



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Παραγωγή Τηλεοπτικού Σήματος

Παντελής Μπαλαούρας
Τμήμα Πληροφορικής & Τηλεπικοινωνιών

Πηγές

[1] Παπαδάκης, Α. 2015. Παραγωγή Τηλεοπτικού Σήματος. *Ψηφιακή τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 2. Διαθέσιμο στο: <http://hdl.handle.net/11419/5007>



Εισαγωγή

Στόχοι ενότητας

- Οι τρόποι αναπαράστασης της πληροφορίας που αφορά τα χρώματα σε μian εικόνα και κατ' επέκταση στο βίντεο.
- Η μετατροπή μεταξύ ισοδύναμων τρόπων αναπαράστασης εικόνων και βίντεο
- Η διαχείριση της υποδειγματοληψίας χρώματος και τη συσχέτιση με τα χαρακτηριστικά του τηλεοπτικού σήματος.
- Οι απαιτήσεις σε εύρος ζώνης και τον τρόπο που αυτές σχετίζονται με τις επιλογές που αφορούν το βίντεο.



Βασικά

- **Βίντεο** = αλληλουχία εικόνων ή αλλιώς πλαίσια
- **Εικόνα**
 - **εικονοστοιχεία** (picture elements- pixels)
 - πλήθος pixel = ανάλυση (resolution)
- Η **ανάλυση** = πλήθους των γραμμών X των στηλών της εικόνας, π.χ. 1920 X 1080 = 2.073.600 pixels
- Η **ταχύτητα**: μέτρηση σε πλαίσια ανά δευτερόλεπτο (frames per second, fps), π.χ. 25 ή 16 fps
 - 51.840.000 pixels / sec



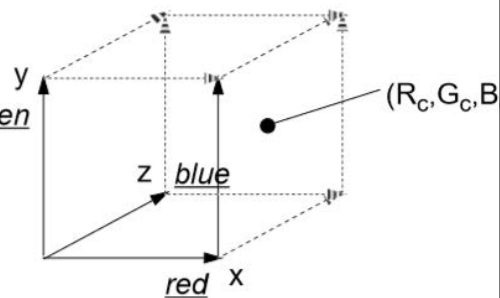
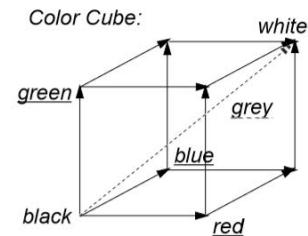
Αναπαράσταση σήματος βίντεο

- **Αναπαράσταση του σήματος βίντεο** με βάση την αναπαράσταση **RGB** (βασικά χρώματα Red, Green, Blue).
 - στη λήψη του σήματος στον εικονολήπτη & και για την αναπαραγωγή αυτού στον δέκτη
- Για την μετάδοσης έχουμε μια μετατροπή από τον χώρο RGB σε έναν άλλο ισοδύναμο τρόπο αναπαράστασης
 - φωτεινότητα
 - μεγέθη που ονομάζουμε χρωμοδιαφορές (Y, Cb, Cr)

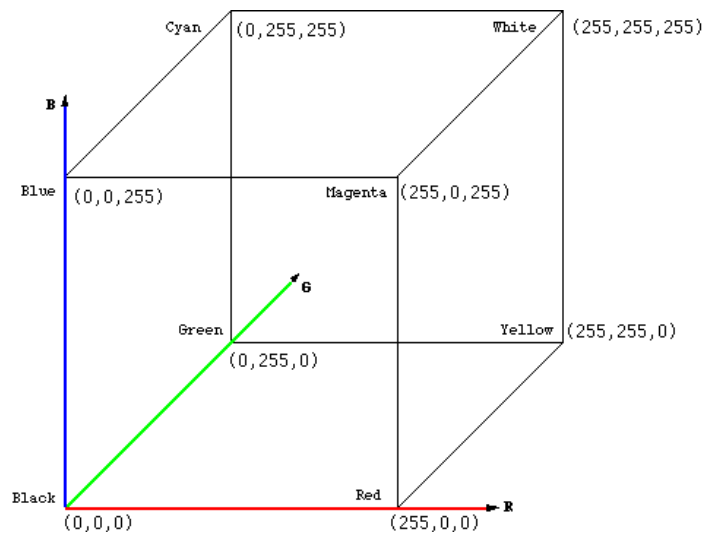


Τριχρωματική Όραση

- Η δυνατότητα παραγωγής οποιουδήποτε χρώματος με μίξη τριών φασματικών χρωμάτων, τα οποία ονομάζουμε κύρια χρώματα και τα οποία είναι το κόκκινο (R), το πράσινο (G) και το μπλε (B).
- Ο χώρος αναπαράστασης RGB χρησιμοποιεί το ποσοστό συμμετοχής των **τριών** αυτών βασικών χρωμάτων για την περιγραφή οποιουδήποτε συνδυασμού.
- **Λευκό:** 100% R, 100% B, 100% G

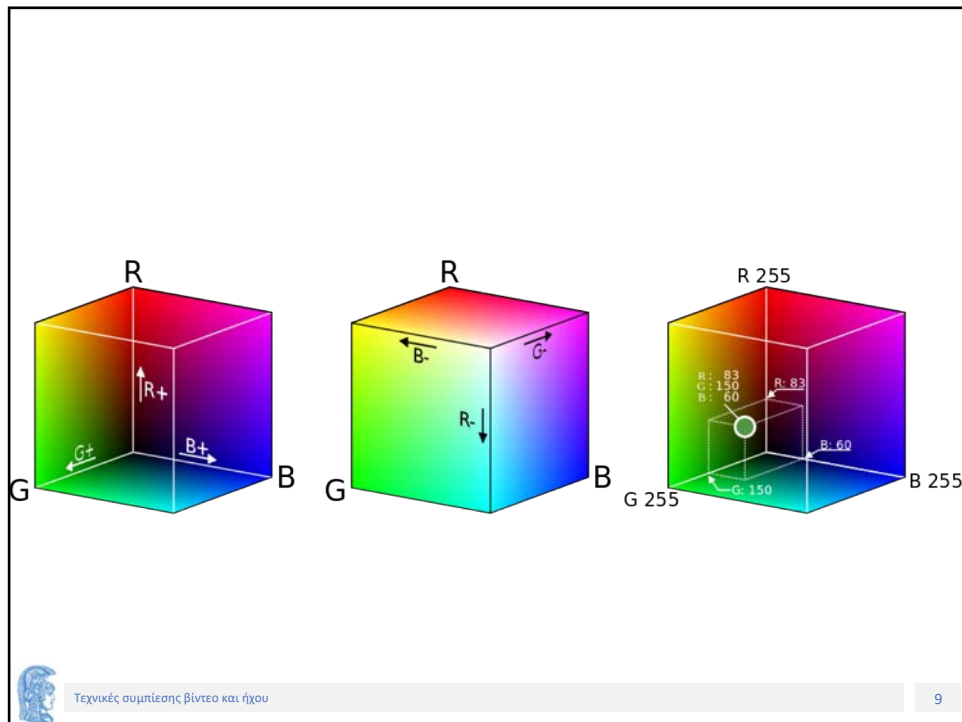


Τεχνικές συμπίεσης βίντεο και ήχου



Τεχνικές συμπίεσης βίντεο και ήχου

8



Πλήθος bits / βασικό χρώμα

- Πλήθος των bits για αναπαράσταση
 - επίπεδα κβαντισμού
- 8 bits (τυπική επιλογή),
 - το εύρος είναι από το 0 έως το 255 (100%)
- 10 bits
 - το εύρος κυμαίνεται από το 0 έως το 1.024 (100%)

Φωτεινότητα και χρωμοδιαφορές

- Ισοδύναμος τρόπος αναπαράστασης

$Y = 0,30R + 0,59G + 0,11B$ (φωτεινότητα, εύρος: λευκό μαύρο)

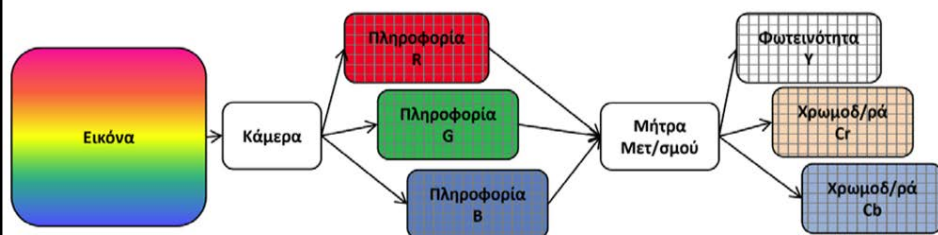
$C_r = R - Y = R - (0,30R + 0,59G + 0,11B) = 0,70R - 0,59G - 0,11B$

$C_b = B - Y = B - (0,30R + 0,59G + 0,11B) = -0,30R - 0,59G + 0,89B$

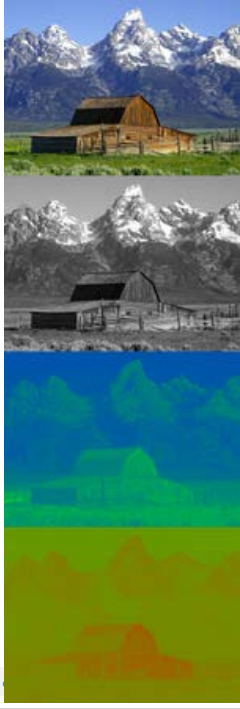
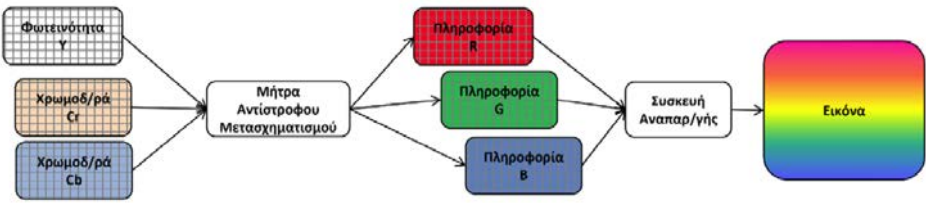
$C_g = G - Y = G - (0,30R + 0,59G + 0,11B) = -0,30R + 0,41G - 0,11B$



Μετατροπή χρωματικής πληροφορίας από τον χώρο RGB στον χώρο YCbCr



Επαναφορά πληροφορίας από τον χώρο YCbCr στον χώρο RGB



Ρυθμός δεδομένων

- Ο ρυθμός δεδομένων του βίντεο εξαρτάται από:
 - την ανάλυση κάθε εικόνας
 - τον αριθμό των bits
 - τη συχνότητα με την οποία λαμβάνονται οι εικόνες
- Παγκοσμίως έχουν επικρατήσει δύο συστήματα:
 - 625 γραμμές ανά εικόνα, 25 Hz (Ευρώπη)
 - 525 γραμμές ανά εικόνα, 30 Hz (Αμερική)



Δημιουργία Ψηφιακού Σήματος

- **Βήμα 1:** λήψη των τιμών των βασικών χρωματικών συνιστωσών κόκκινου, πράσινου και μπλε (Red, Green, Blue).
- **Βήμα 2:** επεξεργασία ώστε να δώσουν το σήμα φωτεινότητας Y και τα σήματα χρωμικότητας, δηλαδή τις χρωμοδιαφορές Cb και Cr.

$$Y = 0,30R + 0,59G + 0,11B \quad (\text{φωτεινότητα, εύρος: λευκό μαύρο})$$

$$R - Y = R - (0,30R + 0,59G + 0,11B) = 0,70R - 0,59G - 0,11B$$

$$B - Y = B - (0,30R + 0,59G + 0,11B) = -0,30R - 0,59G + 0,89B$$

- Οι τιμές της μεταβλητής Y έχουν εύρος τιμών από 0 έως 1
- Η χρωμοδιαφορά **R-Y** από -0,70 έως +0,70 και της **B-Y** από -0,89 έως +0,89.
- Για να αντιστοιχιστούν τα εύρη των τιμών στην μονάδα, γίνεται κλιμάκωση των τιμών και οι χρωμοδιαφορές κυμαίνονται στο διάστημα -0,5 έως και 0,5.

$$Cb = 0,56 * (B - Y)$$

$$Cr = 0,71(R - Y)$$



Φιλτράρισμα και Δειγματοληψία

- Το σήμα στη συνέχεια φιλτράρεται με χρήση φίλτρου χαμηλών συχνοτήτων.
 - το σήμα της φωτεινότητας υφίσταται φιλτράρισμα στα 5,75 MHz
 - τα σήματα των χρωμοδιαφορών στα 2,75 MHz
- Τα σήματα Y, Cb, Cr μετατρέπονται στη συνέχεια σε **σήματα διακριτού χρόνου**, με χρήση δειγματοληψίας.
- Παρατηρούμε ότι η συχνότητα δειγματοληψίας της φωτεινότητας είναι η διπλάσια από αυτή των χρωμοδιαφορών, με αποτέλεσμα το πλήθος των δειγμάτων που αφορούν τη φωτεινότητα να ισούται με το πλήθος των δειγμάτων που αφορούν τις χρωμοδιαφορές.
- Οι συχνότητες δειγματοληψίας για τα σήματα πρέπει να υπερβαίνουν το διπλάσιο της μέγιστης συνιστώσας συχνότητας (σύμφωνα και με το φιλτράρισμα που έχουν υποστεί) ώστε να ικανοποιούν το θεώρημα του Shannon.

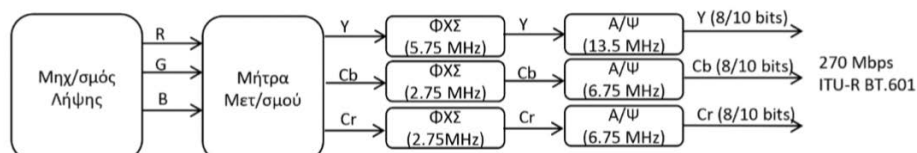


Τεχνικές συμπίεσης βίντεο και ήχου

17

Συχνότητες Δειγματοληψίας

Σήμα	Συχνότητα Δειγματοληψίας (MHz)	Αναλογία	
Y	13,5	2	4
Cr	6,75	1	2
Cb	6,75	1	2



Τεχνικές συμπίεσης βίντεο και ήχου

18

Κβαντισμός

- Αντιστοίχιση κάθε δείγματος σε ακολουθία bits
- Ο κβαντισμός των σημάτων της φωτεινότητας και των χρωμοδιαφορών πραγματοποιείται με χρήση κατ' ελάχιστον
 - οκτώ bits για τη φωτεινότητα
 - τεσσάρων bits για τις χρωμοδιαφορές
 - οι συνήθεις τιμές είναι τα 8 ή τα 10 bits και για τα τρία μεγέθη
- Ομοιόμορφος κβαντισμός με χρήση 8 και 10 bits
 - έχουμε 2^8 και 2^{10} επίπεδα κβαντισμού
 - το πλήθος των τιμών είναι αντίστοιχα 256 και 1.024



Σφάλμα κβαντισμού

- 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 τη στρωγγυλοποίηση που πραγματοποιείται και εξαρτάται από το πλήθος των διαθέσιμων επιπέδων.
- Στην ουσία αφορά τη **διαφορά μεταξύ της πραγματικής τιμής του αναλογικού σήματος και της κοντινότερης κβαντισμένης τιμής.**
- Η στρωγγυλοποίηση αυτή αφορά το σφάλμα κβαντισμού (ή, αλλιώς, τον θόρυβο κβαντισμού).



Ρυθμός μετάδοσης με 8 bits/δείγμα

- Ο ρυθμός μετάδοσης για δειγματοληψία με ρυθμούς 13,5 MHz και 6,75MHz αντίστοιχα και κβαντισμό (α) με χρήση 8 bits και για τις τρεις προκύπτουσες συνιστώσες και (β) 10 bits για τις τρεις συνιστώσες προκύπτει από τις επόμενες σχέσεις. Στην πρώτη περίπτωση έχουμε κβαντισμό με χρήση 8 bits:

$$Y: 13,5 \text{ MHz} * 8 \text{ bits / sample} = 108 \text{ Mbps}$$

$$Cb: 6,75 \text{ MHz} * 8 \text{ bits / sample} = 54 \text{ Mbps}$$

$$Cr: 6,75 \text{ MHz} * 8 \text{ bits / sample} = 54 \text{ Mbps}$$

ΑΘΡΟΙΣΜΑ 261 Mbps



Ρυθμός μετάδοσης με 10 bits/δείγμα

- Στη δεύτερη περίπτωση όπου ο κβαντισμός γίνεται με χρήση 10 bits έχουμε τους επόμενους επιμέρους ρυθμούς:

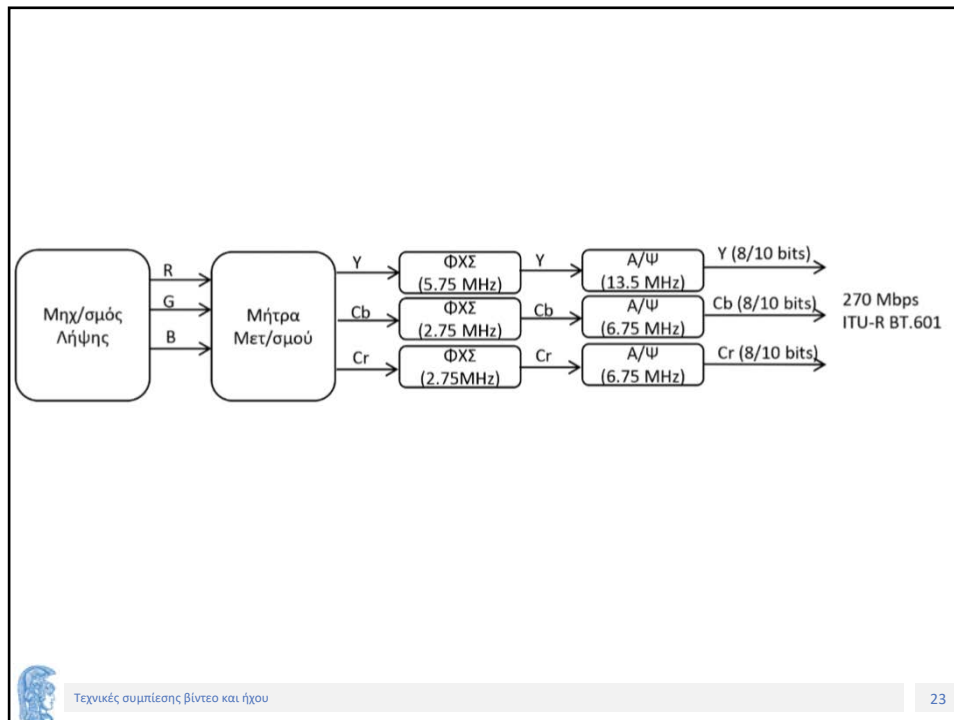
$$Y: 13,5 \text{ MHz} * 10 \text{ bits / sample} = 135 \text{ Mbps}$$

$$Cb: 6,75 \text{ MHz} * 10 \text{ bits / sample} = 67,5 \text{ Mbps}$$

$$Cr: 6,75 \text{ MHz} * 10 \text{ bits / sample} = 67,5 \text{ Mbps}$$

ΑΘΡΟΙΣΜΑ 270 Mbps





Υποδειγματοληψία Χρώματος

- Η εναλλαγή δειγμάτων φωτεινότητας και χρωμοδιαφορών στο σήμα βίντεο είναι:
Y, Cb, Y, Cr, Y, Cb, Y, Cr, ...

- Η αλληλουχία αυτή ονομάζεται **4:2:2** και στην ουσία πρόκειται για μια τεχνική που παραπέμπει στην υποδειγματοληψία του χρώματος, η οποία εκμεταλλεύεται κάποια από τα χαρακτηριστικά του ανθρώπινου συστήματος όρασης.
- Λόγω της αυξημένης ευαισθησίας του ανθρώπινου οφθαλμού στη φωτεινότητα, σε σχέση με την ευαισθησία του στα χρώματα και δεδομένης της ανάγκης για συμπίεση του σήματος μπορούμε να λαμβάνουμε λιγότερα δείγματα χρωματικής πληροφορίας σε σχέση με τα δείγματα τα οποία αφορούν τη φωτεινότητα.

υποδειγματοληψία χρώματος

- Το 4: 2 :2 δεν είναι ο μόνος συνδυασμός, είναι δυνατό να υπάρξουν πολλές περιπτώσεις και συνδυασμοί.
- Χρειάζεται ένας τρόπος περιγραφής αυτής της υποδειγματοληψίας.

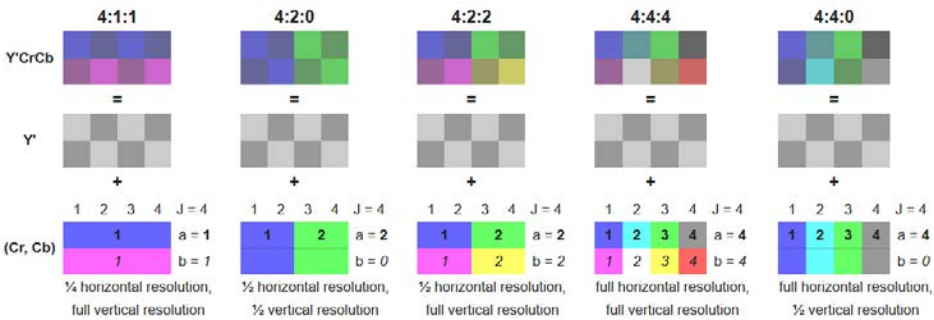


Περιγραφή Υποδειματοληψία Χρώματος

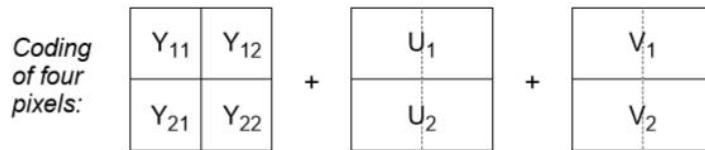
J: a: b

J: horizontal sampling reference (width of the conceptual region). Usually, 4.
a: number of chrominance samples (Cr, Cb) in the first row of *J* pixels.

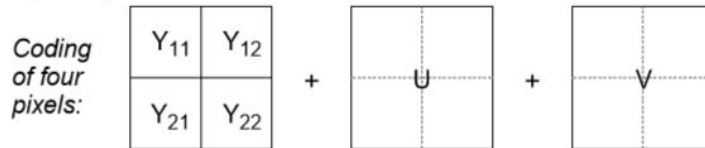
b: number of changes of chrominance samples (Cr, Cb) between first and second row of *J* pixels.

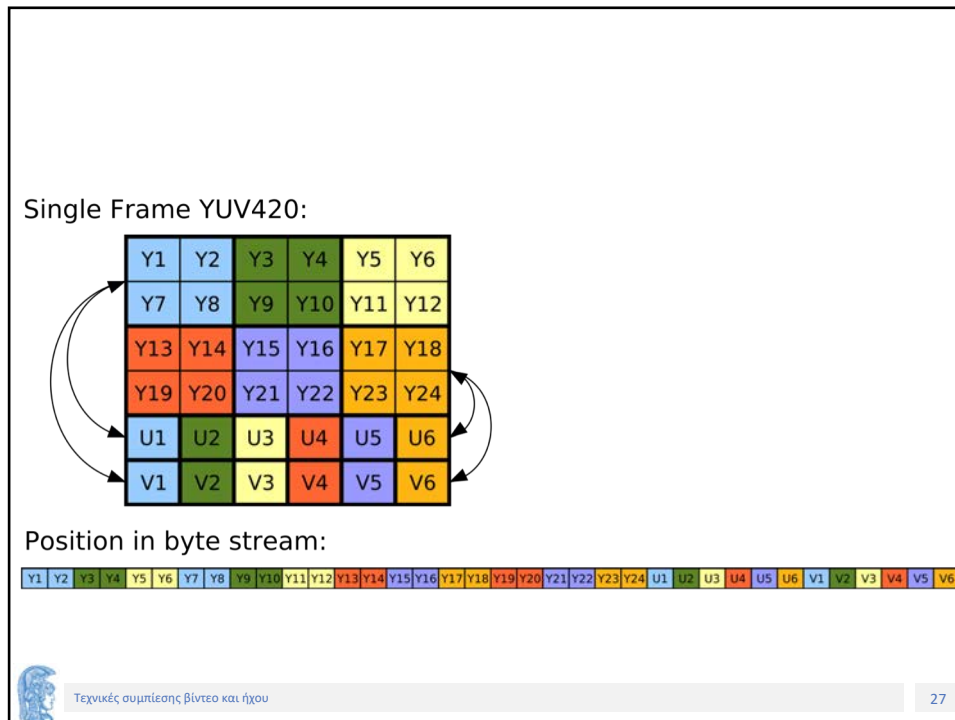


4:2:2: double resolution for luminance



4:1:1: quadrupel resolution for luminance





Ρυθμός μετάδοσης 4:2:0

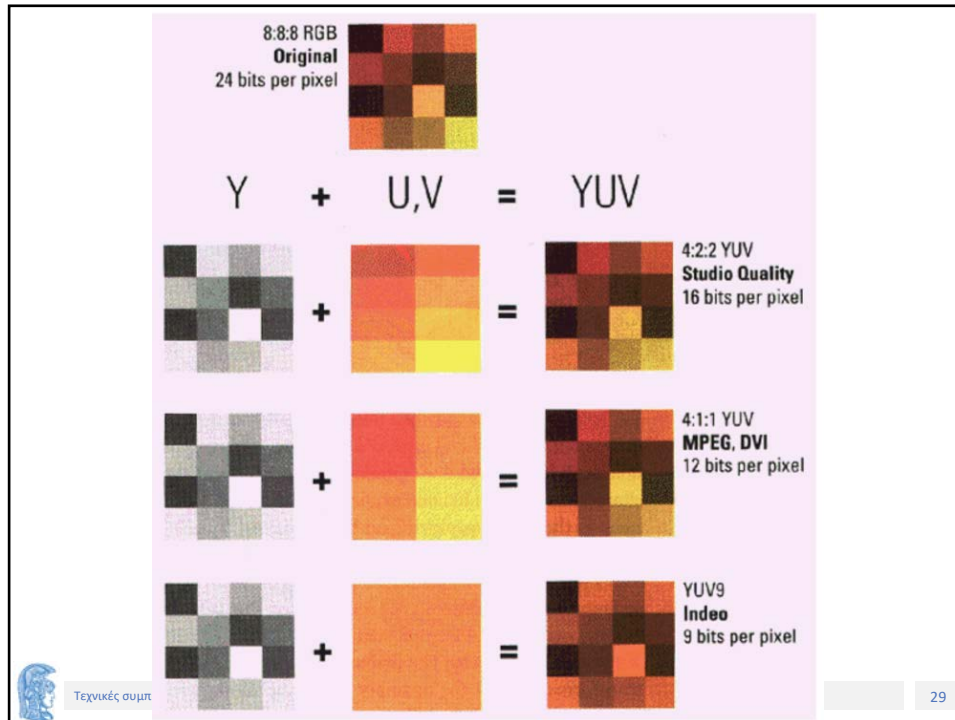
- Η υποδειγματοληψία του χρώματος χαμηλώνει τις απαιτήσεις αναφορικά με τον ρυθμό μετάδοσης.
- Έστω υποδειγματοληψία χρώματος της μορφής **4:2:0** και κβάντιση με 8 bits για τις τρεις συνιστώσες (φωτεινότητα και χρωμοδιαφορές).

$$Y: 13,5 \text{ MHz} * 8 \text{ bits / sample} = 108 \text{ Mbps}$$

$$Cb: (1/4) * 6,75 \text{ MHz} * 8 \text{ bits / sample} = 13,5 \text{ Mbps}$$

$$Cr: (1/4) * 6,75 \text{ MHz} * 8 \text{ bits / sample} = 13,5 \text{ Mbps}$$

ΑΘΡΟΙΣΜΑ: **135 Mbps**



Χαρακτηριστικά δειγματοληψίας και κβαντισμού τηλεοπτικού σήματος κατά το πρότυπο ITU-BT R.601

Παράμετρος	Σύστημα 625 γραμμών 25 πλαίσια / δευτερόλεπτο (50 πεδία / δευτερόλεπτο)	Σύστημα 525 γραμμών 30 πλαίσια / δευτερόλεπτο (60 πεδία / δευτερόλεπτο)
Δείγματα φωτεινότητας ανά γραμμή	864	858
Δείγματα χρωμοδιαφορών ανά γραμμή	432	429
Δομή	Ορθογωνική δομή γραμμής, πεδίου και πλαισίου. Τα δείγματα των χρωμοδιαφορών παρατίθενται μαζί με τα δείγματα φωτεινότητας περιττού αριθμού (1ο, 3ο κ.λπ.).	
Ρυθμός δειγματοληψίας φωτεινότητας	13,5 MHz	
Ρυθμός δειγματοληψίας χρωμοδιαφορών	6,75 MHz	
Κωδικοποίηση	Ομοιόμορφη κωδικοποίηση τύπου PCM με χρήση 8 ή 10 bits (με προτιμητέα την τιμή των 10 bits) για το σήμα της φωτεινότητας και των χρωμοδιαφορών.	
Πλήθος δειγμάτων φωτεινότητας ανά ενεργή ψηφιακή γραμμή	720	
Πλήθος δειγμάτων χρωμοδιαφορών ανά ενεργή ψηφιακή γραμμή	360	
Αντιστοίχιση επιπέδου τιμών σήματος φωτεινότητας και επιπέδων κβαντισμού	220 επίπεδα κβαντισμού, με εκκίνηση το 16 και το λευκό να αντιστοιχεί στο 235.	
Αντιστοίχιση επιπέδου τιμών σημάτων χρωμοδιαφορών και επιπέδων κβαντισμού	225 επίπεδα κβαντισμού, περί την κεντρική τιμή, με το μηδενικό σήμα να αντιστοιχεί στην τιμή 128.	



Συστήματα κατά EBU – TECH 3299

- Πρότυπο EBU – TECH 3299: High Definition (HD), Image Formats for Television Production

Συστήματα	Συντόμευση	Δείγματα φωτεινότητας ανά ενεργή γραμμή	Σύνολο γραμμών ανά πλαίσιο	Πλήθος ενεργών γραμμών	Συχνότητα πλαισίων
Σύστημα 1	720p/50	1.280	750	720	50
Σύστημα 2	1.080i/25	1.920	1.125	1.080	25 (50 για τα πεδία)
Σύστημα 3	1.080p/25	1.920	1.125	1.080	25
Σύστημα 4	1.080p/50	1.920	1.125	1.080	50

- συχνότητα δειγματοληψίας
 - φωτεινότητα: 74,25 MHz ή 72 MHz (ITU-R BT.709)
 - Χρώματα: $0,5 * 74,25 \text{ MHz} = 37.125 \text{ MHz}$ ή 36 MHz (ITU-R BT.709)
 - 4: 2: 2



Ρυθμοί βίντεο HD

$$Y: \quad 74,25 * 10 \text{ Mbit/s} = 742,5 \text{ Mbit/s}$$

$$Cb: \quad 0,5 * 74,25 * 10 \text{ Mbit/s} = 371,25 \text{ Mbit/s}$$

$$Cr: \quad 0,5 * 74,25 * 10 \text{ Mbit/s} = 371,25 \text{ Mbit/s}$$

$$\mathbf{1.485 \text{ Gbit/s}}$$

$$Y: \quad 72 * 10 \text{ Mbit/s} = 720 \text{ Mbit/s}$$

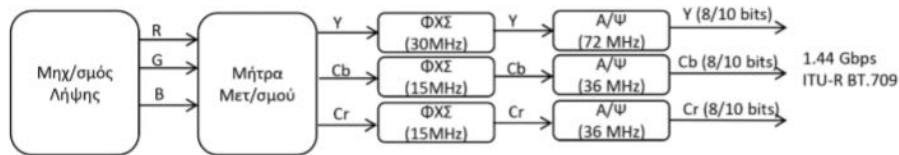
$$Cb: \quad 0,5 * 72 * 10 \text{ Mbit/s} = 360 \text{ Mbit/s}$$

$$Cr: \quad 0,5 * 72 * 10 \text{ Mbit/s} = 360 \text{ Mbit/s}$$

$$\mathbf{1,44 \text{ Gbit/s}}$$



Δημιουργία σήματος HD



Δομή σήματος βίντεο HD (μη συμπιεσμένο)

Βασικά χαρακτηριστικά:

- κατ' ελάχιστον υποστήριξη 720 γραμμών
- αναλογία 16:9
- ανάλυση 1.280 x 720 στα 50 Hz και τα 60 Hz με προοδευτική σάρωση 720p
- ανάλυση 1.980 x 1080 στα 50 Hz και τα 60 Hz για διαπλεγμένη σάρωση 1080i
- υποστήριξη ψηφιακών διεπαφών DVI (Digital Visual Interface) και HDMI
- κρυπτογράφηση τύπου HDCP στις ψηφιακές διεπαφές



Μετάδοση σήματος HD

- Η μετάδοση του σήματος υψηλής ευκρίνειας γίνεται με χρήση του πρωτοκόλλου **HD-SDI** (High Definition Serial Digital Interface)
- Η διεπαφή **DVI** (Digital Visual Interface) χρησιμοποιείται για τις διασυνδέσεις των υπολογιστών (τείνει να αντικαταστήσει τη γνωστή διεπαφή VGA). Υποστηρίζει ρυθμούς μέχρι και 1.65 Gbps.
- Το **HDMI** (High Definition Multimedia Interface) υποστηρίζει ρυθμούς μέχρι και 5 Gbps και μεταφέρει τόσο εικόνα όσο και ήχο.
- Το πρότυπο **HDCP** (High Bandwidth Digital Content Protection) προστατεύει το ψηφιακό υλικό που μεταφέρεται από τις διεπαφές DVI και HDMI από μη επιτρεπόμενη εγγραφή.



Τεχνικές συμπίεσης βίντεο

Παντελής Μπαλαούρας
Τμήμα Πληροφορικής & Τηλεπικοινωνιών

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

- 2 – 7 Mbps
- 38:1 ποσοστό συμπίεσης

Υψηλή ευκρίνεια (Full High Definition – FHD)

- άνω του 1 Gbps

Στόχος

- 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;

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**.
- 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.



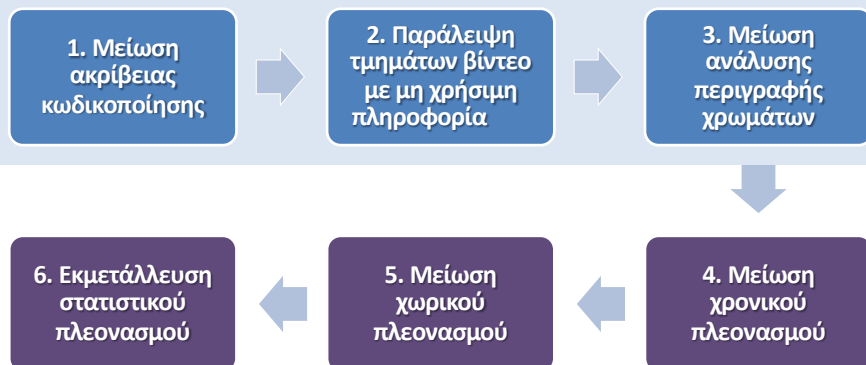
Το προς συμπίεση σήμα

- 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 (Cb και Cr).
- 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-BT.R60**



Βασικά βήματα στη συμπίεση βίντεο

Προετοιμασία σήματος βίντεο

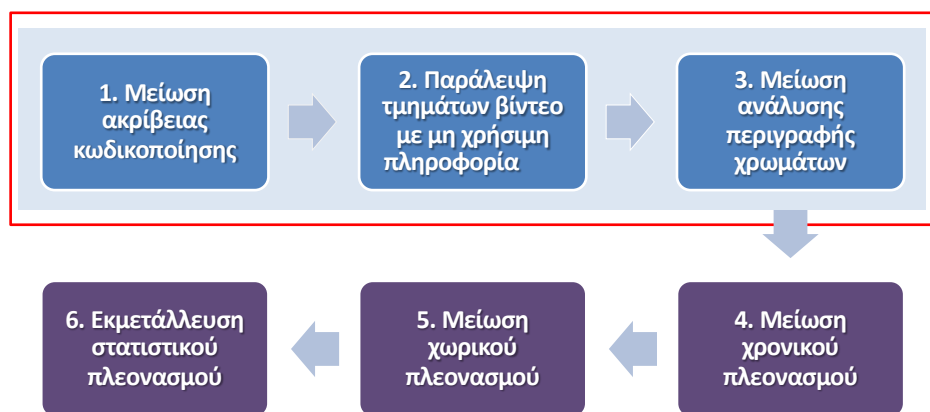


Ο μηχανισμός συμπίεσης του βίντεο

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



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).
- 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.**



Προετοιμασία σήματος βίντεο

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 μsec 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

3. Μείωση ανάλυσης περιγραφής χρωμάτων

- Υποδειγματοληψία 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.

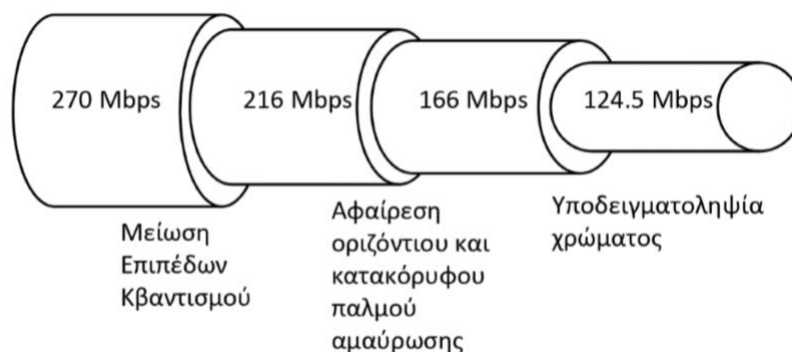
Από την πλευρά του απαιτούμενου ρυθμού έχουμε υποβάθμιση αφού 2*4 δείγματα φωτεινότητας αντιστοιχούν σε 2 δείγματα από τις χρωμοδιαφορές. Η υποβάθμιση του ρυθμού αντιστοιχεί στο 25%.



Προετοιμασία σήματος βίντεο

Συνολικός περιορισμός ρυθμού

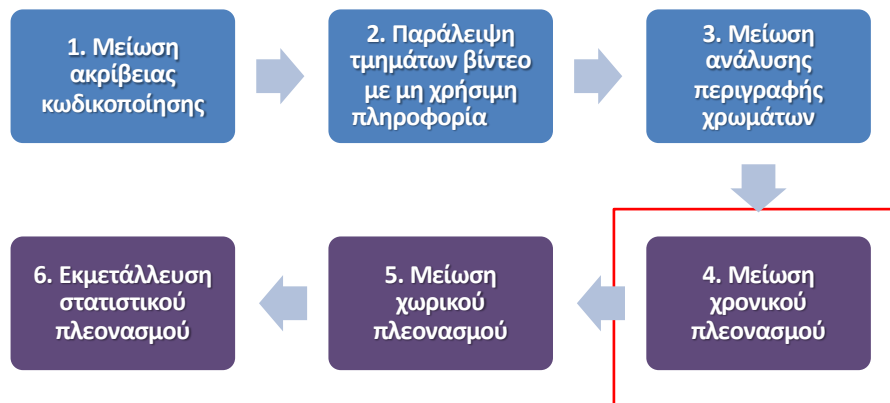
εικονολήπτης



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



4. Μείωση χρονικού πλεονασμού



4. Μείωση χρονικού πλεονασμού

- 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.
- Οι **διαφοροποιήσεις** μεταξύ των εικόνων μπορεί να είναι περιορισμένες και αφορούν το **δυναμικό κομμάτι**, σε αυτό, δηλαδή, που υπάρχει **κάποια κίνηση**
 - οι αλλαγές στις εικόνες μπορεί να προέρχονται από μετακινήσεις αντικειμένων μεταξύ πλαισίων.



Μηχανισμός περιορισμού χρονικού πλεονασμού

- Η αποστολή ολόκληρης της περιγραφής μιας εικόνας οδηγεί στην
 - επαναλαμβανόμενη αποστολή 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 αναγκαία πληροφορία σχετίζεται με τις **περιοχές που υπάρχουν διαφοροποιήσεις σε σχέση με τα προηγούμενα πλαίσια**
 - είναι απαραίτητη η αποστολή της διαφοράς αυτής
- Ο μηχανισμός περιλαμβάνει
 - τον διαχωρισμό της εικόνας σε υποπεριοχές
 - την κωδικοποίηση των εικόνων κατά I, P και B.



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)
 - αντιστοιχούν, δηλαδή, σε μια περιοχή 8x8 για κάθε μια από τις χρωμοδιαφορές Cb και Cr (64 τιμές)
 - το σύνολο των τιμών αυτών (256+64+64) αντιστοιχεί σε ένα **μακρο-μπλοκ**



Τεχνικές συμπίεσης βίντεο και ήχου

63

Κωδικοποίηση των εικόνων κατά I, P, B

- Ορισμένα από τα **πλαίσια κωδικοποιούνται ανεξάρτητα**, χωρίς δηλαδή να λαμβάνεται υπόψη πληροφορία από άλλα πλαίσια
 - I-frames (intracoded frames)
 - κωδικοποιούνται και αποστέλλονται περιοδικά



Τεχνικές συμπίεσης βίντεο και ήχου

64

Κωδικοποίηση των εικόνων κατά I, P, B

- Κάποια πλαίσια κωδικοποιούνται **λαμβάνοντας υπόψη προηγούμενα πλαίσια**
- 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 μακρομπλοκ, δηλαδή σε τμήματα μεγέθους 16x16
- **P-frames (predicted frames)** δηλαδή πλαίσια πρόβλεψης
- 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)
- Η κωδικοποίηση των πλαισίων τύπου P απαιτεί **λιγότερα δεδομένα** σε σχέση με την κωδικοποίηση πλαισίων τύπου I



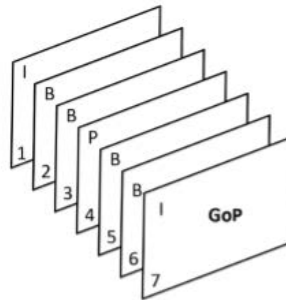
Κωδικοποίηση των εικόνων κατά I, P, B

- Υπάρχουν πλαίσια τα οποία κωδικοποιούνται αξιοποιώντας τον εντοπισμό ομοιοτήτων με **προηγούμενα** αλλά και **επόμενα πλαίσια**
 - από ένα προηγούμενο/επόμενο (χρονικά) πλαίσιο τύπου I ή P
- 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)



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 μίαν 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.



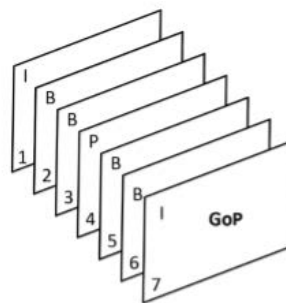
Τεχνικές συμπίεσης βίντεο και ήχου

67

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**.



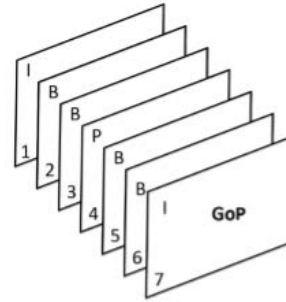
Τεχνικές συμπίεσης βίντεο και ήχου

68

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

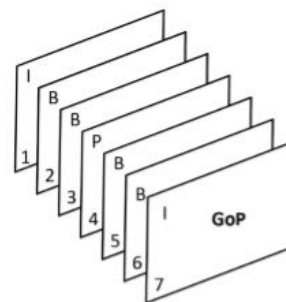


Τεχνικές συμπίεσης βίντεο και ήχου

69

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**.



Τεχνικές συμπίεσης βίντεο και ήχου

70

Αναζήτηση ομοιοτήτων μεταξύ πλαισίων

- Η αναζήτηση γίνεται σε επίπεδο μακρο-μπλοκ με **σύγκριση των μακρο-μπλοκ γειτονικών πλαισίων**.
- Κάθε μακρο-μπλοκ σε μια εικόνα εξετάζεται ώστε να διαπιστωθεί:
 1. **Εάν είναι νέο** και δεν συναντάται στην προηγούμενη εικόνα
 - Κωδικοποιείται εκ νέου
 2. Εάν **είναι όμοιο** με το μακρο-μπλοκ στην αντίστοιχη θέση του προηγούμενου πλαισίου (με κάποιες μικρές διαφοροποιήσεις).
 - Δεν απαιτείται διάνυσμα κίνησης
 - Μετάδοση της διαφοράς τιμών
 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
 - Μετάδοση διαφοράς τιμών



Τεχνικές συμπίεσης βίντεο και ήχου

71

Διάνυσμα κίνησης

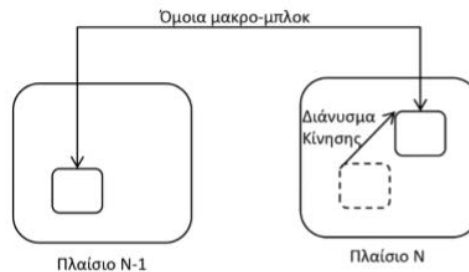
- Για τον υπολογισμό του διανύσματος κίνησης, ο μηχανισμός ξεκινάει από το μακρο-μπλοκ του τρέχοντος πλαισίου
 - εάν είναι **τύπου P** ελέγχει στο προηγούμενο P ή I πλαίσιο για παρόμοια μακρο-μπλοκ
 - εάν είναι **τύπου B** ελέγχει στο προηγούμενο και στο επόμενο πλαίσιο για παρόμοια μακρο-μπλοκ
- Η αναζήτηση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



Τεχνικές συμπίεσης βίντεο και ήχου

72

Διάνυσμα κίνησης



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



Αναζήτηση ομοίων περιοχών

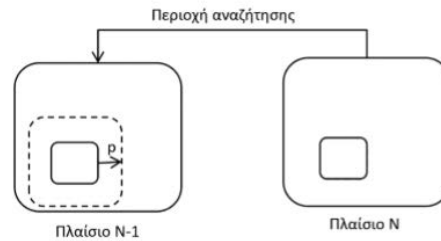
- Σε επίπεδο **εικονοστοιχείου** (pixel) με μεγάλο όμως πλήθος υπολογισμών, γεγονός το οποίο την καθιστά ασύμφορη.
- Σε επίπεδο **τμήματος της εικόνας** (block-based),
 - η πιο απλή περίπτωση και χρησιμοποιείται όπως προαναφέραμε κυρίως σήμερα και μάλιστα με χρήση αλγορίθμων που εκτελούνται με χρήση 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.



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).



Τεχνικές συμπίεσης βίντεο και ήχου

75

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



Τεχνικές συμπίεσης βίντεο και ήχου

76

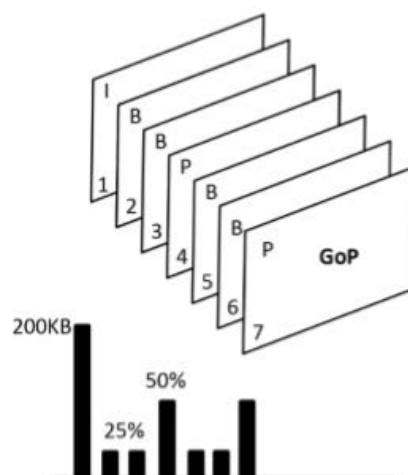
Παράδειγμα διαχείρισης 14 πλαισίων

	Ομάδα Πλαισίων GoP: IBBPBBP													
Αύξαν 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

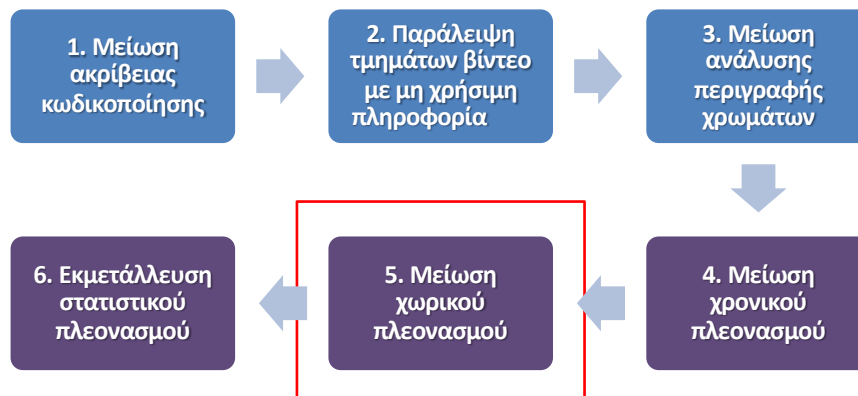
Η επαναφορά των πλαισίων στην ορθή σειρά γίνεται με χρήση του πεδίου χρονοσήμανσης αποκωδικοποίησης (Decoding Time Stamp, DTS) το οποίο βρίσκεται εντός της επικεφαλίδας του πακετοποιημένου στοιχειώδους ρυθμού.



Όφελος από την εκμετάλλευση του χρονικού πλεονασμού



5. Μείωση χωρικού πλεονασμού



5. Μείωση χωρικού πλεονασμού

- Ο **χωρικός πλεονασμός** προέρχεται από τις ομοιότητες σε διαφορετικές περιοχές εντός μιας εικόνας.
- Το φαινόμενο αυτό είναι 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

- Πραγματοποιείται με κωδικοποίηση μετασχηματισμού, ο οποίος μετασχηματίζει την πληροφορία από το **αρχικό πεδίο** σε ένα **ενδιάμεσο πεδίο**, κατάλληλο για συμπίεση.
- Στο δεύτερο αυτό πεδίο
 - επιλέγονται και κωδικοποιούνται οι **συντελεστές του μετασχηματισμού**, οι οποίοι 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.



Διακριτός Μετασχηματισμός Συνημιτόνου

- Ο βασικός αλγόριθμος που χρησιμοποιείται στο JPEG είναι ο Διακριτός Μετασχηματισμός Συνημιτόνου (Discrete Cosine Transform, DCT),
- Ο DCT χρησιμοποιείται και για την κωδικοποίηση των εικόνων 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 με την έννοια ότι πραγματοποιούνται σε μικρό σχετικά αριθμό εικονοστοιχείων, με αποτέλεσμα να αντιστοιχούν σε **υψηλές χωρικές συχνότητες**)
- Αυτό, σε αντιπαράθεση με τις **λιγότερο απότομες διαφοροποιήσεις** εντός της εικόνας, αυτές δηλαδή που γίνονται σε μεγαλύτερες επιφάνειες εικονοστοιχείων (και αντιστοιχούν σε **χαμηλότερες χωρικές συχνότητες**) τις οποίες το ανθρώπινο μάτι είναι σε θέση να αντιληφθεί επακριβώς.



Ιδέα

- Με βάση τα παραπάνω, οι πιο χαμηλές (χωρικές) συχνότητες χρειάζεται να κωδικοποιηθούν με μεγαλύτερη λεπτομέρεια (κυρίως όσον αφορά την κβάντιση), ενώ 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.



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).

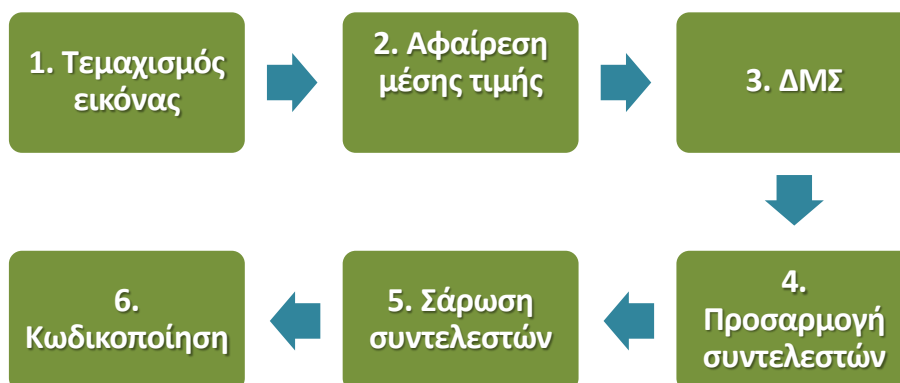


Γιατί ΔΜΣ (DCT);

- 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.

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

4. η προσαρμογή των συντελεστών με βάση την ευαισθησία του ανθρώπινου οφθαλμού, αλλά και του επιθυμητού επιπέδου ποιότητας.
5. η 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.
6. Κωδικοποίηση/συμπύεση



Προσαρμογή των συντελεστών

- Στο βήμα 4 η προσαρμογή των συντελεστών γίνεται με βάση έτοιμους τους **πίνακες, 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.**



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% (Q₅₀)

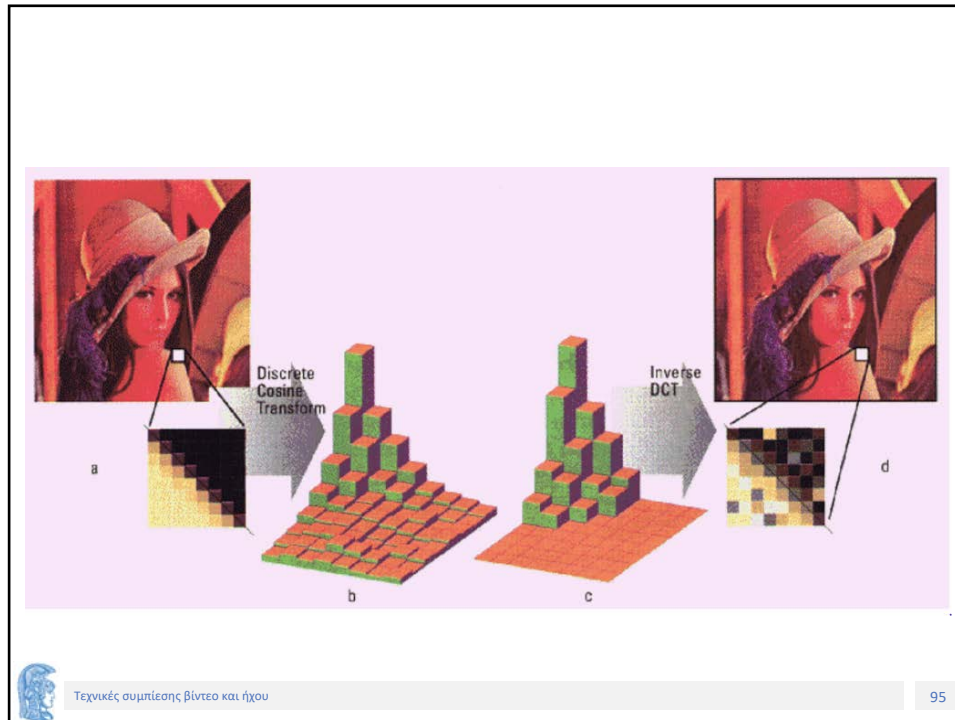


- **Βήμα 4: Κλιμάκωση/προσαρμογή των συντελεστών για ποιότητα 50**

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

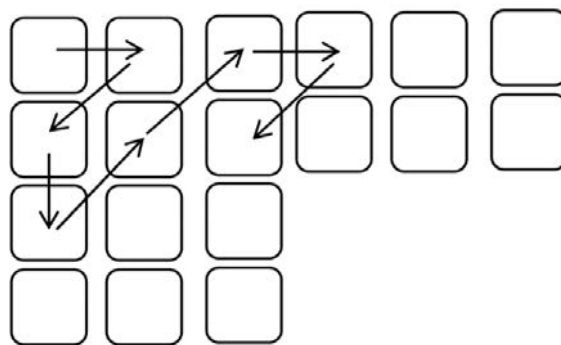
Τιμές μετά την κλιμάκωση





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: Κωδικοποίηση** συντελεστών

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



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

Προκύπτει ο πίνακας φωτεινότητας.



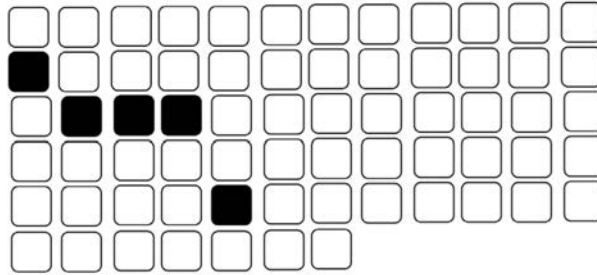
Κωδικοποίηση / συμπίεση συντελεστών (Βήμα 6)

- 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.



- WWWWWWWWWWWWWWWBWWWWWWWWWWWWWWWBBBWWWWWWWWW
WWWWWWWWWWWWWWWWWWWWWBWWWWWWWWWWWWWWWWWWW
- (W, 12), (B, 1), (W, 12), (B, 3), (W, 24), (B, 1), (W, 14)



Τεχνικές συμπίεσης βίντεο και ήχου

105

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



Τεχνικές συμπίεσης βίντεο και ήχου

106



6. Εκμετάλλευση στατιστικού πλεονασμού

- Η εκμετάλλευση του στατιστικού πλεονασμού αποτελεί μια **μη απωλεστική** 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 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

Τεχνικές συμπίεσης βίντεο και ήχου 108

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



Τεχνικές συμπίεσης βίντεο και ήχου

109

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. 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 συμβόλων είναι που εκμεταλλεύεται ο αλγόριθμος Huffman, τον οποίο θα περιγράψουμε στη συνέχεια, αλλά και άλλοι συναφείς αλγόριθμοι.



Τεχνικές συμπίεσης βίντεο και ήχου

110

Κωδικοποίηση Huffman

- Έστω τα k 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 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:

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



- 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

Βήμα 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).

Βήμα 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.

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



Παράδειγμα Βήμα 1

Βήμα 1:

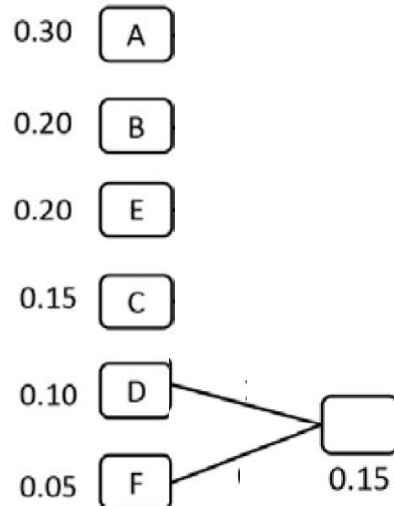
- Δημιουργούνται οι αρχικοί κόμβοι (που αποτελούν τα φύλλα του δένδρου) από τα σύμβολα με αύξουσα ή φθίνουσα σειρά.
- Αναγράφεται κάθε ένα από αυτά το βάρος του.

0.30 0.20 0.20 0.15 0.10 0.05 

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.



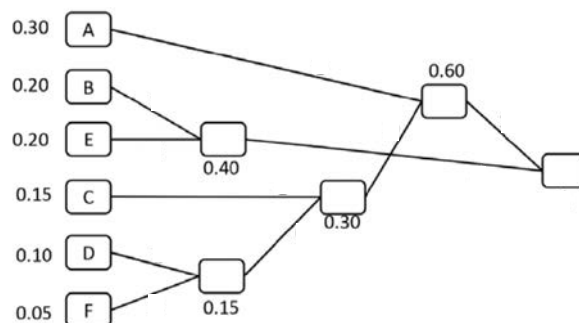
Τεχνικές συμπίεσης βίντεο και ήχου

117

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.



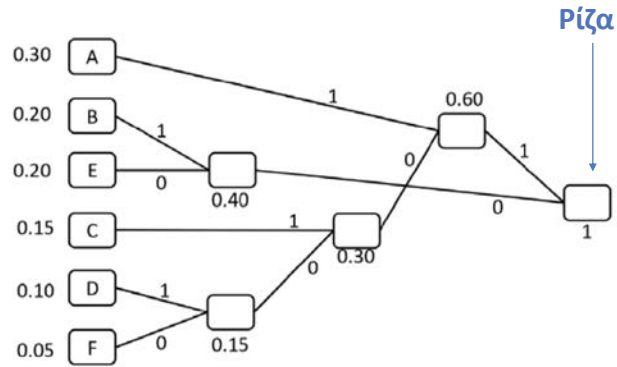
Τεχνικές συμπίεσης βίντεο και ήχου

118

Παράδειγμα Βήμα 4

Βήμα 4:

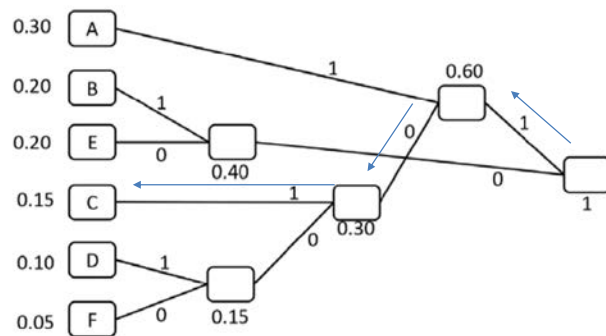
- Ξεκινώντας από τη ρίζα του δένδρου βρίσκουμε τις διαδρομές προς τα φύλλα και σε κάθε κλάδο τοποθετούμε το 0 ή το 1.



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.

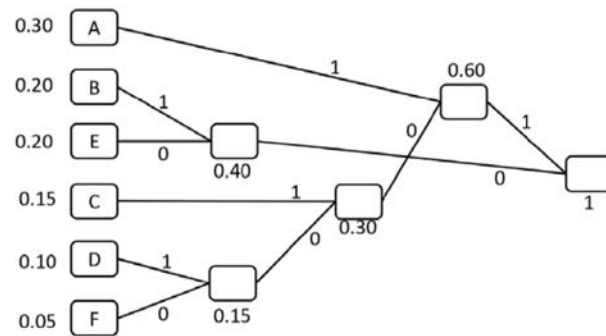


C: 101



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**.



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**.



προφίλς και επίπεδα

Το MPEG έχει δημιουργήσει μια σειρά από προφίλ (profile) και επίπεδα (level) 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** signal είναι το υψηλό προφίλ στο υψηλό επίπεδο **High Profile@High Level**.



Προφίλς

Συν/ση	Όνομασία	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	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



Τεχνικές συμπίεσης βίντεο και ήχου

129

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 χαμηλή σχετικά ποιότητα
- δεύτερο επίπεδο κωδικοποίησης:
 - καλύτερη ποιότητα
 - επαύξηση της ποιότητας επί της βασικής κωδικοποίησης



Τεχνικές συμπίεσης βίντεο και ήχου

130

Κλιμάκωση στο MPEG2

- **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 οι συντελεστές DCT με χαμηλή ανάλυση
 - 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).



Κλιμάκωση στο MPEG2

- **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
- **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 Αναζεύξης 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



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





ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Συμπίεση Ήχου

Παντελής Μπαλαούρας
Τμήμα Πληροφορικής & Τηλεπικοινωνιών

Πηγές

[1] Παπαδάκης, Α. 2015. Παραγωγή Τηλεοπτικού Σήματος. Συμπίεση Ήχου *τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 4. Διαθέσιμο στο: <http://hdl.handle.net/11419/5009>



Εισαγωγή

Εισαγωγή

- Το ανθρώπινο αυτί
 - έχει δυναμικό εύρος ακοής περί τα 140 dB
 - μπορεί να ακούει σε ένα εύρος ζώνης από τις πολύ χαμηλές συχνότητες μέχρι περίπου τα 15 KHz με 20 kHz.
- Τα συστήματα ήχου στα τηλεοπτικά συστήματα πρέπει να καλύπτουν αυτές τις απαιτήσεις.
- Το σήμα του ήχου έχει σημαντικές απαιτήσεις αναφορικά με το εύρος ζώνης οπότε και για το σήμα του ήχου **απαιτείται συμπίεση και κωδικοποίηση.**



Εισαγωγή

- Η συμπίεση, αντίστοιχα με τη λογική που ακολουθήθηκε στην περίπτωση του βίντεο, στηρίζεται στην αφαίρεση
 - της πλεονάζουσας πληροφορίας (**μη απωλεστική**)
 - της πληροφορίας η οποία δεν είναι μεν πλεονάζουσα αλλά η απώλειά της δεν γίνεται εύκολα αντιληπτή από τον χρήστη (**απωλεστική**)
- Οι μέθοδοι συμπίεσης στηρίζονται στο **ψυχοακουστικό μοντέλο**
 - αξιοποιεί κάποια βασικά χαρακτηριστικά του ανθρώπινου αυτιού, αναφορικά με την ακοή.
 - Το πιο σημαντικό ίσως από αυτά είναι η κάλυψη (masking) η οποία συμβαίνει στο πεδίο του χρόνου αλλά και στο πεδίο της συχνότητας.



Τεχνικές συμπίεσης βίντεο και ήχου

139

Κάλυψη (masking)

Κάλυψη στο πεδίο του χρόνου

- Ένας ήχος είναι πολύ κοντινός χρονικά σε κάποιον άλλο. Ο μεταγενέστερος ήχος, υπό ορισμένες συνθήκες, δεν θα γίνει αντιληπτός, θα επικαλυφθεί δηλαδή από αυτόν που προηγήθηκε χρονικά.

Κάλυψη στο πεδίο της συχνότητας

- Έχουμε στην περίπτωση που δύο ήχοι γειτνιάζουν στο πεδίο της συχνότητας (τα φάσματά τους, δηλαδή, έχουν πολύ μικρή απόσταση μεταξύ τους).
- Αυτό έχει σαν αποτέλεσμα ο ένας (αυτός που έχει την υψηλότερη ένταση) να επισκιάζει τον άλλο.
- Για την πρακτική εκμετάλλευση αυτού του χαρακτηριστικού, το φασματικό περιεχόμενο του ήχου χωρίζεται σε μπάντες και εξετάζεται η επίδραση που μπορεί να έχει καθεμία από αυτές στις υπόλοιπες.
- Για την πιο εύκολη διαχείριση του φάσματος, όπως θα δούμε στη συνέχεια, το φάσμα χωρίζεται σε επιμέρους τμήματα της κάθε μπάντας (τις οποίες ονομάζουμε υπομπάντες) το πλήθος των οποίων είναι 32.



Τεχνικές συμπίεσης βίντεο και ήχου

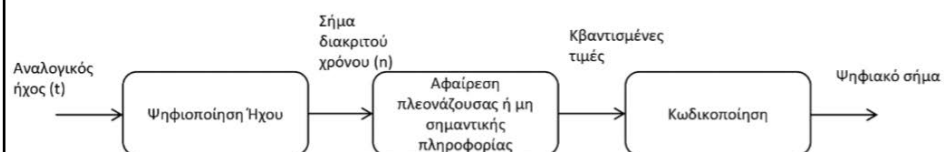
140

Χρήση μετασχηματισμών

- Διακριτού Μετασχηματισμού Συνημιτόνου (Discrete Cosine Transform)
- Τροποποιημένο Διακριτό Μετασχηματισμό Συνημιτόνου (Modified Discrete Cosine Transform)
 - διαφοροποιήσεις στη διαχείριση των άκρων των διαστημάτων όπου εφαρμόζεται



Δημιουργία ψηφιακού ήχου



Πρότυπα

MPEG (Motion Pictures Experts Group) - part 3, θέματα του ήχου

- MPEG-1 audio αποτελείται από τρία επίπεδα (layers) I, II και III. Το επίπεδο III δίνει το γνωστό πρότυπο **mp3** (MPEG-1, layer III).
- MPEG-2
 - 3 επίπεδα του MPEG-1
 - επίπεδο II επεκτάθηκε για να δημιουργήσει το επίπεδο **IIIMC** (multichannel)
 - **AAC** ISO/IEC 13818-7 (Advanced Audio Coding)
- MPEG-4 ISO/IEC 14496-3 **AAC** και **AACPlus**



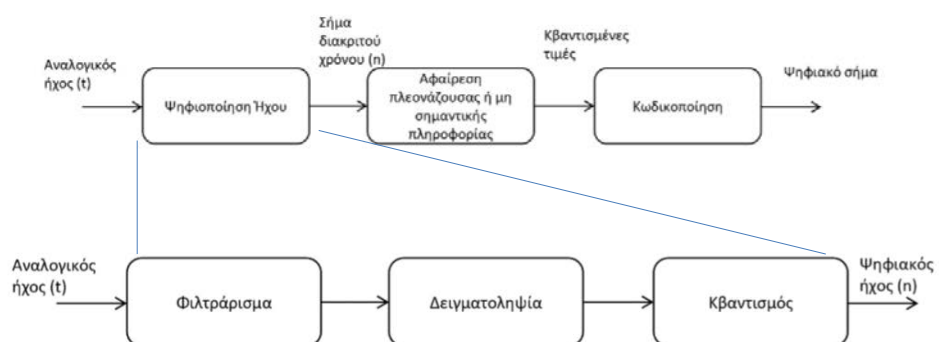
Πρότυπα

- Dolby digital audio (AC-3 audio)
- Από πλευράς ποιότητας, οι δύο βασικές κωδικοποιήσεις (του MPEG και των Dolby Labs) θεωρούνται πλέον ισοδύναμες.
- Οι αποκωδικοποιητές υποστηρίζουν και τις δύο κωδικοποιήσεις.



Ψηφιοποίηση του Ήχου

Βασικά Στάδια Ψηφιοποίησης



Φιλτράρισμα

- Το αναλογικό σήμα περιορίζεται, όσον αφορά το εύρος ζώνης του, με χρήση ενός **Φίλτρου Χαμηλών Συχνοτήτων (ΦΧΣ)**.
- Το ΦΧΣ επιτρέπει τη διέλευση των χαμηλών συχνοτήτων μέχρι κάποια συγκεκριμένη συχνότητα, **τη συχνότητα αποκοπής**, η οποία αποτελεί και βασικό χαρακτηριστικό του φίλτρου.
- Στην περίπτωση του τηλεοπτικού σήματος **η συχνότητα αποκοπής** βρίσκεται στο διάστημα από 15 KHz έως και 20 KHz.
- Η τιμή των 20 KHz προκύπτει από το άνω όριο των συχνοτήτων που είναι (θεωρητικά) σε θέση να ακούει το ανθρώπινο αυτί.



Δειγματοληψία

- Μετατροπή του σήματος από σήμα συνεχούς χρόνου (στον συνεχή χρόνο θεωρούμε τη συνεχή ανεξάρτητη μεταβλητή t) σε σήμα διακριτού χρόνου (στο διακριτό χρόνο θεωρούμε τη διακριτή μεταβλητή n , η οποία παίρνει ακέραιες τιμές)
- Συχνότητα δειγματοληψίας: 32 kHz, 44.1 kHz, 48, 96 kHz
 - 44.1 KHz αντιστοιχεί στην ποιότητα του CD
 - 48 KHz ή 96 KHz δίνουν ποιότητα αντίστοιχη αυτής του τηλεοπτικού στούντιο.
 - 32 KHz δεν θεωρείται επαρκής αλλά εξακολουθεί να διατηρείται για συμβατότητα με το πρότυπο του MPEG



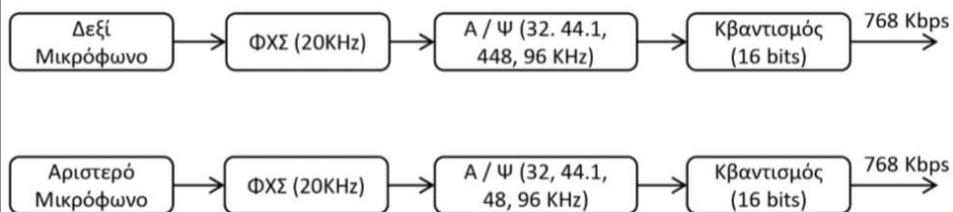
Κβαντισμός

- Η αντιστοίχιση των συνεχών τιμών των δειγμάτων σε διακριτές στάθμες.
- Η ποιότητα του κβαντισμού εξαρτάται από το πλήθος των bits που χρησιμοποιούνται για την αποτύπωση των τιμών.
- Το πλήθος των bits, έστω N , δίνει το πλήθος με τις στάθμες κβαντισμού 2^N . Ο κβαντισμός γίνεται με χρήση **κατ' ελάχιστον 16 bits**, γεγονός το οποίο δημιουργεί 2^{16} στάθμες τιμών.
- Ας υπολογίσουμε τον ρυθμό δεδομένων, C_m , ο οποίος προκύπτει με δειγματοληψία μονοφωνικού ήχου στα 48 KHz και κβάντιση με χρήση 16 bits.

$$C_m = 48 \cdot 10^3 \cdot 16 = 768000 = 768 \text{ Kbps}$$



Ψηφιοποίηση Στερεοφωνικού Ήχου



- Ο στερεοφωνικός ήχος αποτελείται από δύο κανάλια, δύο συσκευές λήψης οι οποίες λειτουργούν ταυτόχρονα, το δεξί και το αριστερό μικρόφωνο.
- Ο ρυθμός διπλασιάζεται, λόγω της ύπαρξης των δύο καναλιών.

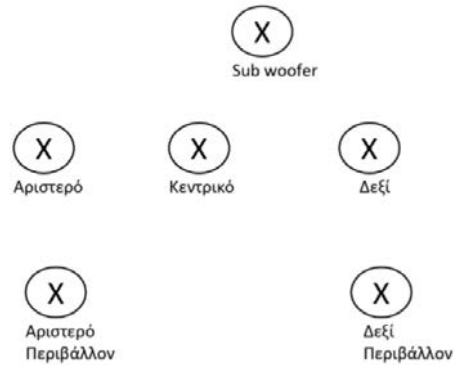
$$C_s = 2 \cdot 768 = 1.536 \text{ Kbps} = 1,5 \text{ Mbps}$$



Πολυκαναλικός Ήχος

Στη μετάδοση ήχου κατά 5.1 μεταδίδονται τα επόμενα κανάλια:

1. αριστερό
2. κέντρο
3. δεξί
4. αριστερό περιβάλλον (surround)
5. δεξί περιβάλλον (surround)
6. το ενισχυτικό των χαμηλών συχνοτήτων (Low Frequency Enhancement, LFE) κανάλι για χρήση σε sub-woofer



Τεχνικές συμπίεσης βίντεο και ήχου

151

Πολυκαναλικός Ήχος

- Η κωδικοποίηση και η συμπίεση γίνονται και μεταξύ των καναλιών που αποστέλλονται. Αυτό σημαίνει ότι τα κανάλια που μεταδίδονται αξιολογούνται για συσχετιζόμενη πληροφορία που δεν συνεισφέρει στο τελικό ηχητικό αποτέλεσμα.
- Πρότυπα:
 - MPEG2 layerII MC (πολυκαναλικός ήχος Multi Channel)
 - Dolby Digital 5.1 Surround.

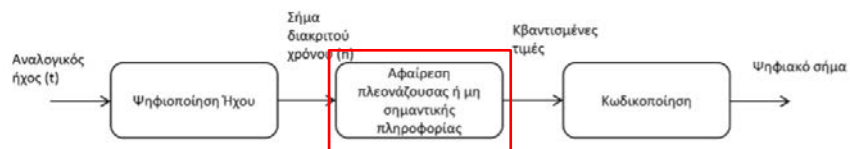


Τεχνικές συμπίεσης βίντεο και ήχου

152

Συμπίεση Ήχου

Στόχος



- Η ροή δεδομένων που προκύπτει από την ψηφιοποίηση του ήχου πρέπει να συμπιεστεί και ο αντίστοιχος ρυθμός να μειωθεί σημαντικά.
- 1.5 Mbps -> 100 Kbps έως 400 Kbps
- Τα ποσοστά συμπίεσης κυμαίνονται
 - από 16:1 (για τη χαμηλή ποιότητα του ήχου)
 - έως και 4:1 για την πιο υψηλή ποιότητα.



Ακουστικό Φάσμα και Ευαισθησία

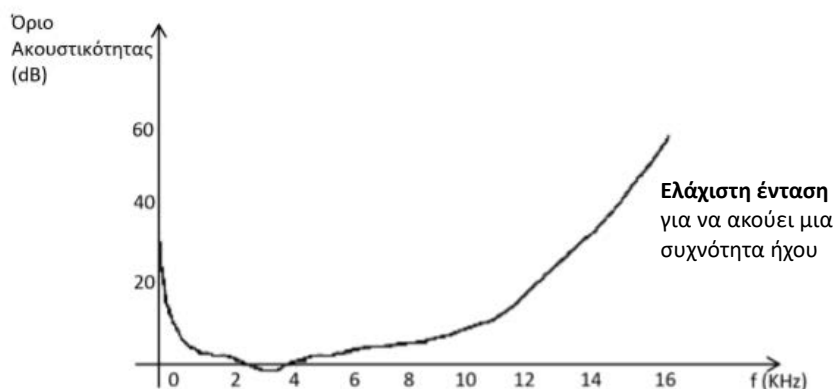
- Στη θεωρία το **εύρος ζώνης των ήχων** που γίνονται αντιληπτοί από το ανθρώπινο αυτί είναι
 - από τα 0 KHz έως και τα 20 KHz
- Σε κάθε περίπτωση ήχοι εκτός αυτού του εύρους ζώνης δεν γίνονται αντιληπτοί από τον άνθρωπο.
- Υπάρχουν περιορισμοί και με την **ένταση των ήχων** οι οποίοι μπορούν να γίνονται αντιληπτοί.
- Το όριο αυτό της έντασης, κάτω από το οποίο ένας ήχος δεν γίνεται αντιληπτός, διαφοροποιείται σε σχέση με τη συχνότητα.
 - Η μέγιστη ευαισθησία του ανθρώπου στον ήχο είναι στο εύρος μεταξύ 3 KHz και 4 KHz.
 - Παρατηρούμε ότι το όριο «ακουστικότητας» στις συχνότητες περί τα 3KHz παρουσιάζει το ολικό ελάχιστο και αυξάνει στις υπόλοιπες. Αυτό σημαίνει ότι στις συχνότητες αυτές μπορούμε να **ακούμε ήχους με την ελάχιστη ένταση**, ενώ για να μπορέσουμε να ακούσουμε ήχους στις υπόλοιπες συχνότητες απαιτείται μεγαλύτερη ένταση.



Τεχνικές συμπύεσης βίντεο και ήχου

155

Καμπύλη ευαισθησίας ανθρώπινου αυτιού σε σχέση με τις συχνότητες των ήχων



Τεχνικές συμπύεσης βίντεο και ήχου

156

Ήχοι που αγνοούνται

- Ήχοι που δεν γίνονται αντιληπτοί από τον άνθρωπο και αυτό συνεπάγεται ότι δεν χρειάζεται να κωδικοποιούνται
 - αγνοηθούν κατά τη διαδικασία της συμπίεσης
- Οι ήχοι των οποίων το φασματικό περιεχόμενο είναι εκτός του εύρους συχνοτήτων [20 Hz, 20 KHz]
 - Φίλτρο Χαμηλών Συχνοτήτων
 - Επιλογή της συχνότητας δειγματοληψίας ώστε να είναι κατ' ελάχιστον το διπλάσιο της μέγιστης τιμής της συχνότητας του ηχητικού σήματος που διέρχεται από το φίλτρο.
 - Π.χ εάν το φίλτρο επιτρέπει τη διέλευση συχνοτήτων μέχρι τα 20 KHz, η συχνότητα δειγματοληψίας πρέπει να είναι κατ' ελάχιστον 40 KHz ώστε να ικανοποιείται το θεώρημα του Shannon
- Οι ήχοι των οποίων η ένταση είναι κάτω από το ελάχιστο όριο ακοής για τις αντίστοιχες συχνότητες στις οποίες εκτείνονται
 - ανάγκη για μια βαθμίδα σύγκρισης της έντασης του εισερχόμενου ήχου με το όριο ακουσικότητας, στην αντίστοιχη συχνότητα.
 - Αυτό σημαίνει ότι εάν η ένταση κάποιου ήχου δεν υπερκαλύπτει αυτό το όριο, ο ήχος μπορεί να αγνοείται.

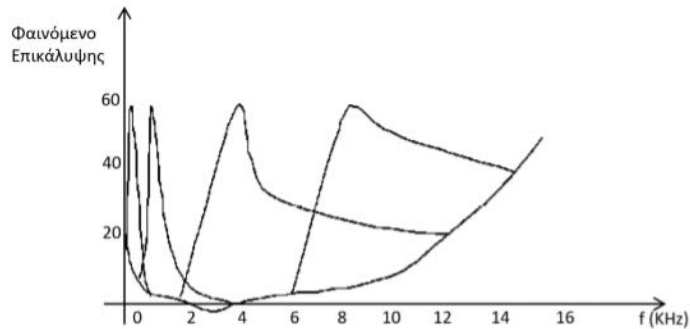


Επικάλυψη στο Πεδίο της Συχνότητας

- Εάν ακούγεται ένας ήχος σε συχνότητα π.χ. f_0 kHz με σταθερό πλάτος, τότε παρατηρήσουμε ότι άλλοι ήχοι στις συχνότητες γύρω από την f_0 για να γίνουν αντιληπτοί πρέπει να υπερβούν κάποιο συγκεκριμένο πλάτος (διαφορετικά επικαλύπτονται).
- Ο ήχος, δηλαδή, στη συχνότητα f_0 επικαλύπτει τους ήχους στις γύρω συχνότητες. Από μετρήσεις στο πλαίσιο πειραμάτων της ανθρώπινης ακοής, προκύπτουν οι επόμενες παρατηρήσεις:
 - Η μορφή της επικάλυψης εξαρτάται από τη συχνότητα του επικαλύπτοντος ήχου.
 - Όσο πιο υψηλή είναι η συχνότητα του επικαλύπτοντος ήχου τόσο μεγαλύτερο το (συχνοτικό) εύρος της επικάλυψης.



Επικάλυψη στο πεδίο της συχνότητας



- Εάν έχουμε έναν ήχο του οποίου η συχνότητα αντιστοιχεί στη συχνότητα του κεντρικού σημείου μιας καμπύλης και η ένταση του ήχου αντιστοιχεί στην κορυφή της καμπύλης, τότε ο ήχος αυτός επικαλύπτει το σύνολο των (τυχόν) ήχων οι οποίοι αντιστοιχούν, όσον αφορά τη συχνότητα και την ένταση στην επιφάνεια κάτω από τη συγκεκριμένη καμπύλη.
- Για να μπορέσει, δηλαδή, να γίνει αντιληπτός ένας ήχος (εάν η συχνότητά του αντιστοιχεί στο εύρος συχνοτήτων που καλύπτει η καμπύλη) πρέπει η ένταση αυτού να υπερβαίνει τον ορίζοντα της καμπύλης.



Τεχνικές συμπίεσης βίντεο και ήχου

159

Επικάλυψη στο Πεδίο του Χρόνου

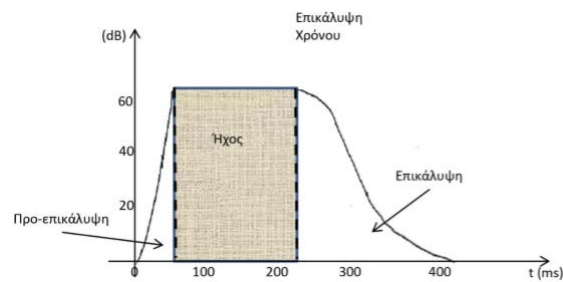
- Ένας έντονος (μεγάλης έντασης) ήχος επικαλύπτει (δεν επιτρέπει να γίνουν αντιληπτοί) ήχοι πριν και μετά από αυτόν εφόσον η ένταση αυτών των ήχων δεν ξεπερνάει κάποιο κατώφλι.
- Η επικάλυψη πριν το χρονικό διάστημα που αναπαράγεται ο ήχος οφείλεται **στην περιορισμένη ευαισθησία του ανθρώπινου αυτιού** (όσον αφορά την κατανόηση των χρονικών διαστημάτων που μεσολαβούν μεταξύ ήχων) αλλά και τον τρόπο με τον οποίο μεταδίδονται τα σήματα από το αυτί στον εγκέφαλο.



Τεχνικές συμπίεσης βίντεο και ήχου

160

Επικάλυψη στο Πεδίο του Χρόνου



Οικονομία λόγω επικάλυψης

- ήχοι που επικαλύπτονται δεν χρειάζεται να κωδικοποιούνται και να αποστέλλονται άρα πρέπει να αγνοούνται κατά τη συμπίεση.
- Προϋπόθεση να γίνει αντιληπτό το **υποσύνολο του ηχητικού περιεχομένου το οποίο επικαλύπτεται.**



Τεχνικές συμπίεσης βάση επικάλυψης

Οι μέθοδοι συμπίεσης του ήχου που χρησιμοποιούνται στους σύγχρονους αλγόριθμους συμπίεσης (μεταξύ των οποίων και στο MPEG) βασίζονται στην επικάλυψη

- κυρίως στο **πεδίο της συχνότητας**
- αλλά και σε αυτό **του χρόνου**

με τις τεχνικές που χρησιμοποιούνται να είναι πολύ παρόμοιες για τους διάφορους κωδικοποιητές.



Κβαντισμός και Θόρυβος Κβαντισμού

- Ο κβαντισμός (ή διακριτοποίηση) αποτελεί την αντιστοίχιση της τιμής του ηχητικού σήματος από την αναλογική μορφή σε συγκεκριμένες διακριτές στάθμες.
- Ο κβαντισμός καθορίζεται από το πλήθος των bits, έστω N , που αξιοποιούνται για την αναπαράσταση των τιμών του σήματος.
- Στην περίπτωση του ήχου, το πλήθος των bits που συνήθως χρησιμοποιείται είναι 16 (ενώ σε κάποιες προγενέστερες συνήθως περιπτώσεις ήταν 8 bits).
- Το πλήθος των bits καθορίζει το πλήθος από τις στάθμες κβαντισμού 2^N και αντίστοιχα τον θόρυβο κβαντισμού.

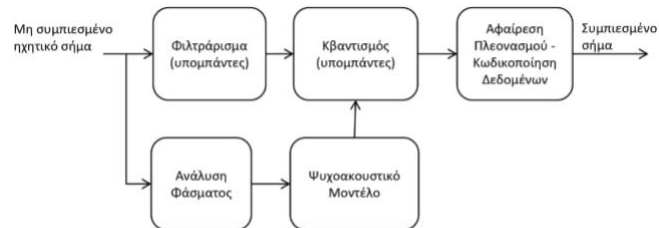


Κωδικοποίηση ήχου



Βήματα κωδικοποίησης ηχητικού σήματος

- Το ψηφιακό ηχητικό σήμα ακολουθεί δύο διακριτές διαδρομές στον κωδικοποιητή, οι οποίες αναπαρίστανται στο επόμενο σχήμα.



- Το σήμα αφού φιλτραρισθεί, οδηγείται **στον αναλυτή φάσματος**, όπου μετασχηματίζεται κατά Διακριτό Μετασχηματισμό Φουριέ (ΔΜΦ), ο οποίος υλοποιείται μέσω του αλγορίθμου του **Ταχέως Μετασχηματισμού Φουριέ**.
- Αφού πραγματοποιηθεί ο Ταχύς Μετασχηματισμός Φουριέ (Fast Fourier Transform, FFT) **ξεχωρίζουν τα τμήματα του ηχητικού σήματος και τα εύρη των συχνοτήτων που αυτά ανήκουν**.



Τεχνικές συμπίεσης βίντεο και ήχου

167

- Παράλληλα το ηχητικό σήμα **υφίσταται φιλτράρισμα** και χωρίζεται σε **υπομπάντες** (οι οποίες αφορούν συγκεκριμένες και περιορισμένες περιοχές συχνοτήτων).
- Οι υπομπάντες αυτές είναι εν γένει πιο εύκολα διαχειρίσιμες, όσον αφορά τη θεώρηση των **φαινομένων κάλυψης**, από το συνολικό ηχητικό σήμα (το οποίο εκτείνεται στο συνολικό εύρος ζώνης).



Τεχνικές συμπίεσης βίντεο και ήχου

168

Ιστορική εξέλιξη – MPEG-1

- Οι δύο τεχνικές, της διαίρεσης σε μπάντες και του μετασχηματισμού, χρησιμοποιήθηκαν στα πρώτα πρότυπα του MPEG που είχαν να κάνουν με τη συμπίεση του ήχου και συγκεκριμένα στο πρότυπο MPEG-1 όπως αυτό διαμορφώθηκε στις αρχές της δεκαετίας του 1990.
- Το πρότυπο MPEG-1 audio αποτελείται από τρία επίπεδα (layers) I, II και III. Το επίπεδο III δίνει το γνωστό πρότυπο mp3 (MPEG1, layer III). Το MPEG-2 υιοθέτησε τα τρία επίπεδα του MPEG-1 όσον αφορά τον ήχο και το επίπεδο II επεκτάθηκε για να δημιουργήσει το επίπεδο II MC, το οποίο είναι και το πολυκαναλικό (multichannel). Το πρότυπο ISO/IEC 13818-3 του MPEG-2 για τον ήχο υιοθετήθηκε το 1994.



Ιστορική εξέλιξη - MPEG-2

- Το MPEG-2 χρησιμοποιείται στην τηλεόραση τυπικής ευκρίνειας (εικόνας και ήχου) αλλά επεκτάθηκε και για την τηλεόραση υψηλής ευκρίνειας.
- Αυτό είχε ως αποτέλεσμα το MPEG-3 να εγκαταλειφθεί και να μην προχωρήσει περαιτέρω.
- Έκτοτε έχουν υπάρξει διαφοροποιήσεις με κυριότερες το MPEG-2 AAC ISO/IEC 13818-7 (όπου AAC είναι το Advanced Audio Coding) και το σύστημα ήχου του MPEG-4 ISO/IEC 14496-3: AAC και AAC Plus.



Ιστορική εξέλιξη

- Το πρότυπο **Advanced Audio Coding (AAC)** έχει σχεδιαστεί ως διάδοχος του προτύπου MPEG 1 & 3 Layer 3.
 - το AAC μπορεί να πετύχει καλύτερη ποιότητα με αντίστοιχο ρυθμό. Το AAC προτυποποιήθηκε από τους οργανισμούς ISO και IEC, ως τμήμα του **MPEG-2 (ως PART 7)** και συγκριμένα ως το πρότυπο ISO/IEC 13818-7:1997) και του **MPEG-4**.
- Αναφορικά με το MPEG 2, το Part 7 ήταν ένα νέο διαφορετικό τμήμα αφού πιο πριν ο ήχος εντάσσεται στο Part 3.
- Το Part 7 δεν είχε συμβατότητα προς τα πίσω. Δεν ήταν, δηλαδή, συμβατό με τα mp1, mp2 και mp3. Για αυτόν τον λόγο αναφερόταν ως μη συμβατό MPEG-2 NBC (Non-Backward Compatible)

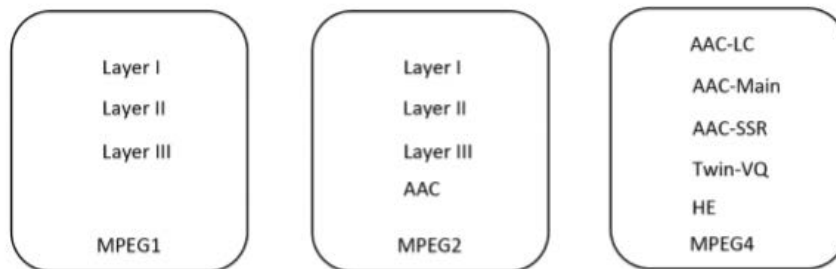


Ιστορική εξέλιξη

- Το MPEG-2 Part 7 όρισε τρία προφίλ:
 - χαμηλής πολυπλοκότητας (Low-Complexity profile, AAC-LC)
 - βασικό προφίλ (Main profile, AAC-Main)
 - κλιμακωτού ρυθμού δειγματοληψίας (Scalable Sampling Rate, AAC-SSR)
- Το MPEG-4 Part 3 περιλαμβάνει
 - τους προηγούμενους κωδικοποιητές MPEG-2 AAC
 - τους νέους MPEG-4 κωδικοποιητές ήχου
 - τον MPEG-4 HE (High Efficiency) AAC,
 - και άλλους όπως ο TWINVQ (Transform-domain Weighted Interleave Vector Quantization).



Εξέλιξη προτύπων κωδικοποίησης ήχου



Τυπικές κωδικοποιήσεις του ήχου

- Σε ένα τυπικό σύστημα ψηφιακής τηλεόρασης, η **επιλογή της δεν είναι αυστηρά προδιαγεγραμμένη** και μάλιστα σε πολλές περιπτώσεις διαφοροποιείται.
- Οι περισσότεροι Audioplayers υποστηρίζουν
 - MPEG-1 Layer III Audio
 - MPEG-2 Layer III Audio
 - AAC MPEG-2 και MPEG-4.



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:

[1] Παπαδάκης, Α. 2015. Παραγωγή Τηλεοπτικού Σήματος. *Ψηφιακή τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 2. Διαθέσιμο στο: <http://hdl.handle.net/11419/5007>

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

[3] Παπαδάκης, Α. 2015. Παραγωγή Τηλεοπτικού Σήματος. Συμπίεση Ήχου *τηλεόραση*. Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. κεφ 4. Διαθέσιμο στο: <http://hdl.handle.net/11419/5009>

