



Αρχές Ψηφιακής Τηλεόρασης ***Ενότητα 3 – Συμπίεση video***

Καθηγητής Δρ. Ευάγγελος Πάλλης
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Ελληνικό Μεσογειακό Πανεπιστήμιο

Περιεχόμενα

- Σύνοψη και μαθησιακά αποτελέσματα/στόχοι
- Εισαγωγή στην Ενότητα 3
- Μηχανισμοί συμπίεσης
- Προετοιμασία σήματος video
 - Μείωση Ακρίβειας video
 - Διάστημα Οριζόντιων και Κατακόρυφων Παλμών
 - Υποδειγματοληψία Χρώματος
 - Περιορισμός Ρυθμού
- Εκμετάλλευση Χρονικού Πλεονασμού
- Εκμετάλλευση Χωρικού Πλεονασμού
 - Διακριτός Μετασχηματισμός Συνημιτόνου (ΔΜΣ)
 - Χρήση ΔΜΣ
 - Σάρωση κατά Ζιγκ Ζαγκ
 - Παραδείγματα
 - Κωδικοποίηση Μήκους Διαδρομής
- Εκμετάλλευση Στατιστικού Πλεονασμού
 - Ομοιόμορφη Κωδικοποίηση
 - Κωδικοποίηση εντροπίας – Huffman
 - Παραδείγματα
- Προφίλ και Επίπεδα

Περιγραφή Ενότητας 3

Σύνοψη

Στην Ενότητα 3 εξετάζουμε τους βασικούς τρόπους με τους οποίους το πρότυπο MPEG-2 προβλέπει τη συμπίεση του σήματος video, εστιάζοντας σε απωλεστικούς και μη απωλεστικούς μηχανισμούς αφαίρεσης πληροφορίας. Οι πρώτοι, «αγνοούν» κάποια πληροφορία κατά τη συμπίεση (την οποία ο δέκτης δεν μπορεί να επαναφέρει) με πρόνοια η πληροφορία αυτή να είναι όσο το δυνατόν πιο περιορισμένης σημασίας για την ποιότητα του περιεχομένου. Οι δεύτεροι, εκμεταλλεύονται κάποιον πλεονασμό της πληροφορίας, όπως είναι ο:

- **χωρικός πλεονασμός**, δηλαδή ομοιότητες μεταξύ των εικονοστοιχείων που δομούν ένα πλαίσιο,
- **χρονικός πλεονασμός**, δηλαδή ομοιότητες μεταξύ διαδοχικών εικόνων ενός video,
- **στατιστικός πλεονασμός**, δηλαδή στατιστικές ιδιότητες της εμφάνισης των διάφορων συμβόλων προς κωδικοποίηση και
- **ψυχο-οπτικός πλεονασμός**, δηλαδή ιδιότητες του ανθρώπινου συστήματος όρασης και στα χαρακτηριστικά της απόκρισης του συστήματος αυτού στα διάφορα ερεθίσματα.

Στα πλαίσια αυτά, η Ενότητα εστιάζεται στη συμπίεση που επιτυγχάνεται στο MPEG-2 μέσω της αφαίρεσης της «πλεονάζουσας» πληροφορίας (π.χ. της πληροφορίας που βρίσκεται στο χρονικό διάστημα που αντιστοιχούν στους παλμούς αμαύρωσης) αλλά και εκμετάλλευσης των παραπάνω ειδών πλεονασμού.

Προαπαιτούμενη γνώση

Η Ενότητα 3 επεξηγεί τους βασικούς μηχανισμούς για τη συμπίεση του video, και συνδέεται με τους μηχανισμούς δημιουργίας του video (βλ. Ενότητα 2) καθώς και με τους μηχανισμούς εισαγωγής της πληροφορίας στον συρμό μεταφοράς (βλ. Ενότητα 4). Υπό αυτήν την έννοια, ενδείκνυται γνώση της γενικής εικόνας του συστήματος και των επιμέρους βαθμίδων που το αποτελούν τόσο στον πομπό όσο και στον δέκτη, ώστε να γίνει αντιληπτή η σημασία της λειτουργικότητας που περιγράφεται σε αυτή την Ενότητα.

Μαθησιακά αποτελέσματα

Με την επιτυχή ολοκλήρωση της Ενότητας 3 ο φοιτητής / τρια θα είναι σε θέση να:

- Κατανοεί τον τρόπο λειτουργίας και οργάνωσης των τηλεοπτικών συστημάτων συμπίεσης πληροφορίας video και τη χρήση τους στην παραγωγή, μετάδοση και λήψη τηλεοπτικών υπηρεσιών.
- Γνωρίζει τις βασικές αρχές συμπίεσης του σήματος video και τους μηχανισμούς που χρησιμοποιούνται στο πρότυπο MPEG-2 για απωλεστική και μη-απωλεστική αφαίρεση πληροφορίας.
- Εφαρμόζει μεθόδους αξιολόγησης και εργαλεία ανάλυσης των επιδόσεων των μηχανισμών συμπίεσης, καθώς και τεχνικές βέλτιστης συμπίεσης για την επίτευξη υψηλής ποιότητας σήματος video.
- Αναλύει και υπολογίζει τα βασικά χαρακτηριστικά συμπίεσης του σήματος video, και του τρόπου εφαρμογής τους σε τηλεοπτικά συστήματα μετάδοσης.
- Προτείνει λύσεις σε θέματα λειτουργίας και συντήρησης των μονάδων συμπίεσης του σήματος video, και τρόπους αξιολόγησης της παρεχόμενης ποιότητας.

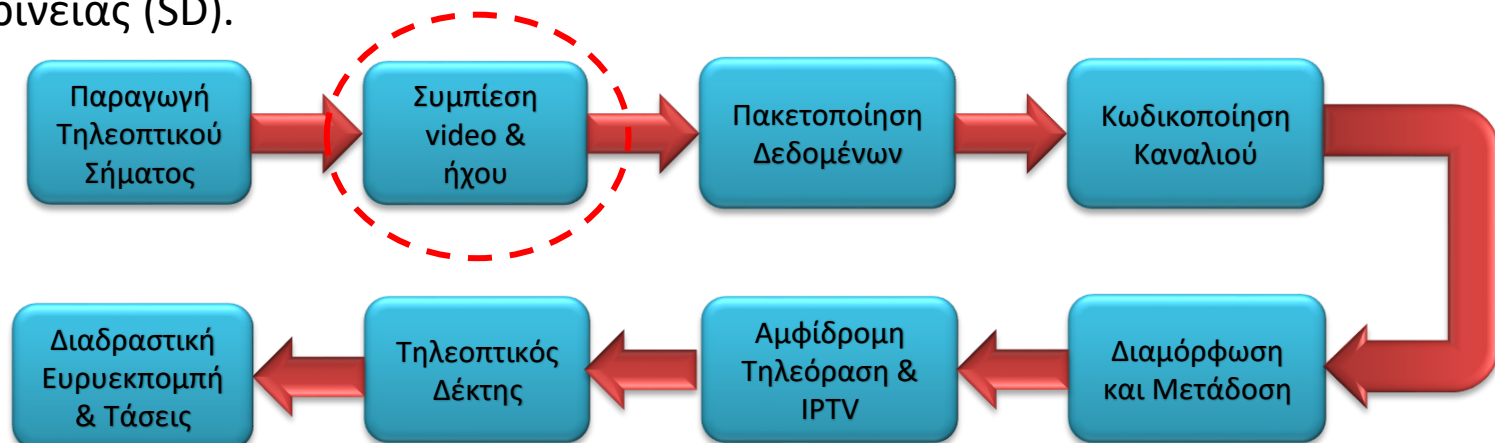
Εισαγωγή

Εισαγωγή στην Ενότητα 3

Όπως είδαμε στην Ενότητα 2, ο ρυθμός που απαιτείται για να μεταδώσουμε το ψηφιοποιημένο video είναι πολύ υψηλός για να υποστηριχθεί από τις τρέχουσες τεχνολογίες μετάδοσης στον τελικό χρήστη (κυρίως μέσω επίγειων ασύρματων υποδομών). Π.χ. για τυπικής ευκρίνειας video (Standard Definition, SD) ο ρυθμός αντιστοιχεί στα 270 Mbps, ενώ για υψηλής ευκρίνειας video ο ρυθμός αντιστοιχεί σε πάνω από 1 Gbps.

Αυτοί οι ρυθμοί μετάδοσης πρέπει να υποβιβαστούν στην τάξη των 2-7 Mbps, άρα απαιτείται ένα σημαντικό ποσοστό συμπίεσης από 38:1 (πιο απλή περίπτωση) έως και 137:1 (πιο σύνθετη περίπτωση). Για τη μετάδοσης σήματος υψηλής ευκρίνειας το συμπιεσμένο σήμα είναι της τάξης των 15 με 20 Mbps.

Σε αυτή την Ενότητα θα εστιάσουμε την περίπτωση συμπίεσης video τυπικής ευκρίνειας (SD).



Εισαγωγή στην Ενότητα 3

Η συμπίεση του video μπορεί να είναι **απωλεστική** (lossy) και **μη-απωλεστική** (lossless).

- Στην απωλεστική συμπίεση υπάρχει απώλεια πληροφορίας.
- Στη μη-απωλεστική συμπίεση δεν έχουμε απώλεια πληροφορίας παρά μόνο απομάκρυνση πλεονάζουσας.

Ως πλεονάζουσα ορίζουμε την πληροφορία η οποία ενδέχεται να:

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

Εισαγωγή στην Ενότητα 3

Μη-απωλεστική αφαίρεση πληροφορίας: παράδειγμα περιορισμού της πλεονάζουσας πληροφορίας είναι οι έξυπνοι τρόποι περιγραφής μιας σειράς από ίδια δεδομένα, όπως στην περίπτωση που αντί να γράψουμε σειριακά το σύμβολο “A” 10 φορές δηλαδή «A, A, ..., A» μπορούμε να πούμε 10 φορές το σύμβολο “A”.

Απωλεστική αφαίρεση πληροφορίας: στόχος είναι η πληροφορία που θα χαθεί (και δεν θα μπορεί να αναπαραχθεί στον δέκτη) να έχει μικρή επίπτωση στην ποιότητα του περιεχομένου και στη γενικότερη αντίληψη του χρήστη. Για παράδειγμα, αφού το ανθρώπινο μάτι είναι πολύ πιο ευαίσθητο στις αλλαγές στις τιμές της φωτεινότητας σε σχέση με τις αλλαγές στις τιμές των χρωμάτων, μπορούμε να δώσουμε την κατάλληλη προτεραιότητα στο σήμα φωτεινότητας έναντι αυτών των χρωμοδιαφορών.

Μηχανισμοί συμπίεσης

Μηχανισμοί συμπίεσης

Τα περισσότερα πρότυπα συμπίεσης ψηφιακού video στηρίζονται στην αξιοποίηση διαφόρων ειδών πλεονασμού της πληροφορίας και στοχεύουν στη μείωσή της, άρα και στη μείωση του όγκου των δεδομένων. Τα είδη πλεονασμού είναι **α)** ο χωρικός πλεονασμός, **β)** ο Χρονικός πλεονασμός, **γ)** ο Στατιστικός πλεονασμός, και **δ)** ο ψυχο-οπτικός πλεονασμός.

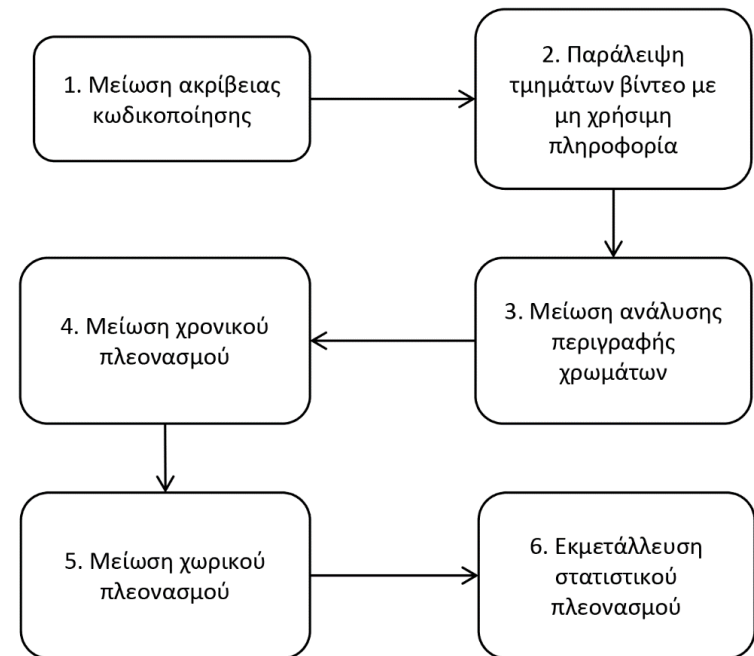
Στη ψηφιακή τηλεόραση, το προς συμπίεση σήμα αποτελείται από τις συνιστώσες των βασικών χρωμάτων (R, G, B) οι οποίες συνδυάζονται ώστε να δώσουν το σήμα της φωτεινότητας (Y) και των χρωμοδιαφορών (Cb και Cr). Στην περίπτωση σήματος 4:2:2, η ανάλυση του χρώματος είναι μισή σε σχέση με την ανάλυση της φωτεινότητας, ενώ στην περίπτωση που έχουμε κβαντισμό 10 bits, ο ρυθμός μπορεί να φτάνει και τα 270 Mbps σύμφωνα και με το πρότυπο του ITU-BT.R601 (21).

- Ο μηχανισμός συμπίεσης των εικόνων βασίζεται στο πρότυπο του **JPEG** (Joint Photographic Experts Group),
- Ο μηχανισμός συμπίεσης του video βασίζεται στο πρότυπο **MPEG** (Motion Pictures Expert Group). Το MPEG-2 έχει χρησιμοποιηθεί και χρησιμοποιείται σε πολλές περιπτώσεις τηλεοπτικής μετάδοσης, ωστόσο σε αρκετές περιπτώσεις (συμπεριλαμβανομένης και της Ελλάδας) χρησιμοποιείται και το πρότυπο MPEG-4 (AVC/H264), το οποίο εξασφαλίζει πιο αποδοτική συμπίεση.

Μηχανισμοί συμπίεσης

Ο μηχανισμός συμπίεσης του video ακολουθεί τα επόμενα έξι (6) βήματα:

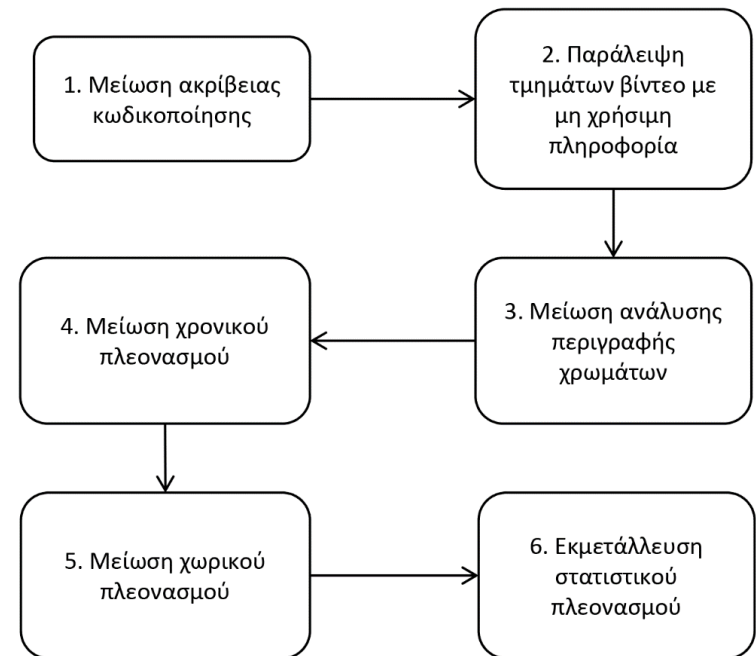
1. Μείωση της ακρίβειας των τιμών της φωτεινότητας και των χρωμοδιαφορών, δηλαδή μείωση του πλήθους των επιπέδων κβαντισμού και συνακόλουθη μείωση των bits.
2. Παράλειψη των τμημάτων του video τα οποία δεν περιέχουν χρήσιμη πληροφορία, (π.χ. οριζόντιους και κατακόρυφους παλμούς αμαύρωσης).
3. Περαιτέρω μείωση της ανάλυσης της περιγραφής των χρωμάτων, πέρα από το πρότυπο 4:2:2 το οποίο χρησιμοποιείται κατά τη δημιουργία του σήματος.



Βασικά βήματα σε μια τυπική διαδικασία συμπίεσης video

Μηχανισμοί συμπίεσης

4. Αξιοποίηση της ομοιότητας η οποία 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.
5. Αξιοποίηση της ομοιότητας που ενδέχεται να υπάρχει μεταξύ των περιοχών που συνθέτουν ένα πλαίσιο. Π.χ. ομοιόμορφες περιοχές του πλαισίου, π.χ. τον ουρανό ή το φόντο.
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.



Βασικά βήματα σε μια τυπική διαδικασία συμπίεσης video

Προετοιμασία σήματος video

Μείωση ακρίβειας video

Στην αναλογική τηλεόραση το κατώφλι του θορύβου στο video είθισται να είναι στα 48dB. Εάν δηλαδή ο σηματοθορυβικός λόγος (Signal to Noise Ratio, SNR) είναι μεγαλύτερος ή ίσος των 48dB, τότε θεωρούμε ότι ο θόρυβος στο video δεν είναι ορατός στο ανθρώπινο μάτι.

Η κβαντοποίηση με χρήση 8 bits παρέχει επίπεδο θορύβου κάτω από τα 48 dB, και για αυτόν τον λόγο ο θόρυβος κβαντοποίησης (SQNR) δεν είναι αντιληπτός από το ανθρώπινο μάτι.

- Υπό αυτήν την έννοια, 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.**

Μείωση ακρίβειας video

Το πλήθος των bits που επιλέγουμε για τον κβαντισμό σχετίζεται με την **ποιότητα του κβαντισμού** και συγκεκριμένα τον λόγο του σήματος προς τον **θόρυβο κβαντισμού** (Signal to Quantization Noise Ratio, SQNR). Το **SQNR σχετίζεται με την στρογγυλοποίηση** που πραγματοποιείται και εξαρτάται από το πλήθος των διαθέσιμων επιπέδων.

- Η στρογγυλοποίηση αυτή αφορά το **σφάλμα κβαντισμού** (ή, αλλιώς, τον **θόρυβο κβαντισμού**).
- Μιας και το σφάλμα κβαντισμού είναι κατά μέγιστο το ήμισυ του διαστήματος κβαντισμού, και υπολογίζεται ως εξής:

$$\text{SQNR} = 20 \log 2^{(N-1)} / (1/2) = 20 * N * \log 2 = 6.02N \text{ (dB)} \quad (1)$$

Όπου N είναι το πλήθος των bits, π.χ. εάν N=8 τότε SQNR=48.16dB, και εάν N=10 τότε SQNR=60.2 dB

Από την παραπάνω σχέση προκύπτει ότι ο θόρυβος κβαντισμού αυξάνεται κατά περίπου 12 dB, με τη μείωση της ακρίβειας από 10 σε 8 bits. Όσον αφορά τον ρυθμό μετάδοσης η μείωση ανέρχεται στο 20%, αφού παραμένει το 80% του αρχικού περιεχομένου.

Διάστημα οριζόντιων και κατακόρυφων παλμών

Το σήμα video το οποίο παράγεται έχει κάποια χρονικά διαστήματα, στα οποία δεν αντιστοιχεί χρήσιμο οπτικό σήμα, και αυτά είναι ο οριζόντιος και ο κατακόρυφος παλμός επιστροφής ή αμαύρωσης (HBI και VBI).

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.

Κέρδος από την μη-αποστολή των HBI και VBI στο Ευρωπαϊκού σύστημα τηλεόρασης.

- VBI: αναφορικά με τον παλμό κατακόρυφης επιστροφής, αυτός αντιστοιχεί σε περίπου 50 γραμμές από τις 625, άρα στο 8% του συνολικού σήματος.

$$50 / 625 = 8\%$$

- HBI: αναφορικά με τον παλμό κατακόρυφης επιστροφής, κάθε γραμμή διαρκεί 64 μs 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.**

Υποδειγματοληψία χρώματος

Κατά τη δημιουργία του σήματος video (βλ. Ενότητα 2), οι χρωμοδιαφορές Cb και Cr λαμβάνονται με μικρότερη συχνότητα δειγματοληψίας σε σχέση με τη φωτεινότητα Y. Συγκεκριμένα, 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 (το σήμα της φωτεινότητας στα 5,75 MHz), και με δεδομένη τη μειωμένη ευαισθησία του ανθρώπινου ματιού στα σήματα των χρωμοδιαφορών, είναι δυνατή η περαιτέρω μείωση του ρυθμού του σήματος των χρωμοδιαφορών κατά το ήμισυ. Συγκεκριμένα, 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.

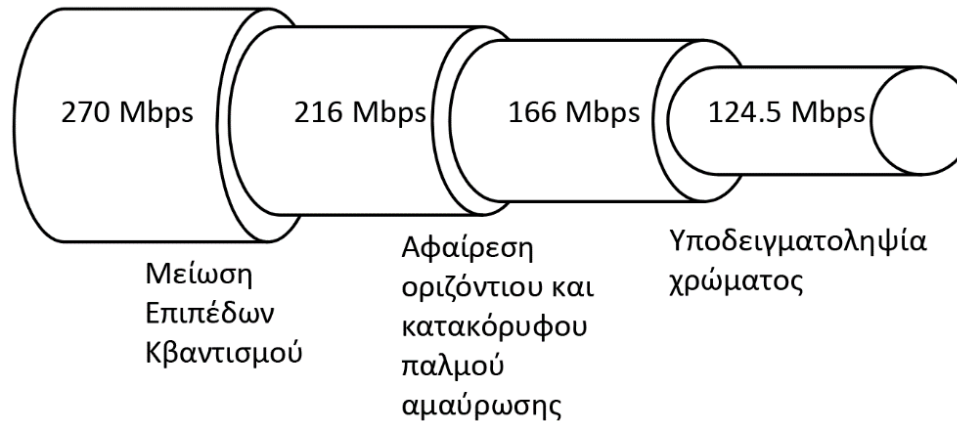
Από την πλευρά του απαιτούμενου ρυθμού έχουμε υποβάθμιση αφού 2×4 δείγματα φωτεινότητας αντιστοιχούν σε 2 δείγματα από τις χρωμοδιαφορές. **Η υποβάθμιση του ρυθμού αντιστοιχεί στο 25%.**

Περιορισμός ρυθμού

Θεωρώντας ότι ο εικονολήπτης (κάμερα) δίνει σήμα video ρυθμού **270 Mbps**, τότε ο περιορισμός αυτού του ρυθμού με βάση α) τη μείωση ακρίβειας video, β) το διάστημα οριζόντιων και κατακόρυφων παλμών και γ) την υποδειγματοληψία χρώματος, είναι

- 20% μικρότερος λόγω της μείωσης των επιπέδων κβαντισμού, άρα γίνεται **216 Mbps**.
- 25% μικρότερος λόγω της αφαίρεσης των παλμών οριζόντιας και κατακόρυφης (HBI και VBI), άρα γίνεται **166 Mbps**.
- 25% μικρότερος λόγω της υποδειγματοληψίας των σημάτων των χρωμοδιαφίρων και τον υποβιβασμό στο πρότυπο 4:2:0, άρα γίνεται **124.5 Mbps**.

Τα επόμενα βήματα έχουν να κάνουν με τον περαιτέρω περιορισμό του ρυθμού στα **2-4 Mbps**. Η επίτευξη αυτού του στόχου γίνεται με πιο περίπλοκους μηχανισμούς.



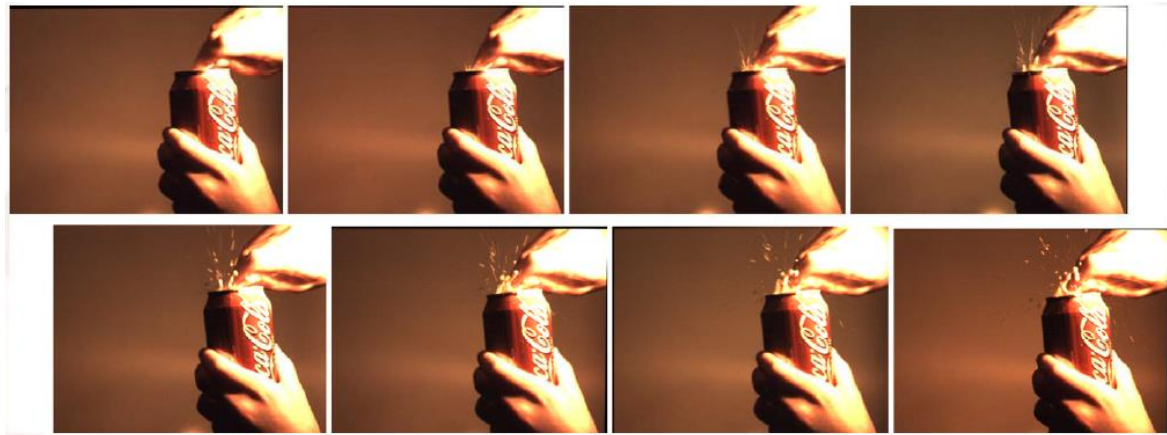
Υποβιβασμός ρυθμού πληροφορίας

Εκμετάλλευση χρονικού πλεονασμού (temporal redundancy)

Εκμετάλλευση χρονικού πλεονασμού

Video: συνεχόμενες εικόνες οι οποίες λαμβάνονται αρκετές φορές σε ένα δευτερόλεπτο.

- Οι εικόνες στις περισσότερες περιπτώσεις μοιάζουν (οπτικά) μεταξύ τους. Το στατικό κομμάτι παραμένει σε μεγάλο βαθμό σταθερό από εικόνα σε εικόνα.
- Οι διαφοροποιήσεις μεταξύ των εικόνων μπορεί να είναι περιορισμένες και αφορούν το δυναμικό κομμάτι (δηλαδή εκεί που υπάρχει κίνηση), αλλά και σε αυτό το κομμάτι οι αλλαγές στις εικόνες μπορεί να προέρχονται από μετακινήσεις αντικειμένων μεταξύ πλαισίων.



Εκμετάλλευση χρονικού πλεονασμού

Το να αποστέλλεται ολόκληρη η περιγραφή μιας εικόνας (ενός πλαισίου) οδηγεί στην επαναλαμβανόμενη αποστολή της ίδιας (πλεονάζουσας) πληροφορίας που θα μπορούσε να δημιουργηθεί εκ νέου στον δέκτη. **Το σκεπτικό του περιορισμού του χρονικού πλεονασμού είναι να αποστέλλεται μόνο η απαραίτητη πληροφορία, η πληροφορία, δηλαδή, που δεν θα μπορούσε να δημιουργηθεί εκ νέου στον δέκτη.**

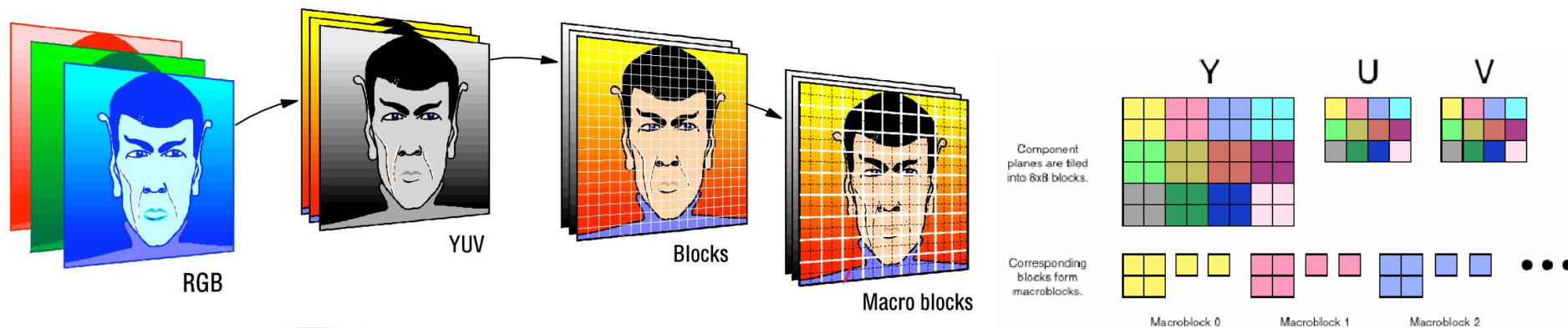
- Η πληροφορία αυτή σχετίζεται με τις περιοχές που υπάρχουν διαφοροποιήσεις σε σχέση με τα προηγούμενα πλαίσια.
- Ο μηχανισμός περιλαμβάνει τον διαχωρισμό της εικόνας σε υποπεριοχές και την κωδικοποίηση των εικόνων κατά **I** (intracoded frames), **P** (predicted frames) και **B** (bidirectional frames).

Εκμετάλλευση χρονικού πλεονασμού

- Μετατροπή από χώρο RGB σε YUV (βλ. την Ενότητα 1 του μαθήματος).
- Χωρισμός της εικόνας σε blocks 8x8 pixels
- Χωρισμός της εικόνας σε macroblocks 16x16 pixels
- 1 macroblock = 4 blocks

Κατά τον διαχωρισμό της εικόνας σε περιοχές, αναζητούνται οι στατικές και οι δυναμικές περιοχές. Η εικόνα χωρίζεται σε ίσες υποπεριοχές, και πιο συγκεκριμένα, σε τμήματα (**macro block**) των **16x16 εικονοστοιχείων (pixels)**.

Στις υποπεριοχές αυτές υπάρχουν 16x16 τιμές φωτεινότητας και σύμφωνα με το πρότυπο 4:2:0, οι τιμές που αντιστοιχούν στα χρώματα είναι σε πλήθος το $2/8=1/4$ αυτών της φωτεινότητας. Αντιστοιχούν, δηλαδή, σε μια **περιοχή 8x8 (block)** για κάθε μια από τις χρωμοδιαφορές Cb και Cr ενώ το σύνολο των τιμών αυτών αντιστοιχεί σε ένα μακρο-μπλοκ. Για πιο εύκολη διαχείριση, η ανάλυση της εικόνας αποτελείται από εικονοστοιχεία τέτοιου πλήθους ώστε να διαιρούνται με το 16 (π.χ. 720 x 576).



Εκμετάλλευση χρονικού πλεονασμού

Με την αποστολή εικόνων κατά I, P και B. αποφεύγουμε την αποστολή επαναλαμβανόμενης πληροφορίας, η οποία δεν είναι απαραίτητη. Αυτό επιτυγχάνεται ως εξής:

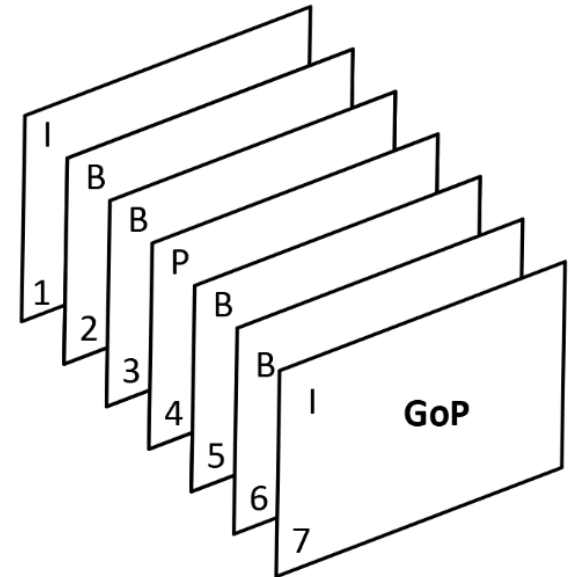
- Ορισμένα από τα πλαίσια κωδικοποιούνται ανεξάρτητα, και αυτούσια (χωρίς δηλαδή να λαμβάνεται υπόψη πληροφορία από άλλα πλαίσια). Τα πλαίσια αυτά ονομάζονται τύπου **I (I-frames, intracoded frames)**. Αυτού του είδους τα πλαίσια κωδικοποιούνται και αποστέλλονται περιοδικά.
- Κάποια πλαίσια κωδικοποιούνται λαμβάνοντας υπόψη προηγούμενα /προγενέστερα πλαίσια. Η κωδικοποίηση γίνεται με εκμετάλλευση των ομοιοτήτων που τυχόν υπάρχουν μεταξύ του τρέχοντος πλαισίου και κάποιου προηγούμενου. Οι ομοιότητες αναζητούνται και εντοπίζονται σε επίπεδο μακρομπλοκ, δηλαδή σε τμήματα μεγέθους 16x16. Αυτού του είδους τα πλαίσια ονομάζονται πλαίσια τύπου **P (P-frames, predicted frames** δηλαδή πλαίσια πρόβλεψης) τα οποία απαιτούν λιγότερα δεδομένα σε σχέση με την κωδικοποίηση πλαισίων τύπου I.
- Υπάρχουν πλαίσια τα οποία κωδικοποιούνται αξιοποιώντας τον εντοπισμό ομοιοτήτων με προηγούμενα αλλά και επόμενα πλαίσια. Εξαρτώνται, δηλαδή, από ένα προηγούμενο (χρονικά) πλαίσιο τύπου I ή P και ένα επόμενο (χρονικά) πλαίσιο ομοίως τύπου I ή P. Τα πλαίσια αυτού του τύπου ονομάζονται τύπου **B (B-frames, bidirectional frames** δηλαδή **πλαίσια αμφίδρομης πρόβλεψης**). Η κωδικοποίηση των πλαισίων τύπου B απαιτεί λιγότερα δεδομένα σε σχέση με την κωδικοποίηση πλαισίων τύπου P (και βέβαια πλαισίων τύπου I).

Εκμετάλλευση χρονικού πλεονασμού

Η αλληλουχία των πλαισίων 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**.

Παράδειγμα GoP συνολικά 7 πλαισίων.

- Αποτελείται από τα εξής πλαίσια: I, B, B, P, B, B και I.
- 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.

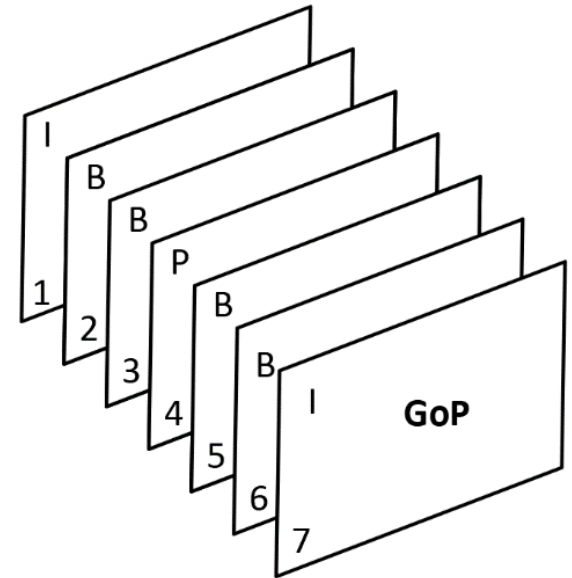


Κλειστή ομάδα πλαισίων (κλειστό GoP)

Εκμετάλλευση χρονικού πλεονασμού

Αλληλεξαρτήσεις στο παράδειγμα του GoP με 7 πλαίσια:

- 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, προηγείται της κωδικοποίησης των 2 και 3.
- Το πλαίσιο 4, ως τύπου P, εξαρτάται από την κωδικοποίηση του πλαισίου 1.
- 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.



Κλειστή ομάδα πλαισίων (κλειστό GoP)

Εκμετάλλευση χρονικού πλεονασμού

Η αναζήτηση γίνεται σε επίπεδο μακρο-μπλοκ. Η βασική σύγκριση αφορά τα μακρο-μπλοκ γειτονικών πλαισίων. Συγκεκριμένα, κάθε μακρο-μπλοκ σε μια εικόνα εξετάζεται ώστε να διαπιστωθεί:

- Εάν είναι όμοιο με το μακρο-μπλοκ στην αντίστοιχη θέση του προηγούμενου πλαισίου. Σε αυτή την περίπτωση, δεν απαιτείται το διάνυσμα κίνησης αλλά είναι δυνατόν να μεταδοθεί η διαφορά μεταξύ των τιμών των μακρο-μπλοκ.
- Εάν έχει προκύψει από μεταφορά αντίστοιχου μακρο-μπλοκ του προηγούμενου πλαισίου (το οποίο μπορεί να βρίσκεται σε διαφορετική θέση). Σε μια τέτοια περίπτωση η περιγραφή του μακρο-μπλοκ μπορεί να δοθεί με τη μετακίνηση του όμοιου του μακρο-μπλοκ το οποίο ανήκει στο προηγούμενο πλαίσιο, και γίνεται με την χρήση του **διανύσματος κίνησης (motion vector)**. Επιπλέον, επειδή ενδέχεται να υπάρχει διαφοροποίηση μεταξύ των δύο μακρο-μπλοκ, υπάρχει η δυνατότητα μεταφοράς και της διαφοράς των δύο μακρο-μπλοκ.
- Εάν είναι νέο και δεν υπάρχει στην προηγούμενη εικόνα/πλαίσιο. Σε αυτή την περίπτωση το μακρο-μπλοκ κωδικοποιείται εκ νέου.

Αξίζει να σημειωθεί ότι αφού μπορεί να έχουμε και πλαίσια τύπου B, τότε το μακρο-μπλοκ αναζητείται όχι μόνο στο προηγούμενο I ή P αλλά και στο επόμενο πλαίσιο.

Εκμετάλλευση χρονικού πλεονασμού

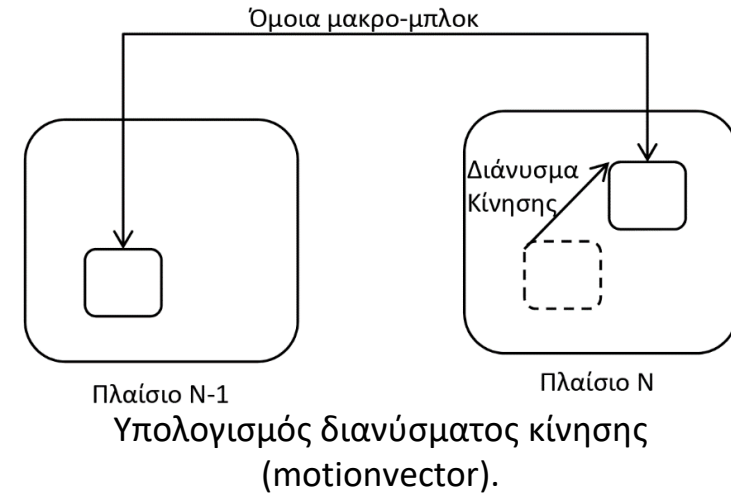
Για τον υπολογισμό του διανύσματος κίνησης, ο μηχανισμός ξεκινάει από το μακρο-μπλοκ του τρέχοντος πλαισίου.

- Εάν το τρέχον πλαίσιο είναι τύπου P, ελέγχει στο προηγούμενο P ή I πλαίσιο για παρόμοια μακρο-μπλοκ.
- Εάν το τρέχον πλαίσιο είναι τύπου B, ελέγχει στο προηγούμενο και στο επόμενο πλαίσιο για παρόμοια μακρο-μπλοκ.

Η αναζήτηση δεν γίνεται σε ολόκληρη την εικόνα αλλά πέριξ της περιοχής του αρχικού μακρο-μπλοκ (μέσω του ταιριάσματος σε επίπεδο μακρο-μπλοκ) προκειμένου να ελαττωθεί η πολυπλοκότητα και το χρονικό διάστημα που απαιτείται για την κωδικοποίηση.

- 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.
- Επίσης υπολογίζεται η διαφορά μεταξύ των δύο μακρο-μπλοκ.

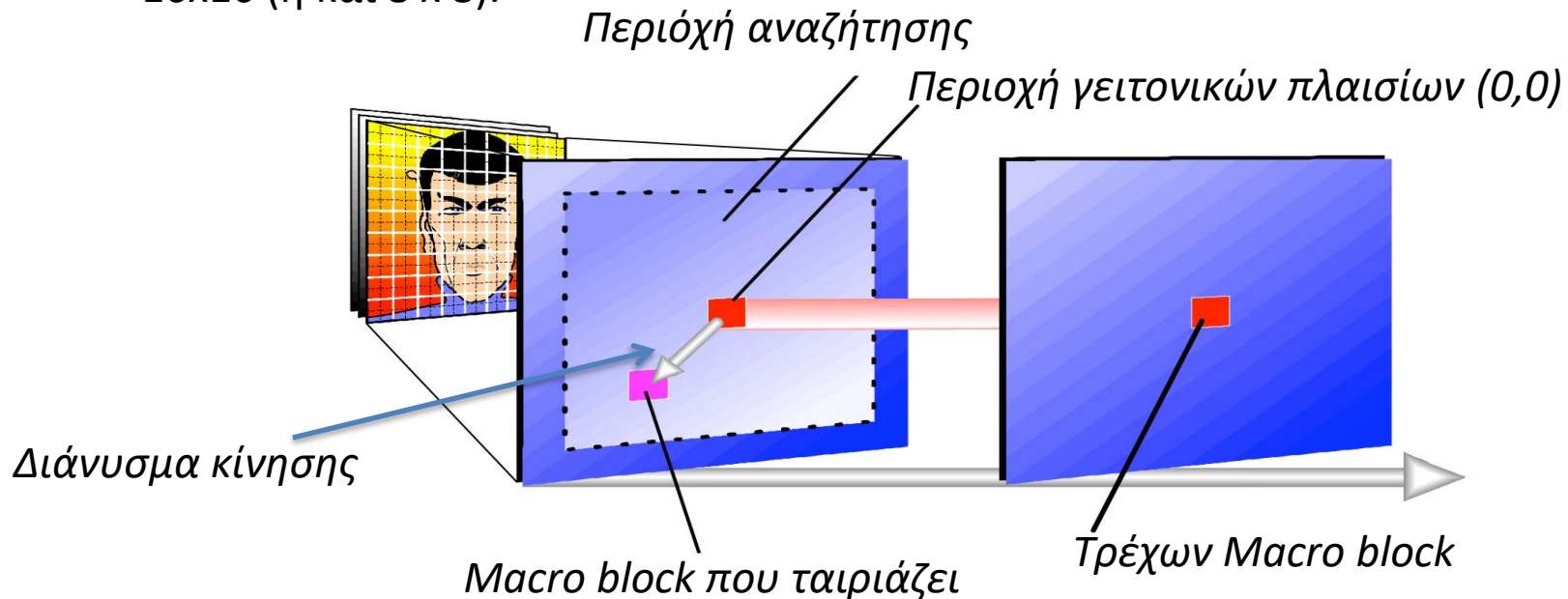
Η διαδικασία στην περίπτωση του πλαισίου τύπου B γίνεται εξετάζοντας τα πλαίσια και προς τις δύο κατευθύνσεις. Παράλληλα με το διάνυσμα κίνησης, μεταδίδονται επιπλέον απαραίτητες πληροφορίες για να καλυφθούν οι διαφοροποιήσεις μεταξύ των δύο πλαισίων.



Εκμετάλλευση χρονικού πλεονασμού

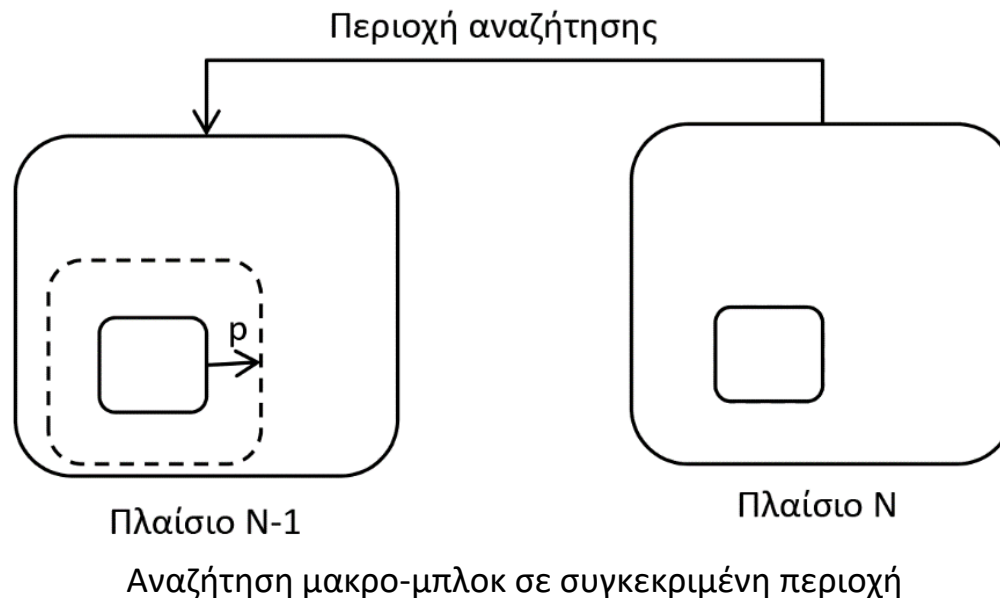
Εκτός από την αναζήτηση όμοιων περιοχών σε επίπεδο μακρο-μπλοκ, μπορεί να έχουμε αναζήτηση σε επίπεδο:

- Εικονοστοιχείου (pixel) με μεγάλο όμως πλήθος υπολογισμών, γεγονός το οποίο την καθιστά ασύμφορη.
- Αντικειμένων εντός της εικόνας (object-based), όπου τα αντικείμενα αναγνωρίζονται αυτόματα εντός της εικόνας και παρακολουθούνται στα διάφορα πλαίσια.
- Τμήματος της εικόνας (block-based), όπου οι τυπικές διαστάσεις του μπλοκ είναι 16x16 (ή και 8 x 8).



Εκμετάλλευση χρονικού πλεονασμού

Η εκτίμηση κίνησης είναι το πιο χρονοβόρο κομμάτι της συμπίεσης. Εάν υποθέσουμε ανάλυση 720x480, το πλήθος των αναζητήσεων πολύ μεγάλο και απαιτείται να περιορίσουμε την αναζήτηση σε συγκεκριμένες περιοχές (συνήθως η περιοχή αναζήτησης εκτείνεται από την τρέχουσα θέση συν p pixels (π.χ. $p=7$). Η τεχνική αυτή παρουσιάζεται στην διπλανή εικόνα.



Εκμετάλλευση χρονικού πλεονασμού

Για την εύρεση του καλύτερου ταιριάσματος χρησιμοποιούνται συνήθως δύο συναρτήσεις:

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

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

$$MSE = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (C_{ij} - R_{ij})^2 \quad (3)$$

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

Τα πλαίσια τύπου **B** βρίσκονται μεταξύ **I** και **P**. Για την αποκωδικοποίηση ενός πλαισίου τύπου **B** απαιτείται η αποκωδικοποίηση των πλαισίων **I** και **P** που το περιβάλλουν. Για το λόγο αυτό δύναται να τροποποιείται κατάλληλα η αλληλουχία των πλαισίων για την κωδικοποίηση, μετάδοση και αποκωδικοποίηση.

Εκμετάλλευση χρονικού πλεονασμού

Παράδειγμα διαχείρισης 14 πλαισίων, για GoP αποτελούμενο από 7 πλαίσια της ακόλουθης μορφής: IBBPBBP.

- Η μνήμη που απαιτείται στον δέκτη στην περίπτωση αυτή είναι 2 θέσεις.
- 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.

	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

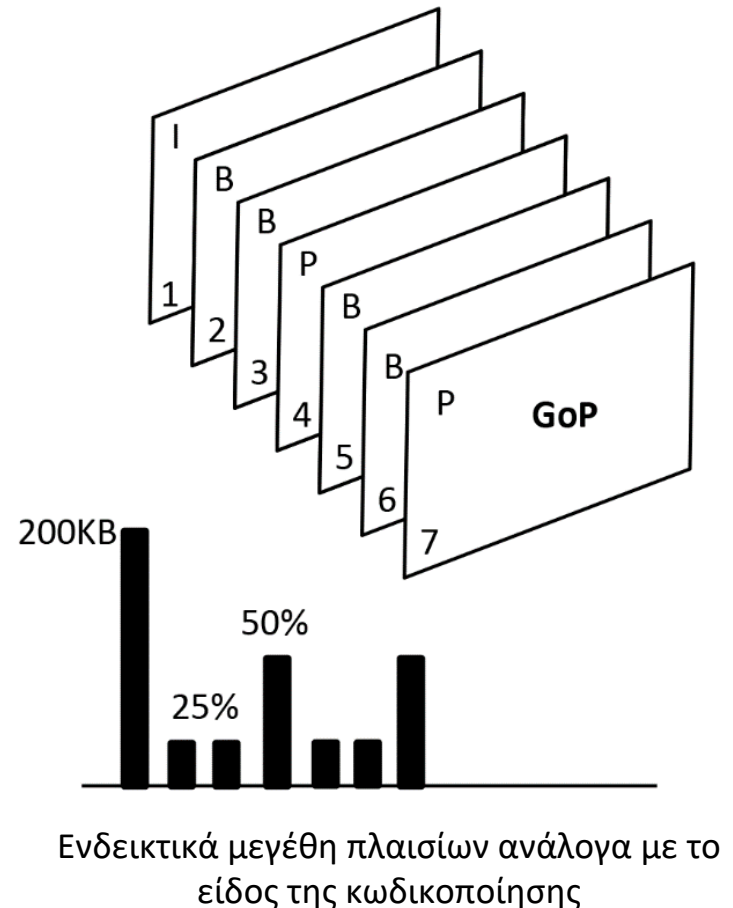
Αλληλουχία πλαισίων σε ένα GoP (7 πλαισίων) κατά τις διάφορες λειτουργίες

Εκμετάλλευση χρονικού πλεονασμού

Όφελος από την εκμετάλλευση του χρονικού πλεονασμού.

- Το μέγεθος ενός πλαισίου που κωδικοποιείται ως I είναι το μεγαλύτερο σε σχέση με τα μεγέθη των P και B.
- Το μέγεθος ενός πλαισίου τύπου P μπορεί να είναι στο ήμισυ αυτού του I,
- Το μέγεθος ενός πλαισίου τύπου B δύναται να φτάσει και στο 25% του I.

Αυτές, βέβαια, είναι ενδεικτικές τιμές, αφού οι αναλογίες είναι εν γένει παραμετροποιήσιμες.



Εκμετάλλευση χρονικού πλεονασμού

Παράδειγμα – 1

Θεωρούμε ένα κωδικοποιημένο video κατά MPEG-2. Το frame rate είναι 25 καρέ/sec. Έστω ότι ένα I-frame έχει μέγεθος 100 Kbytes, ένα P-frame έχει μέγεθος 30 Kbytes και ένα B-frame έχει μέγεθος 20 Kbytes. Θεωρούμε επίσης GOP size=12 (IBBPBBPBBPBB).

- Να υπολογιστεί το bit rate (σε Mbits/sec) του κωδικοποιημένου video
- Αν χρησιμοποιούσαμε GOP size = 6 (IBBPBB), ποιο είναι το bit rate που απαιτείται; Τι κερδίζουμε και τι χάνουμε σε αυτή την περίπτωση;

Απάντηση

- Κάθε GOP χρειάζεται $1 \times 100 + 3 \times 30 + 8 \times 20 = 350$ Kbytes. Κάθε GOP περιέχει 12 frames, οπότε σε 1 sec (25 frames) μεταδίδονται $350 \times (25/12) = 729.1$ Kbytes. Συνεπώς ο ρυθμός μετάδοσης είναι **5.832 Mbits/sec**.
- ii) Κάθε GOP χρειάζεται $1 \times 100 + 1 \times 30 + 4 \times 20 = 210$ Kbytes. Κάθε GOP περιέχει 6 frames, οπότε σε 1 sec (25 frames) μεταδίδονται $210 \times (25/6) = 875$ Kbytes. Συνεπώς ο ρυθμός μετάδοσης είναι **7.0 Mbits/sec**.

Συμπέρασμα: Χρησιμοποιώντας μικρότερο GOP size χάνουμε σε απόδοση κωδικοποίησης, αφού απαιτείται μεγαλύτερο bitrate για τη μετάδοση. Το κέρδος έγκειται στη συχνότερη μετάδοση ανεξάρτητων I-frames (1 σε κάθε 6 frames για GOP=6 αντί 1 σε κάθε 12 frames για GOP=12), κάτι που επιτρέπει στον αποκωδικοποιητή να συνεχίσει την αναπαραγωγή πιο γρήγορα σε περίπτωση λάθους ή απώλειας δεδομένων.

Εκμετάλλευση χρονικού πλεονασμού

Παράδειγμα – 2

Έστω το παρακάτω GOP:

... I B B P B B P B B P B B ...

Αν χαθεί ή παραμορφωθεί το υπογραμμισμένο P frame λόγω λάθους στη μετάδοση, ποια frames είναι πιθανόν να έχουν επίσης πρόβλημα;

Απάντηση

Το I frame είναι ανεξάρτητο και τα δύο επόμενα P frames εξαρτώνται μόνο από αυτό και τα προηγούμενά τους P. Αναφορά στο προβληματικό P frame ενδέχεται να έχουν μόνο τα B frames του GOP, οπότε μόνο αυτά είναι πιθανόν να έχουν επίσης πρόβλημα.

Εκμετάλλευση χωρικού πλεονασμού (spatial redundancy)

Χωρικός πλεονασμός

Ο χωρικός πλεονασμός είναι ένα αρκετά συχνό φαινόμενο που σχετίζεται και αφορά στις ομοιότητες που υπάρχουν σε διαφορετικές περιοχές εντός μιας εικόνας. Η εμφάνιση τέτοιων ομοιοτήτων είναι αρκετά συχνό φαινόμενο και η εκμετάλλευση του χωρικού πλεονασμού γίνεται με τα πρότυπα συμπίεσης της στατικής εικόνας, όπως το JPEG (Joint Pictures Expert Group).

- Στόχος είναι η **αφαίρεση πληροφορίας** η οποία δεν είναι τόσο σημαντική μεταξύ γειτονικών εικονοστοιχείων.
- Πραγματοποιείται με **κωδικοποίηση μετασχηματισμού**, ο οποίος μετασχηματίζει την πληροφορία από το αρχικό πεδίο σε ένα ενδιάμεσο πεδίο (κατάλληλο για συμπίεση) στο οποίο επιλέγονται και κωδικοποιούνται εκείνοι οι συντελεστές του μετασχηματισμού που περιέχουν τη μεγάλη ενέργεια (δηλαδή οι σημαντικοί) – αγνοούνται ή κωδικοποιούνται με μικρότερη ακρίβεια οι υπόλοιποι συντελεστές.
- Το αρχικό σήμα ανακτάται (**με απώλειες**) μέσω του αντιστρόφου μετασχηματισμού.

Χωρικός πλεονασμός

Ο βασικός αλγόριθμος που χρησιμοποιείται στο JPEG είναι ο **διακριτός μετασχηματισμός συνημιτόνου** (Discrete Cosine Transform, DCT).

Αυτός ο τρόπος συμπίεσης εκμεταλλεύεται το γεγονός ότι το ανθρώπινο μάτι δεν έχει μεγάλη ευαισθησία στις υψηλές χωρικές συχνότητες. Πιο συγκεκριμένα

- δεν αντιλαμβάνεται με μεγάλη ακρίβεια τις απότομες διαφοροποιήσεις σε μια εικόνα, δηλαδή αλλαγές που πραγματοποιούνται σε μικρό σχετικά αριθμό εικονοστοιχείων και αντιστοιχούν σε υψηλές χωρικές συχνότητες (spatial frequencies).
- αντιλαμβάνεται επακριβώς τις λιγότερο απότομες διαφοροποιήσεις εντός της εικόνας, δηλαδή αλλαγές που γίνονται σε μεγαλύτερες επιφάνειες εικονοστοιχείων, και οι οποίες αντιστοιχούν σε χαμηλότερες χωρικές συχνότητες.

Οι πιο χαμηλές (χωρικές) συχνότητες κωδικοποιούνται με μεγαλύτερη λεπτομέρεια (κυρίως όσον αφορά την κβάντιση), ενώ οι πιο υψηλές συχνότητες κωδικοποιούνται με λιγότερη λεπτομέρεια, πράγμα που οδηγεί σε 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:** μετασχηματισμός του σήματος στο πεδίο της συχνότητας.

Διακριτός μετασχηματισμός συνημιτόνου - ΔΜΣ

Ο διακριτός μετασχηματισμός συνημιτόνου (ΔΜΣ) αποτελεί ειδική μορφή του διακριτού μετασχηματισμού Φουριέ (ΔΜΦ).

- Πλήθος τιμών μετασχηματίζεται σε σύνολο τιμών ίδιου πλήθους στο πεδίο των συχνοτήτων.
- Μπορεί να χρησιμοποιηθεί είτε μια διάσταση (**μονοδιάστατος ΔΜΣ**), όπου η ανεξάρτητη μεταβλητή είναι συνήθως ο χρόνος, είτε σε δύο διαστάσεις (**δισδιάστατος ΔΜΣ**), όπου οι ανεξάρτητες μεταβλητές είναι το επίπεδο (μήκος και πλάτος).

Μεγάλο πλεονέκτημα στη χρήση του ΔΜΣ έναντι άλλων αλγορίθμων μετασχηματισμού (π.χ. έναντι του ΔΜΦ) είναι η ανθεκτικότητα που επιδεικνύει στην παράλειψη κάποιων από τους συντελεστές που προκύπτουν όταν ο ΔΜΣ εφαρμοστεί στα δεδομένα.

- 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 μοιάζουν σε μεγαλύτερο βαθμό με τα αρχικά σε σχέση με την αντίστοιχη περίπτωση του ΔΜΦ (παράλειψη ίδιους πλήθους συντελεστών).

Μονοδιάστατος ΔΜΣ

Έστω ότι το αρχικό μας σήμα είναι το $x(n)$ και η έξοδος του ΔΜΣ είναι το $y(n)$. Το $y(n)$ δίνεται από την επόμενη σχέση:

$$y(n) = w(n) \sum_{n=1}^N x(n) \cos \frac{\pi(2n-1)(k-1)}{2N} \quad k=1, 2, \dots, N$$

$$w(k) = \begin{cases} \frac{1}{\sqrt{N}}, & k=1 \\ \sqrt{\frac{2}{N}}, & 2 \leq k \leq N \end{cases} \quad (4)$$

Παρατηρούμε ότι N είναι το πλήθος των δεδομένων που πρόκειται να μετασχηματιστούν. Το πλήθος αυτό παραμένει, όπως προαναφέρθηκε, το ίδιο και στις δύο περιπτώσεις.

Μονοδιάστατος ΔΜΣ

Παράδειγμα – 3

Ως παράδειγμα ας εξετάσουμε ένα σύνολο ακεραίων x , από το 1 έως και το 10, το οποίο μετασχηματίζεται κατά ΔΜΣ. Το αποτέλεσμα αυτού του μετασχηματισμού ΔΜΣ είναι:

$$y = [17.3925, -9.0249, 0, -0.9667, 0, -0.3162, 0, -0.1279, 0, -0.0359]$$

Παρατηρούμε ότι

- τα πρώτα (στη σειρά) στοιχεία έχουν υψηλές τιμές (κατ' απόλυτο).
- οι τιμές αυτές μειώνονται (κατ' απόλυτη τιμή) στα επόμενα στοιχεία (χωρίς απαραίτητα να συγκροτούν φθίνουσα ακολουθία).
- οι τιμές δεν ανήκουν στο πεδίο του χρόνου αλλά της συχνότητας.

Σύμφωνα με αυτά που είδαμε προηγουμένως, **οι πρώτες τιμές αφορούν τις χαμηλές συχνότητες, ενώ οι επόμενες σε υψηλότερες.** Η πρώτη τιμή αντιστοιχεί στη μηδενική συχνότητα.

Μονοδιάστατος ΔΜΣ

Εάν θεωρήσουμε ότι το σήμα το οποίο μετασχηματίζεται αντιστοιχεί σε σήμα εικόνας, αποτελείται δηλαδή από ένα πίνακα δύο διαστάσεων, τότε ο ΔΜΣ θα αποτελείται από πίνακα των ίδιων διαστάσεων.

Οι θέσεις του πίνακα αντιστοιχούν ομοίως σε συχνότητες, και στη συγκεκριμένη περίπτωση σε χωρικές συχνότητες (spatial frequencies). Μετασχηματίζοντας δηλαδή κατά ΔΜΣ, μπορούμε να επεξεργαστούμε τους συντελεστές ανάλογα με τις χωρικές συχνότητες στις οποίες αντιστοιχούν. Αυτό μπορεί να γίνει λαμβάνοντας υπόψη τα επίπεδα ευαισθησίας του ανθρώπινου ματιού στις διάφορες συχνότητες.

Έτσι, μετά τον μετασχηματισμό οι συντελεστές διαιρούνται με συγκεκριμένους συντελεστές κβαντοποίησης (οι οποίοι προκύπτουν από την ευαισθησία του ανθρώπινου οφθαλμού στις διάφορες συχνότητες).

- Όσο μεγαλύτερες είναι οι τιμές των συντελεστών κβαντοποίησης (δηλαδή του διαιρέτη) τόσο μειώνονται οι τιμές των συντελεστών προς κβαντοποίηση.
- Σε αρκετές περιπτώσεις οι συντελεστές αυτοί μηδενίζονται (ειδικά καθώς αυξάνουν οι συχνότητες).
- Οι μηδενικοί συντελεστές κωδικοποιούνται με ελάχιστες απαιτήσεις όσον αφορά το πλήθος των bits που χρησιμοποιούνται.

Μονοδιάστατος ΔΜΣ

Είναι προφανές ότι αυτό το είδος της συμπίεσης είναι **απωλεστικό (lossy)**, αφού στον αποκωδικοποιητή δεν μπορούμε να πάρουμε την ίδια εικόνα με την αρχική.

Τα βήματα που ακολουθούνται είναι:

- μετασχηματισμός κατά ΔΜΣ των αρχικών τιμών
- προσαρμογή τιμών με χρήση συντελεστών κβαντισμού
- κβαντισμός των προσαρμοσμένων συντελεστών

Δισδιάστατος ΔΜΣ

Εάν το σήμα το οποίο μετασχηματίζεται αντιστοιχεί σε σήμα εικόνας (αποτελείται δηλαδή από ένα πίνακα δύο διαστάσεων) τότε ο ΔΜΣ θα αποτελείται από πίνακα των ίδιων διαστάσεων. Στην περίπτωση αυτή (δηλ. μετασχηματισμό πινάκων δύο διαστάσεων), η κατάσταση είναι διαφορετική και χρησιμοποιείται ο δισδιάστατος ΔΜΣ.

Στις δύο διαστάσεις οι σχέσεις του μετασχηματισμού ΔΜΣ είναι οι εξής:

$$y(p, q) = a(p)a(q) \sum_{m=1}^M \sum_{n=1}^N x(m, n) \cos \frac{\pi(2m-1)(p-1)}{2M} \cos \frac{\pi(2n-1)(q-1)}{2N} \quad \begin{matrix} p = 1, 2, \dots, M \\ q = 1, 2, \dots, N \end{matrix} \quad (5)$$

όπου $a(p) = \begin{cases} \frac{1}{\sqrt{M}}, & p = 1 \\ \sqrt{\frac{2}{M}}, & 2 \leq p \leq M \end{cases}$ και $a(q) = \begin{cases} \frac{1}{\sqrt{N}}, & q = 1 \\ \sqrt{\frac{2}{N}}, & 2 \leq q \leq N \end{cases}$

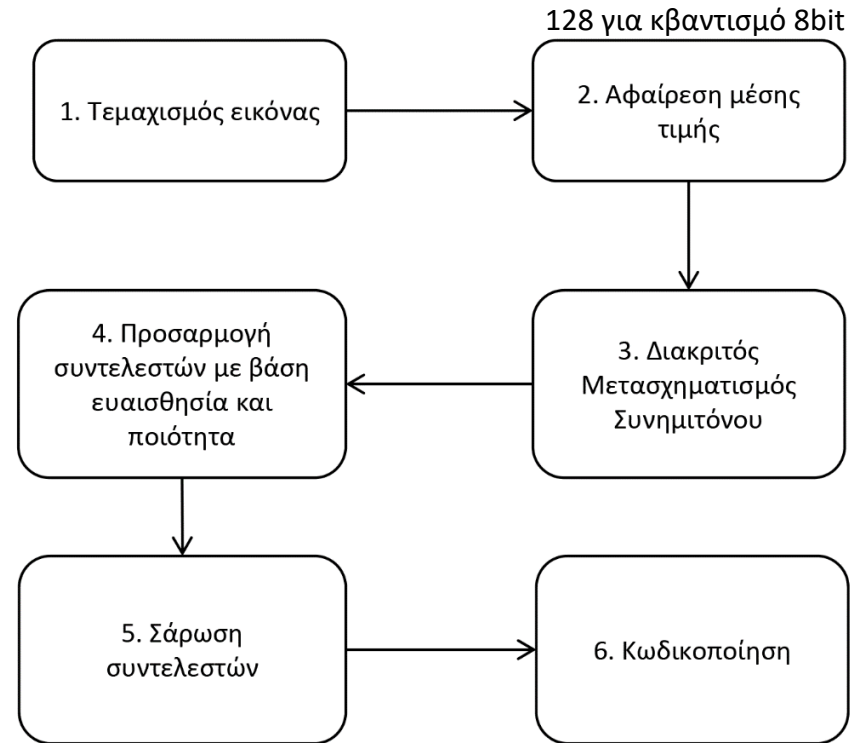
Όπως και στην περίπτωση του μονοδιάστατου μετασχηματισμού ΔΜΣ ο υπολογισμός μπορεί να γίνει με πολλαπλασιασμό πινάκων.

Χρήση ΔΜΣ

Τα βήματα που ακολουθούμε είναι τα εξής:

- τεμαχισμός εικόνας σε τμήματα (μπλοκ) 8x8
- αφαίρεση μέσης τιμής (του 128 στην περίπτωση που ο κβαντισμός θα γίνει με τη βοήθεια 8 bits)
- μετασχηματισμός κατά ΔΜΣ σε κάθε μπλοκ (οι συντελεστές που προκύπτουν ανήκουν στο πεδίο της συχνότητας – χωρικές συχνότητες)
- προσαρμογή συντελεστών με βάση την ευαισθησία του ανθρώπινου οφθαλμού αλλά και του επιθυμητού επιπέδου ποιότητας
- σάρωση συντελεστών με ομαδοποίηση τιμών με την 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)
- Κωδικοποίηση/Κβαντισμός

Η προσαρμογή των συντελεστών γίνεται με βάση έτοιμους πίνακες (MPEG), που μπορούν να προσαρμοστούν ή να τροποποιηθούν από τους κωδικοποιητές ανάλογα με την επιθυμητή ποιότητα και τον βαθμό συμπίεσης. Ο αποκωδικοποιητής λαμβάνει πληροφορίες για αυτούς τους πίνακες καθώς αυτές μεταδίδονται.



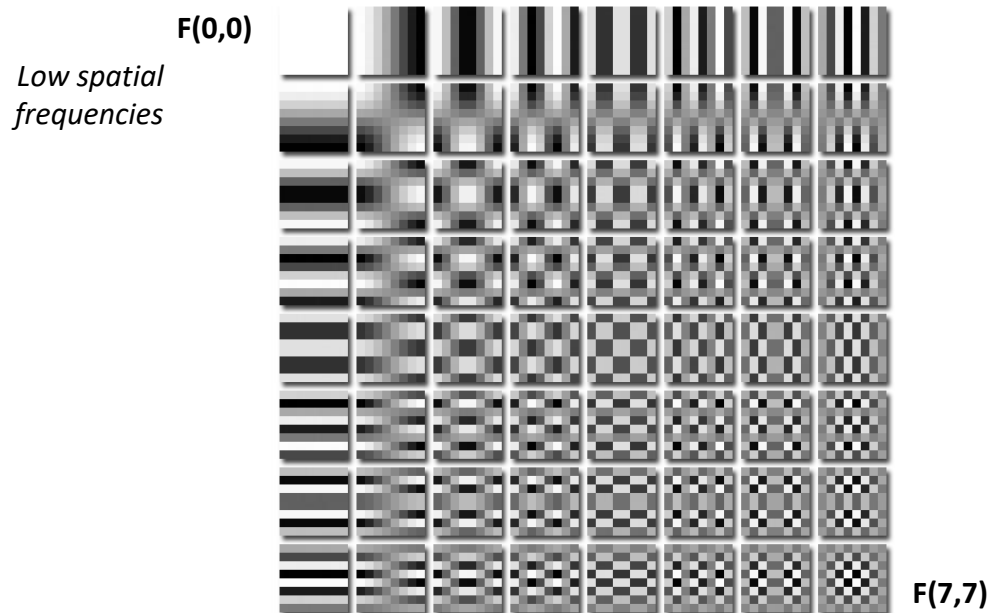
Χρήση του μετασχηματισμού ΔΜΣ στο JPEG.

Χρήση ΔΜΣ

Αναμενόμενα αποτελέσματα του δισδιάστατου ΔΜΣ:

- Το αποτέλεσμα του ΔΜΣ και ο βαθμός συμπίεσης που επιτυγχάνεται εξαρτώνται από την οριζόντια, την κατακόρυφη και τη διαγώνια χωρική συχνότητα (spatial frequencies).
- Για μια εικόνα με ένα μόνο χρώμα, ο ΔΜΣ θα δώσει έναν πίνακα με μεγάλη τιμή για τον πρώτο όρο και μηδενικές για τις υπόλοιπες.
- Για εικόνα με πολλές εναλλαγές ο ΔΜΣ θα δώσει πίνακα με διαφορετικές τιμές.

Στις επόμενες διαφάνειες παρουσιάζονται οι πρακτικές και οι μέθοδοι υπολογισμού του δισδιάστατου ΔΜΣ με χρήση πινάκων.



Οι 64 βασικές χωρικές συχνότητες. Κάθε block, μέσω του μετασχηματισμού ΔΜΣ παρίσταται σαν γραμμικός συνδυασμός των συχνοτήτων αυτών.

High spatial frequencies

F(7,7)

Πρακτικές υπολογισμού

Για ακόμη μεγαλύτερη εξοικονόμηση bits, οι συντελεστές ΔΜΣ δεν μεταδίδονται ως έχουν, αλλά **κβαντίζονται**.

Συγκεκριμένα, γίνονται τα εξής:

1. Πρώτον, κάθε συντελεστής διαιρείται με το αντίστοιχο στοιχείο ενός **Πίνακα Κβάντισης (Quantizer Matrix)** που είναι σταθερός και προκαθορισμένος (περιγράφεται στο πρότυπο MPEG-2).
2. Δεύτερον, το αποτέλεσμα διαιρείται με έναν σταθερό αριθμό που λέγεται **Κλίμακα Κβάντισης (Quantizer Scale)**. Αυτόν τον επιλέγουμε κατά την κωδικοποίηση.
3. Μεταδίδεται τελικά **το τελικό πηλίκο των δύο διαδοχικών διαιρέσεων** (η στρογγυλοποίησή του σε ακέραιο)

Πρακτικές υπολογισμού

Στον δισδιάστατο αποδεικνύεται πως ο μετασχηματισμός ΔΜΣ ενός πίνακα 8x8 ισοδυναμεί με τον πολλαπλασιασμό του αρχικού πίνακα από αριστερά με τον πίνακα των συντελεστών (των συναρτήσεων, δηλαδή, βάσης) έστω T και από δεξιά με τον αντίστροφο του T .

- Με αυτόν τον μετασχηματισμό ΔΜΣ μπορούμε να εξοικονομήσουμε bits γιατί η πληροφορία συνήθως συγκεντρώνεται σε λίγους συντελεστές, κυρίως σε αυτούς που αντιστοιχούν στις χαμηλές χωρικές συχνότητες.
- Οι μηδενικοί συντελεστές **δεν μεταδίδονται**.



Block (8x8)

40	38	45	40	43	54	60	58
39	36	44	32	47	69	77	85
50	40	25	54	66	60	33	32
57	38	38	66	47	11	2	5
59	36	47	62	24	2	9	11
58	41	55	53	6	4	10	1
58	33	57	39	3	5	4	2
64	44	54	35	3	7	3	3

Τιμές των pixel

44	-5	0	-4	0	-1	0	0
12	0	-3	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	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Συντελεστές ΔΜΣ

Διακριτός μετασχηματισμός συνημιτόνου

Πρακτικές υπολογισμού

Στη συνέχεια γίνεται επιλογή των συντελεστών στο πεδίο της συχνότητας (μείωση του πλήθους των συντελεστών).

- Αυτό που θέλουμε να πετύχουμε είναι η μείωση του πλήθους των συντελεστών (με την ελάττωση των τιμών αυτών).

Αντί για έναν ομοιόμορφο κβαντισμό των συντελεστών στο πεδίο της συχνότητας, χρησιμοποιούνται προτυποποιημένοι **Πίνακες Κβάντισης (Quantizer Matrix)** οι οποίοι βασίζονται σε πειράματα στο ανθρώπινο οπτικό σύστημα αναφορικά με την ευαισθησία του στις διάφορες χωρικές συχνότητες (οριζόντια και κατακόρυφα). Οι συντελεστής/στοιχεία του Πίνακα Κβάντισης είναι σταθεροί και προκαθορισμένοι (περιγράφονται στο πρότυπο MPEG-2).

Πρακτικές υπολογισμού

Στη συνέχεια το αποτέλεσμα διαιρείται με έναν σταθερό αριθμό που λέγεται **Κλίμακα Κβάντισης (Quantizer Scale)**. Αυτόν τον επιλέγουμε κατά την κωδικοποίηση.

Δεδομένου ότι υπάρχουν διαφορετικές απαιτήσεις για τον κβαντισμό και τον συνακόλουθο βαθμό συμπίεσης, απαιτείται κάποιος μηχανισμός ώστε να μπορεί να προσαρμόζεται ο πίνακας προσαρμογής.

- Διακρίνουμε τις εξής δυο περιπτώσεις:
 - Υψηλότερη ποιότητα, οπότε πρέπει να διατηρηθούν περισσότεροι συντελεστές και οι τιμές του πίνακα προσαρμογής του κβαντισμού πρέπει να μειωθούν.
 - Χαμηλότερη ποιότητα, οπότε πρέπει να διατηρηθούν λιγότεροι συντελεστές και οι τιμές του πίνακα προσαρμογής του κβαντισμού πρέπει να αυξηθούν.
- Τέλος, μεταδίδεται τελικά **το τελικό πηλίκο των δύο διαδοχικών διαιρέσεων** (η στρογγυλοποίησή του σε ακέραιο).

Πρακτικές υπολογισμού

7842	199	448	362	342	112	31	22
198	151	181	264	59	37	14	3
142	291	218	87	27	88	27	12
111	133	159	119	58	65	36	2
49	85	217	50	8	3	14	12
58	120	60	40	41	11	2	1
30	121	61	22	30	1	0	1
22	28	2	33	24	51	44	81

Αρχικοί Συντελεστές ΔΜΣ

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	48	56	69
27	29	35	38	46	56	69	83

Πίνακας Κβάντισης

980	12	23	16	13	4	1	0
12	9	8	11	2	1	0	0
7	13	8	3	0	2	0	1
5	6	6	4	2	1	0	0
2	3	8	1	0	0	0	0
2	4	2	1	1	0	0	0
1	4	2	1	0	0	0	0
0	0	1	0	0	0	0	0

Συντελεστές που μεταδίδονται

Code	Linear Quant Scale	Non-Linear Quant Scale
1	2	1
8	16	8
16	32	24
20	40	40
24	48	56
28	56	88
31	62	112

Κλίμακα Κβάντισης

Πρακτικές υπολογισμού

Παράδειγμα – 4

Δίνονται οι τιμές των συντελεστών ΔΜΣ ενός block MPEG-2 που έχουν προκύψει μετά από τον μετασχηματισμό ΔΜΣ (Πίνακας Α), καθώς και ο τυποποιημένος πίνακας κβάντισης (Πίνακας Β). Να υπολογίσετε τους συντελεστές ΔΜΣ που τελικά μεταδίδονται χρησιμοποιώντας κλίμακα κβάντισης $q=2$.

Απάντηση

Ο πίνακας των συντελεστών που μεταδίδονται προκύπτει από τη διαίρεση των αρχικών συντελεστών με το αντίστοιχο στοιχείο του πίνακα κβάντισης και στη συνέχεια με την κλίμακα κβάντισης. Το αποτέλεσμα στρογγυλοποιείται στον πλησιέστερο ακέραιο (βλ. Πίνακα Γ).

99	66	32	9	0	0	0	0
54	21	9	5	0	0	0	0
13	1	0	0	0	0	0	0
6	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	0
0	0	0	0	0	0	0	0

Πίνακας Α

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Πίνακας Β

6	2	1	0	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
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Πίνακας Γ

Προσαρμογή Επιπέδου Ποιότητας

- Όσο αυξάνουμε την κλίμακα κβάντισης, τόσο περισσότεροι είναι οι συντελεστές ΔΜΣ που τελικά μηδενίζονται και α μη β δεν μεταδίδονται. Έτσι, εξοικονομούμε bits αλλά χάνουμε σε ποιότητα εικόνας καθώς χάνονται χωρικές συχνότητες.
- Τα block παραμορφώνονται, χάνουν τη λεπτομέρειά τους αλλά και τη συνέχειά τους. Γίνονται δηλαδή ορατά τα όρια των blocks, κάτι που είναι γνωστό ως **blockiness effect**.

Χαμηλή
κλίμακα
κβάντισης



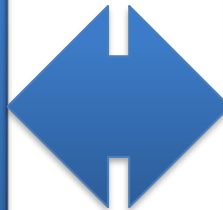
Υψηλή
κλίμακα
κβάντισης

Προσαρμογή Επιπέδου Ποιότητας

Η κλίμακα κβάντισης είναι το κλειδί που ορίζει την ποιότητα αλλά και το bit rate (σε Mbits/s) του κωδικοποιημένου κατά MPEG video.

Χαμηλή Κλίμακα Κβάντισης

1. Υψηλή ποιότητα
2. Δεν χάνονται χωρικές συχνότητες και δεν φαίνονται οι ατέλειες της κωδικοποίησης
3. No blockiness effect
4. Μεγάλος αριθμός bit ανά block
5. Υψηλό bit rate στην έξοδο



Υψηλή Κλίμακα Κβάντισης

1. Χαμηλή ποιότητα
2. Χάνονται χωρικές συχνότητες και φαίνονται οι ατέλειες της κωδικοποίησης
3. Blockiness effect
4. Μικρός αριθμός bit ανά block
5. Χαμηλό bit rate στην έξοδο

Συνήθως στον κωδικοποιητή ορίζουμε το επιθυμητό bit rate και αυτός αυτόματα προσαρμόζει την κλίμακα κβάντισης (μπορεί να αλλάζει σε κάθε καρέ ή και σε κάθε macroblock) προκειμένου να πετύχει στην έξοδο το bit rate που έχουμε ορίσει.

Προσαρμογή Επιπέδου Ποιότητας

Πίνακας κλιμάκωσης (Quantizer Scale) για ποιότητα 50% (Q_{50})

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

Προσαρμογή Επιπέδου Ποιότητας

- Για επιθυμητό επίπεδο ποιότητας ανώτερο του 50, ο πίνακας προσαρμογής της κβαντισμού (Quantizer Scale) προκύπτει πολλαπλασιάζοντας τον αρχικό πίνακα Q_{50} με έναν συντελεστή μικρότερο της μονάδας:

$$Q_{50} * \frac{100 - \text{ΕΕΠ}}{50} \quad (6)$$

- Για επίπεδο ποιότητας κατώτερο του 50, ο πίνακας προσαρμογής κβαντισμού (Quantizer Scale) προκύπτει πολλαπλασιάζοντας τον αρχικό πίνακα Q_{50} με συντελεστή μεγαλύτερο της μονάδας, ως εξής:

$$Q_{50} * \frac{50}{\text{ΕΕΠ}} \quad (7)$$

Όπου ΕΕΠ θεωρούμε το Επιθυμητό Επίπεδο Ποιότητας

Προσαρμογή Επιπέδου Ποιότητας

Με την μέθοδο που περιγράψαμε, ο πίνακας προσαρμογής της κβαντισμού (Quantizer Scale) για **Επιθυμητό Επίπεδο Ποιότητας 90 (Q_{90})** είναι ο εξής:

3	2	2	3	5	8	10	12
2	2	3	4	5	12	12	11
3	3	3	5	8	11	14	11
3	3	4	6	10	17	16	12
4	4	7	11	14	22	21	15
5	7	11	13	16	12	23	18
10	13	16	17	21	24	24	21
14	18	19	20	22	20	20	20

Προσαρμογή Επιπέδου Ποιότητας

Με την μέθοδο που περιγράψαμε, ο πίνακας προσαρμογής της κβαντισμού (Quantizer Scale) για **Επιθυμητό Επίπεδο Ποιότητας 10 (Q_{10})** είναι ο εξής:

80	60	50	80	120	200	255	255
55	60	70	95	130	255	255	255
70	65	80	120	200	255	255	255
70	85	110	145	255	255	255	255
90	110	185	255	255	255	255	255
120	175	255	255	255	255	255	255
245	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Προσαρμογή Επιπέδου Ποιότητας

Όλα εξαρτώνται από το περιεχόμενο

Δεν υπάρχει γενικός κανόνας, δηλ. « X Mbits/s αντιστοιχούν σε καλή ποιότητα και Y Mbits/s αντιστοιχούν σε κακή ποιότητα».

- **Γρήγορες σκηνές** απαιτούν περισσότερα bits για τα διανύσματα κίνησης σε σχέση με τις αργές γι' αυτό και χρειάζονται υψηλότερο bit rate
- **Σκηνές που αλλάζουν απότομα** από καρέ σε καρέ απαιτούν επίσης υψηλότερο bit rate γιατί δεν μπορούν να εκμεταλλευτούν την αντιστάθμιση κίνησης (δεν υπάρχει μεγάλη συσχέτιση μεταξύ γειτονικών καρέ)
- **Σκηνές που έχουν μεγάλη χωρική λεπτομέρεια** απαιτούν επίσης υψηλότερο bit rate γιατί κάθε block έχει μεγάλους συντελεστές ΔΜΣ ακόμη και στις υψηλές χωρικές συχνότητες.

Προσαρμογή Επιπέδου Ποιότητας

Επίδραση λαθών

Λάθη ---> αποτέλεσμα π.χ. διαλείψεων σε ένα ασύρματο κανάλι ή απώλειας πακέτων IP σε ένα δίκτυο δεδομένων λόγω συμφόρησης

- **Απώλεια διανυσμάτων κίνησης** ---> Macroblocks βρίσκονται σε λάθος θέσεις
- **Απώλεια συντελεστών ΔΜΣ** ---> Η εικόνα παραμορφώνεται, χάνονται τα όρια των blocks

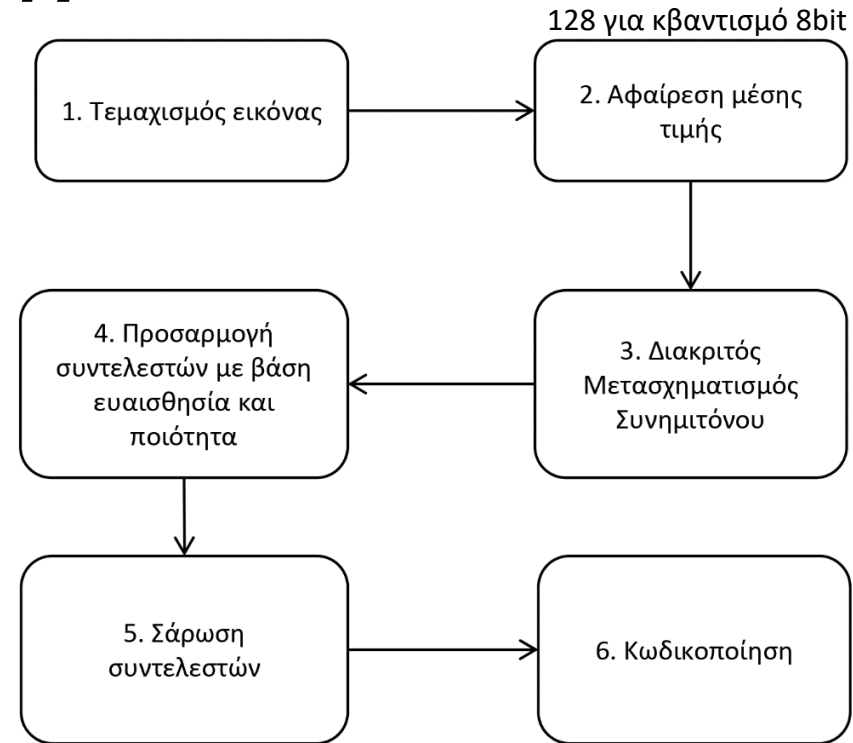


Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Τα βήματα που ακολουθούμε είναι τα εξής:

- τεμαχισμός εικόνας σε τμήματα (μπλοκ) 8×8
- αφαίρεση μέσης τιμής (του 128 στην περίπτωση που ο κβαντισμός θα γίνει με τη βοήθεια 8 bits)
- μετασχηματισμός κατά ΔΜΣ σε κάθε μπλοκ (οι συντελεστές που προκύπτουν ανήκουν στο πεδίο της συχνότητας – χωρικές συχνότητες)
- προσαρμογή συντελεστών με βάση την ευαισθησία του ανθρώπινου οφθαλμού αλλά και του επιθυμητού επιπέδου ποιότητας
- **σάρωση συντελεστών με ομαδοποίηση τιμών με την 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)
- Κωδικοποίηση/Κβαντισμός

Η προσαρμογή των συντελεστών γίνεται με βάση έτοιμους πίνακες (MPEG), που μπορούν να προσαρμοστούν ή να τροποποιηθούν από τους κωδικοποιητές ανάλογα με την επιθυμητή ποιότητα και τον βαθμό συμπίεσης. Ο αποκωδικοποιητής λαμβάνει πληροφορίες για αυτούς τους πίνακες καθώς αυτές μεταδίδονται.



Χρήση του μετασχηματισμού ΔΜΣ στο JPEG.

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Θεωρούμε το παρακάτω block μιας εικόνας. Οι τιμές των pixels (θεωρούμε τη φωτεινότητα) κυμαίνονται από 0 (μαύρο) έως 255 (άσπρο).

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 τμήματος εικόνας

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Αφαιρούμε την τιμή 128 από κάθε pixel (μέση τιμή), οπότε οι τιμές κυμαίνονται από -128 έως 127. Στη συνέχεια πραγματοποιούμε τον μετασχηματισμό ΔΜΣ, και προκύπτει ο πίνακας «Α».

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

Πίνακας «Α». Μετασχηματισμένες κατά ΔΜΣ τιμές

Παρατηρούμε ότι ο πίνακας «Α» αποτελείται από 64 συντελεστές, με αυτούς στο πάνω αριστερά τμήμα να αντιστοιχούν στις χαμηλές συχνότητες. Πηγαίνοντας προς τα κάτω και δεξιά πλησιάζουμε προς τις πιο ψηλές συχνότητες.

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Κάνοντας κλιμάκωση των συντελεστών για ποιότητα 50 (Q_{50}) προκύπτει ο πίνακας «B».

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

Πίνακας «B». Τιμές μετά την κλιμάκωση

Στον πίνακα «B» μόνο οι μη μηδενικές συνιστώσες θα χρησιμοποιηθούν. Οι μη μηδενικές συνιστώσες παρατηρούμε να μαζεύονται στο πάνω αριστερά μέρος του πίνακα.

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Με τη μέθοδο σάρωσης κατά Ζιγκ Ζαγκ προκύπτουν οι εξής τιμές και ο πίνακας («Γ»):

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

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

Πίνακας «Γ». Τιμές μετά την σάρωση Ζιγκ Ζαγκ

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Q50

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

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

Εκμετάλλευση Χωρικού Πλεονασμού - Παράδειγμα

Κατά τη διαδικασία αποκωδικοποίησης (στη λήψη), ξαναδημιουργείται ο κβαντισμένος πίνακας. Κάθε στοιχείο αυτού του πίνακα πολλαπλασιάζεται με το αντίστοιχο στοιχείο του (προτυποποιημένου) πίνακα κβαντίσού και προκύπτει ο πίνακας «Δ» της αποκωδικοποίησης.

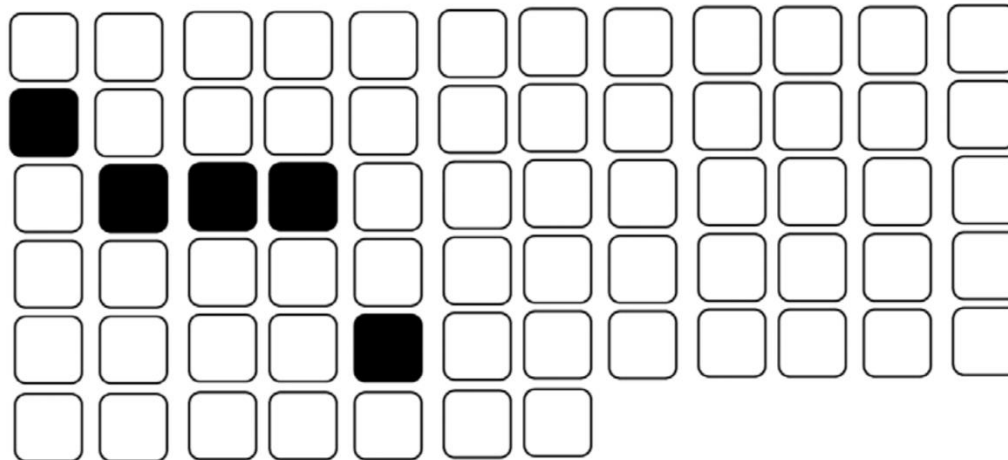
- Στη συνέχεια εφαρμόζεται ο αντίστροφος ΔΜΣ και γίνεται προσέγγιση του πλησιέστερου ακέραιου.
- Επίσης προστίθεται το 128 σε κάθε στοιχείο του προκύπτοντας πίνακα.

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

Πίνακας «Δ». Τιμές μετά την αποκωδικοποίηση

Κωδικοποίηση μήκους διαδρομής

- Η Κωδικοποίηση Μήκους Διαδρομής (Run Length Encoding, RLE) στοχεύει σε μια αποτελεσματική από πλευράς συμπίεσης, περιγραφή αλληλουχίας δεδομένων στις οποίες κάποιες τιμές επαναλαμβάνονται συνεχόμενα.
- Για παράδειγμα, σε μια εικόνα μπορεί να έχουμε πολλά συνεχόμενα άσπρα pixels (W) και κάποια μαύρα (B)



Ενδεικτική ακολουθία ίδιων τιμών σε τμήμα εικόνας

Κωδικοποίηση μήκους διαδρομής

Μια τυπική τεχνική (που εισάγει που εισάγουν έλεγχο - ευφυΐα) είναι ο αλγόριθμος να ενεργοποιείται από 4 ίδιους χαρακτήρες και άνω.

- Αφού γραφτεί το σύμβολο, χρησιμοποιείται ένας ειδικός χαρακτήρας, π.χ., το θαυμαστικό (!) ώστε να ακολουθήσει το πλήθος της εμφάνισης των χαρακτήρων (πλην του πλήθους το οποίο είναι απαραίτητο για την ενεργοποίηση του μηχανισμού).

Με αυτόν τον τρόπο αποφεύγουμε τη χρήση του αλγορίθμου σε περίπτωση συνεχών εναλλαγών.

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

Κωδικοποίηση συμβόλων με χρήση αλγορίθμου μήκους διαδρομής (RLE)

Εκμετάλλευση στατιστικού πλεονασμού

Ομοιόμορφη κωδικοποίηση

Έστω ένα σύνολο από σύμβολα τα οποία αποτελούν ένα αλφάβητο και επιθυμούμε να κωδικοποιήσουμε δυαδικά αυτά τα σύμβολα. Ο πρώτος και πλέον βασικός τρόπος κωδικοποίησης είναι η χρήση της **ομοιόμορφης κωδικοποίησης**, η οποία περιλαμβάνει τα επόμενα βήματα:

1. Με βάση το πλήθος των συμβολών προς κωδικοποίηση, υπολογίζεται το πλήθος των bits που χρειαζόμαστε.
 - Η σχέση υπολογισμού είναι $N \leq 2^n$, όπου N το πλήθος των συμβολών προς κωδικοποίηση και n το πλήθος των bits που απαιτούνται για την κωδικοποίηση κάθε ενός από αυτά τα σύμβολα.
2. Σε κάθε σύμβολο αντιστοιχίζεται ένας διαφορετικός συνδυασμός (όσον αφορά τις τιμές που αυτά έχουν 0 ή 1) από n bits.

Έτσι, για παράδειγμα, εάν επιθυμούμε κωδικοποίηση $N=6$ συμβολών, το πλήθος των bits που απαιτούνται είναι $n=3$, αφού $2^2 < 6$ και $2^3 > 6$. Η ομοιόμορφη κωδικοποίηση θα μας έδινε τις εξής αντιστοιχίσεις:

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

Ομοιόμορφη κωδικοποίηση έξι συμβόλων

Κωδικοποίηση εντροπίας – Huffman

Η εκμετάλλευση του στατιστικού πλεονασμού αποτελεί μια **μη απωλεστική (lossless)** μέθοδο συμπίεσης δεδομένων.

- 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 (με μικρότερες, δηλαδή, σε μήκος λέξεις κωδικοποίησης, όπου το μήκος μεταφράζεται στο πλήθος των bits που χρησιμοποιούνται για την κωδικοποίηση της λέξης), αναμένουμε να έχουμε όφελος σε σχέση με την ομοιόμορφη κωδικοποίηση, η οποία κωδικοποιεί ομοιόμορφα (με ίδιου δηλαδή μήκους λέξεις κωδικοποίησης) όλα τα σύμβολα (χωρίς δηλαδή να λαμβάνει υπόψη τις στατιστικές ιδιότητες εμφάνισης των συμβόλων).

Κωδικοποίηση εντροπίας – Huffman

Αρχή κωδικοποίησης Huffman:

Σε κάποιες περιπτώσεις είναι δυνατόν να υπάρξει πληροφορία (ή εκτίμηση) για τη **συχνότητα εμφάνισης των συμβολών** στις λέξεις που προέρχονται από το συγκεκριμένο αλφάβητο. Συγκεκριμένα, κάποια σύμβολα εμφανίζονται πιο συχνά από τα άλλα.

- Σε μια τέτοια περίπτωση ενδείκνυται η χρήση μικρότερων (σε μήκος) ακολουθιών για τα πιο συχνά χρησιμοποιούμενα σύμβολα.
- Αυτή ακριβώς τη γνώση για τη συχνότητα εμφάνισης των συμβολών είναι που εκμεταλλεύεται ο αλγόριθμος Huffman.

Κωδικοποίηση εντροπίας – Huffman

Το πρόβλημα τίθεται ως εξής:

- Έστω k τα προς κωδικοποίηση σύμβολα (αλφάβητο)

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

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

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

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

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

Κωδικοποίηση εντροπίας – 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:

$$\sum_{i=1}^k w_i n_i \quad (8)$$

Κωδικοποίηση εντροπίας – Huffman

Για την εύρεση της βέλτιστης κωδικοποίησης με τα παραπάνω δεδομένα, ο αλγόριθμος Huffman βασίζεται στην κατασκευή του δυαδικού δένδρου Huffman, σύμφωνα με την επόμενη μεθοδολογία:

1. Δημιουργούνται (απεικονίζονται) οι αρχικοί κόμβοι (που αποτελούν τα φύλλα του δένδρου) από τα προς κωδικοποίηση σύμβολα, όπου σε κάθε ένα από αυτά αναγράφεται το βάρος του (το οποίο αντιστοιχεί στην πιθανότητα εμφάνισής του). Για διευκόλυνση της δημιουργίας του δένδρου τα σύμβολα ενδείκνυται να γράφονται με αύξουσα ή φθίνουσα σειρά (όσον αφορά τα βάρη τους).
2. 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. Το βάρος του νέου κόμβου προκύπτει από το άθροισμα των βαρών των δύο κόμβων από τους οποίους αυτός έχει προκύψει.
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.

Κωδικοποίηση εντροπίας – Huffman

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.
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).

Παρατηρήσεις σχετικά με την κωδικοποίηση Huffman:

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

Κωδικοποίηση εντροπίας – Huffman

Ο αλγόριθμος Huffman συνεργάζεται αρμονικά με τις τεχνικές που συζητήσαμε στις προηγούμενες διαφάνειες. Συγκεκριμένα, αποτελεί το τελευταίο στάδιο της κωδικοποίησης / συμπίεσης των δεδομένων. Η διαδικασία συνοπτικά είναι η εξής:

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

Κωδικοποίηση εντροπίας – Huffman

Παράδειγμα – 5

Έστω τα 6 σύμβολα που είχαμε εξετάσει στο παράδειγμα της ομοιόμορφης κωδικοποίησης, με δεδομένο το πλήθος των εμφανίσεων τους. Από το πλήθος των εμφανίσεων κάθε συμβόλου και το συνολικό πλήθος των εμφανίσεων προκύπτει η αντίστοιχη πιθανότητα εμφάνισης.

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
Άθροισμα πιθανοτήτων εμφάνισης:		1

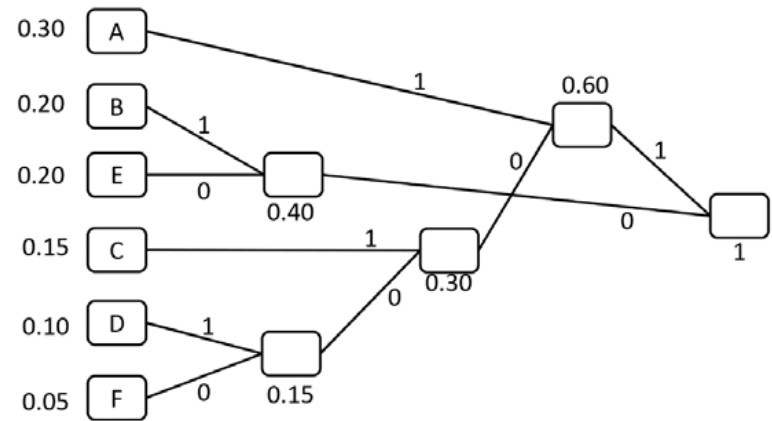
Σύμβολα προς κωδικοποίηση, πλήθος εμφανίσεων και πιθανότητες εμφάνισης

Κωδικοποίηση εντροπίας – Huffman

Με χρήση του αλγόριθμου Huffman, δημιουργείται το «δένδρο» της διπλανής εικόνας. Τα βήματα είναι:

- Τοποθετούνται τα σύμβολα, με φθίνουσα σειρά αναφορικά με τις τιμές των βαρών τους.
- Τα σύμβολα F και D έχοντας τα μικρότερα βάρη συνδυάζονται και δίνουν έναν κόμβο με βάρος $0.05 + 0.10 = 0.15$
- Ο κόμβος που δημιουργήθηκε, με βάρος 0.15 και ο κόμβος C (επίσης με βάρος 0.15) συνδυάζονται και δίνουν τον κόμβο με βάρος 0.30.
- Συνδυάζονται οι κόμβοι B και E με βάρη 0.20 ώστε να δώσουν τον κόμβο με βάρος 0.40.
- Ο κόμβος A και ο κόμβος που δημιουργήθηκε προηγουμένως με βάρος 0.30 συνδυάζονται ώστε να δώσουν τον κόμβο με βάρος 0.60.
- Οι δύο εναπομείναντες κόμβοι με βάρη 0.40 και 0.60, αντίστοιχα, συνδυάζονται ώστε να δώσουν τη ρίζα του δένδρου, με βάρος 1.

Μετά τη δημιουργία του δένδρου αντιστοιχούμε τους κλάδους στα 0 και 1 (ο άνω κλάδος κάθε κόμβου έχει αντιστοιχηθεί στο 1 και ο κάτω κλάδος στο 0). Με βάση τις διαδρομές από τη ρίζα του δένδρου προς κάθε φύλλο, δημιουργείται ο διπλανός πίνακας των κωδικοποιήσεων.



Δημιουργία δένδρου Huffman

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

Σύμβολα και κωδικοποιήσεις

Κωδικοποίηση εντροπίας – Huffman

Σύγκριση της απόδοσης του σε σχέση με την απόδοση της ομοιόμορφης κωδικοποίησης.

- 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:

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

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.

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

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

Προφίλ και επίπεδα

Προφίλ και επίπεδα

Με τη χρήση των μεθόδων που αναφέραμε στις προηγούμενες διαφάνειες για τη μείωση του πλεονασμού, ο ρυθμός του video είναι δυνατό να μειωθεί από τα 270 Mbps (με 4:2:2 χρωματική υποδειγματοληψία) στα 2 Mbps μέχρι και τα 6 Mbps.

- Το πιο σημαντικό βήμα στη διαδικασία της συμπίεσης είναι ο περιορισμός του **χρονικού πλεονασμού** (temporal redundancy) σε συνδυασμό με τον **χωρικό πλεονασμό** (spatial redundancy) και την **εφαρμογή του ΔΜΣ**, καθώς επίσης και η περαιτέρω **υποδειγματοληψία** του χρώματος (με χρήση του μηχανισμού 4:2:0).

Άρα, το επίπεδο συμπίεσης δεν είναι σταθερό αλλά εξαρτάται από παραμέτρους. Το MPEG έχει δημιουργήσει μια σειρά από επίπεδα (level) και προφίλ (profile) με τα οποία ομαδοποιεί τις επιλογές αυτές:

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

Προφίλ και επίπεδα

Προφίλ κατά το MPEG2

Συν/ση	Ονομασία	Τύποι κωδικ/σης εικόνας	Υποδειγμ/ψία Χρώματος	Αναλογία Εικόνας	Κλιμάκωση	Ακρίβεια DCT
SP	Simple Profile	I, P	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι	8, 9, 10
MP	Main Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι	8, 9, 10
SNR	SNR Scalable Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Κλιμάκωση κατά SNR	8, 9, 10
Spatial	Spatially Scalable Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Χωρική κλιμάκωση και κλιμάκωση κατά SNR	8, 9, 10
HP	HP Profile	I, P, B	4:2:2 ή 4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Όχι	8, 9, 10,11
422	4:2:2 Profile	I, P, B	4:2:2 ή 4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Χωρική κλιμάκωση και κλιμάκωση κατά SNR	8, 9, 10,11
MVP	Multi View Profile	I, P, B	4:2:0	Τετράγωνα εικονοστοιχεία 4:3 ή 16:9	Χρονική κλιμάκωση	8, 9, 10

Προφίλ και επίπεδα

Στο πρότυπο MPEG2 η κλιμάκωση θεωρείται σε τρία επίπεδα τα οποία είναι:

- **Σηματοθορυβικό (Signal to Noise Ratio, SNR) επίπεδο**, το οποίο παρέχει σταδιακή αναβάθμιση ή υποβάθμιση της ποιότητας του κωδικοποιημένου βίντεο (με αλλαγές της κλιμάκωσης των τιμών των συντελεστών στο μετασχηματισμό ΔΜΣ και του πλήθους των συντελεστών που λαμβάνονται υπ' όψιν).
- **Χωρικό (spatial) επίπεδο**, το οποίο στοχεύει στην υποστήριξη διαφορετικών συσκευών με διαφορετικές αναλύσεις. Στην περίπτωση αυτή μπορεί, για παράδειγμα, η βασική κωδικοποίηση να γίνει με βάση την τυπική ανάλυση ενώ η ενισχυμένη κωδικοποίηση να γίνει με βάση την υψηλή ανάλυση (SDTV και HDTV αντίστοιχα) .
- **Χρονικό (temporal) επίπεδο**, το οποίο στοχεύει στην κλιμάκωση με στόχο την υποστήριξη διαφορετικών ρυθμών ανανέωσης πλαισίων. Αυτή μπορεί να χρησιμοποιηθεί κυρίως σε ασύρματους διαύλους επικοινωνίας, όπου λόγω φτωχής κάλυψης είναι δυνατόν να μην αποστέλλονται κάποια από τα πλαίσια.

Προφίλ και επίπεδα

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 προφίλ) 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

Επαναληπτικές ασκήσεις – Κριτήρια αξιολόγησης

Κριτήριο αξιολόγησης 1

Υπολογίστε τις δύο πρώτες τιμές του διακριτού μετασχηματισμού συνημιτόνου για το επόμενο σήμα:

$$x(n)=[1 \ 2 \ 3 \ 4]$$

Απάντηση/Λύση

Χρησιμοποιούμε τον τύπο του ΔΜΣ ως εξής:

$$y(1) = 0.5 * (1+2+3+4) = 5$$
$$y(2) = 0.707 * (1*\cos(\pi/8) + 2*\cos(3\pi/8) + 3*\cos(5\pi/8) + 4*\cos(7\pi/8)) = -2.23$$

Κάνοντας τις αντίστοιχες πράξεις μπορούμε να δείξουμε ότι

$$y(3) = 0 \text{ και}$$
$$y(4) = -0.16$$

Κριτήριο αξιολόγησης 2

Άσκηση: Σε ένα video κωδικοποιημένο κατά MPEG-2 αντιστοιχούμε 1 block χρωματικότητας του μπλε (Cb) και 1 block χρωματικότητας του κόκκινου (Cr) για κάθε 4 blocks φωτεινότητας (Y). Να υπολογίσετε σε πόσα blocks συνολικά απαρτίζουν ένα frame από video υψηλής ανάλυσης HD (1280 x 720).

Απάντηση/λύση

- $1280/8 = 160$ και $720/8 = 90$. Άρα κάθε frame έχει $160 \times 90 = 14400$ blocks φωτεινότητας (Y) των 8×8 pixels.
- Επιπλέον επειδή $1 \text{ block Cr} = 1 \text{ block Cb} = 4 \text{ blocks Y}$, απαιτούνται $14400/4 = 3600$ blocks Cb και επίσης 3600 blocks Cr.
- Συνεπώς συνολικά το frame απαρτίζεται από $14400 + 3600 + 3600 = \mathbf{21600 \text{ blocks}}$.

Κριτήριο αξιολόγησης 3

Άσκηση: Για GOP size=12, βρείτε από πόσα GOP συνολικά αποτελείται ένα MPEG-2 video sequence των 25 fps διάρκειας πέντε λεπτών;

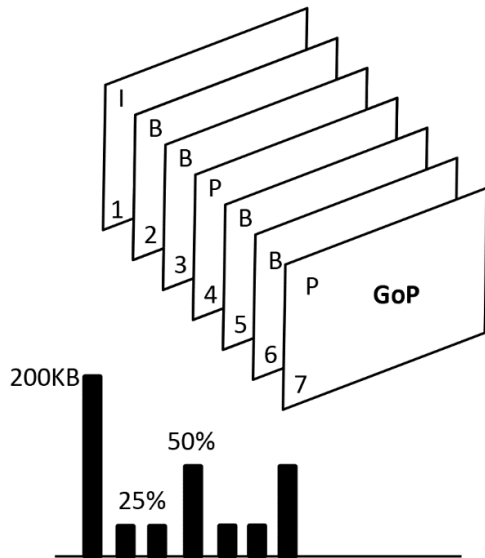
Απάντηση/λύση

Το sequence έχει διάρκεια 5 λεπτά = 300 δευτερόλεπτα
άρα αποτελείται από $300 \times 25 = 7500$ frames.

Συνεπώς, το πλήθος των GOP που απαρτίζει το sequence είναι $7500 / 12 = \mathbf{625\ GOPs}$.

Κριτήριο αξιολόγησης 4

Άσκηση: Όπως εμφανίζεται στην εικόνα το μέγεθος ενός πλαισίου που κωδικοποιείται ως **I** είναι το μεγαλύτερο σε σχέση με τα μεγέθη των **P** και **B**. Το μέγεθος του πλαισίου τύπου **P** μπορεί να είναι στο ήμισυ αυτού του **I**, ενώ του **B** δύναται να φτάσει και στο 25%. Αυτές, βέβαια, είναι ενδεικτικές τιμές αφού οι αναλογίες είναι εν γένει παραμετροποιήσιμες. Υπολογίστε τη συμπίεση που επιτυγχάνουμε με την κωδικοποίηση των πλαισίων σε τύπους **P** και **B** (πέρα από το βασικό τύπου του **I**), με τη βοήθεια της αλληλουχίας πλαισίων που παρουσιάζεται στο επόμενο σχήμα.



Απάντηση: Παρατηρούμε ότι από την αλληλουχία πλαισίων έχουμε ένα πλαίσιο τύπου **I**, δύο πλαίσια τύπου **P** και τέσσερα 4 πλαίσια τύπου **B**.

Κάνουμε τις εξής παρατηρήσεις:

- Εάν δεν είχαμε κωδικοποίηση χρονικού πλεονασμού θα είχαμε 7 πλαίσια τύπου **I**, άρα το σχετικό μέγεθος θα ήταν $7 \cdot I$.
- Στην περίπτωση της ομάδας πλαισίων του σχήματος το μέγεθος υπολογίζεται από το άθροισμα των ποσοστών με βάση τη σύνθεση της ομάδας.

Το άθροισμα αυτό είναι το εξής:

$$1 + 2 \cdot 0.5 \cdot I + 4 \cdot 0.25 \cdot I = 3 \cdot I$$

Προκύπτει, λοιπόν, ότι από τα $7 \cdot I$ πάμε στα $3 \cdot I$ άρα επιτυγχάνουμε συμπίεση 7 προς 3 (2.3 προς 1).

Κριτήριο αξιολόγησης 5

Άσκηση: Δίνονται οι τιμές των συντελεστών ΔΜΣ ενός block MPEG-2 που έχουν προκύψει μετά από τον μετασχηματισμό ΔΜΣ (Πίνακας Α), και ο ο τυποποιημένος πίνακας κβάντισης (Πίνακας Β). Να υπολογίσετε τους συντελεστές ΔΜΣ που τελικά μεταδίδονται χρησιμοποιώντας κλίμακα κβάντισης q όπου i) $q=2$ και ii) $q=4$. Τι κερδίζουμε και τι χάνουμε χρησιμοποιώντας τη μεγαλύτερη κλίμακα κβάντισης;

99	66	32	9	0	0	0	0
54	21	9	5	0	0	0	0
13	1	0	0	0	0	0	0
6	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	0
0	0	0	0	0	0	0	0

Πίνακας Α

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Πίνακας Β

Κριτήριο αξιολόγησης 5

Απάντηση για κλίμακα κβάντισης $q=2$

Ο πίνακας των συντελεστών που μεταδίδονται προκύπτει από τη διαίρεση των αρχικών συντελεστών με το αντίστοιχο στοιχείο του πίνακα κβάντισης και στη συνέχεια με την κλίμακα κβάντισης. Το αποτέλεσμα στρογγυλοποιείται στον πλησιέστερο ακέραιο (βλ. Πίνακας Γ).

99	66	32	9	0	0	0	0
54	21	9	5	0	0	0	0
13	1	0	0	0	0	0	0
6	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	0
0	0	0	0	0	0	0	0

Πίνακας Α

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Πίνακας Β

6	2	1	0	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
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Πίνακας Γ

Κριτήριο αξιολόγησης 5

Απάντηση για κλίμακα κβάντισης $q=4$

Ο πίνακας των συντελεστών που μεταδίδονται προκύπτει από τη διαίρεση των αρχικών συντελεστών με το αντίστοιχο στοιχείο του πίνακα κβάντισης και στη συνέχεια με την κλίμακα κβάντισης. Το αποτέλεσμα στρογγυλοποιείται στον πλησιέστερο ακέραιο (βλ. Πίνακας Δ).

99	66	32	9	0	0	0	0
54	21	9	5	0	0	0	0
13	1	0	0	0	0	0	0
6	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	0
0	0	0	0	0	0	0	0

Πίνακας Α

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

Πίνακας Β

3	1	0	0	0	0	0	0
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	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

Πίνακας Δ

Κριτήριο αξιολόγησης 5

Απάντηση για την σύγκριση των δύο κλιμάκων κβάντισης

Παρατηρούμε ότι για $q=4$ (επειδή οι μηδενικοί συντελεστές δεν μεταδίδονται) μεταδίδονται μόνο 3 τιμές ενώ για $q=2$ μεταδίδονται 5 τιμές. **Άρα για $q=4$ χρειάζονται λιγότερα bytes για την αναπαράσταση του block και γενικότερα μικρότερο bit rate για το video.**

Όμως για $q=4$ μηδενίζονται περισσότεροι ΔΜΣ συντελεστές που σημαίνει ότι τελικά χάνονται χωρικές συχνότητες από το block. Έτσι, παρατηρούμε απώλεια στην λεπτομέρεια της εικόνας και σε ορισμένες περιπτώσεις και “blockiness effect”.

Κριτήριο αξιολόγησης 6

Άσκηση: Μια ψηφιακή εικόνα αποχρώσεων του γκρι αναπαρίσταται όπως στον επόμενο πίνακα (βλ. Πίνακας Α). Υπολογίστε τον πίνακα ιστογράμματος, την κωδικοποίηση Huffman και τα αντίστοιχα σύμβολα κωδικοποίησης ώστε να αποστείλουμε το περιεχόμενο της εικόνας.

20	30	50	30	20
50	30	30	50	20
20	30	50	20	100
150	200	150	20	200

Πίνακας Α

Κριτήριο αξιολόγησης 6

Απάντηση

Ο Πίνακας ιστογράμματος καθώς και οι πιθανότητες εμφάνισης είναι οι εξής (βλ. **Πίνακας Β**).

Αφού σχηματιστεί το δέντρο Huffman προκύπτουν οι ακολουθίες bits για κάθε απόχρωση (βλ. **Πίνακας Γ**)

Σύμβολο	Πλήθος Εμφανίσεων	Πιθανότητα
20	6	0,3
30	5	0,25
50	4	0,2
100	1	0,05
150	2	0,1
200	2	0,1

Πίνακας Β

Σύμβολο	Λέξη Κωδικοποίησης
20	11
30	10
50	00
100	0110
150	010
200	0111

Πίνακας Γ

Βιβλιογραφία και βασικές πηγές

Βιβλιογραφία και βασικές πηγές

Προτεινόμενη βιβλιογραφία

1. Παπαδάκης, Α., 2015. Ψηφιακή τηλεόραση. [ηλεκτρ. βιβλ.] Αθήνα:Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. Διαθέσιμο στο: <http://hdl.handle.net/11419/5005>
2. Κ. Τσαμούταλος και Π. Σαράντης. Αναλογική και Ψηφιακή Τηλεόραση. Σταμούλης, 2003.
3. Π. Βαφειάδης. Αναλογική-ψηφιακή τηλεόραση και video. 7η έκδοση. Π. Βαφειάδης, 2014.
4. International Telecommunication Union, BT.601. Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios. International Telecommunication Union, 2011.
5. International Standards Organization / International Electrotechnical Commission. Information technology — Generic coding of moving pictures and associated audio information: Systems, International Standards Organization / International Electrotechnical Commission ISO/IEC, 13818-1, 2001.
6. International Standards Organization / International Electrotechnical Commission. Information technology — Generic coding of moving pictures and associated audio information: Video, International Standards Organization / International Electrotechnical Commission ISO/IEC, 13818-2, 2000.
7. International Telecommunication Union. H.264: Advanced video coding for generic audiovisual services. International Telecommunication Union, 2008
8. International Telecommunication Union, Information Technology. Digital Compression and Coding of continuous – tone still images. Requirements and Guidelines, Recommendation T.81. International Telecommunication Union, 1992.
9. Fr. Halsall. Multimedia Communications: Applications, Networks, Protocols and Standards. Addison-Wesley, 2001.
10. J. Arnold, M. Frater, M. Pickering. Digital Television Technology and Standards. Wiley Interscience, 2007.

Συναφή επιστημονικά περιοδικά

1. IEEE Communications Magazine, ComSoc
2. IEEE Transactions on Broadcasting