

Συμπίεση/κωδικοποίηση βίντεο

Γιώργος Τζιρίτας
Τμήμα Επιστήμης Υπολογιστών
<http://www.csd.uoc.gr/~tziritas>

Εφαρμογή	Απαίτηση	Παρα- μόρφωση	Μετάδοση	Πρότυπο
Δικτυακό βίντεο κατ'απαίτηση	1,5 Mbps 10 Mbps	Υψηλή Μέση	Internet 100 Mbps	MPEG-1 MPEG-2/4
Βιντεο- τηλέφωνο	64 kbps	Υψηλή	ISDN	H.261
Βιντεο- συνδιάσκεψη	1,5 Mbps 64 kbps	Υψηλή Πολύ υψηλή	Ethernet POTS, Internet	MPEG-1/4, H.263 MPEG-4, H.263
Δορυφορική τηλεόραση	10 Mbps	Χαμηλή	Ειδική υπηρεσία	MPEG-2
CDROM	1,5 Mbps	Μέση	PC	MPEG-1/2/4
DVD	20 Mbps	Χαμηλή	PC	MPEG-2

H.261 : εισαγωγή

Δημοσίευση 1990 (Intern. Telecom. Union)

Βιντεο-διασκέψεις / βιντεο-τηλέφωνο σε γραμμές ISDN

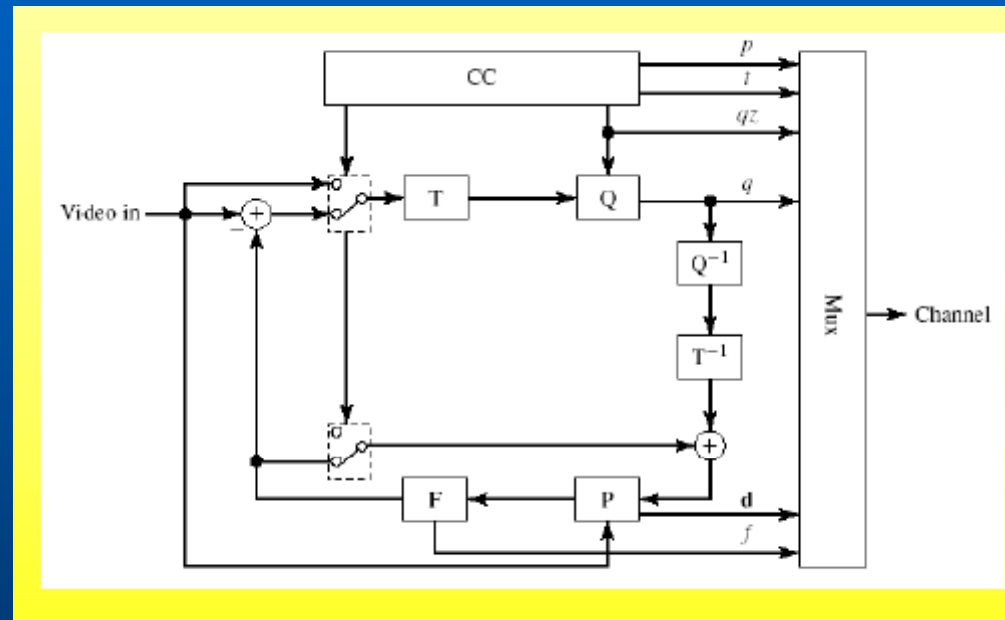
Αρχική μορφή βίντεο : CIF ή QCIF
μικρή ανάλυση, μικρή κίνηση

Παραγόμενος ρυθμός δεδομένων : 37,3 Mbps ή 9,35 Mbps

Ρυθμός συμπιεσμένων δεδομένων : ≥ 384 kbps ή ≥ 64 kbps

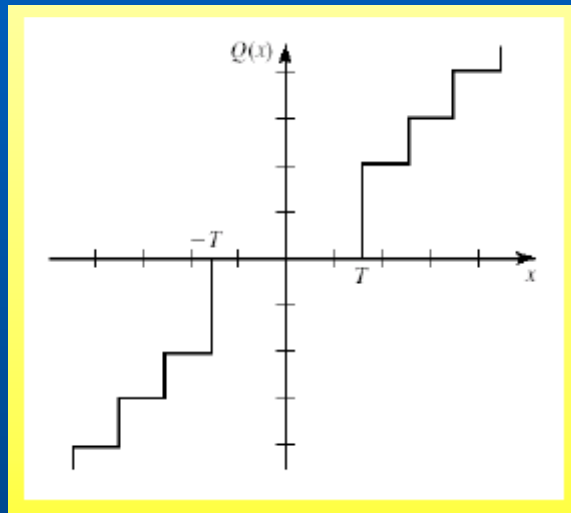
Δυνατότητα παράλειψης εικόνων : 30 fps / 15 fps / 10 fps / 7,5 fps

H.261 : κωδικοποιητής



T : Discrete Cosine Transform (8x8)

H.261 : κβαντιστής



Σταθερός συντελεστής σε
τύπο χωρίς πρόβλεψη :
ομοιόμορφα, βήμα = 8

Υπόλοιποι συντελεστές :
ομοιόμορφα πλην
νεκρής ζώνης, βήμα = 2-62

Νεκρή ζώνη για απαλοιφή
θορύβου και κέρδος στο
βαθμό συμπίεσης

H.261 : αντιστάθμιση κίνησης

Εκτίμηση σε μακρο-μπλοκ (16 x 16)

Ακρίβεια pixel : από -15 σε +15

Διαφορική κωδικοποίηση

Φιλτράρισμα (προαιρετικό) για την αποφυγή
χρονικής διάδοσης των σφαλμάτων πρόβλεψης

Διαχωρίσιμο φίλτρο : $[1/4, 1/2, 1/4]$

H.261 : έλεγχος ρυθμού

Εντροπικός κωδικοποιητής (Huffman)
σε σάρωση αντιδιαγώνια :

(διαστήματα μηδενικών συντελεστών,
μη μηδενικές τιμές συντελεστών)

- Τύπος : με ή χωρίς πρόβλεψη, με ή χωρίς αντιστάθμιση κίνησης
- Χαρακτηρισμός MB : αντισταθμισμένο ή όχι
- Βήμα κβαντισμού
- Χρήση ή όχι φίλτρου

H.261 : χαρακτηριστικά συστήματος

Σύνταξη δεδομένων :

- εικόνα – συγχρονισμός, τύπος εικόνας
- ομάδα από μπλοκ (GOB, 176 x 48)
 - συγχρονισμός, βήμα κβαντισμού
- μακρο-μπλοκ (MB, 16 x 16)
 - τύπος συμπίεσης, μετακίνηση
- μπλοκ (block, 8 x 8)
 - μετασχηματισμός συνημιτόνου

Καθυστέρηση (μέγιστη) : 150 ms

Κώδικας διόρθωσης λαθών : BCH (511,493,2)

H.263 : εισαγωγή

Δημοσίευση : 1995, 1997 (H.263+), 2000 (H.263++)
(Intern. Telecom. Union)

Βιντεο-διασκέψεις / βιντεο-τηλέφωνο
σε γραμμές PSTN ή στο Διαδίκτυο

Στόχος : καλύτερη ποιότητα σε χαμηλότερο ρυθμό

Επίτευγμα :

- 18-24 kbps σε καλύτερη ποιότητα από 64 kbps (H.261)
- Βιντεο-τηλέφωνο σε τηλεφωνικές γραμμές 28,8 kbps
ή ασύρματη επικοινωνία

H.263 : βελτιώσεις ως προς H.261

Καλύτερη εκτίμηση κίνησης :

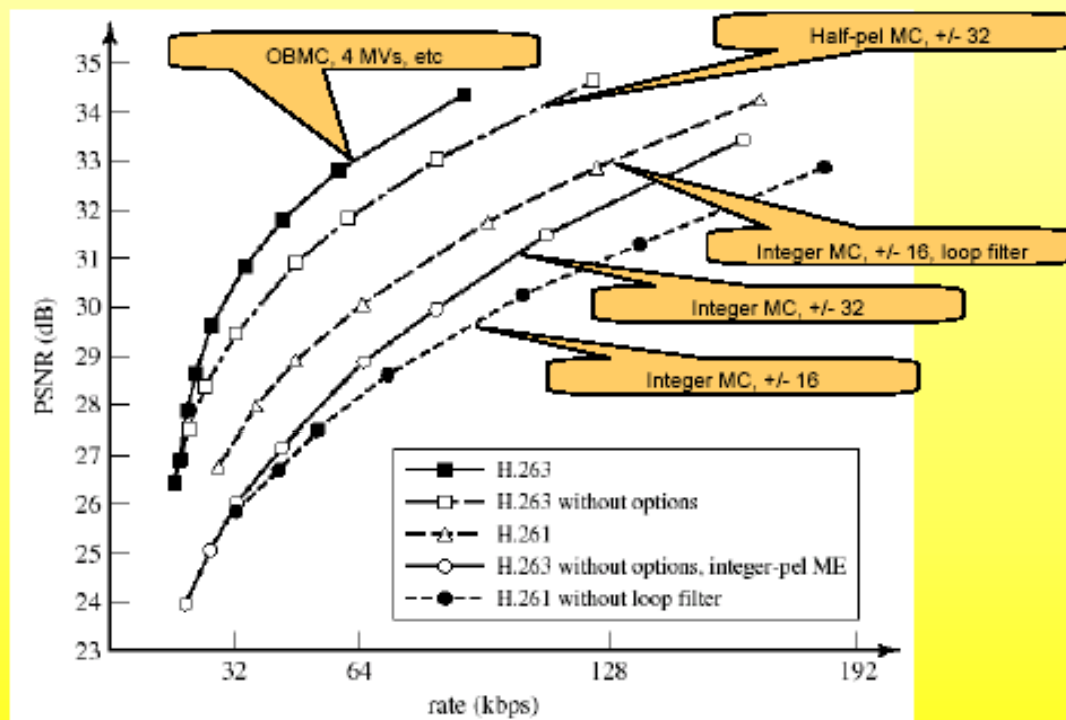
- ακρίβεια εκτίμησης κίνησης $\frac{1}{2}$ pixel
- μεγαλύτερο διάστημα τιμών : από -31,5 έως +31
- προβλεπτική συμπίεση διανυσμάτων κίνησης
- δυνατότητα επικάλυψης μακρο-μπλοκ
- δυνατότητα επιλογής μπλοκ 8 x 8
- δυνατότητα παρεμβολής εικόνας

Καλύτερος εντροπικός κωδικοποιητής

- τριπλά σύμβολα (διάστημα, τιμή, τέλος)
- δυνατότητα χρήσης αριθμητικού κωδικοποιητή

Ομάδα δύο εικόνων, όπου η δεύτερη εικόνα κωδικοποιείται πρώτα με πρόβλεψη. Η πρώτη εικόνα κωδικοποιείται μετά, παρεμβάλλοντας ανάμεσα στην προηγούμενη και την επόμενη.

Η.263 / Η.261: συγκρίσεις



Forman, QCIF, 12.5 Hz

MPEG-1 : εισαγωγή

Δημοσίευση 1990 (Intern. Standardization Organization)

CD-ROM, LAN, WAN, WWW

Βασίσθηκε στα πρότυπα : JPEG, H.261

Αρχική μορφή βίντεο : CIF

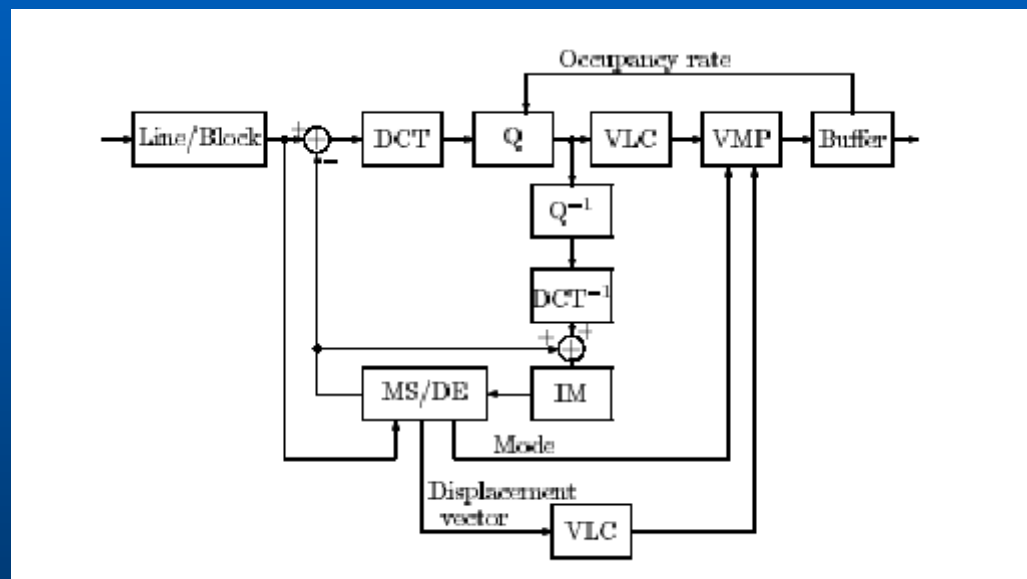
Μέγιστα : 768 x 576 σε 30 fps

Παραγόμενος ρυθμός δεδομένων (CIF): 37,3 Mbps

Ρυθμός συμπιεσμένων δεδομένων : 1,5 Mbps

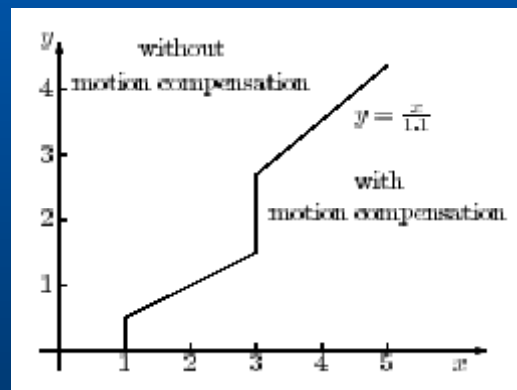
Ασυμμετρία : αποσυμπιεστής μικρής πολυπλοκότητας

MPEG-1 : κωδικοποιητής



MPEG-1 : αντιστάθμιση κίνησης

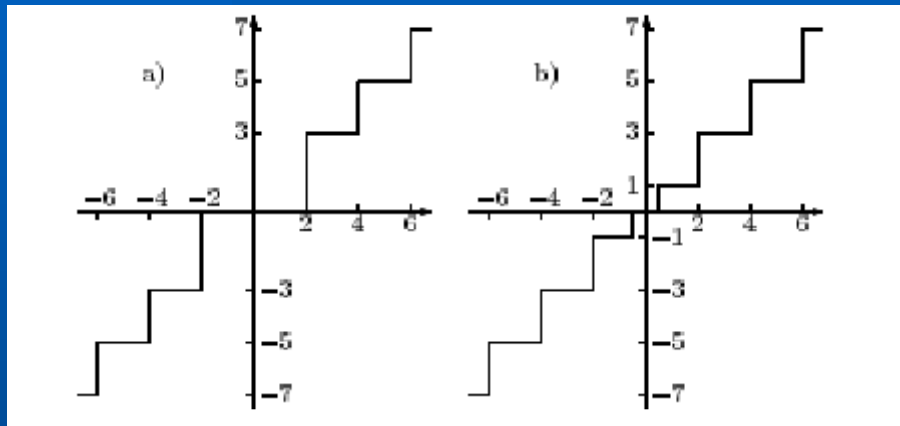
Εκτίμηση σε μακρο-μπλοκ (16 x 16)
Ακρίβεια $\frac{1}{2}$ pixel
Διαφορική κωδικοποίηση



Μέση απόλυτη τιμή
με ή χωρίς
αντιστάθμιση κίνησης

Πρόβλεψη, αν $\sigma_1^2 > \max(\sigma_p^2, 64)$

MPEG-1 : κβαντιστής



με πρόβλεψη

χωρίς πρόβλεψη

πίνακας ορατότητας

Σταθερός συντελεστής σε
τύπο χωρίς πρόβλεψη :
ομοιόμορφα, βήμα = 8

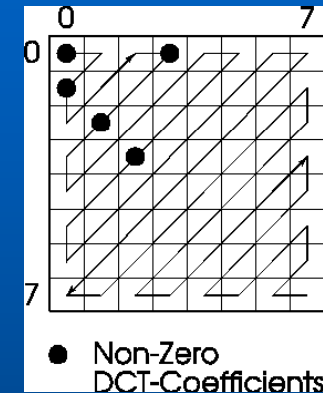
Υπόλοιποι συντελεστές :
ομοιόμορφα πλην ζώνης '0'

	0	1	2	3	4	5	6	7
0	8	16	19	22	26	27	29	34
1	16	16	22	24	27	29	34	37
2	19	22	26	27	29	34	34	38
3	22	22	26	27	29	34	37	40
4	22	26	27	29	32	35	40	48
5	26	27	29	32	35	40	48	58
6	26	27	29	34	38	46	56	69
7	27	29	35	38	46	56	69	83

MPEG-1 : έλεγχος ρυθμού

Εντροπικός κωδικοποιητής (Huffman)
σε σάρωση αντιδιαγώνια :
(διαστήματα μηδενικών συντελεστών,
μη μηδενικές τιμές συντελεστών)

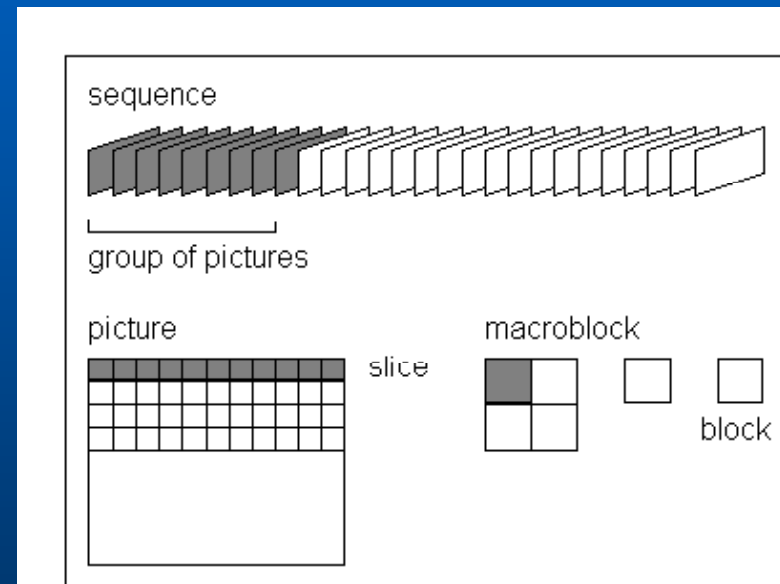
- Τύπος : με ή χωρίς πρόβλεψη, με ή χωρίς αντιστάθμιση κίνησης
- Χαρακτηρισμός MB : αντισταθμισμένο ή όχι
- Βήμα κβαντισμού



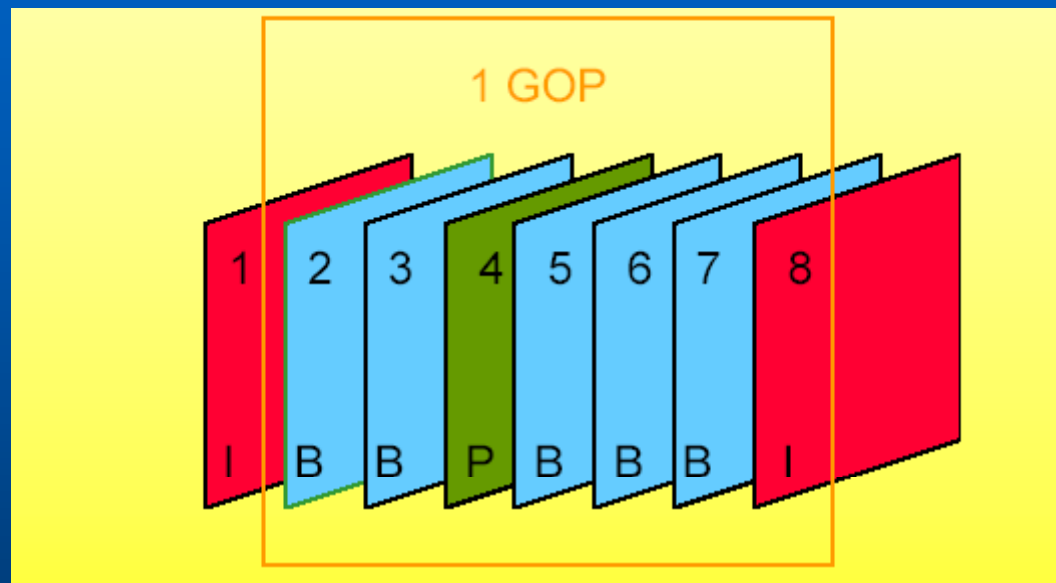
MPEG-1 : δομή δεδομένων

Σύνταξη δεδομένων :

- ακολουθία εικόνων
 - περιεχόμενο, χαρακτηριστικά
- ομάδα εικόνων (GOP) – τυχαία πρόσβαση
- εικόνα – συγχρονισμός, τύπος εικόνας (D, I, P, B)
- «φέτα» εικόνας (slice) – συγχρονισμός
- μακρο-μπλοκ (MB, 16 x 16)
 - τύπος συμπίεσης, μετακίνηση (F, B, A), στατικό
- μπλοκ (block, 8 x 8)
 - μετασχηματισμός συνημιτόνου



MPEG-1 : ομάδα εικόνων



Σειρά κωδικοποίησης : 1 4 2 3 8 5 6 7

MPEG-2 : εισαγωγή

Δημοσίευση 1993 (Intern. Standardization Organization)

Ψηφιακή τηλεόραση, δορυφορική/καλωδιακή τηλεόραση,
υπηρεσίες βίντεο σε δίκτυα, VOD, HDTV, DVD

Βασίσθηκε στο πρότυπο : MPEG-1

Αρχική μορφή βίντεο : CCIR-601

Παραγόμενος ρυθμός δεδομένων (CIF): 160 Mbps

Ρυθμός συμπιεσμένων δεδομένων : 4-15 Mbps

Ασυμμετρία : αποσυμπεστής μικρής πολυπλοκότητας

MPEG-2 : προφίλ και επίπεδα ανάλυσης

	Απλό 4:2:0	Κύριο 4:2:0	Κλιμακωτή ποιότητα	Κλιμακωτό χωρικά	Υψηλό 4:2:0 ή 4:2:2
Υψηλό (60 fps) 1920 x 1152		80 Mbit/s			100 Mbit/s
Υψηλό-1440 (60 fps) 1440 x 1152		60 Mbit/s		60 Mbit/s για 3 στρώματα	80 Mbit/s
Κύριο (30 fps) 720 x 576	15 Mbit/s	15 Mbit/s	15 Mbit/s για 2 στρώματα		20 Mbit/s
Χαμηλό (30 fps) 352 x 288		4 Mbit/s	4 Mbit/s		

MPEG-2 : κλιμακωτή συμπίεση (1/2)

Στρώματα συμπιεσμένου κώδικα :
βασικό και επιπρόσθετα βελτιωτικά

Χρησιμότητα στη μετάδοση σε διαύλους με θόρυβο :
καλύτερη προστασία της πλέον σημαντικής πληροφορίας

Περίπτωση μετάδοσης σε διαύλους μεταβλητού ρυθμού

Προοδευτική μετάδοση : από το πιο σημαντικό
στο πιο λεπτομερειακό

Τρεις τύποι κλιμάκωσης : ποιότητας, χωρική και χρονική

MPEG-2 : κλιμακωτή συμπίεση (2/2)

Κλιμάκωση ποιότητας

Δύο στρώματα κβαντισμού

Χωρική κλιμάκωση

Αποσυμπίεση σε διαφορετικές χωρικές αναλύσεις

Υποδειγματοληψία για τη δημιουργία του βασικού στρώματος

Παρεμβολή για την πρόβλεψη σε μεγαλύτερη ανάλυση

Χρονική κλιμάκωση

Βασικό στρώμα σε χαμηλό ρυθμό

Πρόβλεψη με αντιστάθμιση της κίνησης για βελτίωση

Εισαγωγή MPEG-4

Δημοσίευση 1998 (Intern. Telecom. Union)

Επικοινωνίες με πολυμέσα, Διαδίκτυο,
διαδραστικά παιχνίδια βίντεο, τηλε-παρακολούθηση

- αποτελεσματική βελτιωμένη συμπίεση
- πρόσβαση στα αντικείμενα και αλληλεπίδραση με αυτά
- ενοποιημένο περιβάλλον φυσικών / συνθετικών αντικειμένων
- κλιμάκωση χρόνου, χώρου, ποιότητας, αντικειμένων
- μεγάλη αντοχή σε σφάλματα μετάδοσης

5-64 kbits/s για κινητή τηλεφωνία και

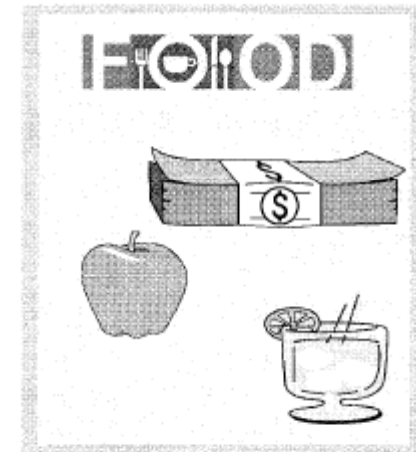
PSTN (public switched telephone network)

2 Mbits/s για τηλεοπτικά προγράμματα και κινηματογραφικές ταινίες

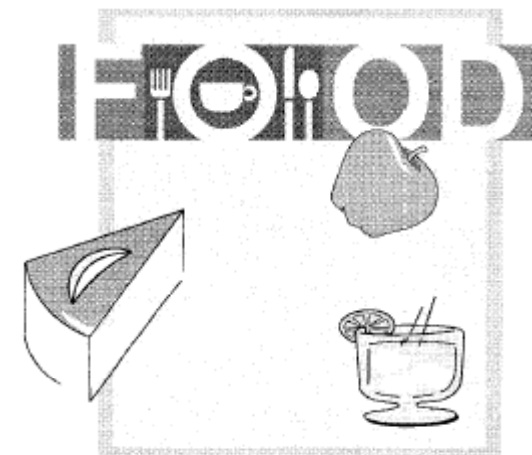
Δείγμα χειρισμού περιεχομένου

Αλλαγή κλίμακας αντικειμένων:

- αφαίρεση
- προσθήκη
- μεγέθυνση/σμίκρυνση
- μετακίνηση



(a)



(b)

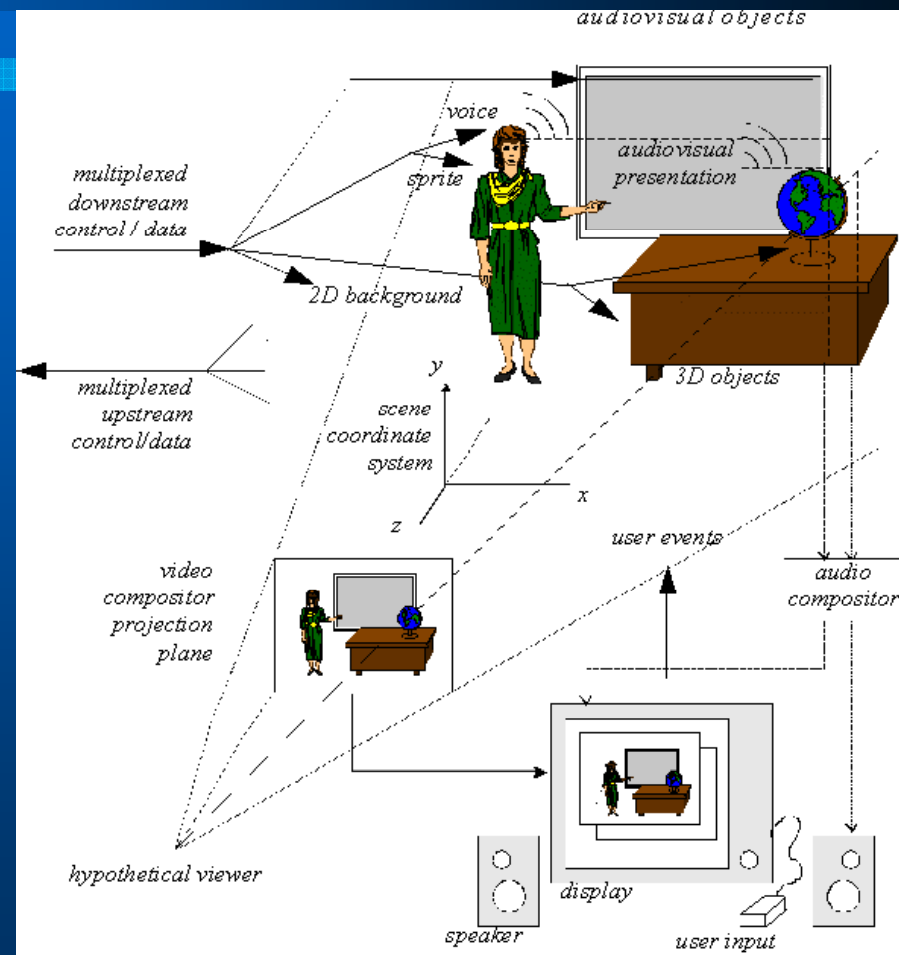
Αναπαράσταση οπτικοακουστικών αντικειμένων

Αρχέτυπα αντικείμενα (AVO):
ακουστικά, οπτικά, οπτικοακουστικά

Σύνθεση οπτικοακουστικών
αντικειμένων σε σκηνή
Περιγραφή σκηνής και αντικειμένων

Ιεραρχική δομή

Ειδικά αντικείμενα:
κείμενο, γραφικά,
ομιλούντα πρόσωπα,
ανθρώπινο σώμα σε κίνηση



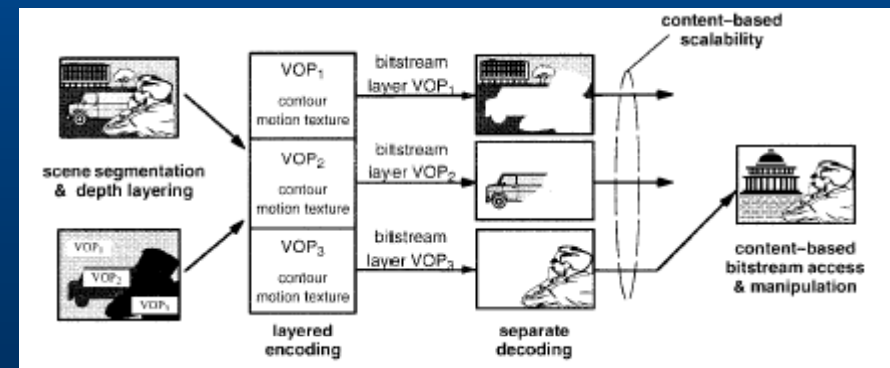
Τύποι κωδικοποίησης βίντεο

Τέσσερις τύποι κωδικοποίησης:

- Κωδικοποίηση αντικειμένων βίντεο οποιουδήποτε σχήματος
Παρόμοια με MPEG-1 (MPEG-2),
κι επιπλέον κωδικοποίηση σχήματος :
χάρτης αντικειμένων (δυαδικός), διαφάνεια αντικειμένων
Δυνατότητα κωδικοποίησης πανοραμάτων
- Κωδικοποίηση δικτυωτού πλέγματος
Τριγωνικό πλέγμα με ευελιξία στη συνθετική κίνηση και
δυνατότητα χρονικών παρεμβολών
- Κωδικοποίηση βασισμένη σε μοντέλα
Πρόσωπο και ανθρώπινο σώμα
- Κωδικοποίηση υφής
Χρήση κυματιδιακού μετασχηματισμού για
φωτεινότητα και χρώμα

Κωδικοποίηση φυσικών αντικειμένων

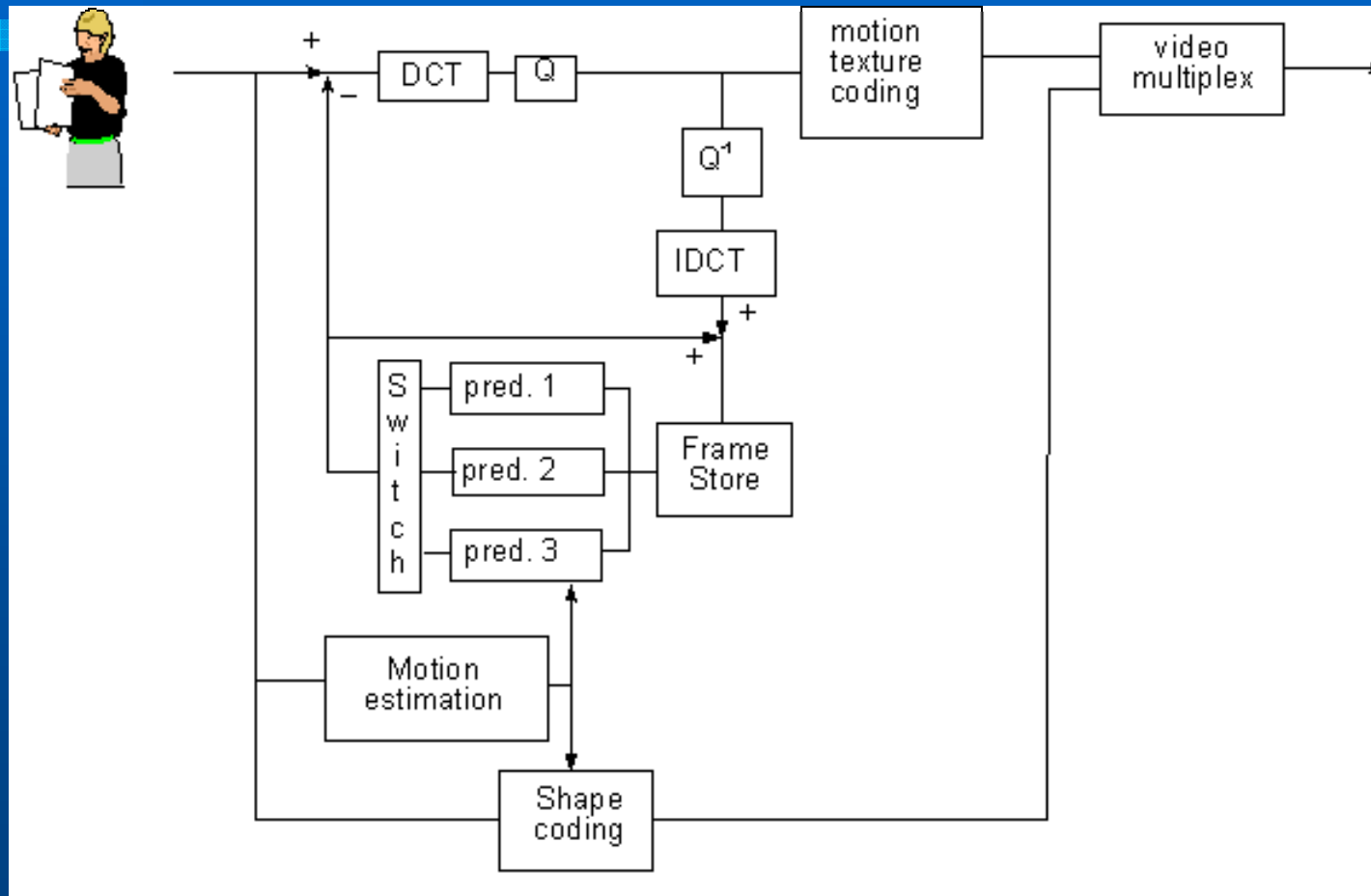
- εικόνα και βίντεο
- υφή για σύνθεση σε δικτυωτά πλέγματα
- 2-Δ δικτυωτό πλέγμα
- γεωμετρική χρονική μεταβολή πλέγματος
- τυχαία πρόσβαση σε κάθε αντικείμενο
- χειρισμός εικόνων και βίντεο
- κλιμάκωση με βάση το περιεχόμενο
- κλιμάκωση χωρική, χρονική και ποιότητας
- αντοχή σε σφάλματα μετάδοσης



Κωδικοποίηση αντικειμένων βίντεο

Τμήμα βίντεο οποιουδήποτε σχήματος (VO)
Στιγμιότυπο: επίπεδο αντικείμενο βίντεο (VOP)
Ορισμός VOP: σχήμα και υφή
Πλήρης ορισμός με την κίνηση του αντικειμένου
Ορθογώνια αντικείμενα: συμβατότητα με
MPEG-1, MPEG-2, H.263

Διάγραμμα κωδικοποιητή αντικειμένων βίντεο

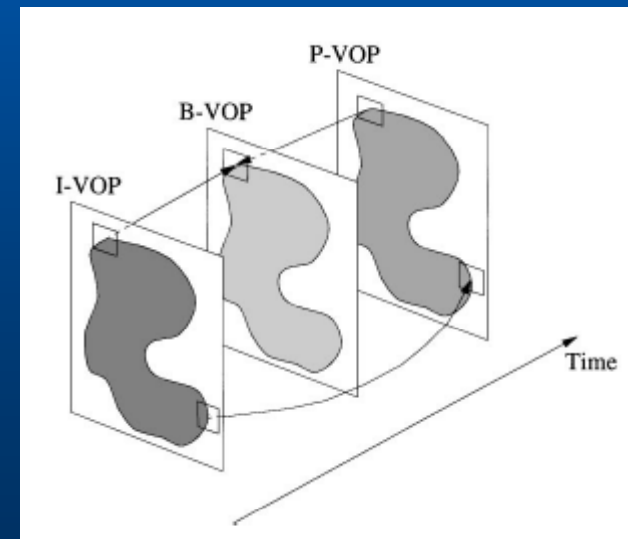
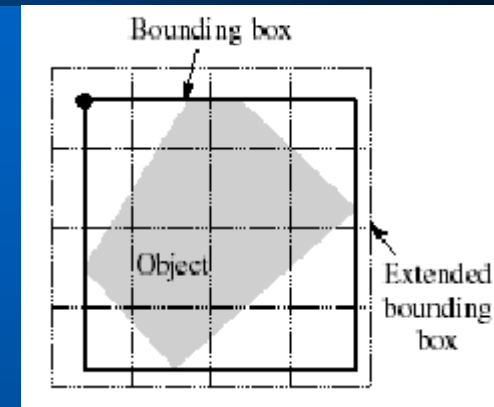


Κωδικοποίηση κίνησης και υφής

Διανύσματα κίνησης ανά μάκρο-μπλοκ
ενδεχόμενα επικαλυπτόμενα, ή διαιρεμένα στα 4
συμπιεσμένα με προβλεπτική κωδικοποίηση

Συμπίεση χωρικών δεδομένων με χρήση
μετασχηματισμού συνημιτόνου

16x16	16x8	8x16	8x8
0	0 1	0 1	0 1 2 3
8x8	8x4	4x8	4x4
0	0 1	0 1	0 1 2 3



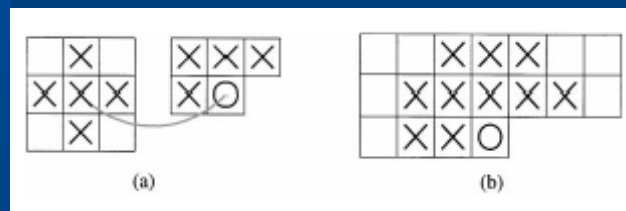
Κωδικοποίηση σχήματος

Πίνακας a (δυναδικός ή με βαθμό διαφάνειας)

Κωδικοποίηση στατική ή χρονική,
με ή χωρίς σφάλματα συμπίεσης

Μπλοκ 16×16 : διαφανή, αδιαφανή, συνοριακά
(δυναδικός πίνακας a)

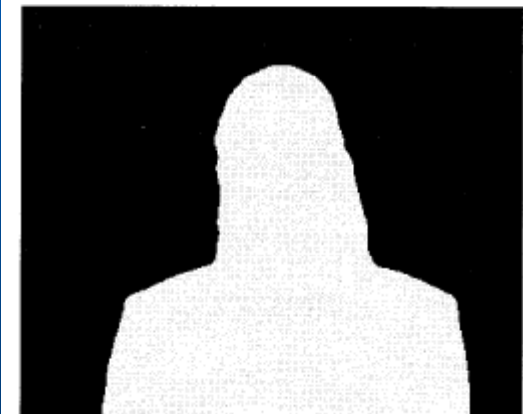
Συμπίεση των δεδομένων του συνόρου για
δυναδικό πίνακα με χρήση αριθμητικού
κωδικοποιητή προσαρμοσμένου στα συμφραζόμενα



Συμπίεση τύπου υφής για πίνακα a με βαθμό διαφάνειας (256 τιμές)



(a)



(b)

Κωδικοποίηση πανοράματος

Κωδικοποίηση φόντου

Αποτελεσματικότητα της συμπίεσης, αφού το φόντο κωδικοποιείται μόνο μία φορά στην αρχή, ενώ η κίνηση μπορεί να περιγραφεί με λίγες παραμέτρους



(a)



(b)



Κωδικοποίηση δικτυωτού πλέγματος

Προέλευση: υπολογιστική γραφιστική
Τριγωνικό πλέγμα

Παραμετρική κίνηση σε κάθε
τριγωνικό τμήμα

Προσαρμογή του τριγωνικού πλέγματος
στο περιεχόμενο

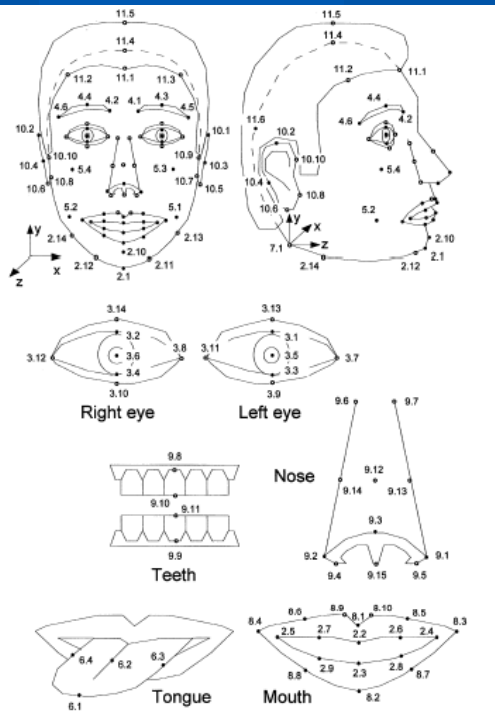
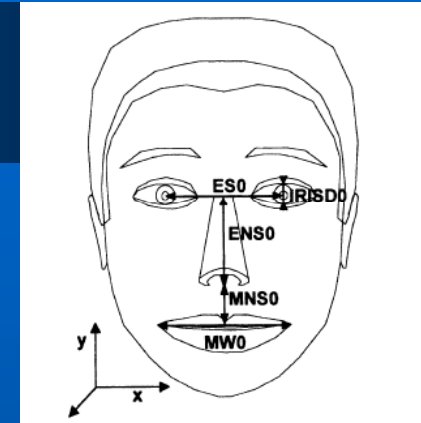


Δυνατότητα μίξης φυσικών και συνθετικών αντικειμένων,
αναπαράστασης φυσικής συνεχούς κίνησης,
παρεμβολής ενδιάμεσων στιγμιότυπων

Κωδικοποίηση κινούμενου προσώπου

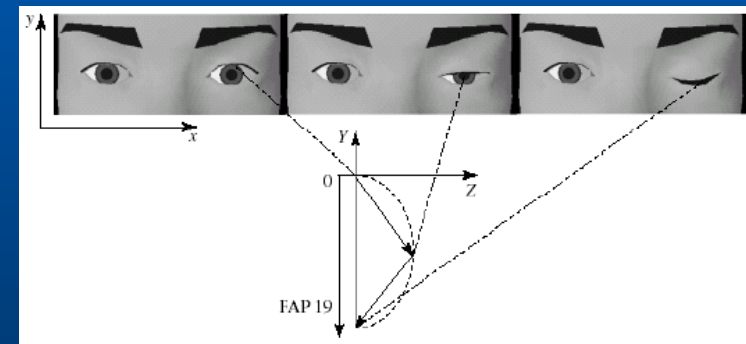
Κίνηση προσώπου

- παράμετροι ορισμού προσώπου
- παράμετροι κίνησης προσώπου



Ορισμός :

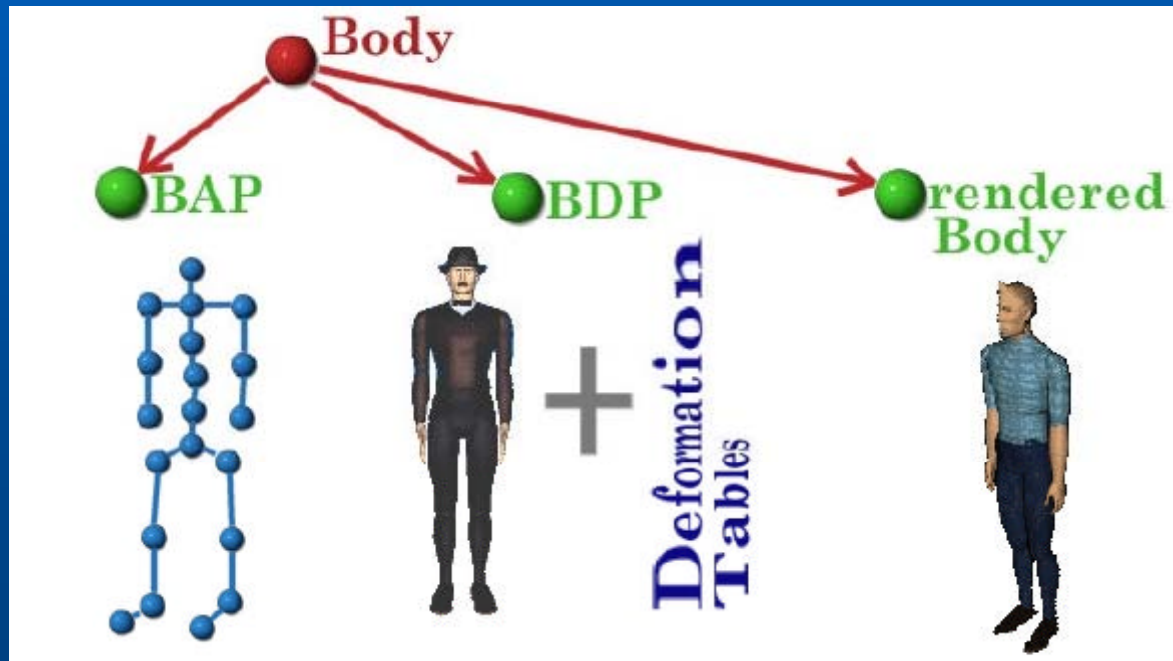
- 84 χαρακτηριστικών σημείων
- 68 απλών κινήσεων
- εκφράσεων προσώπου



Σύνθεση φωνής σε συνδυασμό με συνθετική κίνηση χειλιών (14 κινήσεις στόματος)

Κωδικοποίηση κινούμενου σώματος

- παράμετροι ορισμού 3-Δ μοντέλου σώματος
- παράμετροι κίνησης σώματος



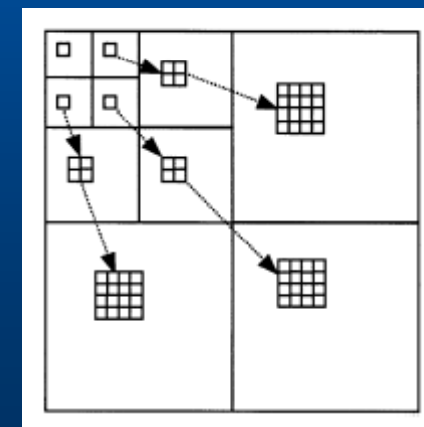
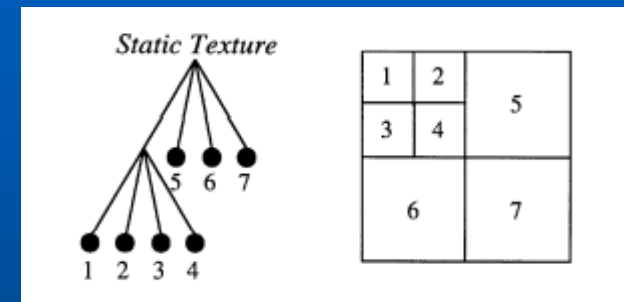
Κωδικοποίηση υφής

2-Δ κυματιδιακή ανάλυση χρησιμοποιώντας διορθωγώνια φίλτρα Daubechies με δυνατότητα προσαρμογής στο σχήμα ενός αντικειμένου

Χωρική πρόβλεψη χρησιμοποιείται για τη συμπίεση της ζώνης χαμηλών συχνοτήτων.

Δένδρα μηδενικών τιμών σε διάφορες κλίμακες χρησιμοποιούνται για τον εντοπισμό των μηδενικών συντελεστών του μετασχηματισμού

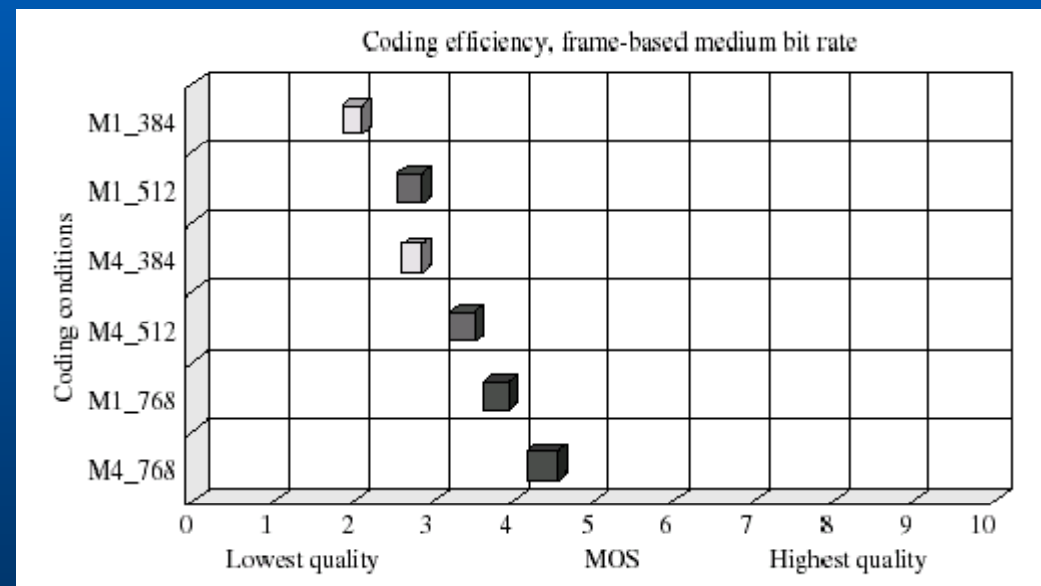
Στη συνέχεια χρησιμοποιείται αριθμητικός κωδικοποιητής



Σύγκριση με MPEG-1

Τρεις ρυθμοί μετάδοσης:

- κάτω από 64 kbits/s
- μεταξύ 64 και 384 kbits/s
- μεταξύ 384 kbits/s και 4 Mbits/s



Εισαγωγή H.264

Δημοσίευση 2003 (Intern. Telecom. Union)

- Μετάδοση δορυφορική, καλωδιακή, επίγεια, DSL
- Αποθήκευση σε οπτικές, μαγνητικές συσκευές, DVD
- Επικοινωνίες ασύρματες, ISDN, LAN, DSL
- Βίντεο κατ' απαίτηση σε ISDN, LAN, DSL

- βελτίωση πρόβλεψης
- προσαρμογή μετασχηματισμού
- βελτίωση κωδικοποίησης
- μεγάλη αντοχή σε σφάλματα μετάδοσης

Βελτίωση πρόβλεψης

Εσωτερικά σε ένα πλαίσιο

- Μπλοκ 16x16 : 4 τύποι
- Μπλοκ 4x4 : 9 τύποι

Με αντιστάθμιση κίνησης

- μεταβλητό μέγεθος : από 16x16 έως 4x4
- ακρίβεια τετάρτου περιόδου
- χρήση φίλτρου μείωσης ορίων των μπλοκ

Σύγκριση

