remaining $\mathbf{V}_i$ positions
0	10	16	13
1	11	18	14
2	13	20	16
3	14	23	18
4	16	25	20
5	18	29	23



Κβαντισμός (QP=0)

72	82	85	79	13107	8066	13107	8066	507	-12	-2	2
74	75	86	82	8066	5243	8066	5243	0	-7	-14	5
84	73	78	80	13107	8066	13107	8066	2	0	-8	-11
77	81	76	84	8066	5243	8066	5243	-1	8	4	3
f				M_f				$\hat{F}$			
10	13	10	13	72	82	85	79	0	0	0	0
13	16	13	16	74	75	86	82	0	0	0	0
10	13	10	13	84	73	78	80	0	0	0	0
13	16	13	16	77	81	76	84	0	0	0	0
V_i				$\tilde{f}$				$\epsilon = f - \tilde{f}$			



Κβαντισμός (QP=18)

72	82	85	79
74	75	86	82
84	73	78	80
77	81	76	84

$\mathbf{f}$

1638	1008	1638	1008
1008	655	1008	655
1638	1008	1638	1008
1008	655	1008	655

$\mathbf{M}_f$

63	-2	0	0
0	-1	-2	1
0	0	-1	-1
0	1	0	0

$\hat{\mathbf{F}}$

80	104	80	104
104	128	104	128
80	104	80	104
104	128	104	128

$\mathbf{V}_i$

70	81	86	78
73	73	85	83
82	75	77	82
77	79	74	85

$\tilde{\mathbf{f}}$

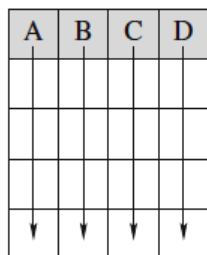
2	1	-1	1
1	2	1	-1
2	-2	1	-2
0	2	2	-1

$\boldsymbol{\epsilon} = \mathbf{f} - \tilde{\mathbf{f}}$

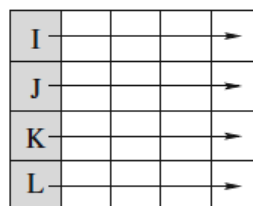


Στατική κωδικοποίηση

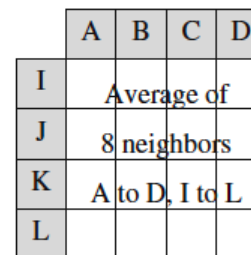
Mode 0: Vertical



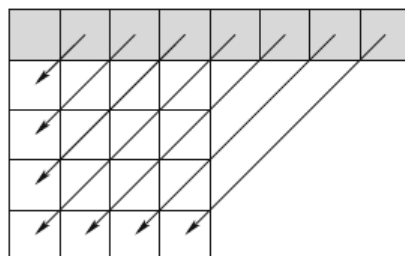
Mode 1: Horizontal



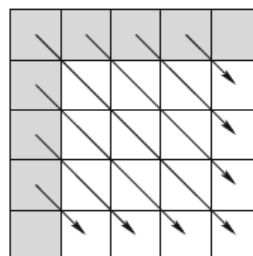
Mode 2: DC



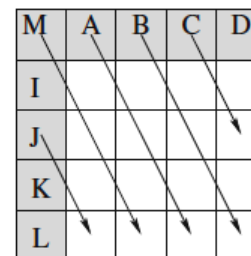
Mode 3: Diagonal Down-Left



Mode 4: Diagonal Down-Right



Mode 5: Vertical-Right



Χωρική πρόβλεψη στατικών μάκρο-μπλοκ

Μεταβλητό μέγεθος μπλοκ

9 τύποι πρόβλεψης για μπλοκ 4 x 4

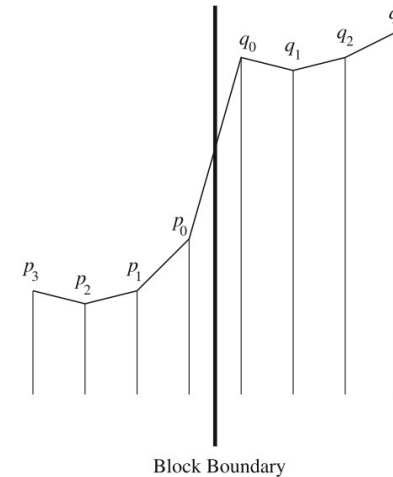
4 τύποι πρόβλεψης για μπλοκ 16 x 16

Ακέραιος μετασχηματισμός 4 x 4



Λείανση παρενεργειών μπλοκ

Γραμμικά φίλτρα λείανσης
Διαφύλαξη πραγματικών ακμών
Μπλοκ 4 x 4



$$|p_0 - q_0| < \alpha(QP),$$

$$|p_0 - p_1| < \beta(QP),$$

$$|q_0 - q_1| < \beta(QP),$$



Κωδικοποίηση Exp-Golomb

Χρησιμοποιείται για διανύσματα κίνησης

[Πρόθεμα] 1 [Επίθεμα]

Κωδικοποίηση του αριθμού N

$$x = \log_2(N/2^k + 1)$$

Πρόθεμα : '0' όσα το ακέραιο μέρος του x (l)

Επίθεμα : Δυαδική αναπαράσταση του $N + 2^k(1 - 2^l)$

Κωδικές λέξεις
για $k=0$

Unsigned N	Signed N	Codeword
0	0	1
1	1	010
2	-1	011
3	2	00100
4	-2	00101
5	3	00110
6	-3	00111
7	4	0001000
8	-4	0001001
...	...	...



Κωδικοποιητές

Προσαρμογή στα δεδομένα
Αντιδιαγώνια σάρωση
CAVLC : πολλοί πίνακες κωδίκων
 κανόνες επιλογής
CABAC : δυαδικοποίηση
 αριθμητική κωδικοποίηση

0	3	0	-1
0	1	1	0
-2	0	0	0
0	0	0	0

HEVC / H.265

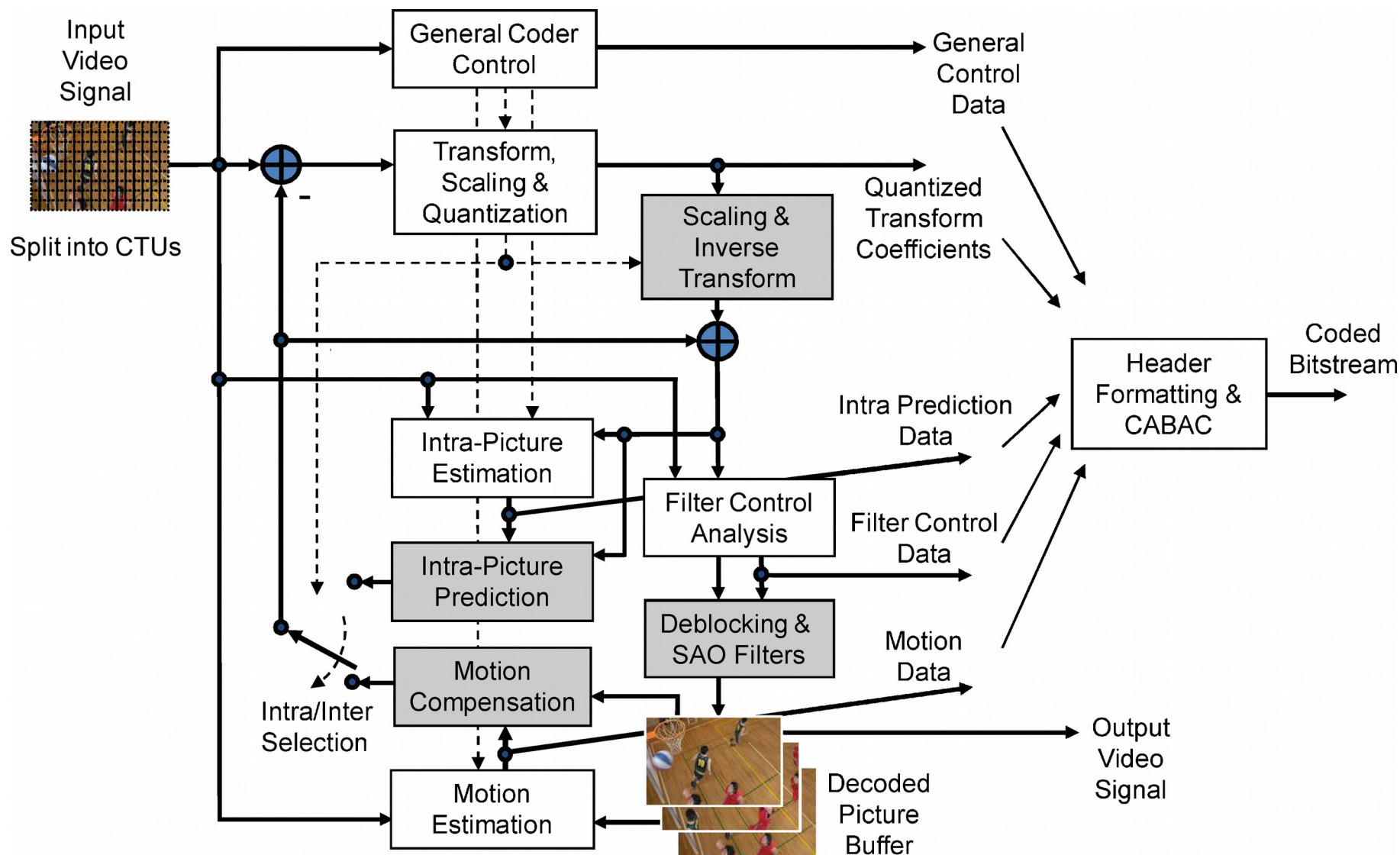
Βελτιωμένες επιδόσεις
Παράλληλοι υπολογισμοί
Πολύ υψηλή ανάλυση
Πολλές όψεις / στερεοσκοπία

Πρώτη έκδοση 2013 (ITU, ISO)

Κλιμακωτή κωδικοποίηση (2014)

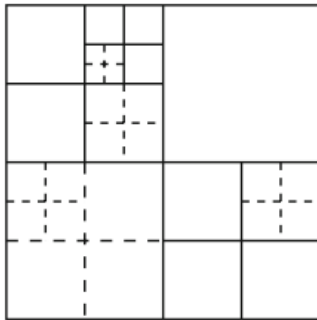
Κωδικοποίηση πολλών όψεων / τρισδιάστατη (2014/2015)

Υβριδικός κωδικοποιητής

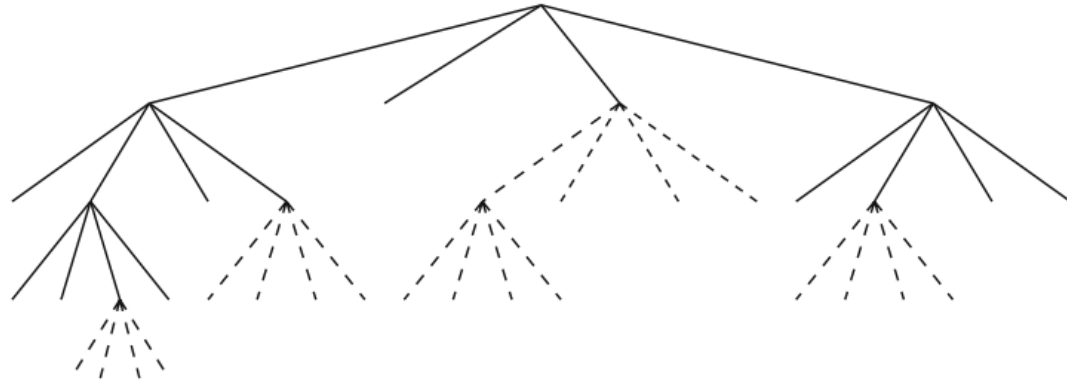


Αντιστάθμιση κίνησης

(a)



(b)



Ιεραρχικό τετραδικό δένδρο
Αρχικό μπλοκ 64 x 64

Ακρίβεια 1/4 pixel
Γραμμική παρεμβολή

Μετασχηματισμός και κβαντισμός

Αριθμητική ακεραίων

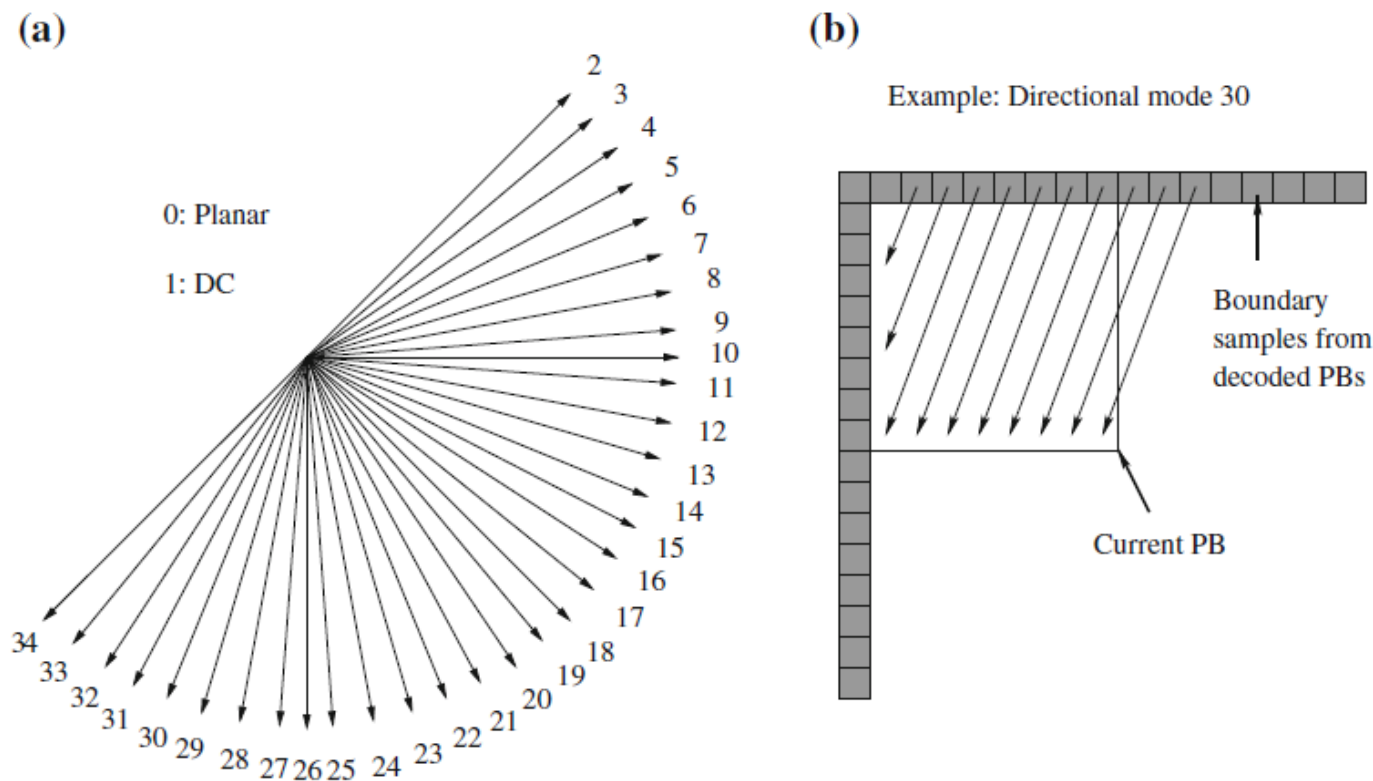
$$\mathbf{F} = \mathbf{T} \times \mathbf{f} \times \mathbf{T}^T$$

Μπλοκ : 4 x 4, 8 x 8, 16 x 16, 32 x 32

$$\mathbf{H}_{16 \times 16} = \begin{bmatrix} 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 \\ 90 & 87 & 80 & 70 & 57 & 43 & 25 & 9 & -9 & -25 & -43 & -57 & -70 & -80 & -87 & -90 \\ 89 & 75 & 50 & 18 & -18 & -50 & -75 & -89 & -89 & -75 & -50 & -18 & 18 & 50 & 75 & 89 \\ 87 & 57 & 9 & -43 & -80 & -90 & -70 & -25 & 25 & 70 & 90 & 80 & 43 & -9 & -57 & -87 \\ 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 & 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 \\ 80 & 9 & -70 & -87 & -25 & 57 & 90 & 43 & -43 & -90 & -57 & 25 & 87 & 70 & -9 & -80 \\ 75 & -18 & -89 & -50 & 50 & 89 & 18 & -75 & -75 & 18 & 89 & 50 & -50 & -89 & -18 & 75 \\ 70 & -43 & -87 & 9 & 90 & 25 & -80 & -57 & 57 & 80 & -25 & -90 & -9 & 87 & 43 & -70 \\ 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 \\ 57 & -80 & -25 & 90 & -9 & -87 & 43 & 70 & -70 & -43 & 87 & 9 & -90 & 25 & 80 & -57 \\ 50 & -89 & 18 & 75 & -75 & -18 & 89 & -50 & -50 & 89 & -18 & -75 & 75 & 18 & -89 & 50 \\ 43 & -90 & 57 & 25 & -87 & 70 & 9 & -80 & 80 & -9 & -70 & 87 & -25 & -57 & 90 & -43 \\ 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 & 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 \\ 25 & -70 & 90 & -80 & 43 & 9 & -57 & 87 & -87 & 57 & -9 & -43 & 80 & -90 & 70 & -25 \\ 18 & -50 & 75 & -89 & 89 & -75 & 50 & -18 & -18 & 50 & -75 & 89 & -89 & 75 & -50 & 18 \\ 9 & -25 & 43 & -57 & 70 & -80 & 87 & -90 & 90 & -87 & 80 & -70 & 57 & -43 & 25 & -9 \end{bmatrix}$$

Απλοποίηση κβαντισμού λόγω καλύτερης προσέγγισης

Στατική κωδικοποίηση



Χωρική πρόβλεψη στατικών μάκρο-μπλοκ
Μεταβλητό μέγεθος μπλοκ (4 x 4 μέχρι 32 x 32)
35 τύποι πρόβλεψης για μπλοκ 4 x 4

Ακέραιος διακριτός μετασχηματισμός ημιτόνου

Στατικά μπλοκ 4 x 4 με πρόβλεψη

$$\mathbf{F} = \mathbf{T} \times \mathbf{f} \times \mathbf{T}^T$$

$$\mathbf{H}_{\text{DST}} = \begin{bmatrix} 29 & 55 & 74 & 84 \\ 74 & 74 & 0 & -74 \\ 84 & -29 & -74 & 55 \\ 55 & -84 & 74 & -29 \end{bmatrix}$$

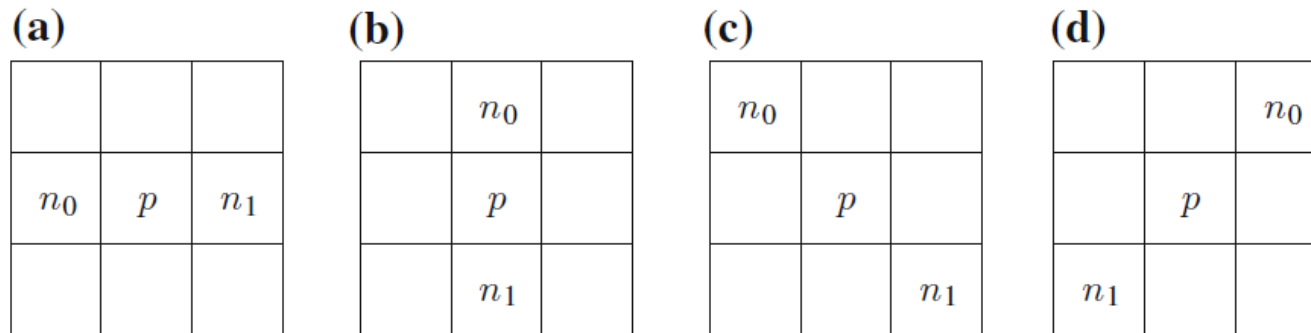
Συνδυασμός χωριστών μετασχηματισμών γραμμές/στήλες

Λείανση παρενεργειών κβαντισμού

Διαφύλαξη πραγματικών ακμών

Φίλτρο στα όρια μπλοκ 8 x 8

Προσαρμοσμένη διόρθωση τιμών



EdgeIdx	Condition	Meaning
0	Cases not listed below	Monotonic area
1	$p < n_0$ and $p < n_1$	Local min
2	$p < n_0$ and $p = n_1$ or $p < n_1$ and $p = n_0$	Edge
3	$p > n_0$ and $p = n_1$ or $p > n_1$ and $p = n_0$	Edge
4	$p > n_0$ and $p > n_1$	Local max

Κωδικοποιητής

Προσαρμογή στα δεδομένα
Τρεις δυνατές σαρώσεις
CABAC : δυαδικοποίηση
αριθμητική κωδικοποίηση

Τύποι κωδικοποίησης

Σύγκριση κωδικοποιητών

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR
INTERACTIVE APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	H.263 CHC	MPEG-4 ASP	MPEG-2 H.262 MP
HEVC MP	40.3%	67.9%	72.3%	80.1%
H.264/MPEG-4 AVC HP	—	46.8%	54.1%	67.0%
H.263 CHC	—	—	13.2%	37.4%
MPEG-4 ASP	—	—	—	27.8%

AVERAGE BIT-RATE SAVINGS FOR EQUAL PSNR FOR
ENTERTAINMENT APPLICATIONS

Encoding	Bit-Rate Savings Relative to			
	H.264/MPEG-4 AVC HP	MPEG-4 ASP	H.263 HLP	MPEG-2/ H.262 MP
HEVC MP	35.4%	63.7%	65.1%	70.8%
H.264/MPEG-4 AVC HP	—	44.5%	46.6%	55.4%
MPEG-4 ASP	—	—	3.9%	19.7%
H.263 HLP	—	—	—	16.2%