

# MPEG-4

# MPEG-4: Βασικά Χαρακτηριστικά

Δ13

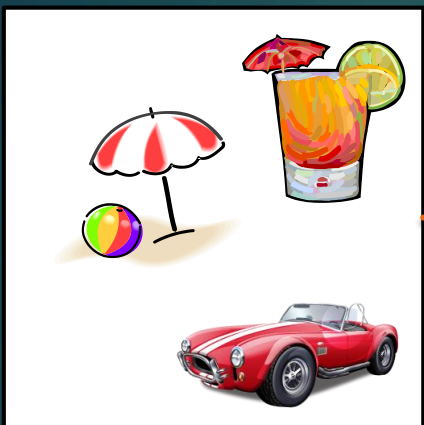
- ▶ Σχεδιάστηκε ώστε να καλύπτει ευρύ φάσμα ρυθμών, από 5 kbps έως 10 Mbps
- ▶ Εκτός από τη συμπίεση δίνει έμφαση και στην αλληλεπίδραση με το χρήστη
- ▶ Χρησιμοποιεί αντικείμενα αντί για μακρομπλόκ
- ▶ Αποδίδει καλύτερα όσον αφορά τη συμπίεση αλλά και το χειρισμό, την αναζήτηση περιεχομένου και την αρχειοθέτηση της εικονοροής

# MPEG-4: Βασικά Χαρακτηριστικά

Δ13

- ▶ Παράδειγμα κωδικοποίησης - αποκωδικοποίησης

Κατάτμηση εικόνας



Κωδικοποίηση

VOP1



VOP2



VOP3



Αποκωδικοποίηση  
Διαχείριση  
Περιεχομένου



VOP4

Περιγραφή  
δομής



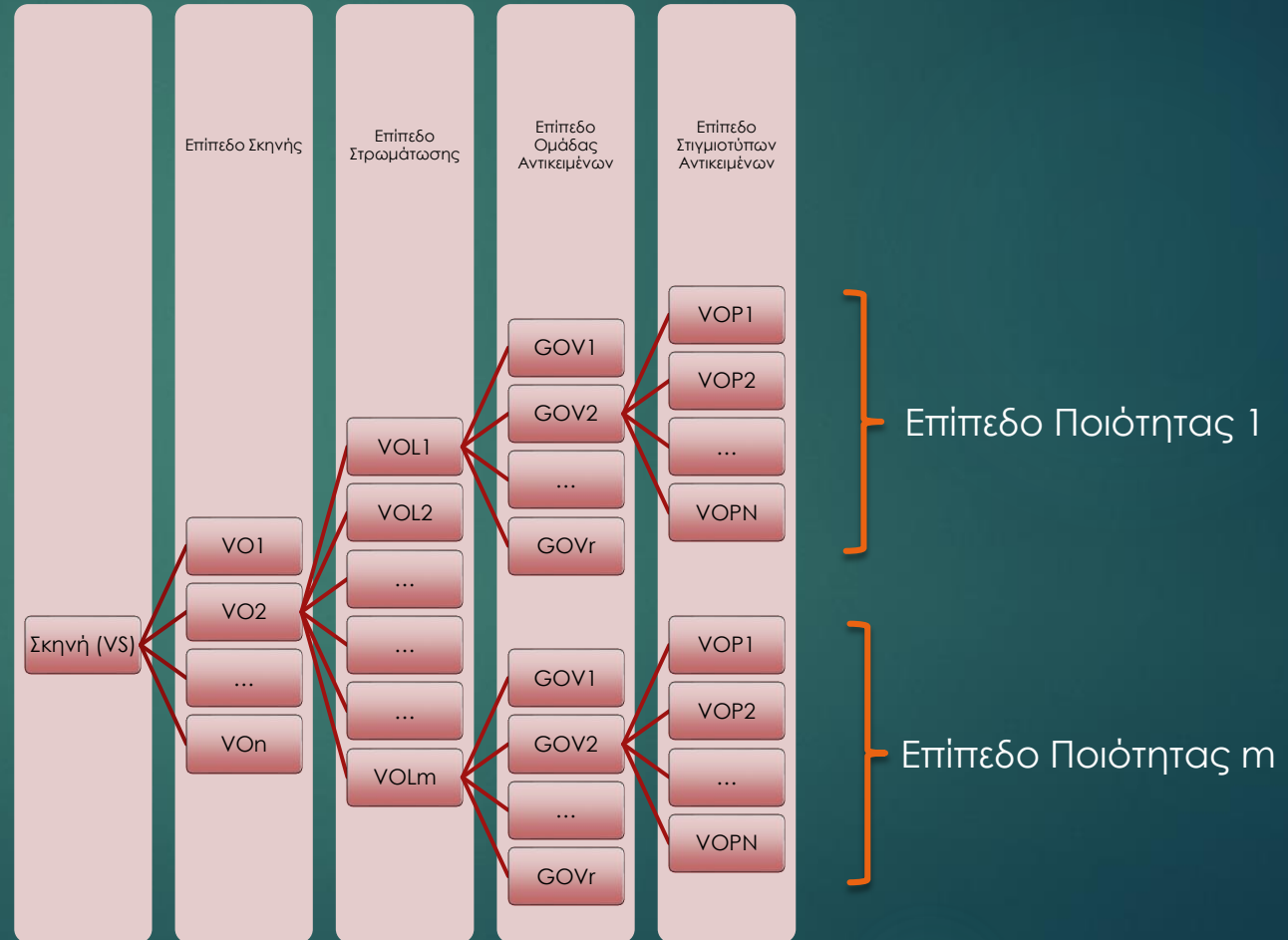
# MPEG-4: Δομή Σκηνης

Δ13

- ▶ Η δομή διαφέρει θεμελιωδώς από τις δομές των προηγούμενων προτύπων όχι λόγω οργάνωσης αλλά λόγω της χρήσης αντικειμενοστραφούς λογικής ως προς την κατάτμηση της εικόνας
- ▶ Διακρίνονται 5 επίπεδα
  - ▶ Επίπεδο Σκηνης (VS)
  - ▶ Επίπεδο Αντικειμένου (VO)
  - ▶ Επίπεδο Στρωμάτωσης Αντικειμένου (VOL)
  - ▶ Επίπεδο Ομάδας Στιγμιότυπων Αντικειμένων (GOV)
  - ▶ Επίπεδο Στιγμιότυπου Αντικειμένου (VOP)

# MPEG-4: Δομή κωδικοποιημένης Σκηνής

Δ13



# MPEG-4: Περιγραφή επιπέδων

Δ13

- ▶ **Επίπεδο Σκηνής (VS):** Αποτελεί την πλήρη περιγραφή μιας σκηνής που εκτυλίσσεται στο χρόνο και η οποία μπορεί να περιέχει πραγματικά ή εικονικά αντικείμενα
- ▶ **Επίπεδο Αντικειμένου (VO):** Είναι η περιγραφή ενός αντικειμένου στη σκηνή το οποίο μπορεί να έχει ακανόνιστο σχήμα
- ▶ **Επίπεδο Στρωμάτωσης Αντικειμένου (VOL):** Είναι ένα επίπεδο που εξυπηρετεί στην υποστήριξη της στρωμάτωσης. Κάθε VO μπορεί να διαθέτει πολλά επίπεδα ποιότητας κάθε ένα από τα οποία ανήκει σε διαφορετικό VOL στη στρωματοποιημένη κωδικοποίηση. Στην περίπτωση που δεν υπάρχει στρωμάτωση υπάρχει αντιστοιχία ένα προς ένα μεταξύ VO και VOL.
- ▶ **Επίπεδο Ομάδας Στιγμιότυπων Αντικειμένων (GOV):** Είναι ένα σύνολο από στιγμιότυπα αντικειμένων που ενσωματώνει πολλά VOP (Αποτελεί αντίστοιχο του GOP για να υποστηρίξει τυχαία προσπέλαση στο MPEG 4)
- ▶ **Επίπεδο Στιγμιότυπου Αντικειμένου (VOP):** Αποτελεί το στιγμιότυπο ενός αντικειμένου μία δεδομένη χρονική στιγμή. Κάθε VOP μπορεί να κωδικοποιείται αυτόνομα ή με πρόγνωση από άλλα VOP

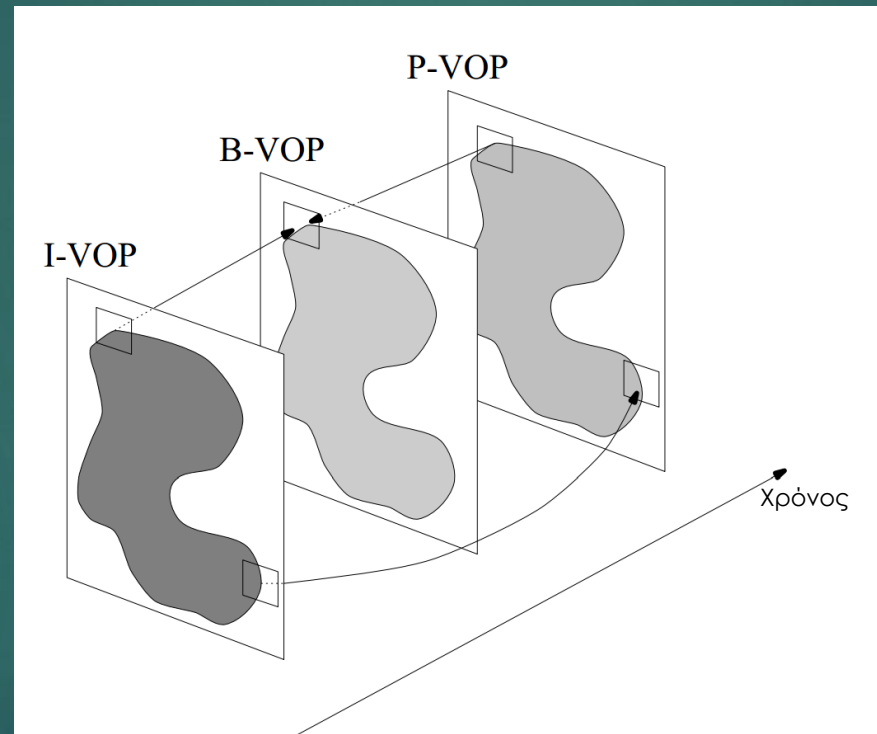
# MPEG-4: Εκτίμηση και αντιστάθμιση κίνησης

Δ13

- ▶ Το MPEG-4 υλοποιεί μια διαδικασία αντιστάθμισης κίνησης μεταξύ αντικειμένων αντί μεταξύ εικονοπλαισίων
- ▶ Ένα εικονοπλάνο αντικειμένων μορφής I καλείται I-VOP
- ▶ Αντίστοιχα ορίζονται τα P-VOP και B-VOP
- ▶ Η κυριότερη δυσκολία έγκειται στα διαφορετικά σχήματα των αντικειμένων
- ▶ Η πληροφορία του σχήματος πρέπει να κωδικοποιηθεί όπως επίσης και η πληροφορία της υφής του σχήματος

# MPEG-4: Εκτίμηση και αντιστάθμιση κίνησης

Δ13

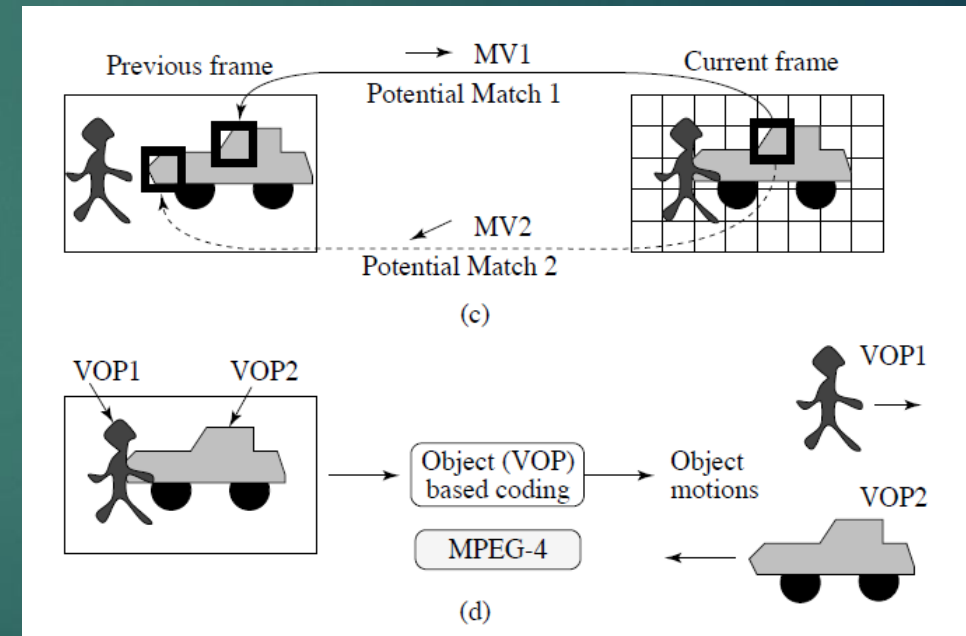
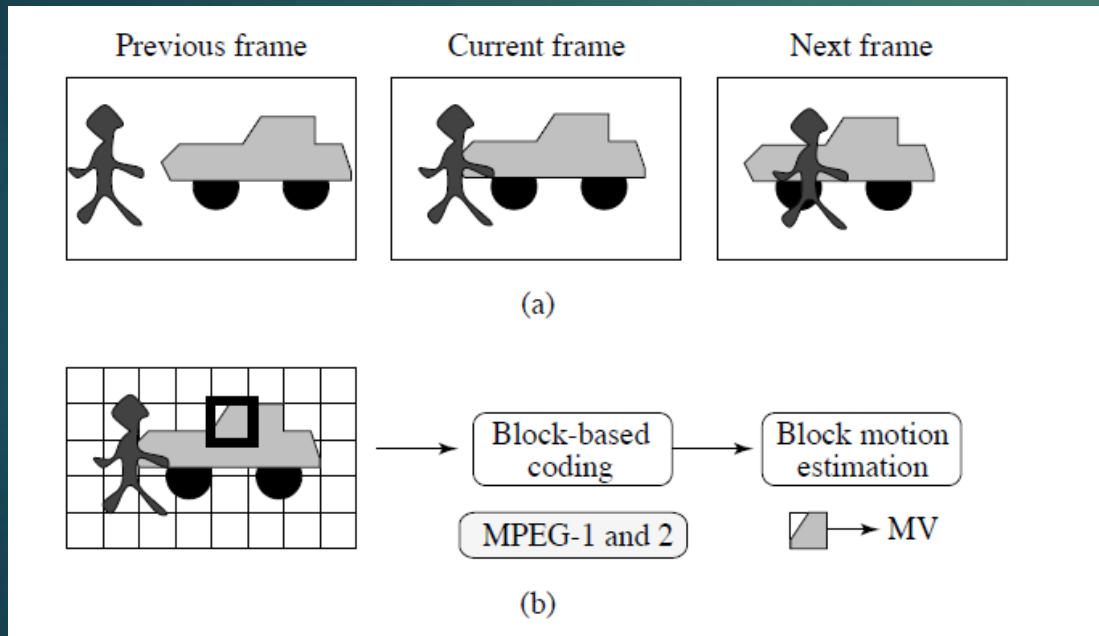


Πηγή: E. Touradj



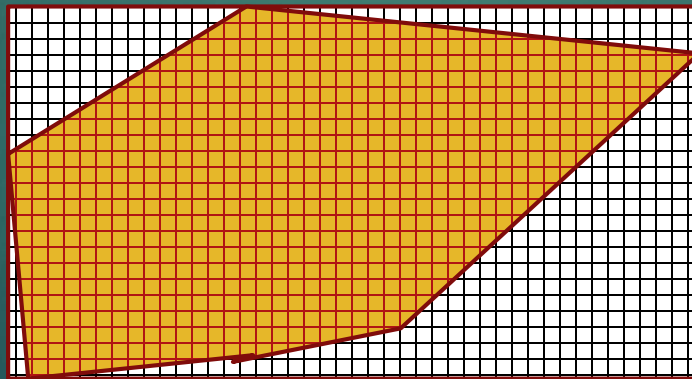
# MPEG-4: Εκτίμηση και αντιστάθμιση κίνησης

Δ13



# MPEG-4: Εκτίμηση και αντιστάθμιση κίνησης

- ▶ Κάθε VOP χωρίζεται σε μικρότερα τμήματα τα οποία μπορούν να χαρακτηριστούν ως μακρομπλόκ μεγέθους Y: 16x16, CbCr: 8x8
- ▶ Τα μακρομπλόκ των ορίων των VOP έχουν ειδικό χειρισμό
- ▶ Ορίζεται ένα περιεγραμμένο παραλληλόγραμμο που περικλείει το VOP με μέγεθος ακέραιο πολλαπλάσιο του μεγέθους του μακρομπλόκ του καναλιού Y.



# MPEG-4: Εκτίμηση και αντιστάθμιση κίνησης

Δ13

- ▶ Τα εσωτερικά μακρομπλόκ κάθε VOP εκτιμώνται με τις τεχνικές που αναπτύχθηκαν για τα MPEG 1,2
- ▶ Τα μακρομπλόκ των ορίων εκτιμώνται χρησιμοποιώντας μόνο τα εικονοστοιχεία που ανήκουν στο VOP χρησιμοποιώντας τεχνικές επέκτασης
- ▶ Για κάθε ΔΚ ενός μακρομπλόκ γίνεται πρόγνωση από τρεις γείτονες όπως στο H.263
- ▶ Υποστηρίζεται ακρίβεια  $\frac{1}{2}$  και  $\frac{1}{4}$  του εικονοστοιχείου κατά την αναζήτηση της βέλτιστης εκτίμησης

# MPEG-4: Κωδικοποίηση

Δ13

- ▶ Πραγματοποιείται ξεχωριστή κωδικοποίηση για την υφή, το περίγραμμα του VOP, την υφή των στατικών αντικειμένων και των λοιπών δομών

# MPEG-4: Κωδικοποίηση

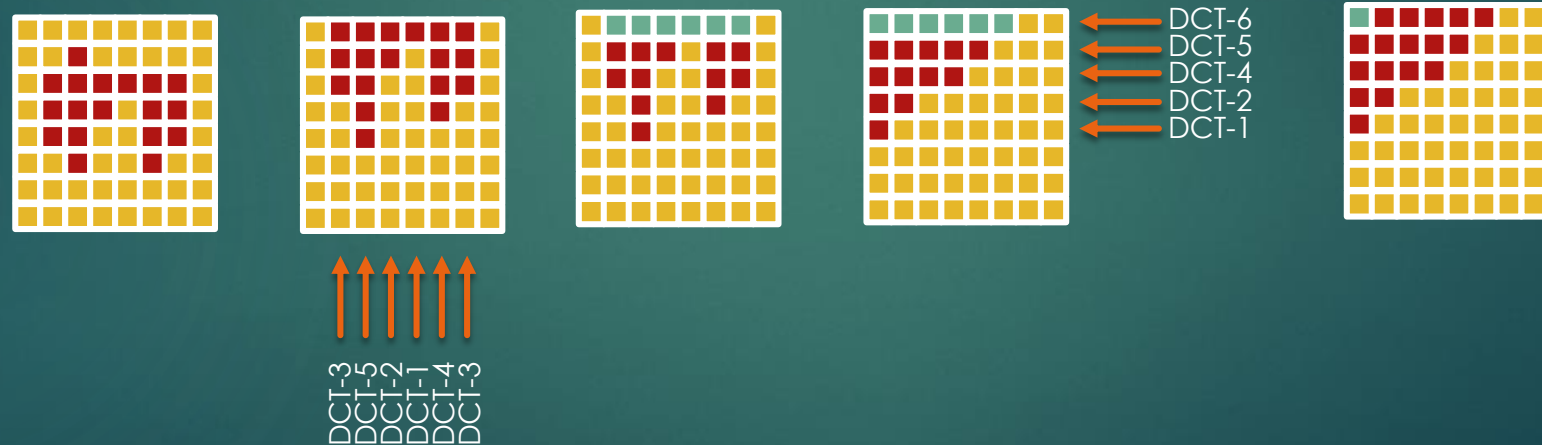
Δ13

- ▶ Η κωδικοποίηση Υφής βασίζεται στο μετασχηματισμό DCT σταθερού ή μετασχηματισμό DCT μεταβλητού σχήματος
- ▶ DCT Σταθερού σχήματος:
  - ▶ Για τα εσωτερικά μακρομπλόκ κωδικοποιούνται τα 8x8 μπλόκ ή τα υπολειμματικά 8x8 μπλόκ με παρόμοιο τρόπο με ότι έχει αναφερθεί στα προηγούμενα πρωτόκολλα
  - ▶ Για τα μακρομπλόκ των ορίων τα εικονοστοιχεία που δεν ανήκουν στο VOP αντικαθίστανται με την τιμή μηδέν

# MPEG-4: Κωδικοποίηση

Δ13

- ▶ DCT Μεταβλητού σχήματος:
  - ▶ Περιλαμβάνει μετακινήσεις των εικονοστοιχείων ώστε να δημιουργηθούν συμπαγείς περιοχές
  - ▶ Μονοδιάστατους μετασχηματισμούς μεταβλητού μήκους και ολισθήσεις
  - ▶ Μεταβιβάζεται και επιπλέον πληροφορία για την ορθή ανασύσταση του μπλόκ



# H.264

# H.264: Βασικά Χαρακτηριστικά

Δ13

- ▶ Μεγάλη αύξηση αποδοτικότητας έως 50% σε σχέση με τους προηγούμενους κωδικοποιητές για την ίδια ποιότητα εικόνας
- ▶ Αποτελεί το πρότυπο εκλογής για όλων των ειδών τις εφαρμογές (από κινητά έως HDTV)
- ▶ Χρησιμοποιεί τεχνικές πρόγνωσης και μετασχηματισμού για τη μείωση του πλεονασμού
- ▶ Χρησιμοποιεί μετασχηματισμό ακεραίων σε μέγεθος μπλόκ 4x4 (ή 8x8)
- ▶ Χρησιμοποιεί μεταβλητού μεγέθους μακρομπλόκ από 16x16 έως 4x4



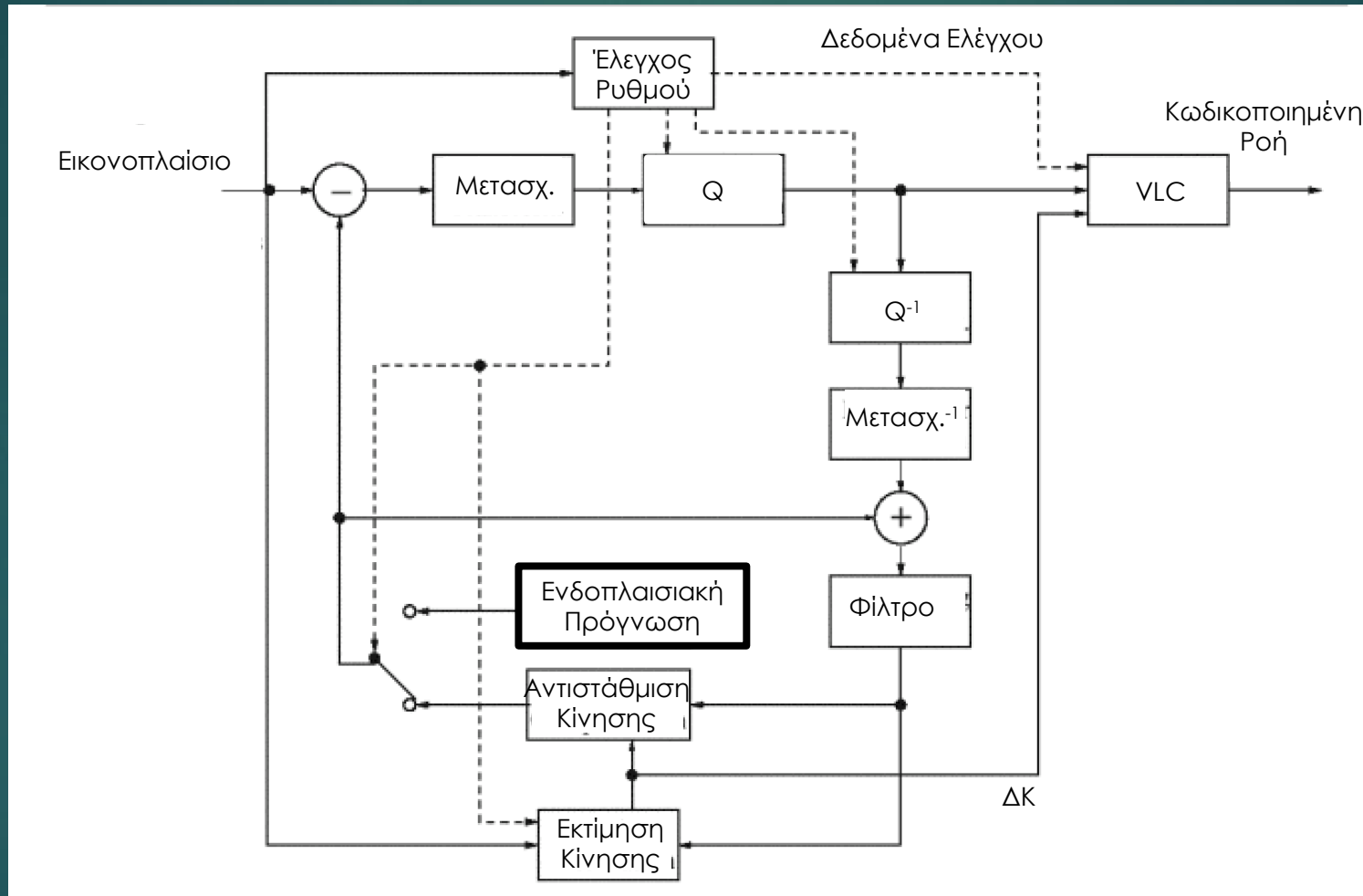
# H.264: Βασικά Χαρακτηριστικά

Δ13

- ▶ Ακρίβεια αναζήτησης έως  $\frac{1}{4}$  του εκονοστοιχείου με χρήση τεχνικών 2Δ παρεμβολής
- ▶ Χρήση περισσοτέρων εικονοπλαισίων αναφοράς για την εκτίμηση πλαισίων πρόγνωσης
- ▶ Χρήση κατευθυντικής ενδοπλαισιακής πρόγνωσης για εικονοπλαίσια τύπου I
- ▶ Χρήση φίλτρων για μείωση του φαινομένου μπλόκ
- ▶ Χρήση προσαρμοστικών κωδικών μεταβλητού μήκους για την κωδικοποίηση
- ▶ Χρήση εξελιγμένων τεχνικών συγχρονισμού για την ανάκαμψη από λάθη

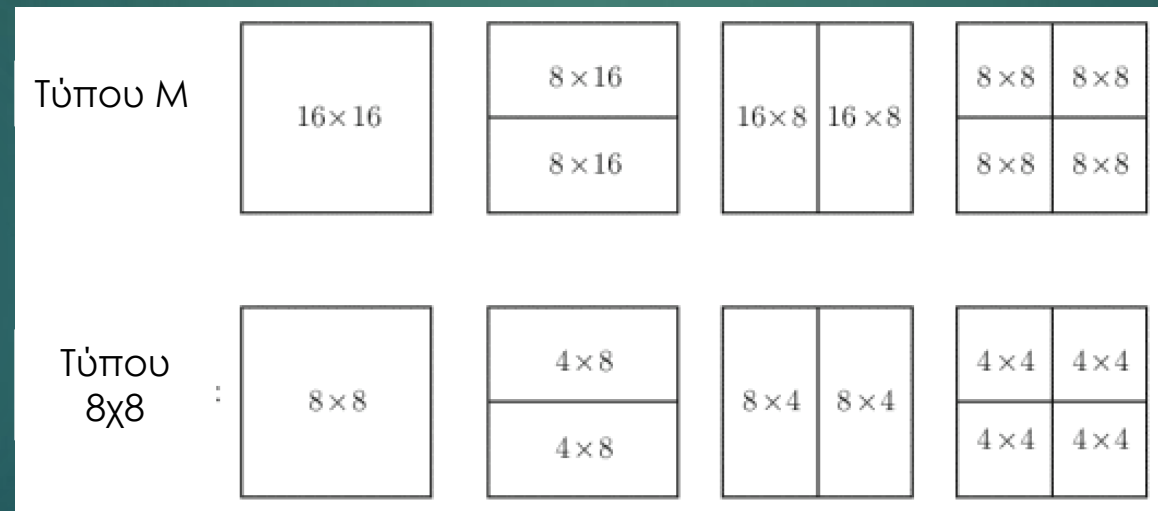
# H.264: Αρχιτεκτονική κωδικοποιητή

Δ13



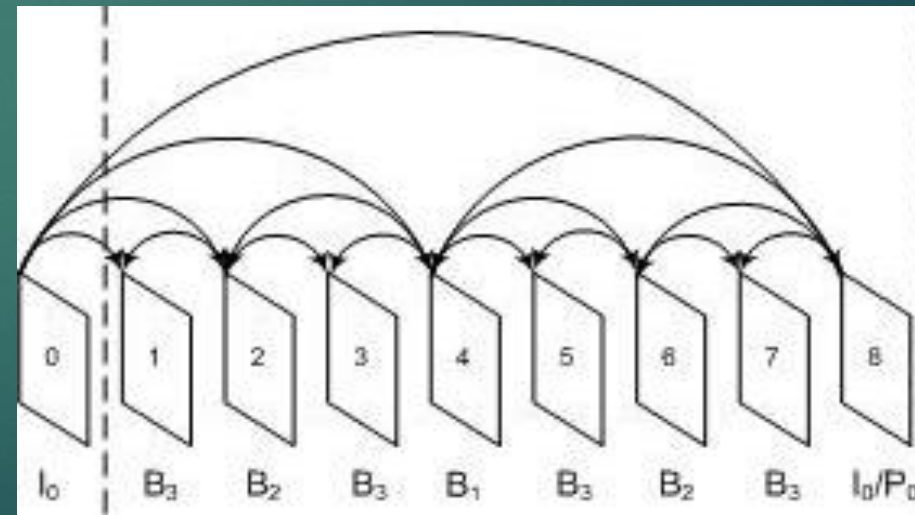
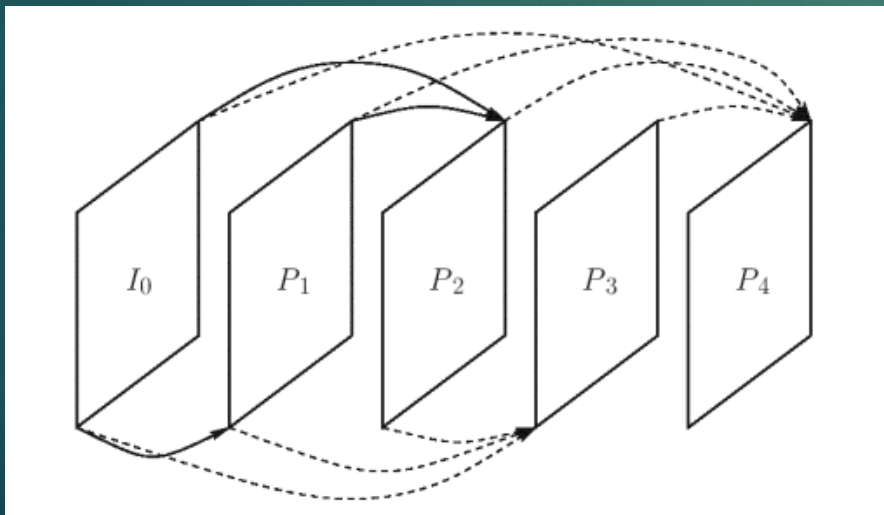
# H.264: Εκτίμηση κίνησης

- ▶ Παρόμοια με MPEG-2, H.263
- ▶ Χρησιμοποιεί μακρομπλόκ 16x16 (Τύπος M) αλλά και μικρότερα (τύπος 8x8)



# H.264: Είδη ομάδων πλαισίων (GOP)

- ▶ Επεκτείνεται η δομή ενός GOP
  - ▶ Υποστηρίζει GOP χωρίς B εικονοπλαίσια
  - ▶ Υποστηρίζει πολλαπλή πρόγνωση για P εικονοπλαίσια
  - ▶ Υποστηρίζει ιεραρχική πρόγνωση και δομή GOP



# H.264: Μετασχηματισμός

Δ13

- ▶ Η χρήση κλασσικού DCT θα εισήγαγε σφάλματα λόγω κινητής υποδιαστολής και αύξηση υπολογιστικής πολυπλοκότητας
- ▶ Αυξημένη ακρίβεια πρόγνωσης σε I και P,B μειώνει το συσχετισμό μεταξύ γειτονικών εικονοστοιχείων
- ▶ Αντικατάσταση με μετασχηματισμό ακεραίων 4x4 (ή 8x8)
  - ▶ ακριβής αντίστροφος
  - ▶ Ανεξάρτητος από το υλικό
  - ▶ Αποτελεί προσέγγιση του DCT
  - ▶ Αποδοτική υλοποίηση

# H.264: Μετασχηματισμός

- ▶ Προκύπτει από τον 4x4 πίνακα του DCT:

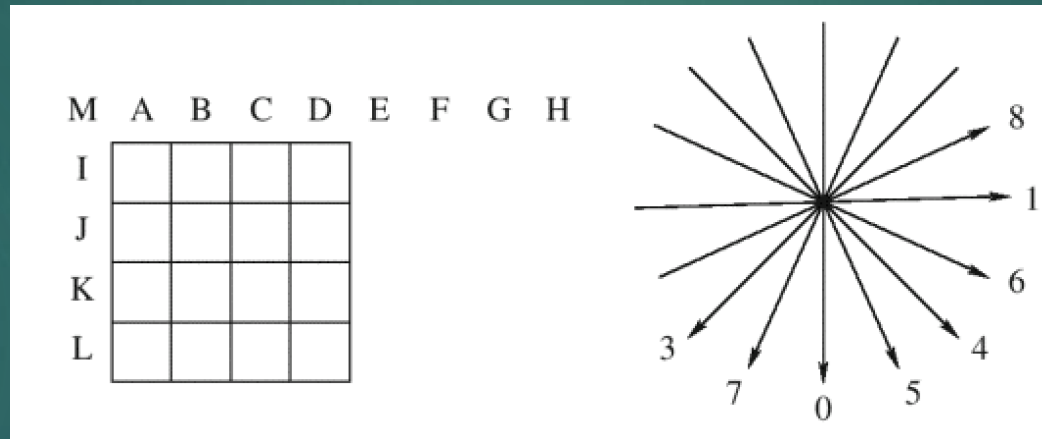
$$\mathbf{T}_4 = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix}$$

- ▶ Όπου  $a = 1/2$ ,  $b = \sqrt{\frac{1}{2}} \cos \frac{\pi}{8}$ ,  $c = \sqrt{\frac{1}{2}} \cos \frac{3\pi}{8}$
- ▶ Ο ακέραιος μετασχηματισμός προκύπτει ως:  $H = \text{round}(w \cdot T_4)$
- ▶ Για μεγάλες τιμές του  $w$  αυξάνει το μέγεθος των συντελεστών και δεν καλύπτει το μετασχηματισμό ακρίβεια 16 bit οπότε προτιμώνται μικρές τιμές (π.χ.  $w=2.5$ )
- ▶ Ο πίνακας μετασχηματισμού που προκύπτει είναι ορθογώνιος ( $H^{-1} = H^T$ ) αλλά όχι απαραίτητα κανονικοποιημένος (κατά γραμμές)
- ▶ Σε περίπτωση που απαιτείται κανονικοποίηση αυτή εκτελείται κατά το στάδιο της κβάντισης μεταβάλλοντας κατάλληλα τους συντελεστές κβάντισης ανά γραμμή
- ▶ Αντίστοιχη μέριμνα για την κανονικοποίηση λαμβάνεται υπόψη κατά την αποκωδικοποίηση

# H.264: Κωδικοποίηση εικονοπλαισίων I

Δ13

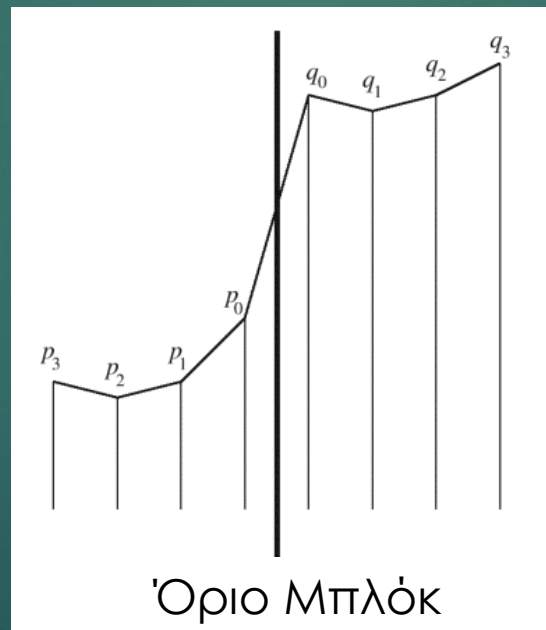
- ▶ Κάθε εικονοστοιχείο ενός 4x4 μπλόκ μπορεί να προγνωστεί από γείτονες του με χρήση 8 κατευθύνσεων πρόγνωσης
- ▶ Σε περίπτωση χρήσης μπλόκ 16x16 χρησιμοποιούνται μόνο 4 κατευθύνσεις πρόγνωσης



Li & Brew

# H.264: Φίλτρο μείωσης φαινομένου μπλόκ

- ▶ Φίλτρο εξομάλυνσης που εφαρμόζεται στα όρια των μπλόκ για να εξασφαλίσει ομαλή μεταβολή των τιμών
- ▶ Λαμβάνει υπόψη την κβάντιση για να προστατεύσει πραγματικές ακμές





# H.264: Στρωμάτωση

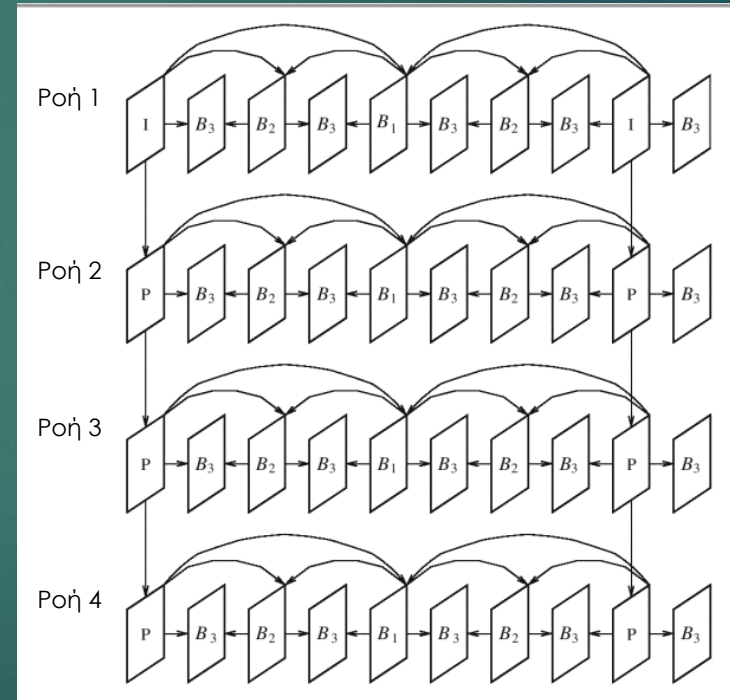
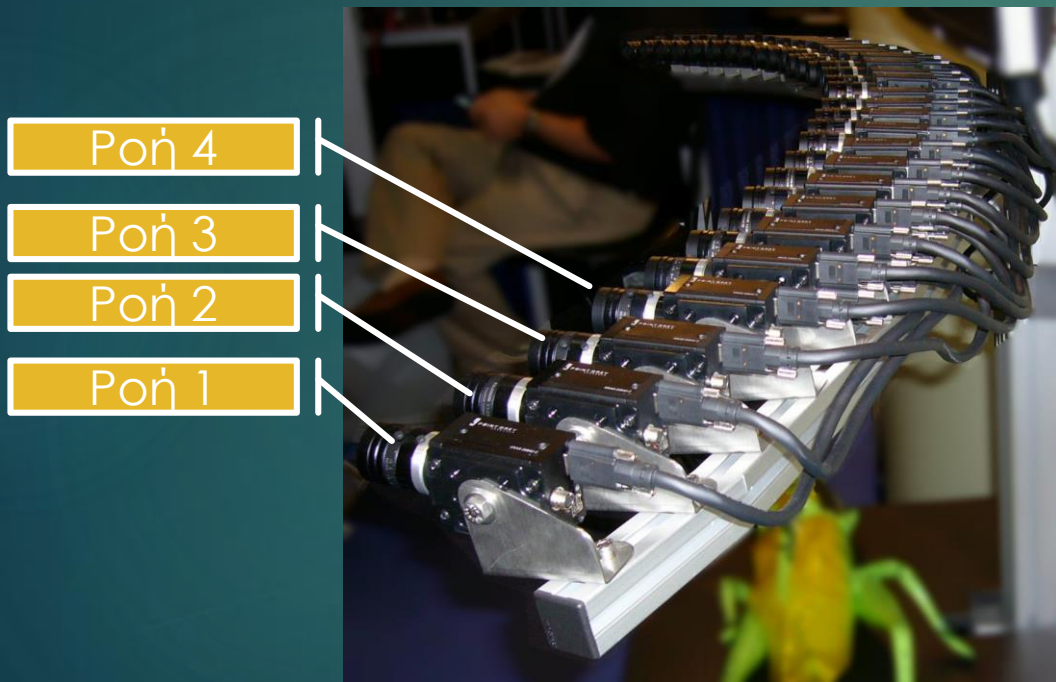
Δ13

- ▶ Υποστηρίζεται στρωμάτωση στο χρόνο , στην ανάλυση των εικονοπλαισίων και στο επίπεδο ποιότητας όπως και συνδυασμούς τους
- ▶ Παρόμοια με το MPEG-2 και MPEG-4

# H.264: 3Δ εικονοροές

Δ13

- ▶ Υποστηρίζονται 3Δ εικονοροές για μονοδιάστατες διατάξεις συσκευών λήψης



# H.265 (HEVC)- MPEG-H Part 2

# H.265: Βασικά Χαρακτηριστικά

Δ13

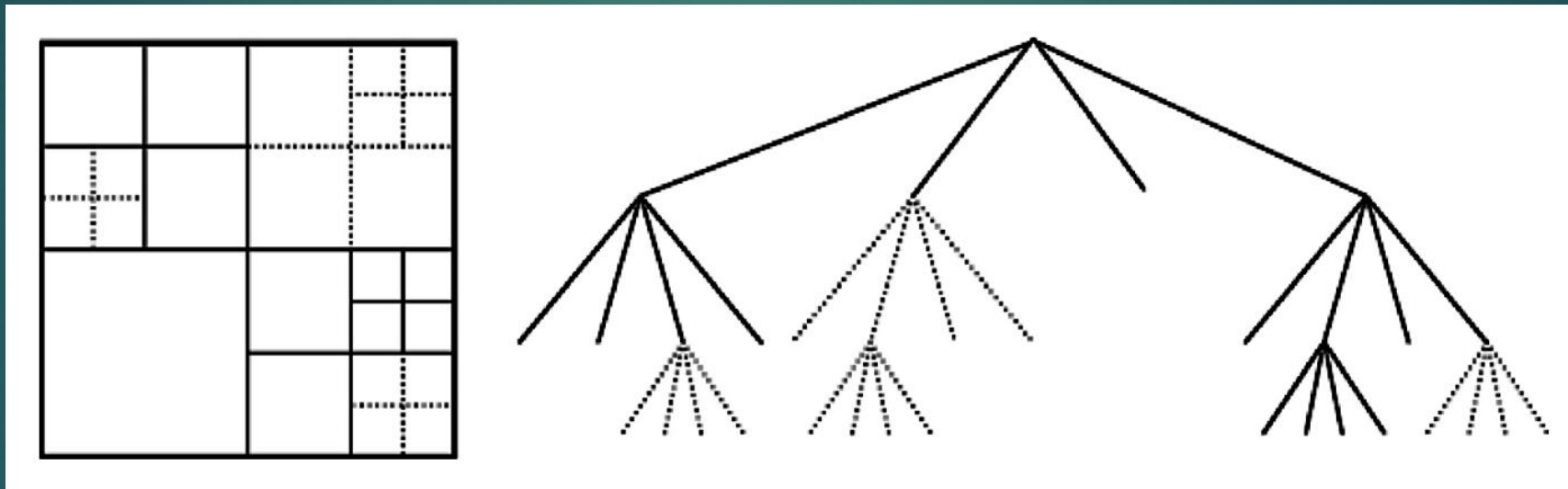
- ▶ Μεταβλητό μέγεθος μπλόκ για εκτίμηση κίνησης από 4x4 έως 64x64
- ▶ Τα μακρομπλόκ έχουν αντικατασταθεί από μία δομή τετραδικού δέντρου
- ▶ Αξιοποιούνται οι δυνατότητες παράλληλης επεξεργασίας του σύγχρονου υλικού
- ▶ Χρησιμοποιούνται μετασχηματισμοί ακεραίων μεγέθους 4x4, 8x8, 16x16, 32x32
- ▶ Χρησιμοποιείται ακρίβεια  $\frac{1}{4}$  εικονοστοιχείου με εξελιγμένες μεθόδους παρεμβολής
- ▶ Χρησιμοποιούνται 33 κατευθύνσεις πρόγνωσης για τα εικονοπλαίσια τύπου I
- ▶ Υπάρχει η δυνατότητα χρήσης του μετασχηματισμού DST σε συγκεκριμένες περιπτώσεις για το κανάλι Y
- ▶ Χρησιμοποιούνται φίλτρα για αφαίρεση του φαινομένου μπλόκ

# H.265: Δομή μακρομπλόκ (Τετραδικό δέντρο)

- ▶ CTU: Κάθε CTU περιέχει μία δομή για κάθε χρώμα η οποία ονομάζεται CTB. Το CTB για το κανάλι Y έχει μέγεθος  $N \times N$  όπου  $N=16,32,64$  και κάθε CTB για τα Cr,Cb έχει μέγεθος  $N/2 \times N/2$
- ▶ CB είναι οι μονάδες που οργανώνονται στο δέντρο του CTB με ελάχιστο μέγεθος  $8 \times 8$  (Y) και  $4 \times 4$  (Cb,Cr) και μέγιστο όσο το μέγεθος του CTB. Ένα CU περιλαμβάνει ένα CB του καναλιού Y και δύο CB των καναλιών Cb,Cr
- ▶ PB: Κάθε CB μπορεί να διαιρεθεί περαιτέρω σε PBs για την εκτέλεση πρόγνωσης. Το μοντέλο πρόγνωσης για κάθε PU, Το οποίο αντιστοιχεί στο CU, είναι σταθερό.
- ▶ TB: Κάθε CB μπορεί να διαιρεθεί περαιτέρω σε TBs για κωδικοποίηση με μεγέθη από  $32 \times 32$  έως  $4 \times 4$ . Τα TB ενός καναλιού Y και των αντίστοιχων Cb, Cr δομούν ένα TU.

# H.265: Δομή μακρομπλόκ (Τετραδικό δέντρο)

Δ13



# H.265: Δομή περιοχών κωδικοποίησης

- ▶ Ένα εικονοπλαίσιο χωρίζεται σε περιοχές οι οποίες περιέχουν πολλαπλά CTU
- ▶ Οι Περιοχές μπορεί να είναι:
  - ▶ τυχαίας μορφής και μεγέθους (Slices)
  - ▶ τετράγωνης μορφής (Tiles)
  - ▶ γραμμικής μορφής (σειρά από CTU) για χρήση σε παράλληλα συστήματα

# H.265: Μετασχηματισμός και κωδικοποίηση

- ▶ Ορίζεται ένας μετασχηματισμός ακεραίων μεγέθους 32x32 και οι υπόλοιποι δομούνται επιλέγοντας τις αρχικές N τιμές μιας γραμμής και υποδειγματοληπτώντας τις γραμμές
- ▶ Οι τιμές του μετασχηματισμού είναι αρκετά κοντά σε αυτές του DCT και δεν απαιτούν κανονικοποίηση
- ▶ Χρησιμοποιείται ο πίνακας κβάντισης του H.264 με ένα συντελεστή ποιότητας στο διάστημα [0-51]
- ▶ Οι κβαντισμένοι συντελεστές κωδικοποιούνται με κωδικοποιητή εντροπίας



# H.265: Σύνοψη του πρωτοκόλλου

Δ13

- ▶ <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6316136>