



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Τεχνικές συμπίεσης βίντεο κατά MPEG

Παντελής Μπαλαούρας
Τμήμα Πληροφορικής & Τηλεπικοινωνιών
2020 - 2021

Πηγές

[1] Παπαδάκης, Α. 2015. Συμπίεση Βίντεο. *Ψηφιακή τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 3. Διαθέσιμο στο: <http://hdl.handle.net/11419/5008>



Εισαγωγή

Γιατί απαιτείται η συμπίεση

Τυπική ευκρίνεια (Standard Definition –SD)

- 270 Mbps
- δεν υποστηρίζεται από τις τρέχουσες τεχνολογίες μετάδοσης

Στόχος συμπίεσης

- 2 – 7 Mbps
- 38:1 ποσοστό συμπίεσης

Υψηλή ευκρίνεια (Full High Definition – FHD)

- άνω του 1 Gbps

Στόχος συμπίεσης

- 15 -20 Mbps
- 137:1 ποσοστό συμπίεσης



Είδη συμπίεσης

- Μη απωλεστική
 - Απομάκρυνση πλεονάζουσας πληροφορίας
- Απωλεστική
 - Απώλεια πληροφορίας



Πλεονάζουσα πληροφορία

Πλεονάζουσα πληροφορία είναι αυτή που

- Εμφανίζεται πολλές φορές στη ροή
 - είναι δυνατή η απομάκρυνσή της σε κάποιες από αυτές τις εμφανίσεις
- Αφορά δεδομένα τα οποία δεν περιέχουν χρήσιμη πληροφορία
 - δεδομένα που χρησιμοποιούνται για να γεμίσουν κενές θέσεις στη ροή και να αποκτήσει αυτή συγκεκριμένη μορφή
- Δεδομένα τα οποία μπορούν να ανασκευαστούν στον δέκτη με τη βοήθεια μαθηματικών υπολογισμών



Τεχνικές περιορισμού της πλεονάζουσας πληροφορίας

- ΑΑΑΑΑΑΑΑ 8 φορές το Α
- Οι πλέον χρησιμοποιούμενες λέξεις έχουν περιορισμένο μήκος π.χ. τα άρθρα



Είναι πάντα εφικτό να πετυχαίνουμε την επιθυμητή συμπίεση, μόνο με **μη απωλεστικό τρόπο**;

Όχι !

Είναι δυνατόν η πληροφορία που θα χαθεί να έχει **μικρή επίπτωση στην ποιότητα του περιεχομένου** και στη γενικότερη αντίληψη του χρήστη;

Υπάρχει όμως τέτοιου είδους πληροφορία;

Ναι !



- Κατάλληλη προτεραιότητα στα διάφορα είδη πληροφορίας, βασιζόμενοι στη **γνώση της φυσιολογίας της όρασης και ακοής**, π.χ.
 - ο ανθρώπινος οφθαλμός είναι πολύ πιο ευαίσθητος στις αλλαγές στις τιμές της **φωτεινότητας** σε σχέση με τις αλλαγές στις τιμές των **χρωμάτων**.
 - δεν είναι ευαίσθητος στις **υψηλές συχνότητες** στην οπτική πληροφορία χρωμάτων που αντιστοιχούν στις **συνεχόμενες αλλαγές στις τιμές των χρωμάτων μεταξύ γειτονικών εικονοστοιχείων**



Μηχανισμοί συμπίεσης βίντεο

Είδη πλεονασμού

Η πλειονότητα των προτύπων κωδικοποίησης και συμπίεσης ψηφιακού βίντεο στηρίζονται στην **αξιοποίηση διαφόρων ειδών πλεονασμού στοχεύοντας στη μείωσή του**, άρα και στη μείωση του όγκου των δεδομένων.

Τα **είδη πλεονασμού** είναι τα επόμενα:

- **Χωρικός πλεονασμός**, ο οποίος οφείλεται στις ομοιότητες μεταξύ των εικονοστοιχείων που δομούν ένα πλαίσιο.
- **Χρονικός πλεονασμός**, ο οποίος οφείλεται στις ομοιότητες μεταξύ διαδοχικών εικόνων ενός βίντεο.
- **Στατιστικός πλεονασμός**, ο οποίος οφείλεται στις στατιστικές ιδιότητες της εμφάνισης των διαφόρων συμβόλων προς κωδικοποίηση.
- **Ψυχοοπτικός πλεονασμός**, ο οποίος οφείλεται στις ιδιότητες του ανθρώπινου συστήματος όρασης και στα χαρακτηριστικά της απόκρισης του συστήματος όρασης στα διάφορα ερεθίσματα.



Χωρική & Χρονική κωδικοποίηση

- video: ακολουθία εικόνων που εμφανίζονται με σταθερό ρυθμό
 - π.χ. 24 εικόνες/sec
- Κωδικοποίηση/συμπύεση: χρησιμοποίησε πλεονασμό **εντός** και **μεταξύ** των εικόνων για να μειωθεί ο # bits που χρησιμοποιούνται για κωδικοποίηση της εικόνας
 - χωρικό (εντός της εικόνας)
 - χρονικό (από μια εικόνα στην άλλη)



πλαίσιο i

παράδειγμα χωρικής κωδικοποίησης: αντί να στείλει N τιμές ίδιου χρώματος (όλα μωβ), στέλνει μόνο 2 τιμές: τιμή χρώματος (μωβ) και αριθμό επαναλαμβανόμενων τιμών (N)



πλαίσιο $i+1$

παράδειγμα χρονικής κωδικοποίησης: αντί να στείλει ολόκληρο το πλαίσιο στο $i+1$, στέλνει μόνο τις διαφορές από το πλαίσιο i



Το προς συμπίεση σήμα

- Στην περίπτωση της ψηφιακής τηλεόρασης το προς συμπίεση σήμα είναι αυτό που παράγεται από τον εικονολήπτη (κάμερα).
- Οι συνιστώσες των βασικών χρωμάτων, τα οποία είναι το **κόκκινο**, το **πράσινο** και το **μπλε (R, G, B)** συνδυάζονται ώστε να δώσουν το σήμα της φωτεινότητας Y και των χρωμοδιαφορών μπλε και κόκκινου (C_b και C_r).
- Η ανάλυση του χρώματος είναι μισή σε σχέση με την ανάλυση της φωτεινότητας, στην περίπτωση σήματος 4:2:2.
- Στην περίπτωση που έχουμε κβαντισμό **10 bits**, ο ρυθμός μπορεί να φτάνει και τα **270 Mbps** σύμφωνα και με το πρότυπο του ITU-R BT.601



Βασικά βήματα στη συμπίεση βίντεο

Προετοιμασία σήματος βίντεο

1. Μείωση ακρίβειας κωδικοποίησης



2. Παράλειψη τμημάτων βίντεο με μη χρήσιμη πληροφορία



3. Μείωση ανάλυσης περιγραφής χρωμάτων



6. Εκμετάλλευση στατιστικού πλεονασμού



5. Μείωση χωρικού πλεονασμού



4. Μείωση χρονικού πλεονασμού

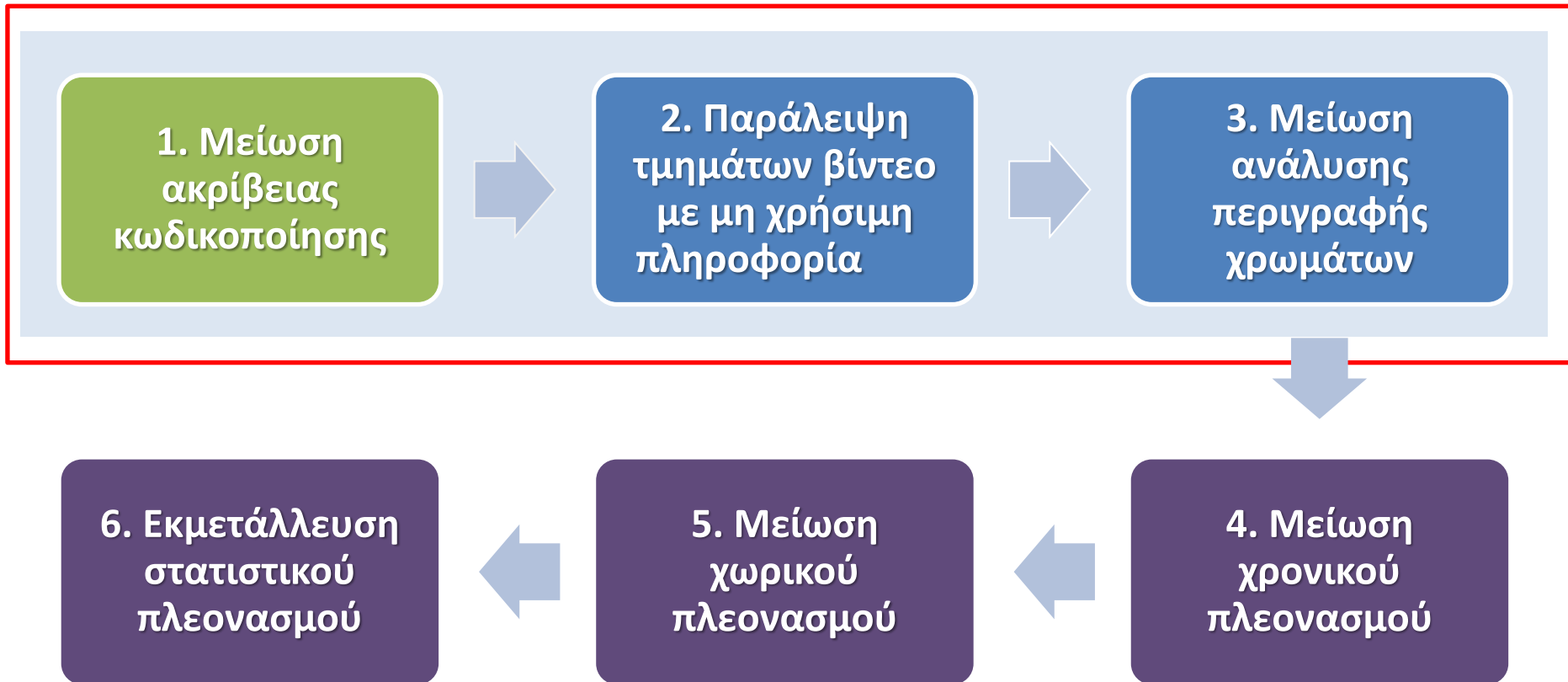


Ο μηχανισμός συμπίεσης του βίντεο

1. Μείωση της ακρίβειας των τιμών της φωτεινότητας και των χρωμοδιαφορών, δηλαδή μείωση του πλήθους των επιπέδων κβαντισμού και συνακόλουθη μείωση των bits που χρησιμοποιούνται για την αναπαράσταση των τιμών από 10 σε 8.
2. Παράλειψη των τμημάτων του βίντεο τα οποία δεν περιέχουν χρήσιμη πληροφορία, όπου με την έννοια χρήσιμη πληροφορία εννοούμε το τμήμα που δεν αντιστοιχεί σε οπτικό σήμα αλλά στους οριζόντιους και κατακόρυφους παλμούς αμαύρωσης.
3. Περαιτέρω μείωση της ανάλυσης της περιγραφής των χρωμάτων, πέρα από το πρότυπο 4:2:2 το οποίο χρησιμοποιείται κατά τη δημιουργία του σήματος. Είναι δηλαδή δυνατή η χρήση διαφορετικών προτύπων στις γραμμές του σήματος, ώστε να εκμεταλλευτούμε τη σχετικά περιορισμένη ευαισθησία του ανθρώπινου οφθαλμού στην πληροφορία που αφορά τα χρώματα.
4. Αξιοποίηση της ομοιότητας η οποία τυπικά εμφανίζεται μεταξύ χρονικά διαδοχικών πλαισίων. Η ιδέα είναι να γίνει χρήση της περιγραφής ενός υπάρχοντος πλαισίου για την υποβοήθηση της περιγραφής ενός από τα επόμενα πλαίσια. Χρησιμοποιούνται, δηλαδή, εικόνες αναφοράς και αξιοποιείται η κωδικοποίηση μόνο των διαφορών μεταξύ των διαδοχικών εικόνων. Η ομοιότητα αυτή είναι πιο έντονη στα χρονικά διαστήματα που δεν έχουμε έντονη κίνηση στο βίντεο.
5. Αξιοποίηση της ομοιότητας που ενδέχεται να υπάρχει μεταξύ των περιοχών που συνθέτουν ένα πλαίσιο. Η ομοιότητα αυτή είναι αρκετά τυπική σε πολλές περιπτώσεις όπου έχουμε σχεδόν ομοιόμορφες περιοχές, όπως, για παράδειγμα, τον ουρανό ή άλλες επιφάνειες.
6. Αξιοποίηση του γεγονότος ότι ορισμένα από τα σύμβολα εμφανίζονται πιο συχνά από άλλα. Ότι το περιεχόμενο, δηλαδή, μπορεί να έχει σταθερές στατιστικές ιδιότητες αναφορικά με τη συχνότητα εμφάνισης των συμβόλων. Τέτοιου είδους ιδιότητες είναι χρήσιμες αναφορικά με την κωδικοποίηση των συμβόλων.



Προετοιμασία σήματος βίντεο



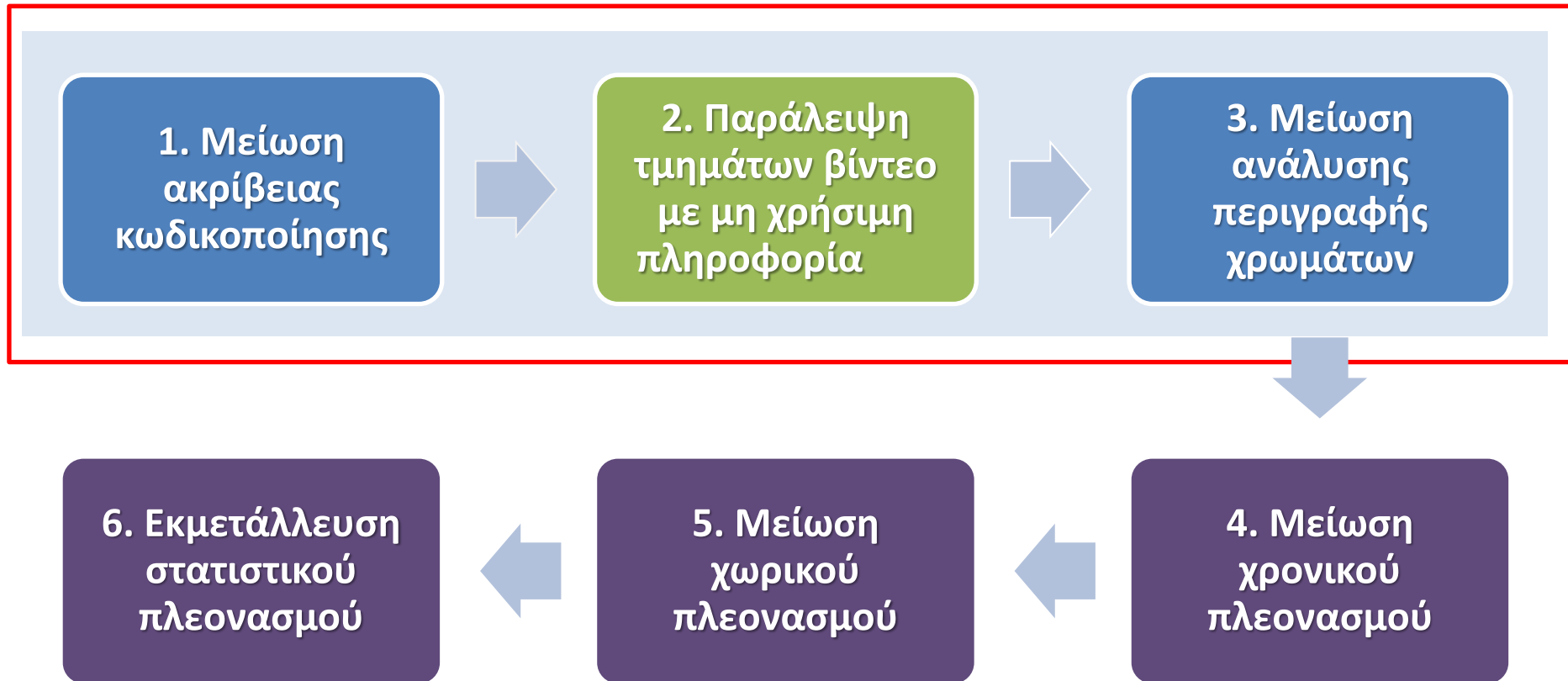
1. Μείωση ακρίβειας βίντεο

Μείωση της ακρίβειας των τιμών της φωτεινότητας και των χρωμοδιαφορών, δηλαδή μείωση του πλήθους των **επιπέδων κβαντισμού** και συνακόλουθη μείωση των bits που χρησιμοποιούνται για την αναπαράσταση των τιμών από 10 σε 8.

- Οι συνιστώσες Y, Cb, Cr μπορούν να έχουν ακρίβεια **10 bits** όταν χρησιμοποιούνται εντός του στούντιο αλλά η συγκεκριμένη ακρίβεια δεν είναι απαραίτητη κατά τη μετάδοση του σήματος.
- Όσον αφορά τον ρυθμό μετάδοσης η **μείωση ανέρχεται στο 20%**, παραμένει το 80% του αρχικού περιεχομένου (λόγω του υποβιβασμού της ακρίβειας κβαντισμού από τα 10 στα 8 bits).



Προετοιμασία σήματος βίντεο



2. Παράλειψη μη χρήσιμης πληροφορία

- Το σήμα βίντεο το οποίο παράγεται έχει κάποια χρονικά διαστήματα, στα οποία δεν αντιστοιχεί **χρήσιμο οπτικό σήμα**
 - για λόγους συμβατότητας με τα αναλογικά συστήματα τηλεόρασης (τα οποία βασίζονται στον καθοδικό σωλήνα, είτε όσον αφορά τη λήψη του σήματος είτε την αναπαραγωγή του),
- Αυτά είναι τα διαστήματα που αντιστοιχούν στον **οριζόντιο** και τον **κατακόρυφο** παλμό επιστροφής (ή αμαύρωσης). Τα διαστήματα αυτά δεν φέρουν χρήσιμη πληροφορία.
 - ενώ στην αναλογική τηλεόραση κατά τη διάρκεια του κατακόρυφου παλμού αμαύρωσης μεταδίδεται το Teletext, στην περίπτωση της ψηφιακής τηλεόρασης, οι συγκεκριμένες πληροφορίες μεταδίδονται με διαφορετικό τρόπο, ως διακριτή ψηφιακή ροή.
 - παρά τη δυνατότητα να μεταδίδονται άλλου είδους πληροφορίες, όπως, για παράδειγμα, ο ήχος, στην περίπτωση της ψηφιακής τηλεόρασης δεν γίνεται εκμετάλλευση αυτών των διαστημάτων (αφού και ο ήχος μεταδίδεται ξεχωριστά ως διακριτή ροή).



2. Παράλειψη μη χρήσιμης πληροφορία

- Αυτό σημαίνει ότι στην κωδικοποίηση MPEG, τα χρονικά διαστήματα που αντιστοιχούν στους παλμούς επιστροφής δεν αξιοποιούνται.
- Επειδή στην πλευρά του δέκτη είναι δυνατή η επαναδημιουργία τους, **δεν απαιτείται η αποστολή τους.**



2. Παράλειψη μη χρήσιμης πληροφορία

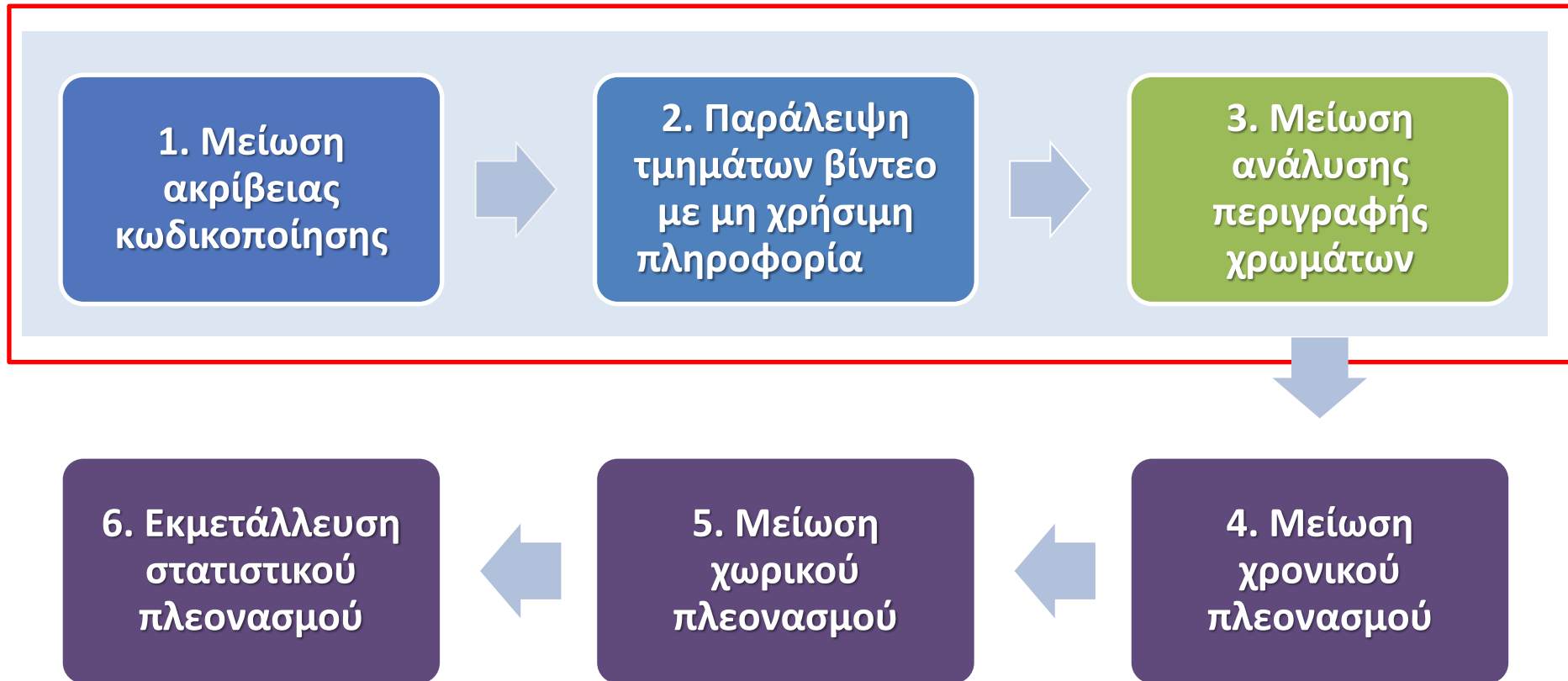
Ποιο είναι το κέρδος στην περίπτωση που δεν αποστέλλονται τα συγκεκριμένα χρονικά διαστήματα, για την περίπτωση του ευρωπαϊκού συστήματος;

- Αναφορικά με τον **παλμό κατακόρυφης επιστροφής / αμαύρωσης**, αυτός αντιστοιχεί σε περίπου **50 γραμμές από τις 625, άρα στο 8%** του συνολικού σήματος $50 / 625 = 8\%$
- Επιπλέον, υπενθυμίζουμε ότι κάθε γραμμή διαρκεί **64 μσεκ των οποίων τα 12 μs** αφιερώνονται στην επιστροφή της δέσμης από δεξιά προς τα αριστερά. Αυτό αντιστοιχεί στο ποσοστό που δίνεται από την επόμενη σχέση: $12 / 64 = 19\%$

Αθροιστικά, τα δύο παραπάνω ποσοστά δίνουν **περί το 25%** του σήματος, λόγω και της επικάλυψης που υπάρχει μεταξύ των παλμών.



Προετοιμασία σήματος βίντεο



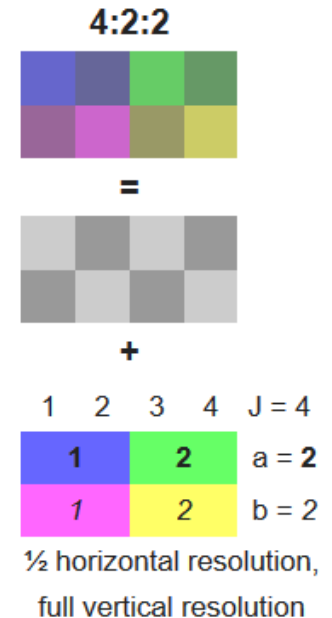
Προετοιμασία σήματος βίντεο

3. Μείωση ανάλυσης περιγραφής χρωμάτων

- **Υποδειματοληψία Χρώματος:** Οι συνιστώσες που αφορούν τα χρώματα και συγκεκριμένα οι χρωμοδιαφορές Cb, Cr λαμβάνονται με μικρότερη συχνότητα δειγματοληψίας σε σχέση με τη φωτεινότητα Y.
- Η συχνότητα δειγματοληψίας είναι η μισή (κατά την οριζόντια κατεύθυνση) και σε κάθε 4 δείγματα φωτεινότητας αντιστοιχούν 2 δείγματα από τις χρωμοδιαφορές. Έχουμε, δηλαδή, το πρότυπο:

4 : 2 : 2

- Αυτό συνεπάγεται τη μείωση του εύρους ζώνης του σήματος των χρωμοδιαφορών στα 2,75 MHz (σε σχέση με το σήμα της φωτεινότητας το οποίο έχει εύρος ζώνης 5,75 MHz).



Προετοιμασία σήματος βίντεο

3. Μείωση ανάλυσης περιγραφής χρωμάτων

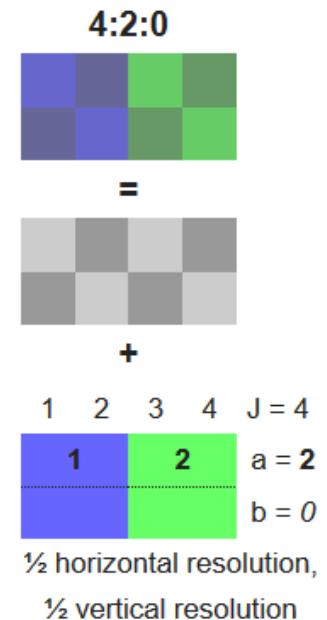
Είναι δυνατή η περαιτέρω μείωση του ρυθμού του σήματος των χρωμοδιαφορών κατά το ήμισυ.

Συγκεκριμένα, το σκεπτικό είναι να λαμβάνονται δείγματα χρωμοδιαφορών ανά δύο γραμμές (στη μία, δηλαδή, λαμβάνονται και στην άλλη όχι). Αυτό μας δίνει το πρότυπο:

4 : 2 : 0

το οποίο δεν επιβαρύνει ιδιαίτερα την αντιλαμβανόμενη από τον άνθρωπο ποιότητα.

Από την πλευρά του απαιτούμενου ρυθμού έχουμε υποβάθμιση αφού 2×4 δείγματα φωτεινότητας αντιστοιχούν σε 2 δείγματα από τις χρωμοδιαφορές. Η υποβάθμιση του ρυθμού αντιστοιχεί **στο 25%**.



Προετοιμασία σήματος βίντεο

Συνολικός περιορισμός ρυθμού

1. Μείωση ακρίβειας κωδικοποίησης

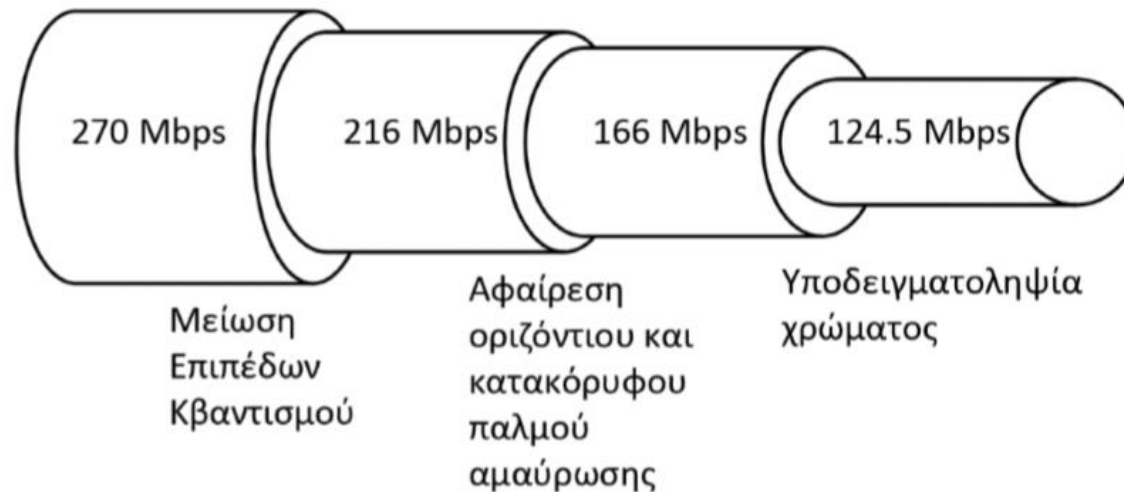


2. Παράλειψη τμημάτων βίντεο με μη χρήσιμη πληροφορία



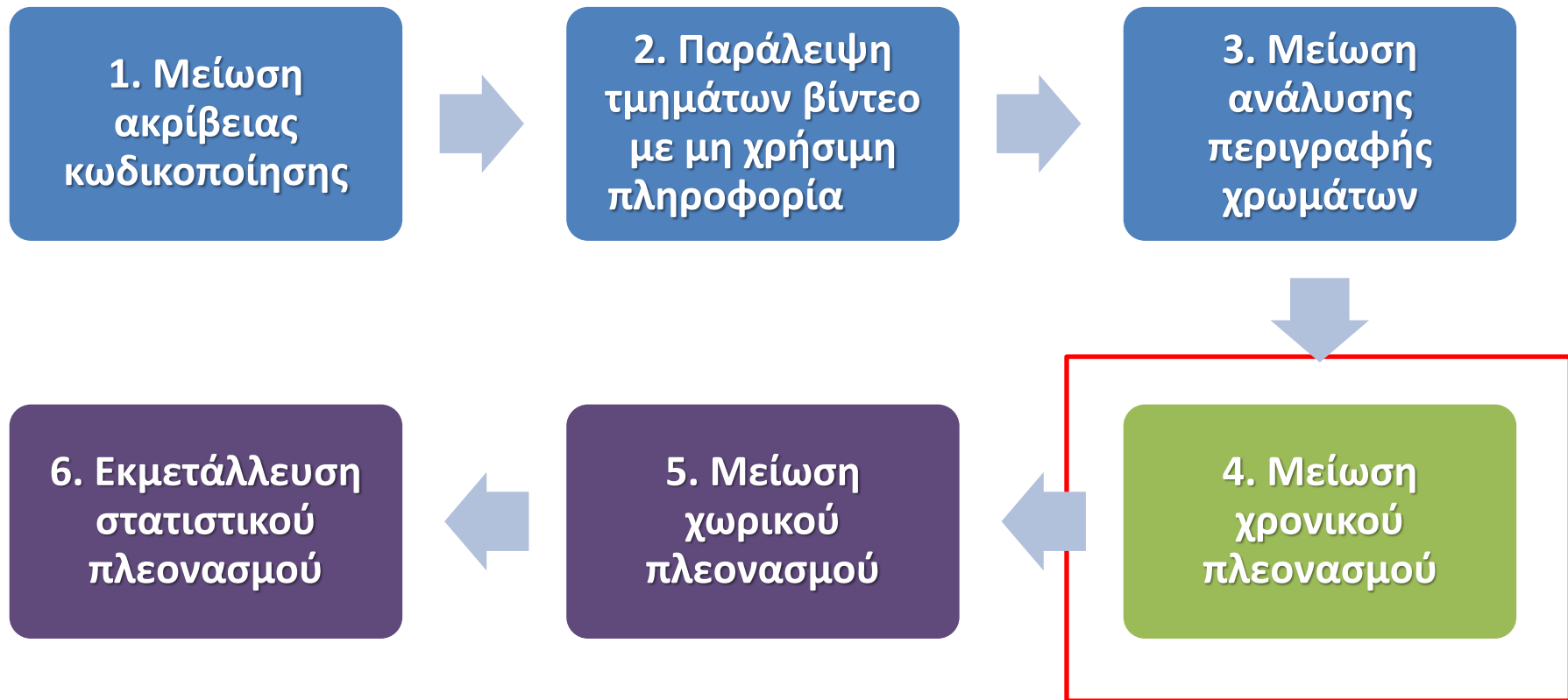
3. Μείωση ανάλυσης περιγραφής χρωμάτων

εικονολήπτης



Επόμενος στόχος: πως θα περιοριστεί ο ρυθμός στα 2 - 4 Mbps

4. Μείωση χρονικού πλεονασμού



4. Μείωση χρονικού πλεονασμού (1)

- Το βίντεο είναι **συνεχόμενες εικόνες** οι οποίες λαμβάνονται αρκετές φορές σε ένα δευτερόλεπτο π.χ. 25 πλαίσια/δευτ.
- Οι εικόνες αυτές στις **περισσότερες περιπτώσεις μοιάζουν (οπτικά) μεταξύ τους**
 - το **στατικό κομμάτι** παραμένει σε μεγάλο βαθμό σταθερό από εικόνα σε εικόνα
 - το **δυναμικό κομμάτι**, σε αυτό, δηλαδή, που υπάρχει **κάποια κίνηση**



πλαίσιο i



πλαίσιο $i+1$

Οι διαφοροποιήσεις μεταξύ των εικόνων μπορεί να είναι **περιορισμένες** και οι αλλαγές στις εικόνες μπορεί να προέρχονται από **μετακινήσεις αντικειμένων** μεταξύ πλαισίων.

Μηχανισμός περιορισμού χρονικού πλεονασμού

- Η μετάδοση **ολόκληρης της περιγραφής** μιας εικόνας οδηγεί στην
 - επαναλαμβανόμενη μετάδοση της ίδιας πληροφορίας
 - ή πληροφορίας που μπορεί να δημιουργηθεί εκ νέου στον δέκτη
- Το σκεπτικό του περιορισμού του χρονικού πλεονασμού **είναι να αποστέλλεται μόνο η απαραίτητη πληροφορία**
 - πληροφορία που δεν θα μπορούσε να δημιουργηθεί εκ νέου



Μηχανισμός περιορισμού χρονικού πλεονασμού

- Η απαραίτητη πληροφορία σχετίζεται με τις **περιοχές που υπάρχουν διαφοροποιήσεις σε σχέση με τα προηγούμενα πλαίσια**
 - είναι απαραίτητη η μετάδοση της **διαφοράς μεταξύ των πλαισίων**
- Ο μηχανισμός περιορισμού περιλαμβάνει
 - A. τον διαχωρισμό της εικόνας σε υποπεριοχές
 - B. την κωδικοποίηση των εικόνων κατά
 - I. I πλαίσια – κωδικοποιούνται αυτόνομα & ανεξάρτητα
 - II. P πλαίσια – κωδικοποιούνται οι διαφορές με τα I
 - III. B πλαίσια - κωδικοποιούνται οι διαφορές με τα I & P



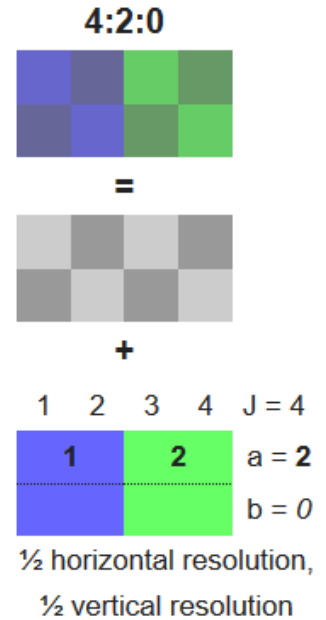
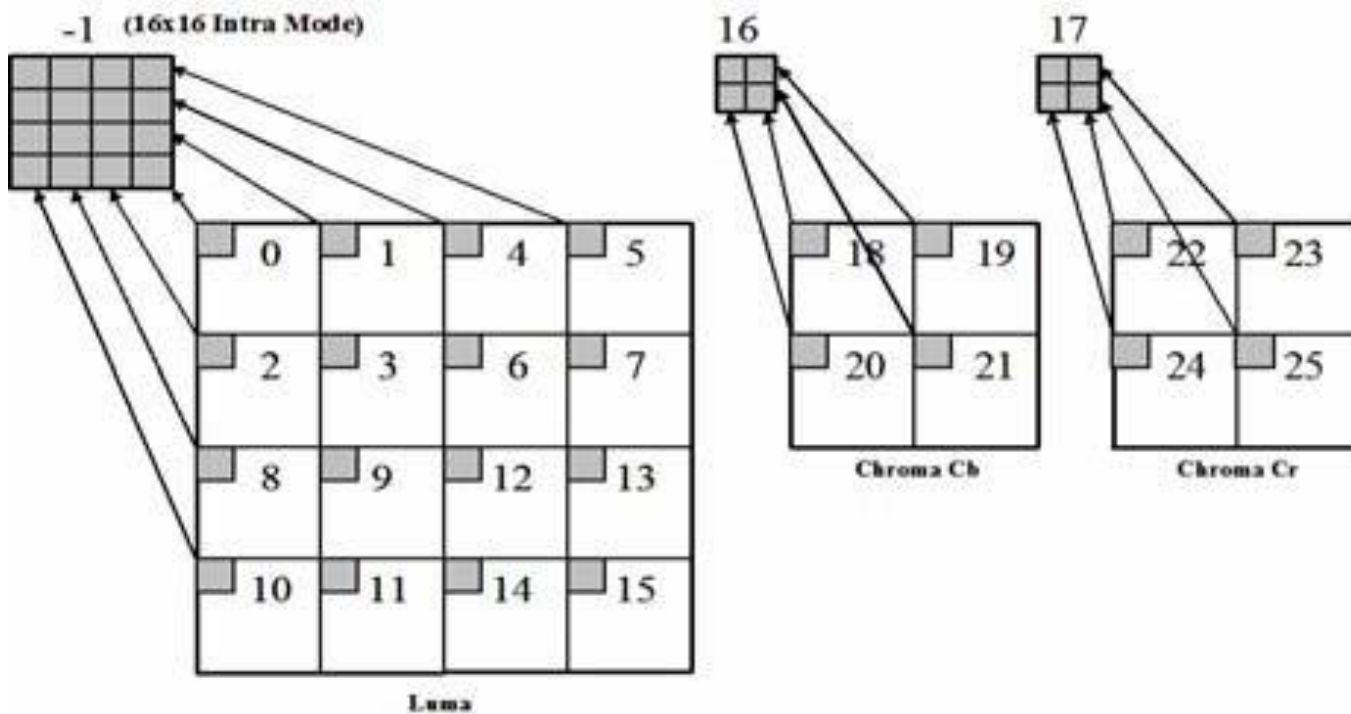
A. Διαχωρισμός σε Υποπεριοχές (blocks)

- Κατά τον διαχωρισμό της εικόνας σε **περιοχές**, αναζητούνται
 - οι **στατικές** και
 - οι **δυναμικές** περιοχές
- Για να είναι διαχειρίσιμη η αναζήτηση των περιοχών αυτών εντός μιας εικόνας, αυτή χωρίζεται σε **ίσες υποπεριοχές**
 - τμήματα (block) των 16x16 εικονοστοιχείων (pixels)
 - 16x16 τιμές φωτεινότητας (256)
 - χρώματα είναι σε πλήθος το $2/8=1/4$ αυτών της φωτεινότητας (πρότυπο 4:2:0)
 - 64 τιμές για τα C_b & C_r (32 & 32)
 - αντιστοιχούν, δηλαδή, σε μια περιοχή 8x8 για κάθε μια από τις χρωμοδιαφορές C_b και C_r (64 τιμές)
 - το σύνολο των τιμών αυτών (256+32+32) αντιστοιχεί σε ένα **μακρο-μπλοκ**



μακρο-μπλοκ

- 18, 22 -> 0, 1, 2, 3



B-I. Κωδικοποίηση των εικόνων κατά I

- Ορισμένα από τα **πλαίσια κωδικοποιούνται ανεξάρτητα**, χωρίς δηλαδή να λαμβάνεται υπόψη πληροφορία από άλλα πλαίσια
 - I- **πλαίσια** (intracoded frames) ενδοπλαισιακή συμπίεση
 - κωδικοποιούνται ανεξάρτητα από άλλα πλαίσια
 - Χρησιμοποιούν τεχνικές συμπίεσης που εφαρμόζονται στην πληροφορία μόνον μέσα στο ίδιο το πλαίσιο
 - αποστέλλονται περιοδικά
 - αποτελούν πλαίσια-κλειδιά (key frames) στην αλληλουχία των συμπιεσμένων πλαισίων που δημιουργεί το MPEG.



B-II Κωδικοποίηση των εικόνων κατά P

- Κάποια πλαίσια κωδικοποιούνται **λαμβάνοντας υπόψη προηγούμενα πλαίσια** - διαπλαισιακή (interframe) συμπίεση
 - Η κωδικοποίηση γίνεται με **εκμετάλλευση των ομοιοτήτων** που τυχόν υπάρχουν μεταξύ του **τρέχοντος πλαισίου** και κάποιου (**συγκεκριμένου**) **προηγούμενου**
 - Οι ομοιότητες αναζητούνται και εντοπίζονται σε επίπεδο μακρο-μπλοκ, δηλαδή σε τμήματα μεγέθους 16x16
- **P-frames (predicted frames)** δηλαδή πλαίσια πρόβλεψης



B-II Κωδικοποίηση των εικόνων κατά P

- Επισημαίνεται ότι ο εντοπισμός της τυχόν ομοιότητας σε ένα πλαίσιο τύπου P, γίνεται αξιοποιώντας το **προηγούμενο πλαίσιο τύπου I ή P** (ανάλογα με το πιο είναι πιο κοντινό χρονικά)
- Αποθηκεύεται στο πλαίσιο P μόνο η πληροφορία που δείχνει το πόσο έχει **μεταβληθεί αυτό το πλαίσιο σε σχέση με ένα προηγούμενο πλαίσιο αναφοράς** (συνήθως ένα προηγούμενο $-I$ ή $-P$ πλαίσιο).
- Έτσι το πλαίσιο που αντιστοιχεί στο πλαίσιο P αναδημιουργείται προσθέτοντας στο πλαίσιο αναφοράς (I ή P) την **πληροφορία μεταβολής** (πλαίσιο P).
- Η κωδικοποίηση των πλαισίων τύπου P απαιτεί **λιγότερα δεδομένα** σε σχέση με την κωδικοποίηση πλαισίων τύπου I



P Frame

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

I Frame (Frame 1)

minus

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0

Frame 3

=

		0	0		
		0	1	1	
			1	1	

P Frame

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

I Frame (Frame 1)

minus

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0

Frame 3

=

		0	0		
		0	1	1	
			1	1	

P Frame



Δημιουργία P Frame

Frame I (1)

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Frame 3

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0

Δ

=

Frame P

		0	0		
		0		1	
			1	1	



Αναπαραγωγή Frame 3 P Frame

Frame I (1)

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Frame P

		0	0		
		0		1	
			1	1	

Frame 3

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0



B- III Κωδικοποίηση των εικόνων κατά B

- Υπάρχουν πλαίσια τα οποία κωδικοποιούνται αξιοποιώντας τον εντοπισμό ομοιοτήτων με **προηγούμενα** αλλά και **επόμενα πλαίσια**
 - από ένα προηγούμενο και ένα επόμενο (χρονικά) πλαίσιο
 - I και P ή P και P
- B-frames (bidirectional frames) δηλαδή **πλαίσια αμφίδρομης πρόβλεψης**
- Αποθηκεύουν μόνον την πληροφορία μεταβολής τους σε σχέση με ένα προηγούμενο και με ένα επόμενο πλαίσιο αναφοράς
- Το πλαίσιο που αντιστοιχεί στο πλαίσιο-B αναδημιουργείται λαμβάνοντας υπόψη και τα δύο αυτά πλαίσια (πριν & μετά)
- Η κωδικοποίηση των πλαισίων τύπου B απαιτεί λιγότερα δεδομένα σε σχέση με την κωδικοποίηση πλαισίων τύπου P (και βέβαια πλαισίων τύπου I)



B Frame

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

I Frame (Frame 1)

minus

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0

Frame 2

		0		1	
		0		1	

		0		1	
		0		1	

		0	0		
		0		1	
			1	1	

P Frame

				1	

$$\text{B Frame} = \text{Frame 1} - \text{Frame 2} - \text{P Frame}$$



Δημιουργία P Frame

Frame I (1)

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Frame 2

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0

=

Ενδιαμεσο

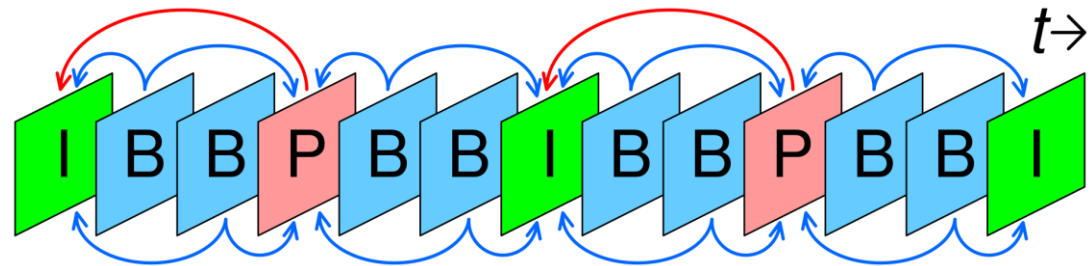
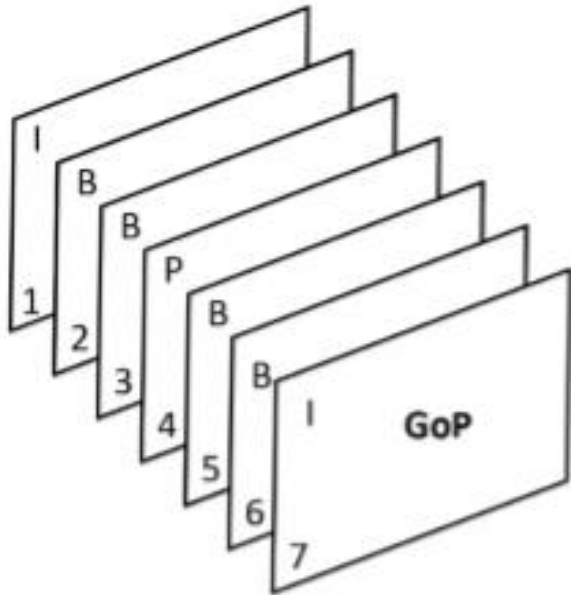
		0		1	
		0		1	

		0	0		
		0		1	
			1	1	



Group of Pictures (GoP)

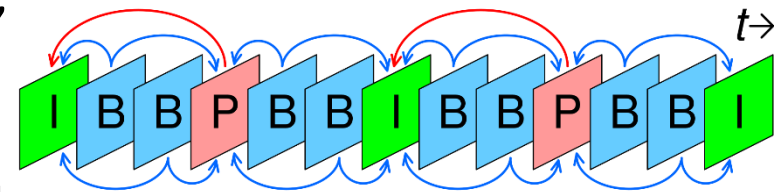
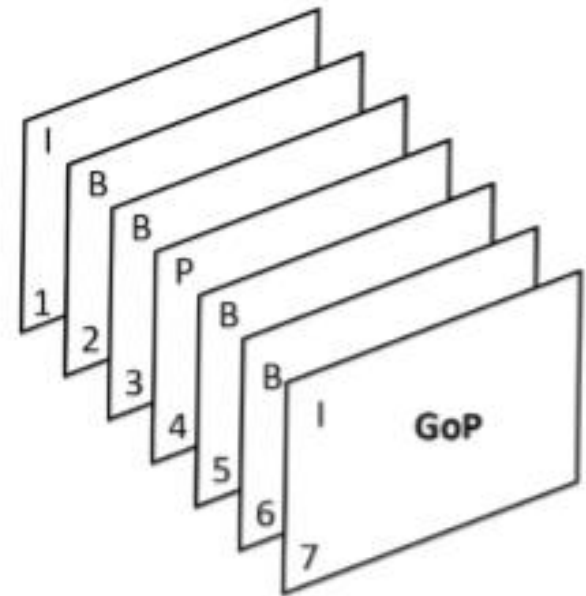
- Η αλληλουχία των πλαισίων I, P, B και ο τρόπος που εναλλάσσονται είναι δεδομένος και αμετάβλητος στο χρόνο.
- Συγκεκριμένα, ορίζεται η ομάδα εικόνων, **Group of Pictures (GoP)**, η οποία αποτελείται από μια αλληλουχία διαφορετικών τύπων πλαισίων, I, P και B.
- Ένα παράδειγμα GoP συνολικά 7 πλαισίων: Παρατηρούμε ότι αποτελείται από τα εξής πλαίσια: I, B, B, P, B, B και I.



Group of Pictures (GoP)

Κάποιες παρατηρήσεις:

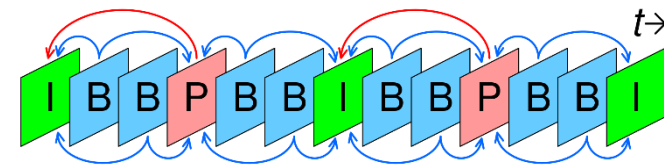
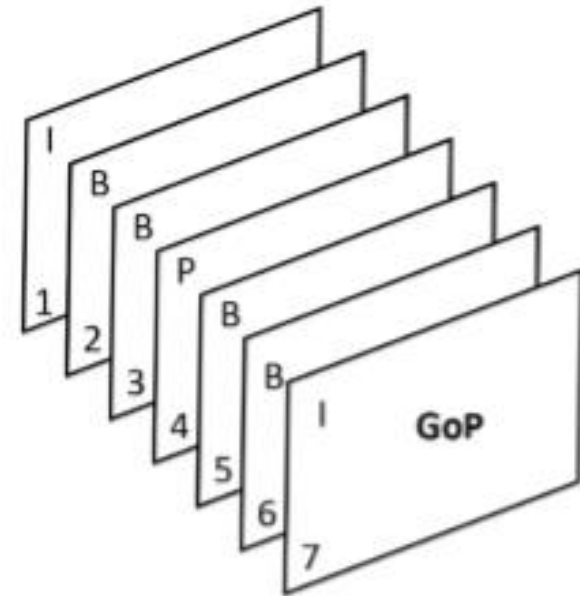
- Ένα GoP ανοίγει με πλαίσιο τύπου I, αφού η κωδικοποίηση αυτού πρέπει να είναι αυτόνομη και να μην έχει εξαρτήσεις από προηγούμενα και εν γένει γειτονικά GoP.
- Ένα GoP μπορεί να κλείσει με πλαίσιο τύπου P ή I.
 - Δεν δύναται να κλείσει με πλαίσιο τύπου B, αφού αυτό θα συνεπάγεται εξάρτηση από επόμενο πλαίσιο.
- Εάν ένα GoP κλείσει με πλαίσιο τύπου I, τότε λέμε ότι το **GoP είναι κλειστό**.



Group of Pictures (GoP)

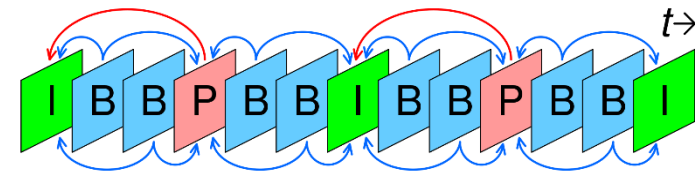
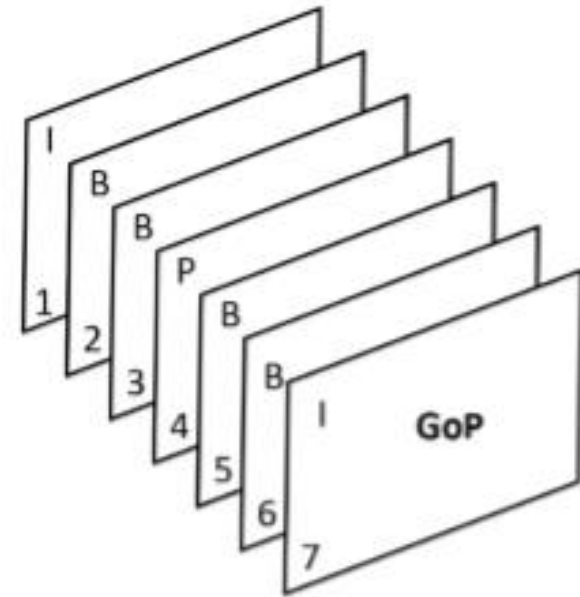
Στο παράδειγμα έχουμε τις εξής αλληλεξαρτήσεις:

- Το πλαίσιο 1 είναι **τύπου I** και κωδικοποιείται αυτόνομα.
- Τα πλαίσια 2 και 3 είναι **τύπου B** και κωδικοποιούνται λαμβάνοντας υπόψη τις κωδικοποιήσεις των πλαισίων **τύπου I** και **P** που είναι εκατέρωθέν τους, δηλαδή των πλαισίων 1 και 4. Αυτό σημαίνει ότι η κωδικοποίηση του πλαισίου 4 (*P frame*), προηγείται της κωδικοποίησης των 2 και 3.
- Το πλαίσιο 4, ως **τύπου P**, εξαρτάται από την κωδικοποίηση του πλαισίου 1 (*I frame*).



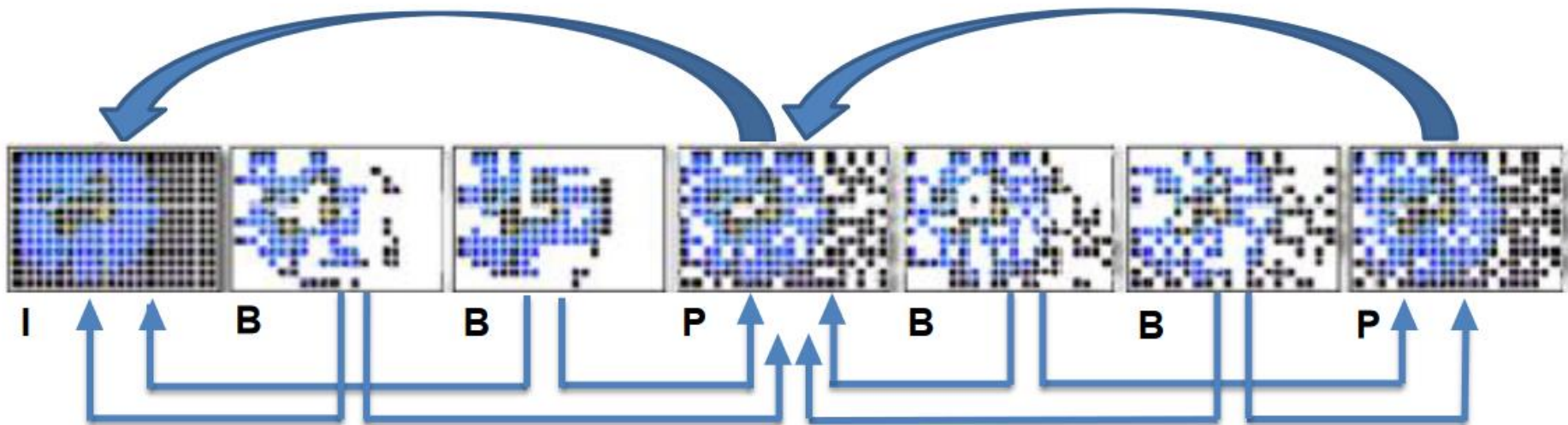
Group of Pictures (GoP)

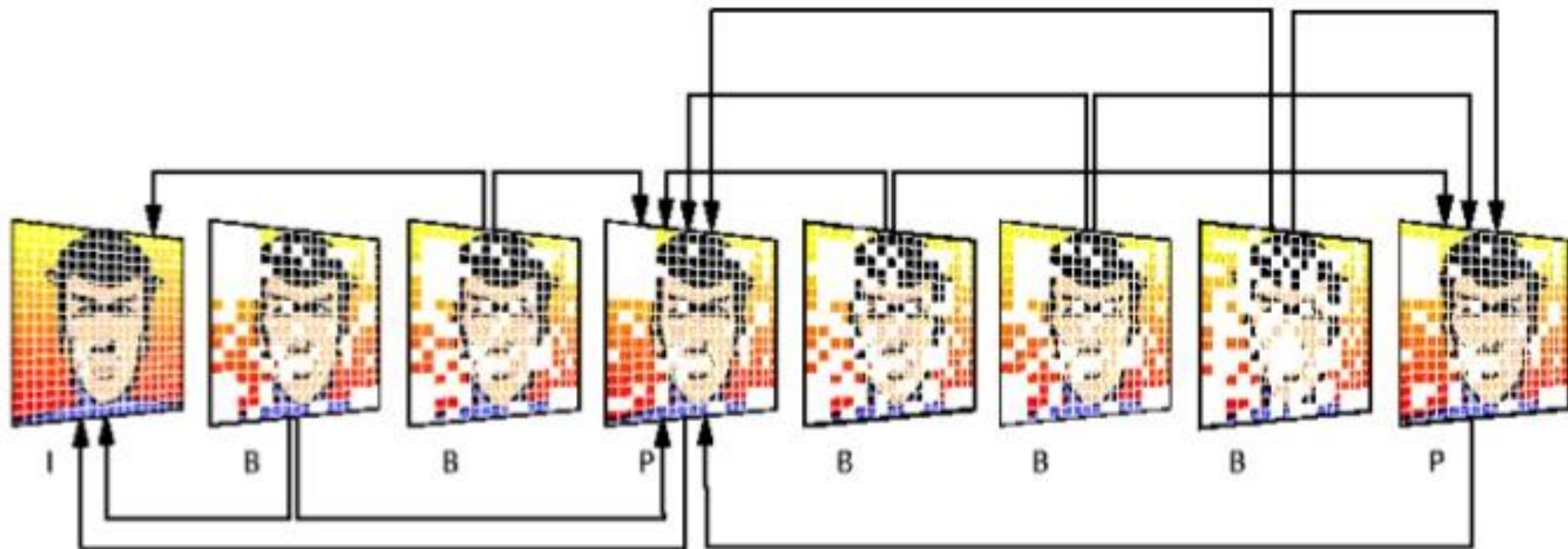
- Τα πλαίσια 5 και 6, ως **τύπου B**, εξαρτώνται από τις κωδικοποιήσεις των πλαισίων 4 και 7 (P και I αντίστοιχα).
- Το πλαίσιο 7 κωδικοποιείται αυτόνομα και μάλιστα πριν από τις κωδικοποιήσεις των 5 και 6 λόγω των σχετικών αλληλεξαρτήσεων.
- Η επόμενη ομάδα πλαισίων επαναλαμβάνει ακριβώς την ίδια δομή.
 - Εκκινεί, δηλαδή, με πλαίσιο τύπου I, το οποίο σημαίνει ότι έχουμε δύο συναπτά πλαίσια τύπου I.
- Εάν η ομάδα πλαισίων κλείνει με πλαίσιο το οποίο κωδικοποιείται ως P, τότε λέμε ότι η **ομάδα πλαισίων είναι ανοικτή**.



Αλληλουχία I, B, P πλαισίων

- I: πλαίσιο αναφοράς (ενδοπλαισιακά)
- P: προβλεπόμενο πλαίσιο (διαφορά με I)(forward prediction)
- B: διπλής κατεύθυνσης (διαφορά με P & I --ή P και P)
 - (forward & backward prediction, bidirectional)





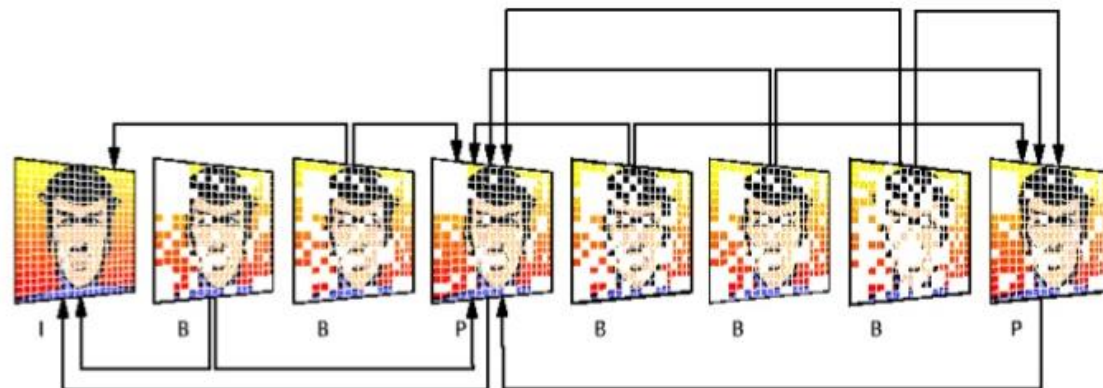
- *I* frames are independently encoded
- *P* frames are based on previous *I*, *P* frames
 - Can send motion vector plus changes
- *B* frames are based on previous and following *I* and *P* frames
 - In case something is uncovered



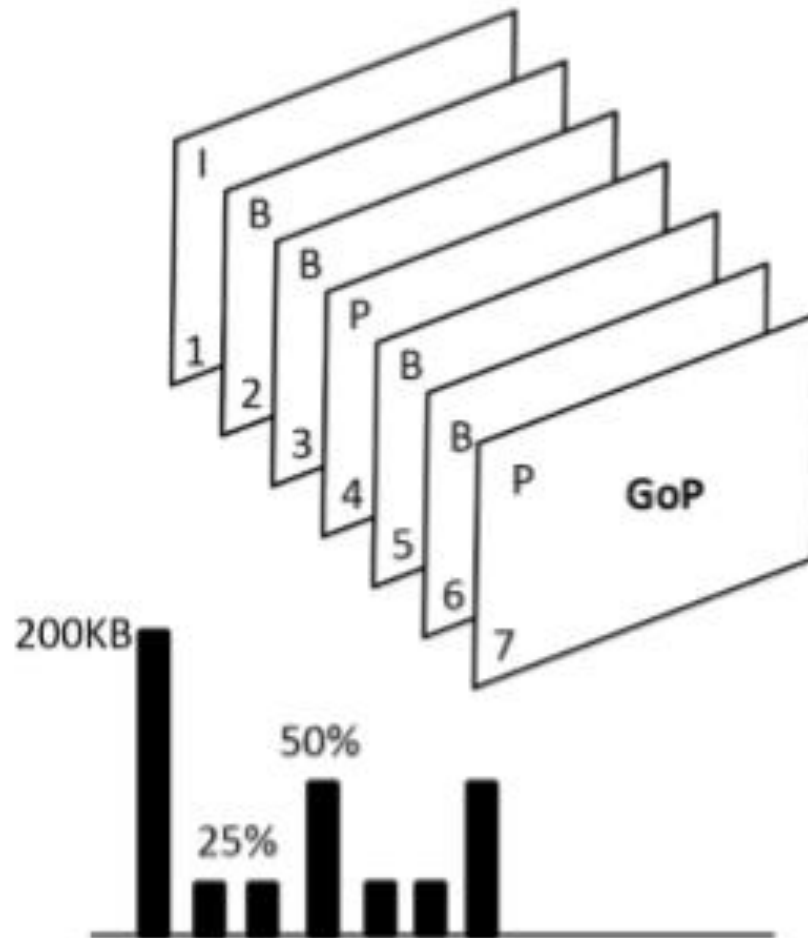
Παράδειγμα διαχείρισης 14 πλαισίων

	Ομάδα Πλαισίων GoP: IBBPBBP													
Αύξων αριθμός πλαισίου	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Αλληλουχία Κωδικοποίησης	1 (I)	4(P)	2 (B)	3 (B)	7 (P)	5 (B)	6 (B)	8 (I)	11 (P)	9 (B)	10 (B)	14 (P)	12 (B)	12 (B)
Αλληλουχία Μετάδοσης & Αποκωδ/σης	1 (I)	4(P)	2 (B)	3 (B)	7 (P)	5 (B)	6 (B)	8 (I)	11 (P)	9 (B)	10 (B)	14 (P)	12 (B)	12 (B)
Αλληλουχία Αναπαραγωγής	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Η επαναφορά των πλαισίων στην ορθή σειρά γίνεται με **χρήση του πεδίου χρονοσήμανσης αποκωδικοποίησης (Decoding Time Stamp, DTS)** το οποίο βρίσκεται εντός της επικεφαλίδας του πακετοποιημένου στοιχειώδους ρυθμού.



Όφελος από την εκμετάλλευση του χρονικού πλεονασμού



Frame Number

To be compressed as:

Frame 1

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

Intraframe (I)

Frame 2

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0

Bidirectional (B)

Frame 3

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0

Predictive (P)



0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

I Frame (Frame 1)

minus

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0

Frame 3

=

		0	0		
		0	1	1	
			1	1	

P Frame

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0

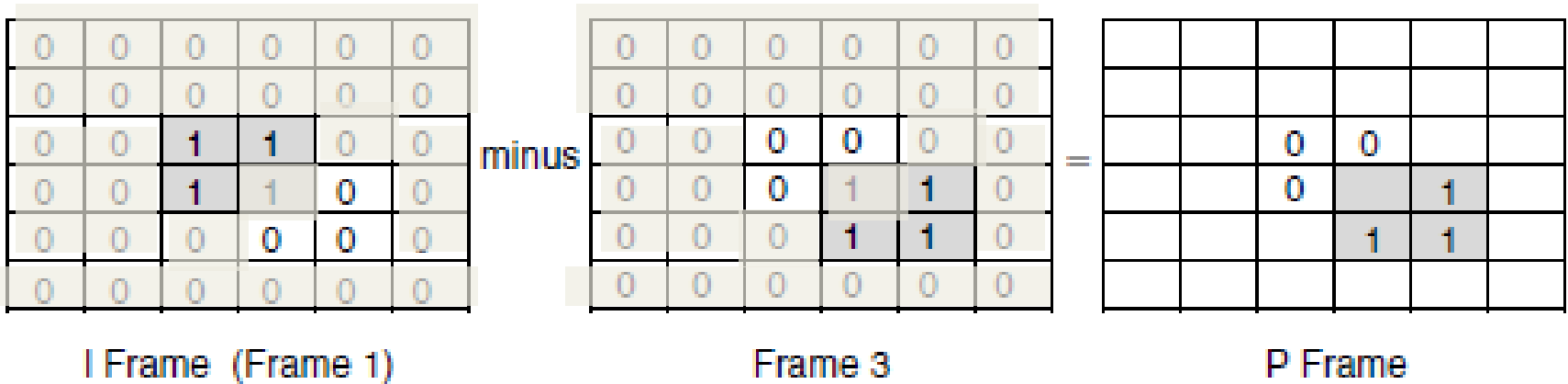
Frame 2

				1	

B Frame = Frame 1 – Frame 2 – P Frame



P Frame



B Frame

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	0	0
0	0	1	1	0	0
0	0	0	0	0	0
0	0	0	0	0	0

I Frame (Frame 1)

minus

0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	1	1	0
0	0	0	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0

Frame 2

		0		1	
		0		1	

		0		1	
		0		1	

		0	0		
		0		1	
			1	1	

P Frame

				1	

$$\text{B Frame} = \text{Frame 1} - \text{Frame 2} - \text{P Frame}$$



Αναζήτηση ομοιοτήτων μεταξύ πλαισίων

- Η αναζήτηση γίνεται σε επίπεδο μακρο-μπλοκ με **σύγκριση των μακρο-μπλοκ γειτονικών πλαισίων**.
- Κάθε μακρο-μπλοκ σε μια εικόνα P ή B εξετάζεται ώστε να διαπιστωθεί:
 1. **Εάν είναι νέο** και δεν συναντάται στην εικόνα αναφοράς
 - Μετάδοση της πληροφορίας, κωδικοποιείται εκ νέου
 2. **Εάν είναι όμοιο** με το μακρο-μπλοκ στην αντίστοιχη θέση του προηγούμενου πλαισίου (με κάποιες μικρές διαφοροποιήσεις).
 - Μετάδοση της διαφοράς (τιμές)
 - Δεν απαιτείται διάνυσμα κίνησης
 3. **Εάν έχει προκύψει από μεταφορά αντίστοιχου μακρο-μπλοκ** του προηγούμενου πλαισίου (το οποίο μπορεί να βρίσκεται σε διαφορετική θέση).
 - Με **διάνυσμα κίνησης** περιγράφεται η μετακίνηση του ομοίου μακρο-μπλοκ
 - ένα ζευγάρι αριθμών (πχ. 0,0, ή 12, 4 ή 25, -4 κλπ.) που δηλώνει τη κίνηση του macroblock
 - Μετάδοση διαφοράς τιμών



Διάνυσμα κίνησης – Υπολογισμός (1/2)

- Για τον υπολογισμό του διανύσματος κίνησης (μόνο στα P, B), ο μηχανισμός ξεκινάει από το μακρο-μπλοκ του τρέχοντος πλαισίου
 - εάν είναι **τύπου P** ελέγχει στο προηγούμενο P ή I πλαίσιο για παρόμοια μακρο-μπλοκ
 - εάν είναι **τύπου B** ελέγχει στο προηγούμενο και στο επόμενο πλαίσιο για παρόμοια μακρο-μπλοκ
- Η αναζήτηση γίνεται μέσω του ταιριάσματος σε επίπεδο μακρο-μπλοκ όχι pixel.

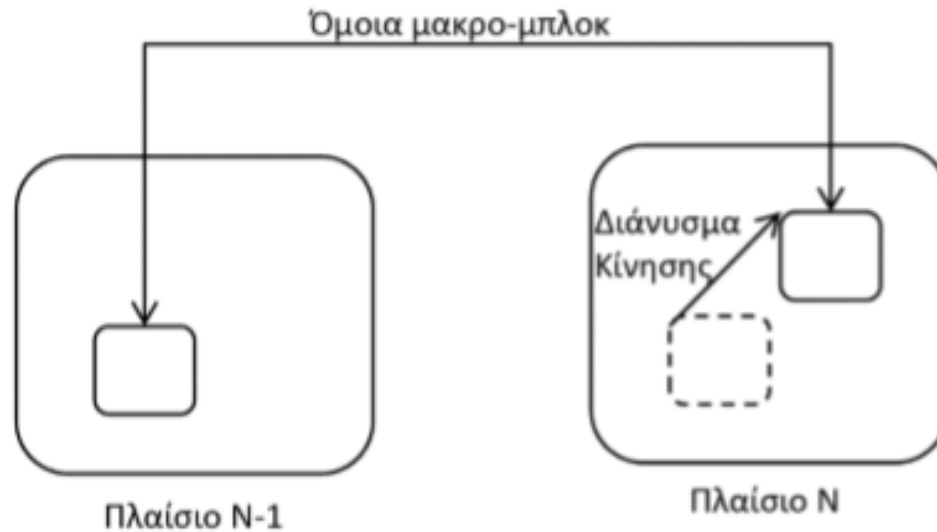


Διάνυσμα κίνησης – Υπολογισμός (2/2)

- Η αναζήτηση δεν γίνεται σε ολόκληρη την εικόνα **αλλά πέριξ της περιοχής** του αρχικού μακρο-μπλοκ.
 - Ο περιορισμός αυτός της περιοχής αναζήτησης γίνεται για να ελαττωθεί
 - η πολυπλοκότητα του συνολικού συστήματος
 - το χρονικό διάστημα που απαιτείται για την κωδικοποίηση
- Υπενθύμιση
 - Εάν βρεθεί κάποιο αντίστοιχο σε τιμές μακρο-μπλοκ, τότε **υπολογίζεται το διάνυσμα κίνησης**, το οποίο αντιπροσωπεύει τη **διαφορά στη θέση του αρχικού** με το εντοπισμένο μακρο-μπλοκ.
 - Υπολογίζεται η διαφορά μεταξύ των δύο μακρο-μπλοκ.



Διάνυσμα κίνησης



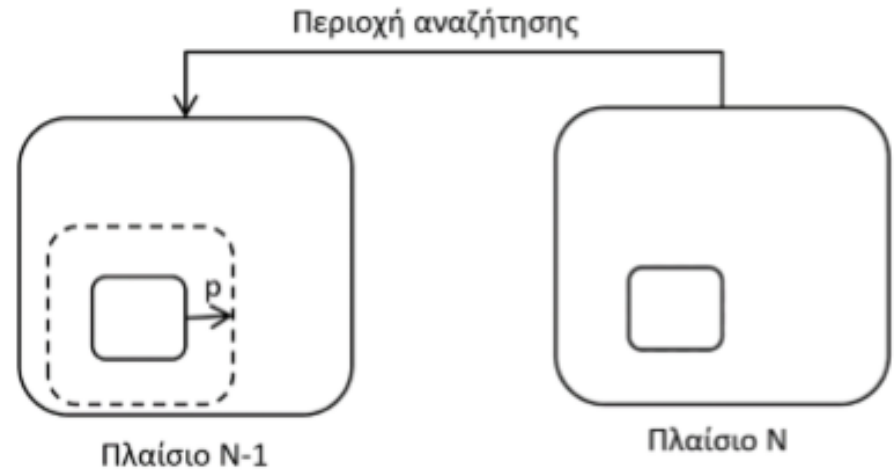
- Παράλληλα με το **διάνυσμα κίνησης**, μεταδίδονται επιπλέον απαραίτητες πληροφορίες για να καλυφθούν οι διαφοροποιήσεις μεταξύ των δύο πλαισίων.



Εκτίμηση κίνησης

Η εκτίμηση κίνησης είναι **το πιο χρονοβόρο κομμάτι της συμπίεσης**. Οι σχετικοί αλγόριθμοι αναζητούν για κάθε μακρο-μπλοκ, το αντίστοιχο μακρο-μπλοκ σε προηγούμενο (ή επόμενο) πλαίσιο που ταιριάζει.

- Εάν υποθέσουμε ανάλυση 720x480, το πλήθος των αναζητήσεων είναι απαγορευτικό από πλευράς απόδοσης.
- Απαιτείται να περιορίσουμε την αναζήτηση σε συγκεκριμένες περιοχές.
- Συνήθως η περιοχή αναζήτησης εκτείνεται από την **τρέχουσα θέση συν p pixels** (π.χ. $p=7$).



Αναζήτηση όμοιων περιοχών

- Σε επίπεδο **εικονοστοιχείου** (pixel) με μεγάλο όμως πλήθος υπολογισμών, γεγονός το οποίο την καθιστά ασύμφορη.
- Σε επίπεδο **τμήματος της εικόνας** (block-based),
 - η πιο απλή περίπτωση και χρησιμοποιείται όπως προαναφέραμε κυρίως σήμερα και μάλιστα με χρήση αλγορίθμων που εκτελούνται με χρήση υλικού για να εξασφαλιστεί καλύτερη απόδοση σε σχέση με τη χρήση λογισμικού.
 - Τυπικές διαστάσεις του μακρο-μπλοκ είναι 16x16 (ή και 8 x 8).
- Σε επίπεδο **αντικειμένων** εντός της εικόνας (object-based), όπου τα αντικείμενα αναγνωρίζονται αυτόματα εντός της εικόνας και παρακολουθούνται στα διάφορα πλαίσια.



Συναρτήσεις καλύτερου ταιριάσματος

- Mean Absolute Difference (MAD) – Μέση τιμή των απόλυτων διαφορών
- Mean Squared Error (MSE) – Μέση τιμή τετραγώνων των σφαλμάτων

$$MAD = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |C_{ij} - R_{ij}|$$

$$MSE = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (C_{ij} - R_{ij})^2$$

όπου N είναι η διάσταση (του τετραγωνικού) μακρο-μπλοκ, C τιμή στο τρέχον πλαίσιο και R στο πλαίσιο αναφοράς

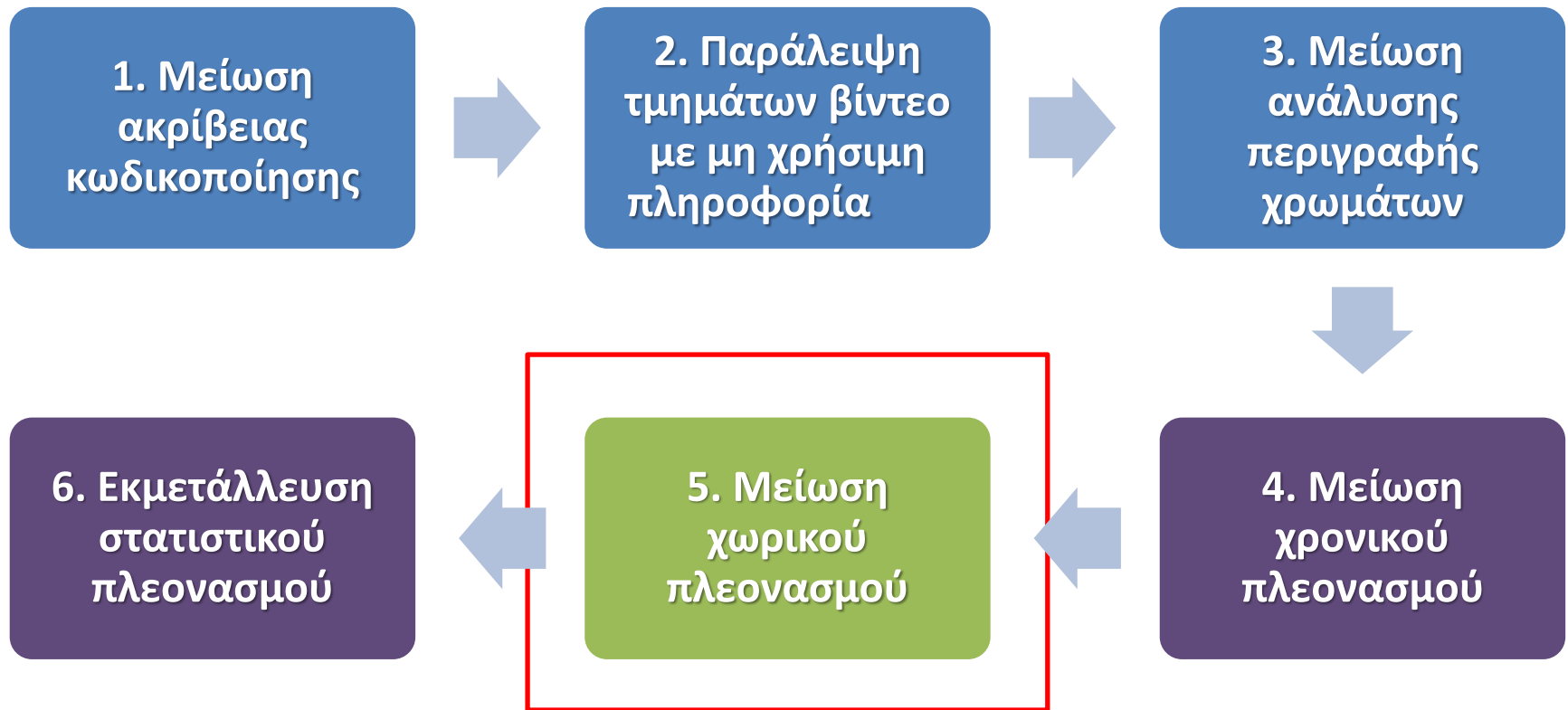


Τα βήματα του MPEG

1. Αναλύεται το αρχείο video για να καθοριστεί το ποια πλαίσια θα κωδικοποιηθούν ως πλαίσια-I, ποια ως πλαίσια-P και ποια ως πλαίσια-B.
2. Τα πλαίσια διαιρούνται σε macroblocks (16x16 pixels)
3. Το σήμα μετατρέπεται από RGB σε YUV ενώ ταυτόχρονα γίνεται και υποδειγματοληψία χρώματος.
4. Εφαρμόζεται η τεχνική εκτίμησης κίνησης σε κάθε πλαίσιο-P και -B
5. Σε κάθε πλαίσιο I,-P και B τα macroblocks συμπιέζονται **κατά JPEG**,
6. Εφαρμογή συμπίεσης κατά Huffman στις τελικές τιμές



5. Μείωση χωρικού πλεονασμού



5. Μείωση χωρικού πλεονασμού

- Ο **χωρικός πλεονασμός** προέρχεται από τις ομοιότητες σε διαφορετικές περιοχές εντός μιας εικόνας.
- Οι ομοιότητες είναι φαινόμενο αρκετά συχνό και γίνεται εκμετάλλευση του χωρικού πλεονασμού με τα **πρότυπα συμπίεσης της στατικής εικόνας**.
 - Το πλέον διαδεδομένο αυτών των προτύπων είναι το JPEG (Joint Pictures Expert Group).
- Η εκμετάλλευση του χωρικού πλεονασμού συνίσταται στην **αφαίρεση πληροφορίας, η οποία δεν είναι τόσο σημαντική μεταξύ γειτονικών εικονοστοιχείων**.



Μετασχηματισμός πληροφορίας & κωδικοποίησή της

- **Βήμα 1:** Μετασχηματισμός πληροφορίας
 - από το **αρχικό πεδίο** σε ένα **ενδιάμεσο πεδίο**
 - Το ενδιάμεσο πεδίο είναι πιο κατάλληλο για συμπίεση
- **Βήμα 2:** Σε αυτό το πεδίο
 - επιλέγονται και κωδικοποιούνται οι **συντελεστές του μετασχηματισμού**, οι οποίοι περιέχουν τη μεγάλη ενέργεια (*αυτοί που εμφανίζονται πιο συχνά*, δηλαδή οι σημαντικοί)
 - αγνοούνται ή κωδικοποιούνται με μικρότερη ακρίβεια οι υπόλοιποι συντελεστές.
- Το αρχικό σήμα ανακτάται (με απώλειες) μέσω του αντιστρόφου μετασχηματισμού.



Ιδέα

- Οι πιο **χαμηλές** (χωρικές) **συχνότητες** χρειάζεται να κωδικοποιηθούν με **μεγαλύτερη λεπτομέρεια** (κυρίως όσον αφορά την κβάντιση),
- Οι πιο **υψηλές συχνότητες** **δεν** απαιτείται να κωδικοποιηθούν με πολύ μεγάλη λεπτομέρεια.
- Η μείωση αυτή στη λεπτομέρεια είναι δυνατό να προκαλέσει σημαντική εξοικονόμηση στον απαιτούμενο ρυθμό μετάδοσης.
- **Ερώτημα:** *Πώς θα ξεχωρίσουμε την πληροφορία που αφορά τις υψηλές (χωρικές) συχνότητες από την πληροφορία που αφορά τις χαμηλές συχνότητες;*
- **Απάντηση:** Με μετασχηματισμό του σήματος στο πεδίο της συχνότητας.



Διακριτός Μετασχηματισμός Συνημιτόνου

- Ο βασικός αλγόριθμος που χρησιμοποιείται ο Διακριτός Μετασχηματισμός Συνημιτόνου (Discrete Cosine Transform, DCT)
 - κωδικοποίηση στατικών, μεμονωμένων εικόνων στο JPEG
 - κωδικοποίηση των εικόνων και στην περίπτωση του βίντεο
- Αυτός ο τρόπος συμπίεσης βασίζεται στο γεγονός ότι **ο ανθρώπινος οφθαλμός**
 - δεν αντιλαμβάνεται με μεγάλη ακρίβεια τις **απότομες διαφοροποιήσεις** σε μια εικόνα
 - απότομες γιατί πραγματοποιούνται σε μικρό σχετικά αριθμό εικονοστοιχείων - αντιστοιχούν σε **υψηλές χωρικές συχνότητες**
 - αντιλαμβάνεται επακριβώς τις **λιγότερο απότομες** διαφοροποιήσεις εντός της εικόνας
 - αυτές δηλαδή που γίνονται σε μεγαλύτερες επιφάνειες εικονοστοιχείων - αντιστοιχούν σε **χαμηλότερες χωρικές συχνότητες**)



Τι κάνει ο DCT

- Ο DCT μετασχηματίζει ένα πίνακα 8 x 8 που περιλαμβάνει τιμές έντασης (φωτεινότητας, χρωμοδιαφοράς) στο διάστημα [0, 255] σε έναν πίνακα 8 x 8 που περιλαμβάνει **τιμές συχνότητας**, που ονομάζονται **χωρικές συχνότητες** και κάθε θέση αποτελεί ένα συντελεστή ή συνιστώσα DCT.
- Οι συντελεστές αυτοί **δηλώνουν τη σχετική βαρύτητα κάθε συχνότητας που συμβάλει στη σύνθεση του διακριτού σήματος** της εικόνας και έχουν πραγματικές ακέραιες τιμές
- Οι πάνω και αριστερά θέσεις αντιστοιχούν **στις πιο σημαντικές (μεγαλύτερη βαρύτητα – συχνότητα εμφάνισης των τιμών έντασης)**, με το F(0,0) να αντιστοιχεί στη **μέση φωτεινότητα της αρχικής εικόνας**.
- Οι θέσεις πηγαίνοντας από αριστερά -> δεξιά και πάνω -> κάτω πηγαίνουν
 - από τις χαμηλές χωρικές συχνότητες, αντιστοιχούν σε ήπιες αλλαγές έντασης μεταξύ γειτονικών εικονοστοιχείων)
 - στις πιο υψηλές συχνότητες (πιο απότομες αλλαγές έντασης)

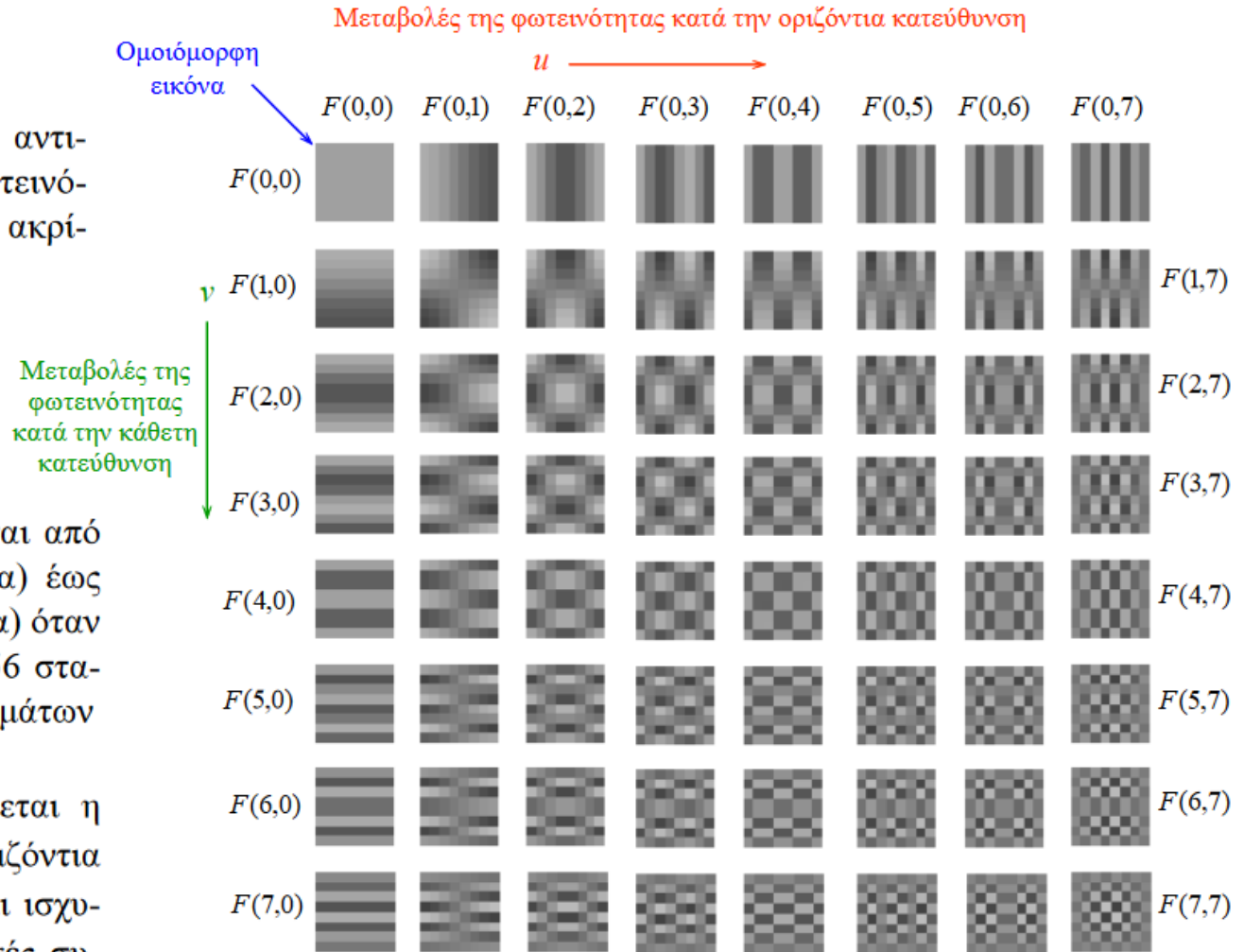


Οι συναρτήσεις βάσης για πίνακα 8×8 στο DCT

Ο συντελεστής $F(0, 0)$, αντιπροσωπεύει τη μέση φωτεινότητα της εικόνας (για την ακρίβεια είναι αντίστοιχη)

Η τιμή της μπορεί να είναι από $F(0, 0) = 0$ (μαύρη εικόνα) έως $F(0,0) = 255$ (λευκή εικόνα) όταν έχουμε αναπαράσταση 256 σταθμών φωτεινότητας ή χρωμάτων

Όσο ταχύτερα μεταβάλλεται η φωτεινότητα κατά την οριζόντια κατεύθυνση τόσο γίνονται ισχυρότερες οι μεγάλες χωρικές συχότητες (π.χ. $F(7,0) > F(7,7)$)



Επεξεργασία εικόνας με DCT (1)

- Είσοδος
 - Μονάδες δεδομένων 8×8 εικονοστοιχείων
 - Κάθε μονάδα μετασχηματίζεται χωριστά
 - Δεν έχει σημασία η ταξινόμηση
- Έξοδος
 - Μονάδες δεδομένων 8×8 συντελεστών



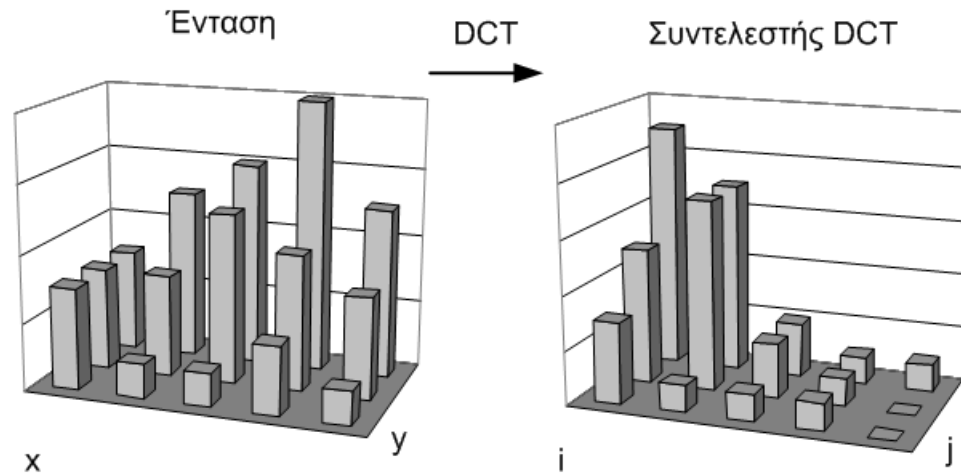
Επεξεργασία εικόνας με DCT (2)

$$F[i, j] = \frac{1}{4} C(i)C(j) \sum_{x=0}^7 \sum_{y=0}^7 P[x, y] \cos \frac{(2x+1)i\pi}{16} \cos \frac{(2y+1)j\pi}{16}$$

- Ευθύς μετασχηματισμός DCT (FDCT)
 - $[0, 255] \rightarrow [-128, 127]$ με αφαίρεση
 - Αρχικά: Τιμές έντασης $P[x, y]$ με $x, y = [0, 7]$ (64 τιμές)
 - Τελικά: $F[i, j]$ με $i, j = [0, 7]$ (64 τιμές)
- Τα συνημίτονα δεν εξαρτώνται από $P[x, y]$
 - Υπολογίζονται προκαταβολικά



Επεξεργασία εικόνας με DCT (3)



- Συντελεστής $F[0,0]$: συντελεστής DC
 - Μέση τιμή της μονάδας δεδομένων
- Υπόλοιποι συντελεστές: συντελεστές AC
- Συγκέντρωση υψηλών τιμών κοντά στο DC
 - Υψηλές τιμές στις χαμηλές χωρικές συχνότητες



Επεξεργασία εικόνας με DCT (4)

$$P[x, y] = \frac{1}{4} \sum_{i=0}^7 \sum_{j=0}^7 C(i)C(j)F[i, j] \cos \frac{(2x+1)i\pi}{16} \cos \frac{(2y+1)j\pi}{16}$$

- Αντίστροφος μετασχηματισμός DCT (IDCT)
 - Υπολογισμός $P[x, y]$ από $F[i, j]$
 - Τα συνημίτονα πάλι δεν εξαρτώνται από $F[i, j]$
 - Υπολογίζονται προκαταβολικά
 - Θεωρητικά, πλήρης ανακατασκευή εικόνας
 - Στην πράξη, σφάλματα στρογγυλοποίησης



Γιατί DCT στο JPEG;

- Φυσικές εικόνες
 - Μεγάλες περιοχές με παρόμοια χρώματα
 - Πολλοί συντελεστές AC σχεδόν μηδενικοί
 - Δείχνουν την απόσταση των εναλλαγών
 - Όταν δεν υπάρχουν, οι συντελεστές είναι μηδέν
 - Διαφοροποίηση σε σχέση με γραφικά
 - Δεν υπάρχουν απότομες αλλαγές χρωμάτων οι οποίες παράγουν μη μηδενικούς συντελεστές AC



Ο Διακριτός Μετασχηματισμός Συνημιτόνου (ΔΜΣ), Discrete Cosine Transform (DCT)

- Ο ΔΜΣ (DCT), είναι μια ειδική μορφή του Διακριτού Μετασχηματισμού Φουριέ (Discrete Fourier Transform).
- Το πλήθος τιμών μετασχηματίζεται σε σύνολο τιμών ίδιου πλήθους στο πεδίο των συχνοτήτων.
- Ο μετασχηματισμός συνημιτόνου χρησιμοποιείται σε μία ή σε περισσότερες διαστάσεις.
 - στη μία διάσταση η ανεξάρτητη μεταβλητή είναι ο χρόνος
 - στις δύο διαστάσεις οι ανεξάρτητες μεταβλητές είναι το επίπεδο (μήκος και πλάτος).

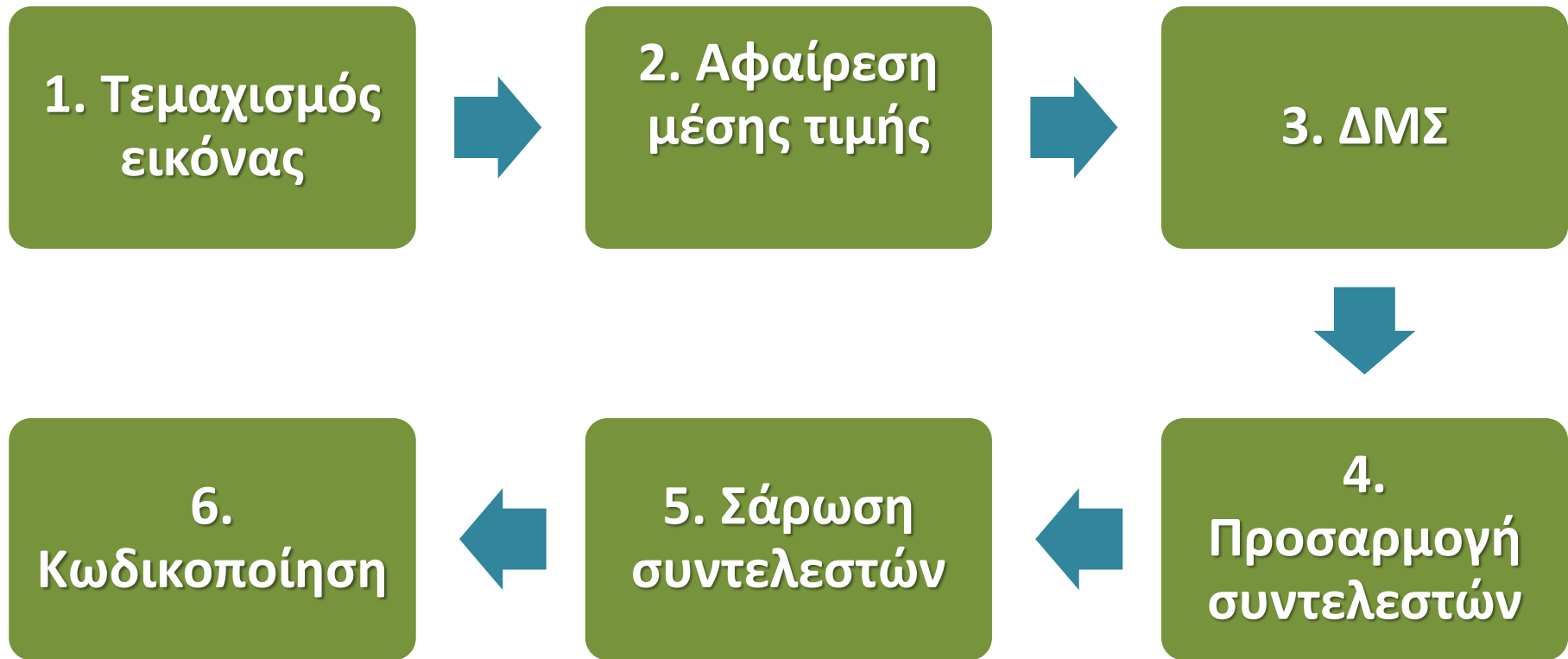


Γιατί ΔΜΣ (DCT);

- Ο ΔΜΣ επιδεικνύει **ανθεκτικότητα** στην παράλειψη ενός ποσοστού από τους συντελεστές που προκύπτουν αφού εφαρμοστεί σε κάποια δεδομένα, σε σχέση με άλλους ισοδύναμους αλγορίθμους μετασχηματισμού όπως ο **Διακριτός Μετασχηματισμός Φουριέ**.
 - Η ανθεκτικότητα αυτή είναι μεγαλύτερη από αυτή του ΔΜΦ και οδηγεί σε σημαντικό βαθμό συμπίεσης.
 - Η ανθεκτικότητα αφορά το γεγονός ότι μετά τον αντίστροφο μετασχηματισμό τα δεδομένα που προκύπτουν ομοιάζουν σε μεγαλύτερο βαθμό με τα αρχικά σε σχέση με την αντίστοιχη περίπτωση (παράλειψη ίδιου πλήθους συντελεστών) του ΔΜΦ.



Τα 6 βήματα για τη μείωση του χωρικού πλεονασμού



Βήματα 1 έως 3

Τα βήματα τα οποία ακολουθούνται στην περίπτωση του JPEG και του MPEG περιλαμβάνουν

1. τον τεμαχισμό της εικόνας σε μπλοκ, διαστάσεων συνήθως 8x8 εικονοστοιχείων
2. την αφαίρεση της μέσης τιμής (του 128 στην περίπτωση που ο κβαντισμός θα γίνει με τη βοήθεια 8 bits) και στη συνέχεια
3. το μετασχηματισμό κατά Διακριτό Μετασχηματισμό Συνημιτόνου (ΔΜΣ).

Οι συντελεστές που προκύπτουν ανήκουν στο πεδίο της συχνότητας (αντιστοιχούν στις διάφορες χωρικές συχνότητες).



Αναμενόμενα αποτελέσματα του δισδιάστατου DCT

- Το αποτέλεσμα του μετασχηματισμού και ο βαθμός συμπίεσης που επιτυγχάνεται εξαρτώνται από την οριζόντια, την κατακόρυφη και τη διαγώνια χωρική συχνότητα.
- Μια εικόνα με **ένα μόνο χρώμα** θα δίνει σαν αποτέλεσμα έναν πίνακα με μεγάλη τιμή για τον πρώτο όρο και μηδενικές για τις υπόλοιπες.
- Μια εικόνα με **πολλές εναλλαγές** θα δίνει **πίνακα με διαφορετικές τιμές**.



Παράδειγμα

154	123	123	123	123	123	123	136
192	180	136	154	154	154	136	110
254	198	154	154	180	154	123	123
239	180	136	180	180	166	123	123
180	154	136	167	166	149	136	136
128	136	123	136	154	180	198	154
123	105	110	149	136	136	180	166
110	136	123	123	123	136	154	136

Τιμές φωτεινότητας Y ενός τμήματος (*block*) εικόνας 8×8 .

- Οι τιμές των pixels κυμαίνονται από 0 (μαύρο) έως 255 (άσπρο).



- **Βήμα 2:** Αφαίρεση μέσης τιμής 128 από κάθε pixel, οπότε οι τιμές κυμαίνονται από -128 έως 127.
- **Βήμα 3:** Πραγματοποιούμε τον μετασχηματισμό DCT. Χαμηλές συχνότητες

Βασικό
χρώμα
του
μπλοκ

162	41	20	72	30	12	-20	-11
30	108	20	32	28	-15	18	-2
-94	-60	12	-43	-31	6	-3	7
-39	-83	-5	-22	-13	15	-1	3
-31	18	-5	-12	14	-6	11	-6
-1	-12	13	0	28	13	8	3
5	-2	12	7	-19	-13	8	12
-10	11	8	-16	21	0	6	11

Μετασχηματισμένες κατά DCT τιμές



Βήματα 4 έως 6

- **Βήμα 4: Προσαρμογή των συντελεστών** με βάση την ευαισθησία του ανθρώπινου οφθαλμού, αλλά και του επιθυμητού επιπέδου ποιότητας.
- **Βήμα 5: Σάρωση των συντελεστών** με τέτοιο τρόπο, ώστε να ομαδοποιούνται αποδοτικά οι μεγαλύτερες και οι μικρότερες τιμές.
 - τεχνική του ζιγκ ζαγκ, οι μη μηδενικές (και εν γένει οι μεγαλύτερες τιμές) βρίσκονται στο άνω αριστερά τμήμα του πίνακα, ενώ οι μικρότερες τιμές βρίσκονται στο κάτω δεξιά τμήμα του πίνακα.
- **Βήμα 6: Κωδικοποίηση/συμπίεση**



Προσαρμογή/κλιμάκωση των συντελεστών

- Στο βήμα 4 η προσαρμογή των συντελεστών (διαιρούνται) γίνεται με βάση έτοιμους τους **πίνακες κλιμάκωσης, προτυποποιημένους** κατά το MPEG.
- Οι πίνακες αυτοί μπορούν να προσαρμοστούν ή να τροποποιηθούν από τους κωδικοποιητές στις εκάστοτε ανάγκες (αναφορικά με την ποιότητα και τον βαθμό συμπίεσης).
- Ο αποκωδικοποιητής λαμβάνει πληροφορίες για αυτούς τους πίνακες **καθώς αυτές μεταδίδονται.**



162	41	20	72	30	12	-20	-11
30	108	20	32	28	-15	18	-2
-94	-60	12	-43	-31	6	-3	7
-39	-83	-5	-22	-13	15	-1	3
-31	18	-5	-12	14	-6	11	-6
-1	-12	13	0	28	13	8	3
5	-2	12	7	-19	-13	8	12
-10	11	8	-16	21	0	6	11

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
14	17	22	29	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Προτυποποιημένος Πίνακας κλιμάκωσης DCT για ποιότητα 50% (Q_{50})



- **Βήμα 4:** Κλιμάκωση/προσαρμογή των συντελεστών για ποιότητα 50

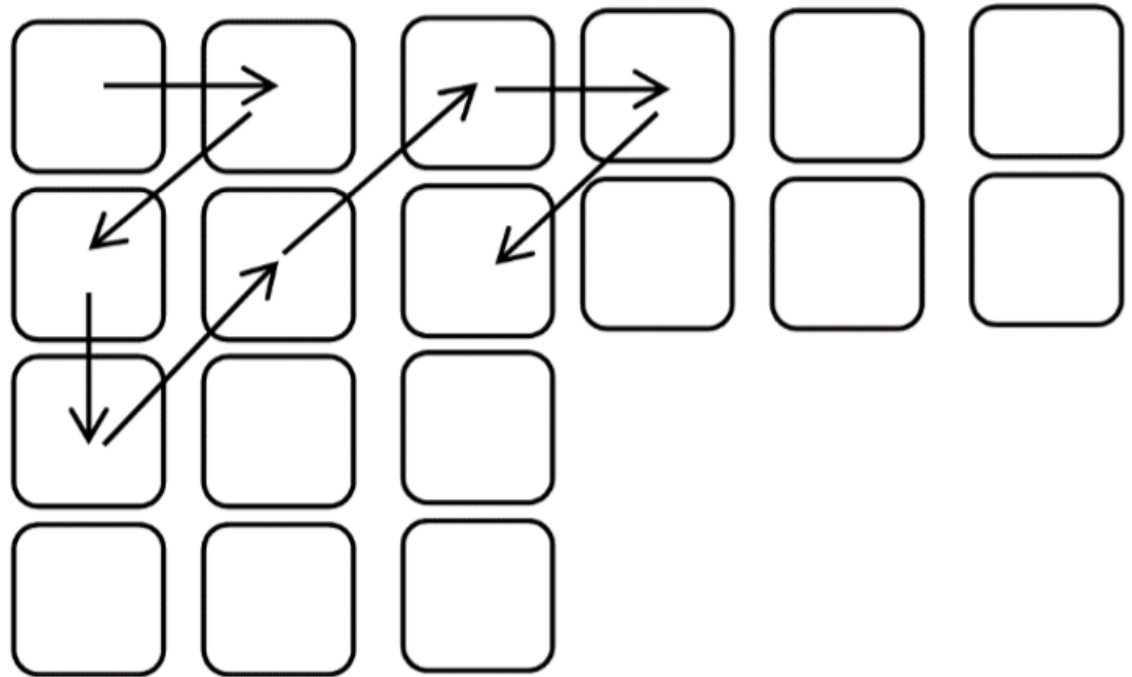
10	4	2	5	1	0	0	0
3	9	1	2	1	0	0	0
-7	-5	1	-2	-1	0	0	0
-3	-5	0	-1	0	0	0	0
-2	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Τιμές μετά την κλιμάκωση



Σάρωση συντελεστών κατά Ζιγκ Ζαγκ

- Ομαδοποίηση των τιμών και κατά το δυνατόν συγκέντρωση των μηδενικών τιμών, η σάρωση γίνεται κατά **ζιγκ ζαγκ**, ξεκινώντας **από το άνω αριστερά τμήμα** και προχωρώντας προς τα **δεξιά και κάτω**.



- **Βήμα 5:** Σάρωση συντελεστών κατά ζιγκ ζαγκ

10, 4, 3, -7, 9, 2, 5, 1, -5, -3, -2, -5, 1, 2, 1, 0, 1, -
2, 0, 1, 0, 0, 0, 0, -1, -1, 0, 0, ...0



- Δείτε Πρακτικές Υπολογισμού δισδιάστατου DCT και Προσαρμογή Επιπέδου Ποιότητας στο [1]



- **Βήμα 6: Κωδικοποίηση** συντελεστών που έχουν προκύψει από τη σάρωση κατά Ζιγκ Ζαγκ

Θα το εξετάσουμε σε λίγο ...



Αποκωδικοποίηση

Ανάστροφη διαδικασία

- **Βήμα:** Αποκωδικοποίηση (βήμα 6 – κωδικοποίηση)

Περίπου:

10, 4, 3, -7, 9, 2, 5, 1, -5, -3, -2, -5, 1, 2, 1, 0, 1, -
2, 0, 1, 0, 0, 0, 0, -1, -1, 0, 0, ...0

- **Βήμα:** Δημιουργία πίνακα



- **Βήμα:** Πολλαπλασιασμός κάθε τιμής με το αντίστοιχο στοιχείο του προτυποποιημένου πίνακα

160	44	20	80	24	0	0	0
36	108	14	38	26	0	0	0
-98	-65	16	-48	-40	0	0	0
-42	-85	0	-29	0	0	0	0
-36	22	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0



160	44	20	80	24	0	0	0
36	108	14	38	26	0	0	0
-98	-65	16	-48	-40	0	0	0
-42	-85	0	-29	0	0	0	0
-36	22	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

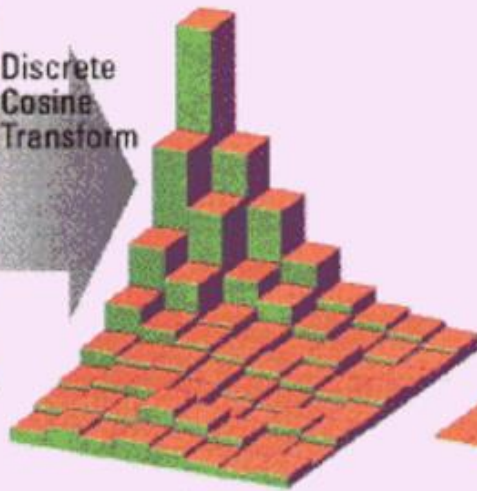
162	41	20	72	30	12	-20	-11
30	108	20	32	28	-15	18	-2
-94	-60	12	-43	-31	6	-3	7
-39	-83	-5	-22	-13	15	-1	3
-31	18	-5	-12	14	-6	11	-6
-1	-12	13	0	28	13	8	3
5	-2	12	7	-19	-13	8	12
-10	11	8	-16	21	0	6	11





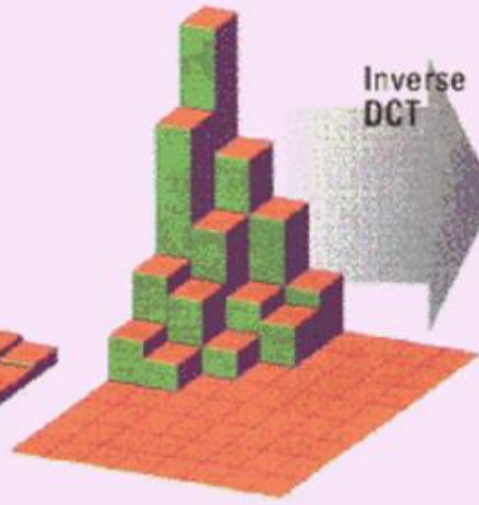
a

Discrete
Cosine
Transform



b

Inverse
DCT



c



d



- **Βήμα:** Εφαρμογή του αντίστροφου DCT με προσέγγιση πλησιέστερου ακεραίου
- **Βήμα:** Πρόσθεση της μέσης τιμής 128 σε κάθε στοιχείο του πίνακα

Προκύπτει ο πίνακας φωτεινότητας.



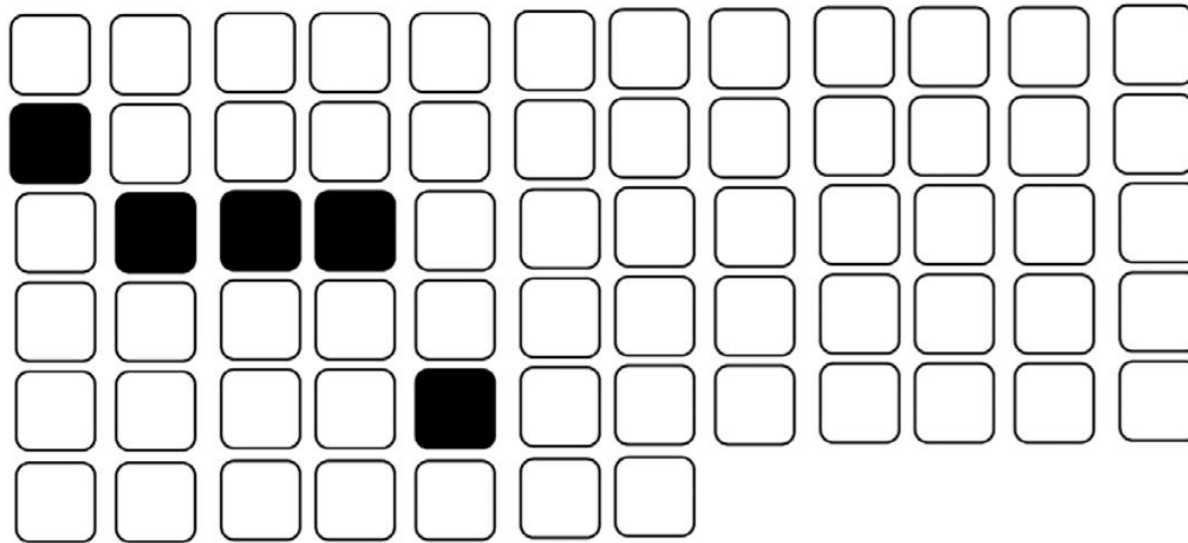
Κωδικοποίηση / συμπίεση συντελεστών (Βήμα 6)

- Εφαρμόζεται στο αποτέλεσμα της σάρωσης κατά Ζιγκ-Ζαγκ
- Τεχνική:
 - **Κωδικοποίηση Μήκους Διαδρομής (Run Length Encoding – RLE)**



Κωδικοποίηση Μήκους Διαδρομής (Run Length Encoding – RLE)

- Στοχεύει σε μια αποτελεσματική από πλευράς συμπίεσης, περιγραφή αλληλουχίας δεδομένων στις οποίες κάποιες τιμές επαναλαμβάνονται συνεχόμενα.



- WWWWWWWWWWWWWBWWWWWWWWWWWWWWWWBWWWWWWWWWW
WW
- (W, 12), (B, 1), (W, 12), (B, 3), (W, 24), (B, 1), (W, 14)



Ζητήματα απόδοσης του RLE

- Αποδοτικός ο RLE όταν πρέπει να έχουμε συνεχόμενες ίδιες τιμές
- Αν οι τιμές δεν είναι συνεχόμενες ο όγκος μπορεί να αυξηθεί
- Τεχνικές που εισάγουν ευφυΐα
- Πχ. ο αλγόριθμος να ενεργοποιείται από 4 ίδιους χαρακτήρες και άνω
 - σύμβολο, ειδικός χαρακτήρας, (!) πλήθος της εμφάνισης των επιπλέον χαρακτήρων

Αρχική Σειρά Συμβόλων	Κωδικοποίηση
A	A
AA	AA
AAA	AA
AAAA	A!0
AAAAA	A!1
AAAAAA	A!2
AAAAAAA	A!3



6. Εκμετάλλευση στατιστικού πλεονασμού



6. Εκμετάλλευση στατιστικού πλεονασμού

- Η εκμετάλλευση του στατιστικού πλεονασμού αποτελεί μια **μη απωλεστική** μέθοδο συμπίεσης δεδομένων.
- Μέθοδος **κωδικοποίησης εντροπίας**.
- Αναπτύχθηκε το 1952 από τον David Huffman και η λογική αυτής της μεθόδου στηρίζεται στην εκμετάλλευση του γεγονότος ότι σε **οποιοδήποτε αλφάβητο κάποια από τα σύμβολα εμφανίζονται πιο συχνά από κάποια άλλα**.
- Αυτό σημαίνει ότι αν τα πιο συχνά εμφανιζόμενα σύμβολα κωδικοποιηθούν πιο αποδοτικά αναμένουμε να έχουμε όφελος σε σχέση με την **ομοιόμορφη κωδικοποίηση**
 - με μικρότερες, δηλαδή, σε αριθμό bits λέξεις κωδικοποίησης,
 - ομοιόμορφα: όλα τα σύμβολα με ίδιου δηλαδή μήκους λέξεις κωδικοποίησης



Ομοιόμορφη Κωδικοποίηση

- Υπολογίζεται το πλήθος των bits με βάση το πλήθος των συμβόλων προς κωδικοποίηση
- Η σχέση υπολογισμού είναι $N \leq 2^n$
 - N το πλήθος των συμβόλων προς κωδικοποίηση
 - n το πλήθος των bits που απαιτούνται για την κωδικοποίηση
- Σε κάθε σύμβολο αντιστοιχίζεται ένας διαφορετικός συνδυασμός από n bits.
- Παράδειγμα
 - 6 σύμβολα,
 - πλήθος των bits: 3 ($2^2 < 6$ και $2^3 > 6$)

Σύμβολα	Κωδικοποίηση
A	000
B	001
C	010
D	011
E	100
F	101



Κωδικοποίηση Huffman

- Σε κάποιες περιπτώσεις είναι δυνατόν να υπάρξει πληροφορία (ή εκτίμηση) για τη συχνότητα εμφάνισης των συμβόλων στις λέξεις που προέρχονται από το συγκεκριμένο αλφάβητο
 - κάποια σύμβολα εμφανίζονται πιο συχνά από τα άλλα
- Ενδείκνυται η χρήση μικρότερων (σε μήκος) ακολουθιών για τα πιο συχνά χρησιμοποιούμενα σύμβολα.
- Αυτή ακριβώς τη γνώση για τη συχνότητα εμφάνισης των συμβόλων είναι που εκμεταλλεύεται ο **αλγόριθμος Huffman** αλλά και άλλοι συναφείς αλγόριθμοι.



Κωδικοποίηση Huffman

- Έστω τα k προς κωδικοποίηση **σύμβολα** (αλφάβητο)

$$\mathbf{A} = \{a_1, a_2, \dots, a_k\}$$

- Και έστω τα βάρη τα οποία αντιστοιχούν στα ανωτέρω σύμβολα, τα οποία αντιστοιχούν στην **πιθανότητα εμφάνισης** **κάθε συμβόλου**

$$\mathbf{W} = \{w_1, w_2, \dots, w_k\},$$

- Επιθυμούμε να υπολογίσουμε τις **λέξεις κωδικοποίησης** (τις ακολουθίες δηλαδή bits όπου κάθε ένα από αυτά έχει συγκεκριμένη τιμή)

$$C(\mathbf{A}, \mathbf{W}) = \{c_1, c_2, \dots, c_k\}, \text{ αποτελούμενες από } N = \{n_1, n_2, \dots, n_k\} \text{ bits}$$



Κωδικοποίηση Huffman

- Οι λέξεις κωδικοποίησης αντιστοιχούν στα σύμβολα του συνόλου A και επιλέγονται έτσι ώστε να ελαχιστοποιείται το άθροισμα των γινομένων του πλήθους των bits που χρησιμοποιούνται για την κωδικοποίηση κάθε συμβόλου επί την πιθανότητα εμφάνισης του συμβόλου αυτού (το βάρος του).
- Επιθυμούμε, δηλαδή, να ελαχιστοποιηθεί το εξής άθροισμα:

$$\sum_{i=1}^k w_i n_i$$



- Ο αλγόριθμος Huffman για την εύρεση της βέλτιστης κωδικοποίησης με τα παραπάνω δεδομένα βασίζεται στην **κατασκευή του δυαδικού δένδρου Huffman.**



Κατασκευή του δυαδικού δένδρου Huffman (1)

Βήμα 1:

1. Δημιουργούνται (απεικονίζονται) οι **αρχικοί κόμβοι** (που αποτελούν τα φύλλα του δένδρου) από τα **προς κωδικοποίηση σύμβολα**,
2. Τα σύμβολα ενδείκνυται να γράφονται με **αύξουσα ή φθίνουσα σειρά** με βάση τα βάρη τους.
3. Σε κάθε ένα από αυτά αναγράφεται το **βάρος του** (το οποίο αντιστοιχεί στην πιθανότητα εμφάνισής του).

Βήμα 2:

1. Επιλέγοντας τους **δύο κόμβους με τα μικρότερα βάρη**, δημιουργείται ένα νέος εσωτερικός κόμβος, ο οποίος συνδέει (έχει, δηλαδή, σαν παιδιά του) τους δύο επιλεγμένους κόμβους.
2. Το βάρος του νέου κόμβου προκύπτει από το άθροισμα των βαρών των δύο κόμβων από τους οποίους αυτός έχει προκύψει.



Κατασκευή του δυαδικού δένδρου Huffman (2)

Βήμα 3: Το βήμα (2) επαναλαμβάνεται μέχρι να εξαντληθούν όλοι οι κόμβοι (είτε φύλλα είτε σύνθετοι κόμβοι). Με την εξάντληση των κόμβων δημιουργείται τελικά η ρίζα του δένδρου που έχει βάρος το συνολικό άθροισμα των βαρών, δηλαδή μονάδα.

Βήμα 4: Ξεκινώντας από τη ρίζα του δένδρου (τον τελευταίο κόμβο, δηλαδή, της προηγούμενης διαδικασίας) βρίσκουμε τις διαδρομές προς τα φύλλα και σε κάθε κλάδο τοποθετούμε το 0 ή το 1.

Βήμα 5: Ακολουθώντας τις διαδρομές που σχηματίστηκαν στο προηγούμενο βήμα (βήμα 4) υπολογίζεται η κωδικοποίηση που αντιστοιχεί σε κάθε φύλλο (αρχικό, δηλαδή, σύμβολο), ως συνδυασμός των 0 και 1 που αντιστοιχούν στους ενδιάμεσους κλάδους (που συνθέτουν τη διαδρομή).



Παράδειγμα

Σύμβολα	Πλήθος εμφανίσεων	Πιθανότητα εμφάνισης
A	6	0,3
B	4	0,2
C	3	0,15
D	2	0,1
E	4	0,2
F	1	0,05
	20	1

Σύμβολα	Πλήθος εμφανίσεων	Πιθανότητα εμφάνισης
A	6	0,3
B	4	0,2
E	4	0,2
C	3	0,15
D	2	0,1
F	1	0,05
	20	1



Παράδειγμα Βήμα 1

Βήμα 1:

- Δημιουργούνται οι αρχικοί κόμβοι (που αποτελούν τα φύλλα του δένδρου) από τα σύμβολα με αύξουσα ή φθίνουσα σειρά.
- Αναγράφεται κάθε ένα από αυτά το βάρος του.

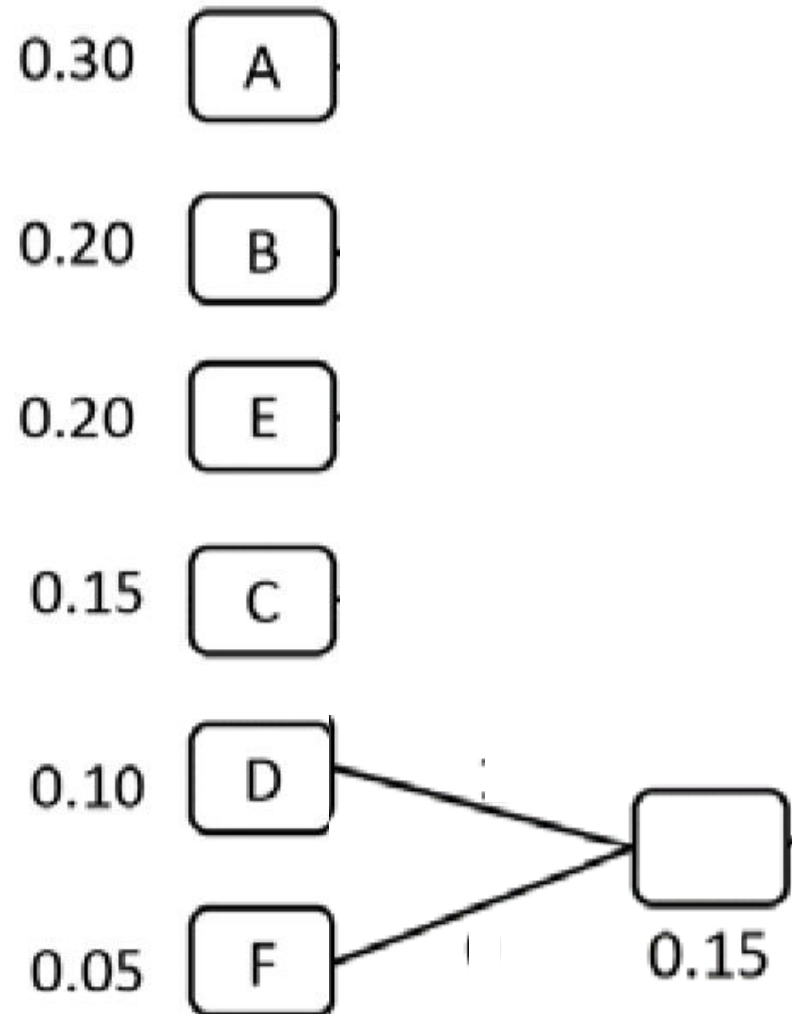
0.30	A
0.20	B
0.20	E
0.15	C
0.10	D
0.05	F



Παράδειγμα Βήμα 2

Βήμα 2:

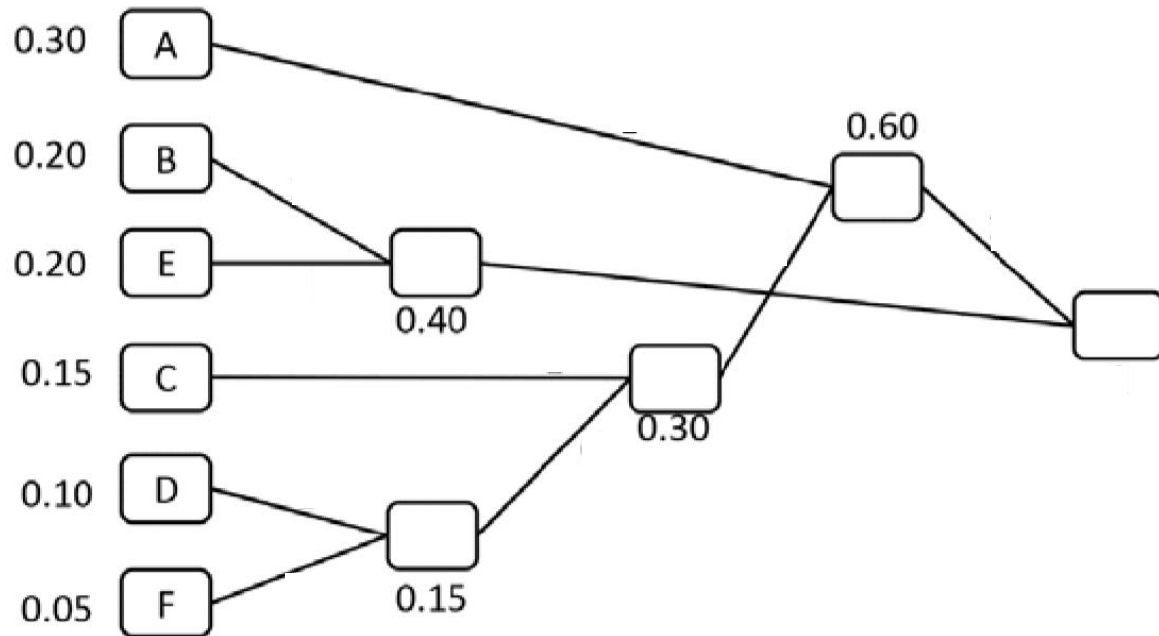
- Επιλέγονται οι δύο κόμβους με τα μικρότερα βάρη
- Δημιουργείται ένα νέος εσωτερικός κόμβος, ο οποίος συνδέει (έχει, δηλαδή, σαν παιδιά του) τους δύο κόμβους.
- Το βάρος του νέου κόμβου προκύπτει από το άθροισμα των βαρών των δύο κόμβων.



Παράδειγμα Βήμα 3

Βήμα 3:

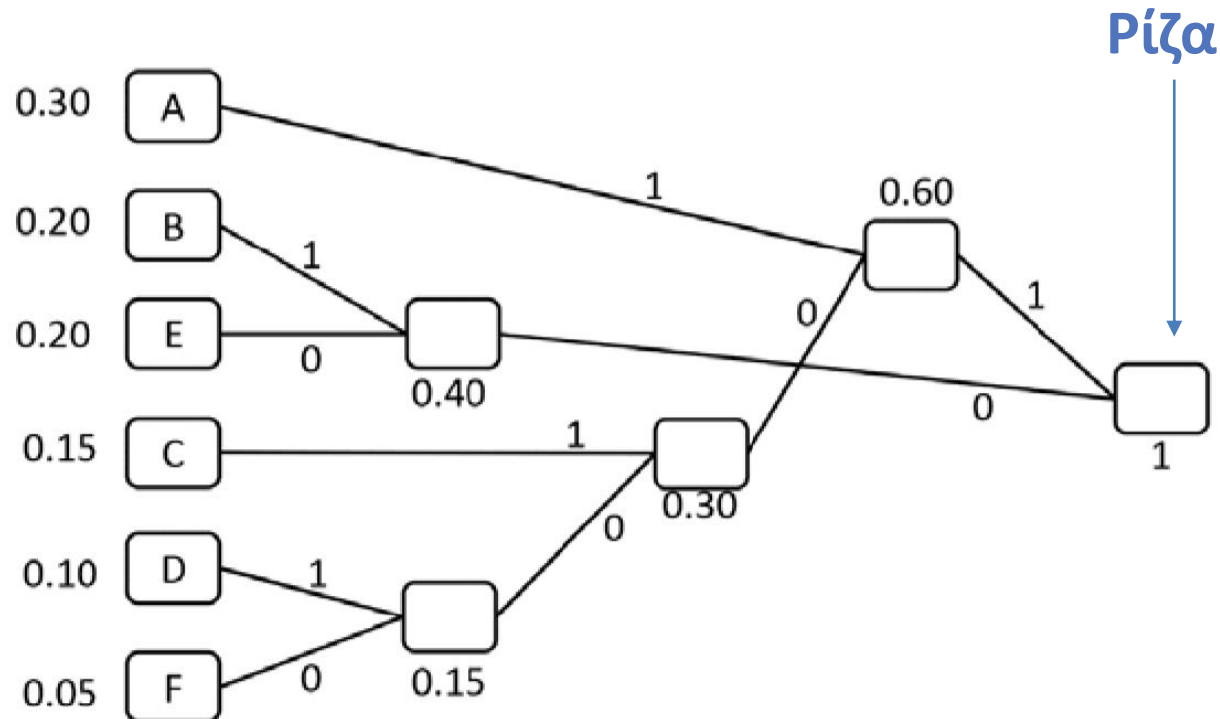
- Το βήμα (2) επαναλαμβάνεται μέχρι να εξαντληθούν όλοι οι κόμβοι (είτε φύλλα είτε σύνθετοι κόμβοι).
- Δημιουργείται τελικά η ρίζα του δένδρου που έχει βάρος το συνολικό άθροισμα των βαρών, δηλαδή μονάδα.



Παράδειγμα Βήμα 4

Βήμα 4:

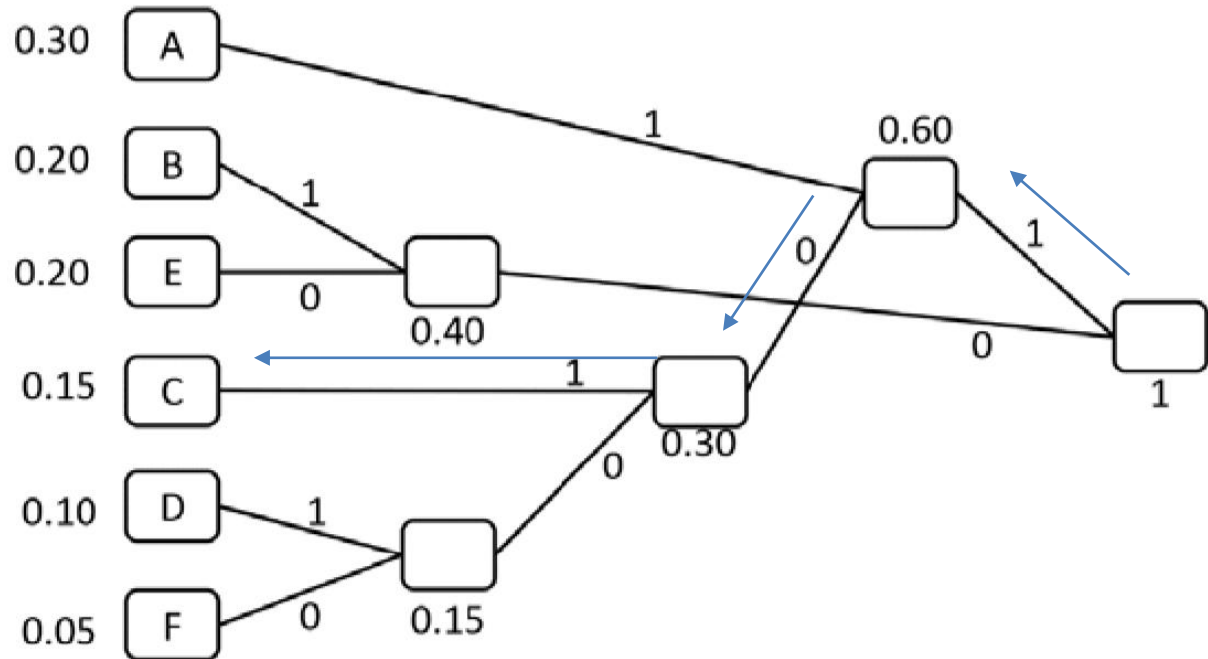
- Ξεκινώντας από τη ρίζα του δένδρου βρίσκουμε τις διαδρομές προς τα φύλλα και σε κάθε κλάδο τοποθετούμε το 0 ή το 1.



Παράδειγμα Βήμα 5

Βήμα 5:

- Ξεκινώντας από τη **ρίζα** υπολογίζεται η κωδικοποίηση που αντιστοιχεί σε κάθε φύλλο ως συνδυασμός των 0 και 1 που αντιστοιχούν στους κλάδους της διαδρομής.

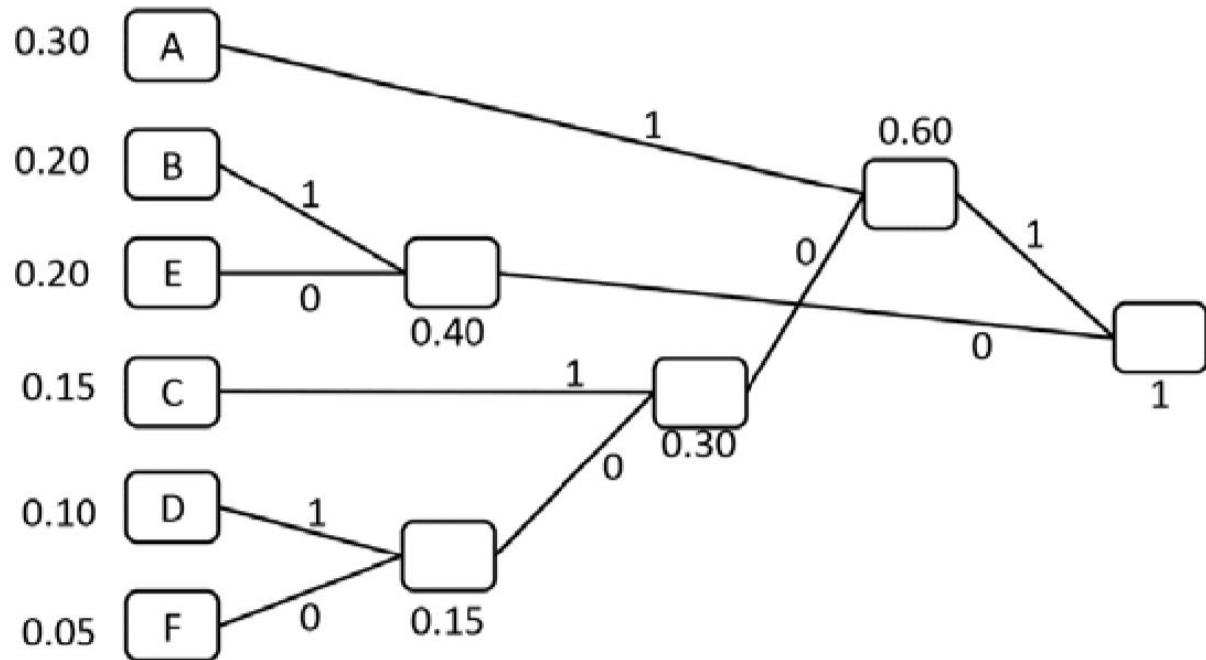


C: 101



Πίνακας κωδικοποιήσεων

Σύμβολα	Κωδικοποίηση
A	11
B	01
C	101
D	1001
E	00
F	1000



Παρατηρήσεις

- Κάθε σύμβολο αντιστοιχεί σε **μία μοναδική κωδικοποίηση.**
- Η κωδικοποίηση οποιουδήποτε συμβόλου **δεν είναι πρόθεμα της κωδικοποίησης κάποιου άλλου συμβόλου.**



Απόδοση κωδικοποίησης Huffman

- Απόδοση κωδικοποίησης Huffman:

$$\sum_{i=1}^k w_i n_i = 0.3*2 + 0.2*2 + 0.15*3 + 0.10*4 + 0.20*2 + 0.05*4 = 2.45$$

- Απόδοση ομοιόμορφης κωδικοποίησης:

$$\sum_{i=1}^k w_i n_i = 3*1 = 3$$

- Προκύπτει λοιπόν ότι, στο συγκεκριμένο παράδειγμα, η κωδικοποίηση Huffman είναι πιο αποδοτική από την ομοιόμορφη κωδικοποίηση κατά περίπου 22%, αφού κάθε σύμβολο στην κωδικοποίηση Huffman κωδικοποιείται με 2.45 bits, ενώ στην ομοιόμορφη με 3 bits.



Πότε χρησιμοποιείται

- Ο αλγόριθμος Huffman συνεργάζεται αρμονικά με τις τεχνικές που συζητήσαμε στις προηγούμενες ενότητες.
- Συγκεκριμένα, αποτελεί **το τελευταίο στάδιο της κωδικοποίησης / συμπίεσης των δεδομένων.**

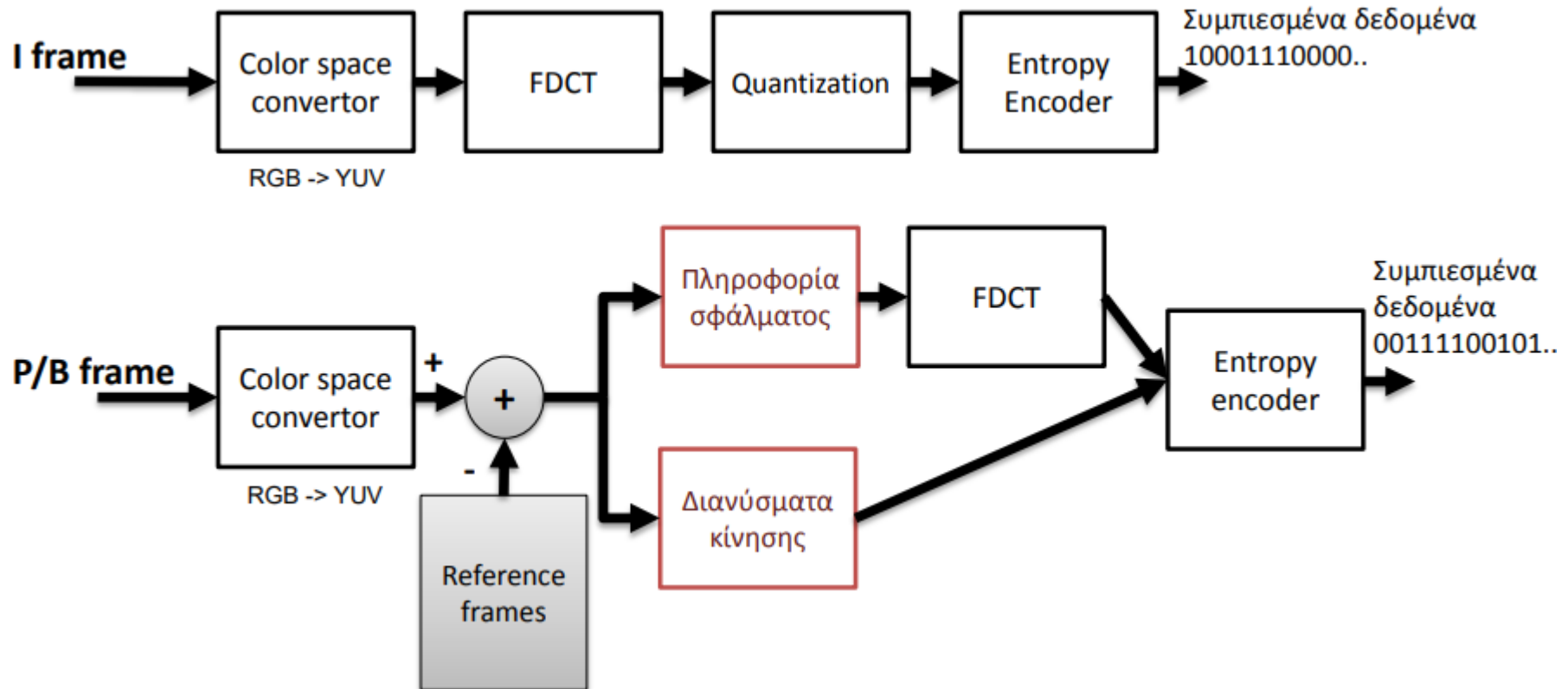


Πότε εφαρμόζεται ο αλγόριθμος Huffman

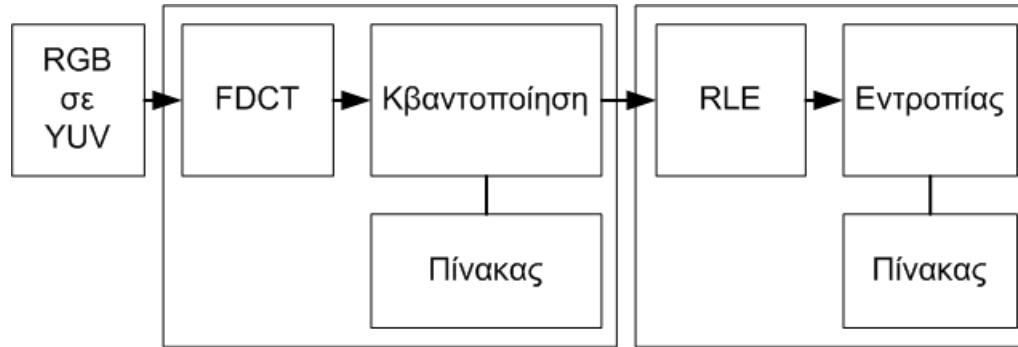
- Αρχικά υπολογίζεται ο **χωρικός πλεονασμός** σε επίπεδο μακρο-μπλοκ.
- Η περιγραφή κάποιων εξ αυτών μπορεί να βασιστεί στις περιγραφές των προηγούμενων και για αυτόν τον λόγο δημιουργείται το **διάνυσμα κίνησης**.
- Στα υπόλοιπα μακρο-μπλοκ (χωρίς διάνυσμα κίνησης) αλλά και για τις διαφορές του προβλεπόμενου μακρο-μπλοκ από το πραγματικό εφαρμόζεται ο **Διακριτός Μετασχηματισμός Συνημιτόνου**.
- Μετά την **κβάντιση των συντελεστών** του ΔΜΣ, γίνεται **κωδικοποίηση μήκους διαδρομής**.
- Στα **αποτελέσματα της κβάντισης** αλλά και στα **διανύσματα κίνησης** **εφαρμόζεται ο αλγόριθμος Huffman**.



Συγκεντρωτικό Διάγραμμα Κωδικοποίησης MPEG



Πλαίσιο-1 (1 από 2)



- Πλαίσιο -I (I-frames)
 - Κωδικοποίηση χωρίς αναφορές σε άλλα καρέ
 - Κατάλληλα για τυχαία προσπέλαση
 - JPEG με απωλεστικό ακολουθιακό ρυθμό
 - Μικρές αποκλίσεις για απλότητα

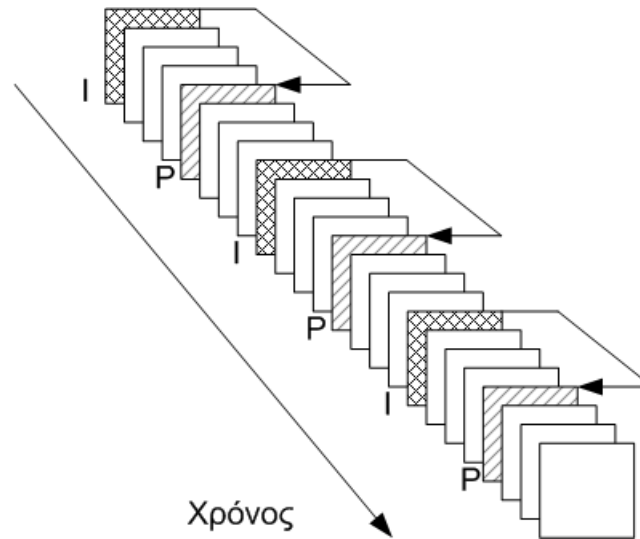


Πλαίσιο -I (2 από 2)

- Είσοδος: Μπλοκ (8x8)
 - Μετασχηματισμός DCT
 - Κβαντοποίηση συντελεστών
 - DPCM για συντελεστές DC (Διαφορική Παλμοκωδική διαμόρφωση)
 - Διάταξη ζιγκ-ζαγκ για συντελεστές AC
 - Κωδικοποίηση μήκους σειρών (RLE)
 - Κωδικοποίηση εντροπίας παρόμοια με Huffman



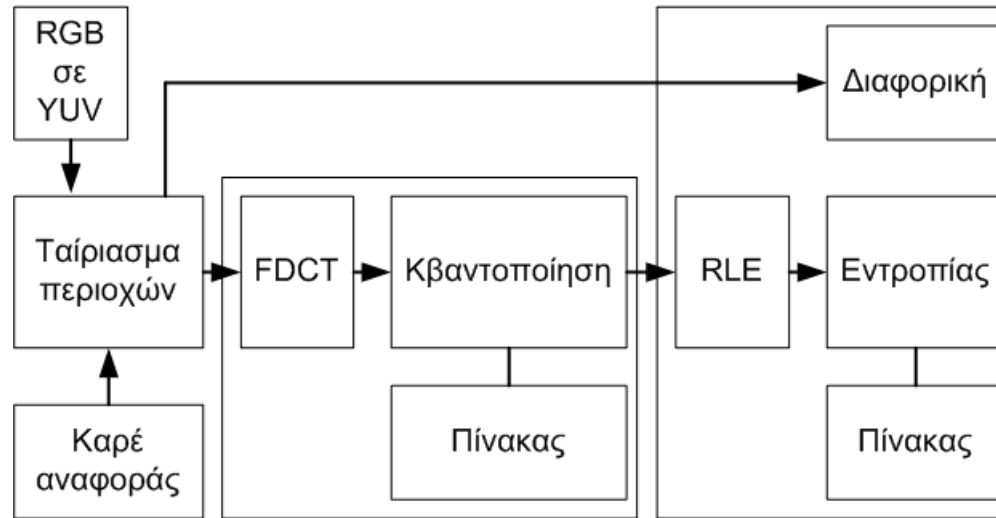
Πλαίσια -P (1 από 2)



- Πλαίσιο -P (P-frames)
 - Επανόρθωση κίνησης με πλαίσιο αναφοράς
 - Αναφορά: αμέσως προηγούμενο πλαίσιο -I ή -P
 - Εντοπισμός πλέον παρόμοιου μακρομπλόκ
 - Επιτρέπονται και μακρομπλόκ χωρίς πρόβλεψη



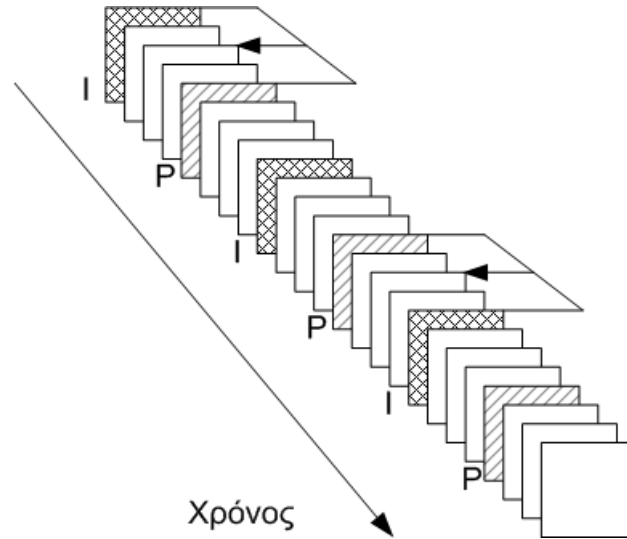
Καρέ-P (2 από 2)



- Διάνυσμα κίνησης
 - DPCM σε γειτονικά μακρομπλόκ
- Διαφορές μπλοκ: μετασχηματισμός DCT
 - Ταυτόχρονη κωδικοποίηση/κβαντοποίηση AC και DC



Καρέ-B



- Καρέ-B (B-frames)
 - Πρόβλεψη με δύο καρέ αναφοράς
 - Ίδια μέθοδος κωδικοποίησης με τα καρέ-P
 - Περιοδική χρήση καρέ-I και καρέ-P
 - Κωδικοποίηση ενδιάμεσων καρέ-B



MPEG Προφίλ και επίπεδα

Παράμετροι ορίζουν τα επίπεδα συμπίεσης

- Ο ρυθμός του βίντεο είναι δυνατό να μειωθεί από τα 270 Mbps (με 4:2:2 χρωματική υποδειγματοληψία) στα 2 Mbps μέχρι και τα 6 Mbps.
- Με τη χρήση των μεθόδων που αναφέραμε για τη **μείωση του πλεονασμού**, καθώς και κάποιας πληροφορίας η **παράλειψη** της οποίας δεν είναι τόσο σημαντική από την άποψη της ποιότητας.
- Το πιο σημαντικό βήμα στη διαδικασία της συμπίεσης είναι ο περιορισμός του **χρονικού πλεονασμού** σε συνδυασμό με τον **χωρικό πλεονασμό** και την εφαρμογή του **ΔΜΣ**, καθώς επίσης και η περαιτέρω **υποδειγματοληψία του χρώματος** (με χρήση του μηχανισμού 4:2:0).
- Από τα παραπάνω μπορούμε να διαπιστώσουμε ότι **το επίπεδο συμπίεσης δεν είναι σταθερό** αλλά εξαρτάται από παραμέτρους τις οποίες μπορεί κάποιος **να επιλέξει και να αλλάξει**.



Προφίλ και Επίπεδα

Το MPEG έχει δημιουργήσει μια σειρά από **προφίλ** (profile) και **επίπεδα** (level) με τα οποία ομαδοποιεί τις επιλογές βάσει διαφορετικών παραμέτρων. Τα **κυριότερα** είναι:

- Η **τυπική ευκρίνεια (SD)** με χρήση του **προτύπου (4:2:0)** ονομάζεται κύριο προφίλ του κύριου επιπέδου **Main Profile@Main Level**
- Η **τυπική ευκρίνεια (SD)** με χρήση του **(4:2:2)** ονομάζεται υψηλό προφίλ του κύριου επιπέδου **High Profile@Main Level**
- Η **υψηλή ευκρίνεια (HD)** με χρήση του **4:2:0** αντιστοιχεί στο κύριο προφίλ του υψηλού επιπέδου **Main Profile@High Level**
- Η **υψηλή ευκρίνεια (HD)** με χρήση του **4:2:2** signal είναι το υψηλό προφίλ στο υψηλό επίπεδο **High Profile@High Level**



Προφίλς

Συν/ση	Ονομασία	Τύποι κωδικ/σης εικόνας	Υποδειγμ/ψία Χρώματος	Αναλογία Εικόνας	Κλιμάκωση Ποιότητας
SP	Simple Profile	I, P	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι
MP	Main Profile	I, P, <u>B</u>	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι
SNR	SNR Scalable Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	<u>Κλιμάκωση κατά SNR</u>
Spatial	Spatially Scalable Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	<u>Χωρική κλιμάκωση</u> και κλιμάκωση κατά SNR
HP	HP Profile	I, P, B	<u>4:2:2</u> ή 4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι
422	4:2:2 Profile	I, P, B	4:2:2 ή 4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	<u>Χωρική κλιμάκωση</u> και κλιμάκωση
MVP	Multi View Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	<u>Χρονική κλιμάκωση</u>



MPEG-2 Profiles

Abbr.	Name	Picture Coding Types	Chroma Format	Scalable modes	Intra DC Precision
SP	Simple profile	I, P	4:2:0	none	8, 9, 10
MP	Main profile	I, P, B	4:2:0	none	8, 9, 10
SNR	SNR Scalable profile	I, P, B	4:2:0	SNR^[a]	8, 9, 10
Spatial	Spatially Scalable profile	I, P, B	4:2:0	SNR ^[a] , spatial^[b]	8, 9, 10
HP	High profile	I, P, B	4:2:2 or 4:2:0	SNR ^[a] , spatial ^[b]	8, 9, 10, 11
422	4:2:2 profile	I, P, B	4:2:2 or 4:2:0	none	8, 9, 10, 11
MVP	Multi-view profile	I, P, B	4:2:0	Tempora^[c]	8, 9, 10



Κλιμάκωση ποιότητας

Η ανάγκη για κλιμάκωση στην ποιότητα της κωδικοποίησης προκύπτει από τις **διαφορετικές δυνατότητες στα υποκείμενα δίκτυα** π.χ. IPTV.

- **βασική κωδικοποίηση:**
 - κωδικοποίηση του βίντεο με βασική χαμηλή σχετικά ποιότητα
- **δεύτερο επίπεδο κωδικοποίησης:**
 - καλύτερη ποιότητα
 - επαύξηση της ποιότητας επί της βασικής κωδικοποίησης



Κλιμάκωση στο MPEG2

1. Σηματοθορυβικό επίπεδο (Signal to Noise Ratio, SNR)

- Αλλαγές της κλιμάκωσης των τιμών των συντελεστών στο μετασχηματισμό DCT και του πλήθους των συντελεστών που λαμβάνονται υπ' όψιν.
- Στη **βασική έκδοση** κβαντίζονται/κλιμακώνονται οι συντελεστές DCT με χαμηλή ανάλυση π.χ Q_{50} .
- Για **επαύξηση της ποιότητας** υπολογίζονται οι διαφορές των συντελεστών DCT από τους κβαντισμένους και κβαντίζονται οι διαφορές.
- Η βασική και η ενισχυμένη κωδικοποίηση αντιστοιχούν στην ίδια χωρική ανάλυση του βίντεο (διαστάσεις).



162	41	20	72	30	12	-20	-11
30	108	20	32	28	-15	18	-2
-94	-60	12	-43	-31	6	-3	7
-39	-83	-5	-22	-13	15	-1	3
-31	18	-5	-12	14	-6	11	-6
-1	-12	13	0	28	13	8	3
5	-2	12	7	-19	-13	8	12
-10	11	8	-16	21	0	6	11

160	44	20	80	24	0	0	0
36	108	14	38	26	0	0	0
-98	-65	16	-48	-40	0	0	0
-42	-85	0	-29	0	0	0	0
-36	22	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0



Κλιμάκωση στο MPEG2

2. Χωρικό (spatial) επίπεδο

- υποστήριξη διαφορετικών συσκευών με **διαφορετικές αναλύσεις**
 - η βασική κωδικοποίηση με βάση την τυπική ανάλυση
 - η ενισχυμένη κωδικοποίηση να γίνει με βάση την υψηλή ανάλυση

3. Χρονικό (temporal) επίπεδο

- υποστήριξη διαφορετικών **ρυθμών ανανέωσης πλαισίων**. π.χ. σε ασύρματους διαύλους επικοινωνίας, όπου λόγω φτωχής κάλυψης είναι δυνατόν να μην αποστέλλονται κάποια από τα πλαίσια.



Επίπεδα κατά το MPEG2

Συν/ση	Ονομασία	Ρυθμός Ανανέωσης Πλαισίων	Μέγιστη Οριζόντια Ανάλυση	Μέγιστη Κατακόρυφη Ανάλυση	Μέγιστο πλήθος Δειγμάτων Φωτεινότητας ανά δευτερόλεπτο	Μέγιστος Ρυθμός (σε κύριο προφύ.) σε Mbps
LL	Low Level	23.976, 24, 25, 29.97, 30	352	288	3.041.280	4
ML	Main Level	23.976, 24, 25, 29.97, 30	720	576	10.368.000, εκτός από το Highprofile όπου ο περιορισμός γίνεται 14.475.600 για 4:2:0 και 11.059.200 για 4:2:2	15
H-14	High 1440	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1440	1152	47.001.600 εκτός από το High profile με 4:2:0 όπου είναι 62.668.800	60
HL	High Level	23.976, 24, 25, 29.97, 30, 50, 59.94, 60	1920	1152	62.668.800 εκτός από το High profile με 4:2:0 όπου είναι 83.558.400	80

https://en.wikipedia.org/wiki/H.262/MPEG-2_Part_2



Συνήθεις συνδυασμοί προφίλ και επιπέδου

- MP@ML, 720x480, 30, 4:2:0, 4 Mbps
 - ψηφιακή τηλεόραση τυπικής ανάλυσης
- MP@ML, 720x576, 25, 4:2:0, 4 Mbps
 - ψηφιακή τηλεόραση τυπικής ανάλυσης
- MP@HL, 1920x1080, 30, 4:2:0, 80 Mbps
 - ψηφιακή τηλεόραση υψηλής ευκρίνειας
- MP@HL, 1920x720, 60, 4:2:0, 80 Mbps
 - ψηφιακή τηλεόραση υψηλής ευκρίνειας



DVB-AVC (Audio & Video Coding)

- **Specification for the use of Video and Audio Coding in Broadcast and Broadband Applications**
- <https://dvb.org/?standard=specification-for-the-use-of-video-and-audio-coding-in-broadcast-and-broadband-applications>
- Video encoding in DVB broadcast bitstreams:
 - MPEG-2 video, or H.264/AVC video, or HEVC video, or VC-1 video



DVB Video Coding: MPEG-2

- MPEG-2 Main Profile at Main Level is used for MPEG-2 encoded SDTV.
- MPEG-2 Main Profile at High Level is used for MPEG-2 encoded HDTV.
- The 25/30 Hz MPEG-2 SDTV IRD

Χαρακτηριστικά

- MPEG-2 HDTV pictures have 16:9 or 2.21:1 aspect ratio; IRDs support 16:9 and optionally 2.21:1 aspect ratio. • MPEG-2 IRDs support the use of pan vectors to allow a 4:3 monitor to give a full-screen display of a 16:9 coded picture of SDTV resolution



DVB Video Coding: H.264/ACV

- H.264/AVC Main Profile at Level 3 is used for H.264/AVC SDTV.
- H.264/AVC High Profile at Level 4 is used for 25 Hz and 30 Hz H.264/AVC HDTV.
- H.264/AVC High Profile at Level 4.2 is used for 50 Hz and 60 Hz H.264/AVC HDTV.
- H.264/AVC Scalable High Profile at Level 4 is used for 25 Hz and 30 Hz SVC HDTV.
- H.264/AVC Stereo High Profile at Level 4 is used for 25 Hz and 30 Hz MVC Stereo HDTV.
- H.264/AVC Scalable High Profile at Level 4.2 is used for 50 Hz and 60 Hz SVC HDTV.
- 25/30 Hz H.264/AVC SDTV/HDTV IDR
 - The 25 Hz H.264/AVC SDTV IDR, 25 Hz H.264/AVC HDTV IDR, 30 Hz H.264/AVC SDTV IDR, 30 Hz H.264/AVC HDTV IDR
- H.264/AVC HDTV pictures have 16:9 aspect ratio; IRDs support 16:9 aspect ratio.
- IDR (Instantaneous Decoding Refresh)
 - IRDs may also optionally support the use of the Active Format Description (refer to annex B of the present document) as part of the logic to control the processing and positioning of the reconstructed image for display.
 - IRDs may also optionally support frame compatible plano-stereoscopic 3DTV services (see annex H).
 - IRDs may also optionally support service frame compatible plano-stereoscopic 3DTV services with HEVC coding (see annex J).



DVB Video Coding: HEVC

- HEVC Main or Main 10 Profile at Level 4.1 is used for HEVC HDTV.
- HEVC Main 10 Profile at Level 5.1 is used for HEVC UHDTV and HEVC HDR UHDTV.
- HEVC Main 10 Profile at Level 5.2 is used for HEVC HFR UHDTV.
- 25/30 Hz HEVC SDTV/HDTV IDR
 - The 25 Hz HEVC SDTV IRD, 25 Hz HEVC HDTV IRD, 30 Hz HEVC SDTV IRD, 30 Hz H.264/AVC HDTV IRD
- IDR
 - The HEVC UHDTV IRD and HEVC HDR UHDTV IRD support frame rates of 24 000/1 001, 24, 25, 30 000/1 001, 30, 50, 60 000/1 001 and 60 Hz.
 - The HEVC HDR HFR UHDTV IRD supports frame rates of 24 000/1 001, 24, 25, 30 000/1 001, 30, 50, 60 000/1 001, 60, 100, 120 000/1 001 and 120 Hz



DVB Video Coding: VC-1

- VC-1 Advanced Profile at Level 1 / 3 are used for VC-1 SDTV.
- 25/30 Hz VC-1 SDTV/HDTV IDR



DVB audio codings

Audio content complies with

- MPEG-1 Layer I, MPEG-1 Layer II,
- MPEG-2 Layer II backward compatible,
- AC-3, Enhanced AC-3,
- AC-4,
- DTS Audio, DTS-HD, DTS-UHD,
- MPEG-4 AAC, MPEG-4 HE AAC, MPEG-4 HE AAC v2 or MPEG-H LC audio.
- Other
 - MPEG-1 Layer II, MPEG-4 AAC, MPEG-4 HE AAC and MPEG-4 HE AAC v2 audio streams may optionally include MPEG Surround data. •
 - Sampling rates of 32 kHz, 44,1 kHz and 48 kHz are supported by IRDs.





Τέλος Ενότητας

Σημειώματα

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση 1.0.



Σημείωμα Αναφοράς

Copyright Εθνικών και Καποδιστριακών Πανεπιστημίων Αθηνών, Παντελής Μπαλαούρας, «Συστήματα Ψηφιακής Ευρυεκπομπής, Παραγωγή Τηλεοπτικού Σήματος, Τεχνικές Συμπίεσης Βίντεο και Ήχου.». Έκδοση: 1.0. Αθήνα, 2019. Διαθέσιμο από τη δικτυακή διεύθυνση: <https://eclass.uoa.gr/courses/D476/> .



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.



Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων (1/2)

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων και ουσιαστικά αποτελεί μετεγγραφή των ακόλουθων έργων σε διαφάνειες:

[2] Παπαδάκης, Α. 2015. Συμπίεση Βίντεο. *Ψηφιακή τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 3. Διαθέσιμο στο:
<http://hdl.handle.net/11419/5008>

