

Ψηφιακή Επεξεργασία Σήματος & Εικόνας

Μάθημα 24^ο

Ελληνικό Μεσογειακό Πανεπιστήμιο
Τμήμα Ηλεκτρονικών Μηχανικών

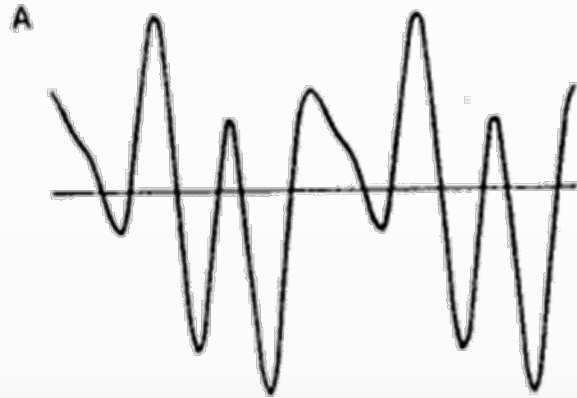
Ψηφιακή Επεξεργασία Εικόνας
(Μέρος 3^ο)

Α. Μπακλέζος

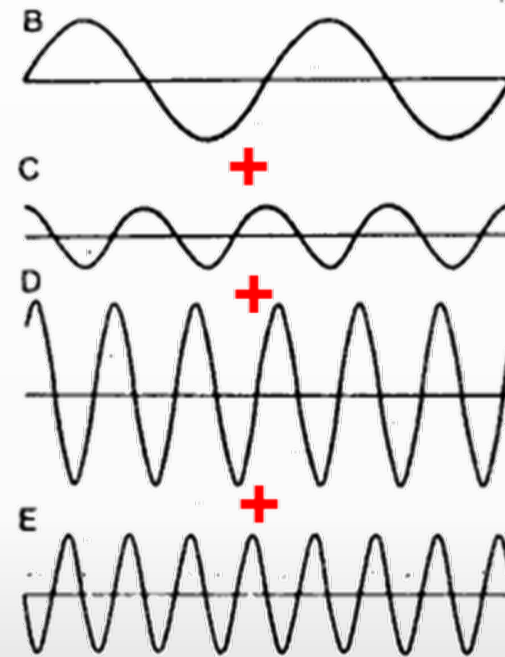
abaklezos@hmu.gr

Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Ανάλυση Fourier : κάθε σήμα (διδιάστατη εικόνα) μπορεί να εκφραστεί σαν άθροισμα ενός αριθμού απλών ημιτονοειδών σημάτων.

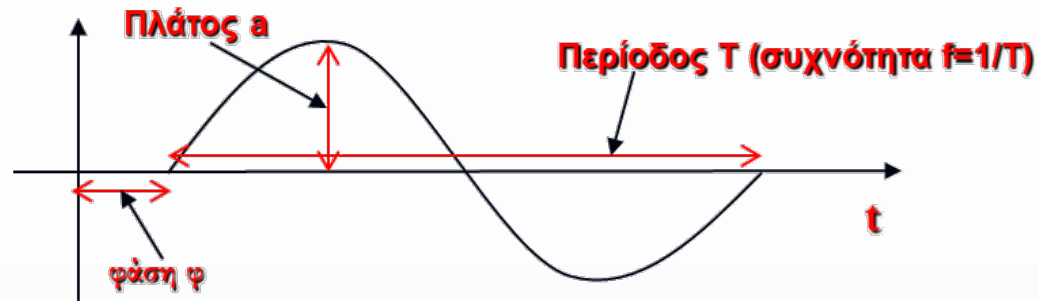


=

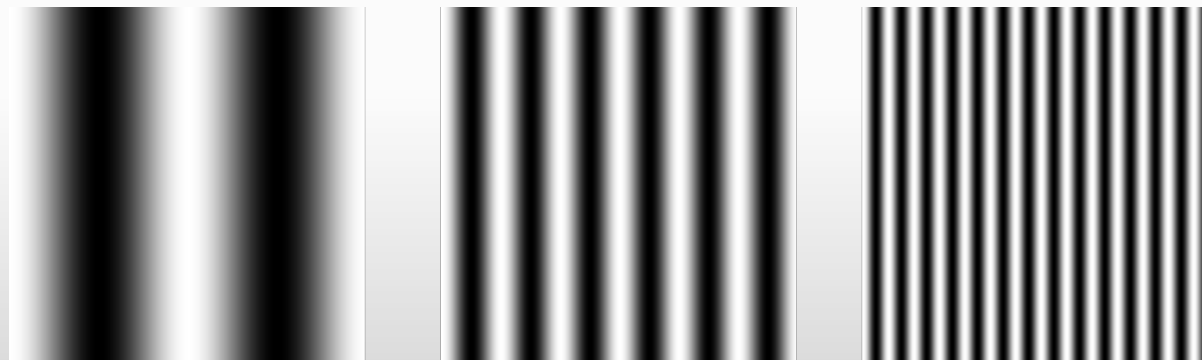


Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Ένα ημιτονοειδές σήμα περιγράφεται από την φάση, το πλάτος και την περίοδο (συχνότητα) του.



- Στην περίπτωση δισδιάστατων εικόνων ένα ημιτονοειδές σήμα εκφράζει μεταβολή φωτεινότητας στο χώρο της εικόνας (επίπεδο)



Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Ο DFT είναι ο δειγματοληπτημένος M/T Fourier και περιέχει μόνο τις συχνότητες με αρκετά μεγάλο πλάτος που αρκούν για να περιγραφεί ο χώρος της εικόνας.
- Το πλήθος των συχνοτήτων αντιστοιχούν στο πλήθος των pixel, δηλαδή η εικόνα στο χώρο των pixel και των συχνοτήτων έχουν το ίδιο μέγεθος.
- Για μια τετράγωνη εικόνα μεγέθους $N \times N$ pixel, ο 2-D DFT είναι:

$$F(k, l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-j2\pi(\frac{ki}{N} + \frac{lj}{N})}$$

- όπου $f(i, j)$ είναι η εικόνα στο χώρο των pixel (spatial domain) και ο εκθετικός όρος είναι η συνάρτηση βάσης που αντιστοιχεί σε κάθε σημείο $F(k, l)$ στο χώρο Fourier (συχνότητες k, l)
- Οι συναρτήσεις βάσης είναι ημίτονα και συνημίτονα με αυξανόμενες συχνότητες
- $F(0,0)$ είναι η DC συνιστώσα της εικόνας που αντιστοιχεί στην μέση φωτεινότητα
- $F(N - 1, N - 1)$ αντιστοιχεί στην μεγιστη συχνότητα

Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Η αντίστροφη διαδικασία για την επιστροφή στο χώρο της εικόνας (pixels) από το χώρο της συχνότητας γίνεται μέσω του αντίστροφου IDFT:

$$f(a, b) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k, l) e^{i2\pi(\frac{ka}{N} + \frac{lb}{N})}$$

- Ο όρος κανονικοποίησης $\frac{1}{N^2}$ σε ορισμένα βιβλία βρίσκεται στο ευθύ M/T αντί για τον αντίστροφο
- Χωρίς αποδειξη, ο DFT μπορεί να σπάσει σε δύο βήματα

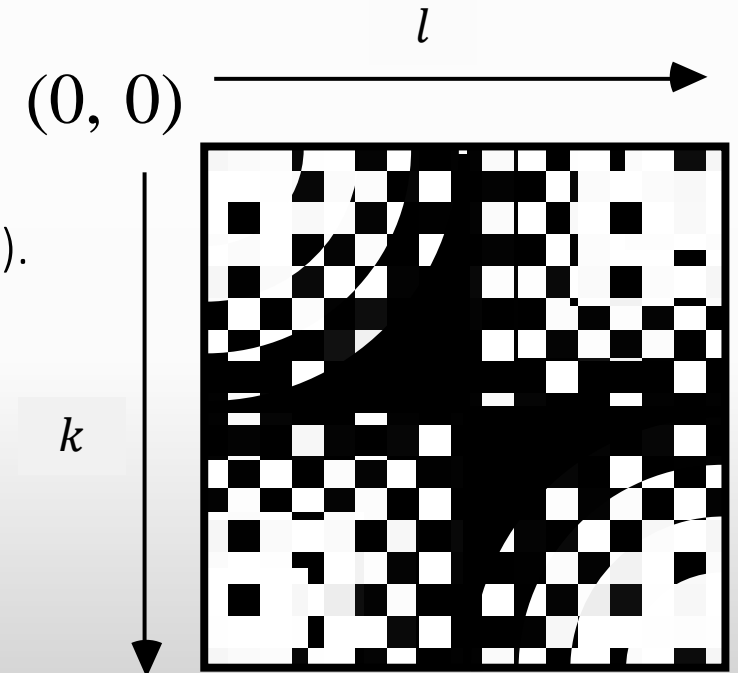
$$F(k, l) = \frac{1}{N} \sum_{b=0}^{N-1} P(k, b) e^{-i2\pi \frac{lb}{N}}$$

$$P(k, b) = \frac{1}{N} \sum_{a=0}^{N-1} f(a, b) e^{-i2\pi \frac{ka}{N}}$$

- τα οποία απαιτούν συνολικά $2N$ 1-D DFT και μειώνουν το πλήθος των συνολικών υπολογισμών.

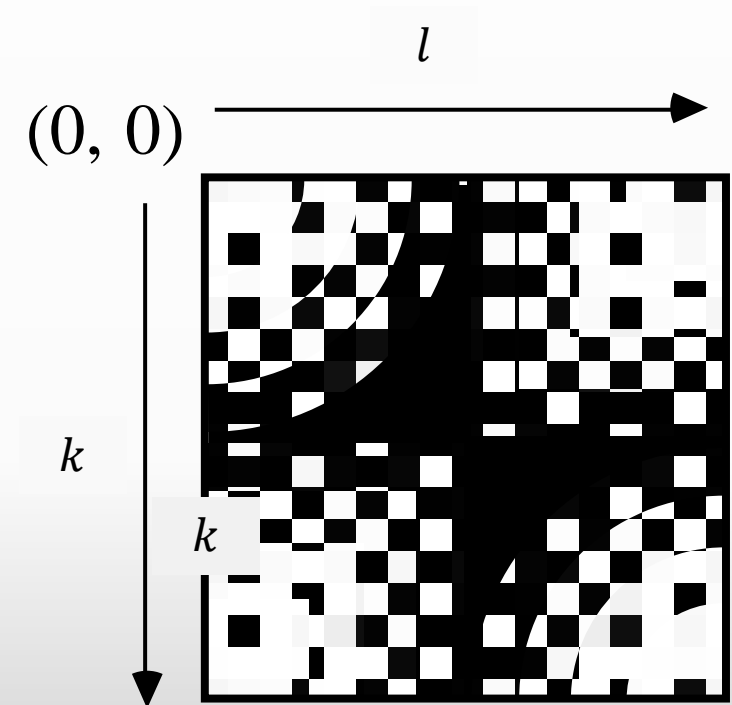
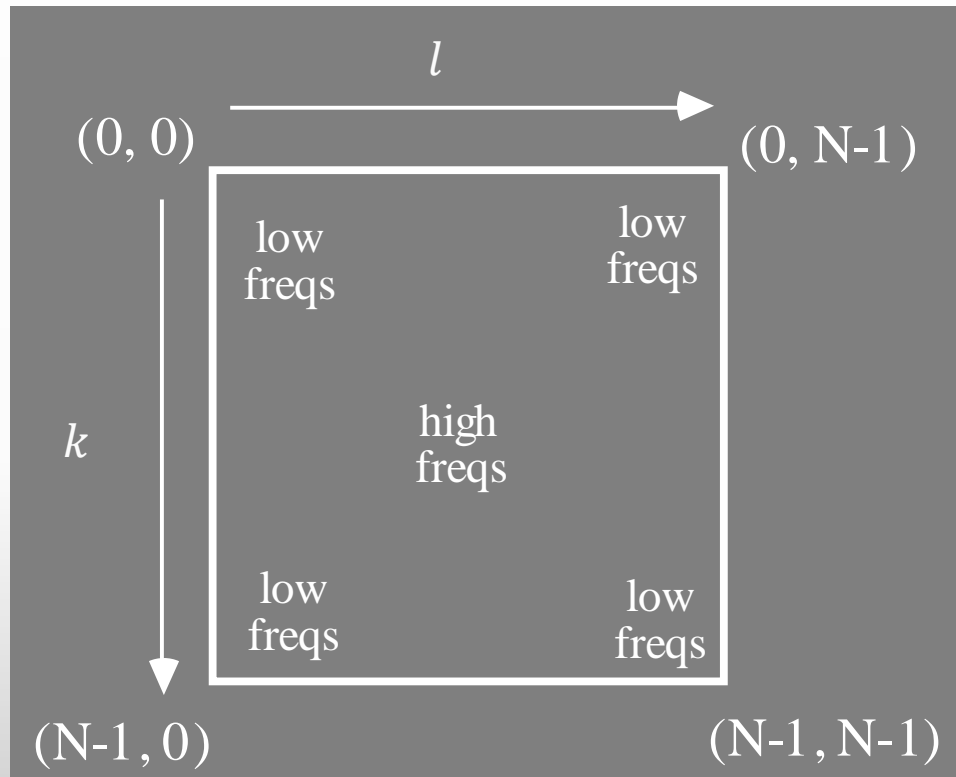
Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Ο DFT επιστρέφει μια «μιγαδική εικόνα» (έναν πίνακα μιγαδικών) που αντιστοιχεί σε (πίνακα) πραγματικού μέρους και (πίνακα) φανταστικού μέρους ή (πίνακα) μέτρο και (πίνακα) φάση.
- Συνήθως για να δούμε τα χαρακτηριστικά μιας εικόνας στο πεδίο της συχνότητας αρκεί ο πίνακας μέτρου αλλά για να μπορέσουμε να επιστρέψουμε στο χώρο της εικόνας (pixels) χρειαζόμαστε και την φάση.
- Υλοποίηση μέσω του αλγορίθμου FFT
- Πλάτος DFT:
Συμμετρία και στις δύο κατευθύνσεις (k -κατεύθυνση και l -κατεύθυνση).
Οι μονάδες μέτρησης είναι cycles/image (κύκλοι/εικόνα).



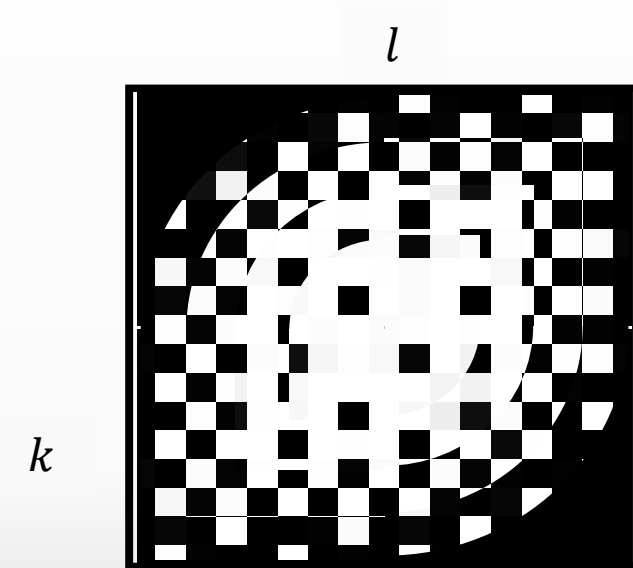
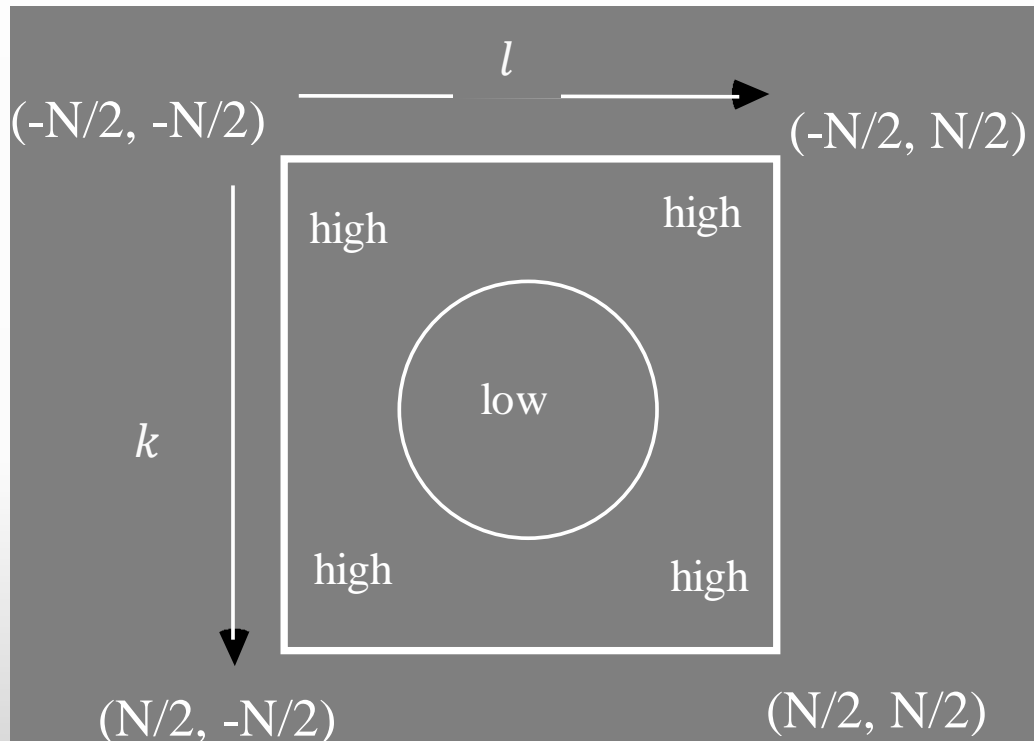
Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

Οι ψηλότερες συχνότητες αναπαριστούνται κοντά στο $(k, l) = (N/2, N/2)$, δηλαδή στο κέντρο.



Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

- Λόγω συμμετρίας του DFT, με μια μετατόπιση στο μισό μήκος και στις δύο διευθύνσεις έχουμε τη τιμή DC να προβάλλεται στο κέντρο της εικόνας.
- Έτσι όσο απομακρυνόμαστε από το κέντρο τόσο υψηλότερη είναι η αντίστοιχη συχνότητα.

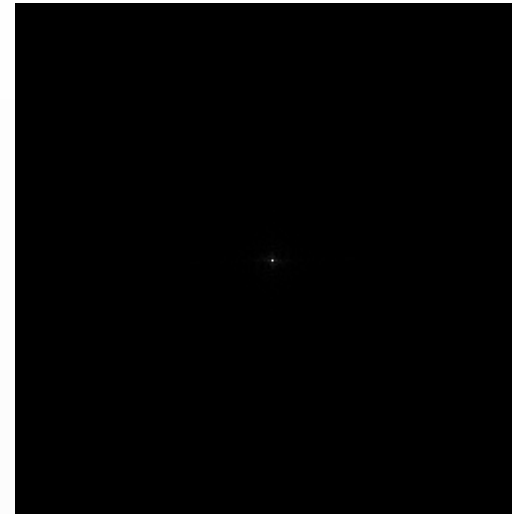


Centered DFT

Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

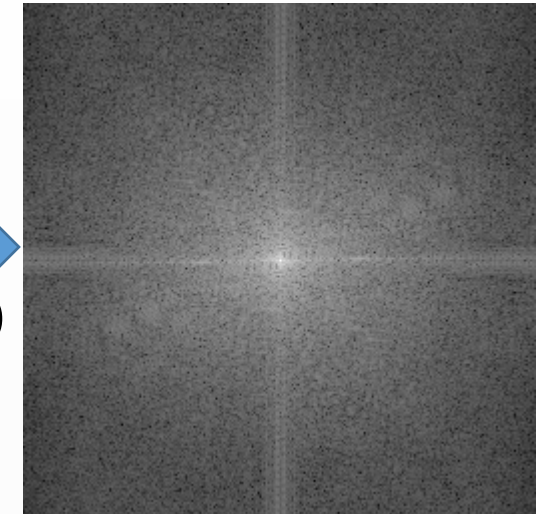
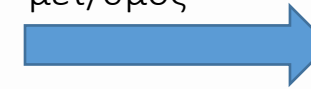


Μέτρο DFT



Λογαριθμικός
μετ/σμός

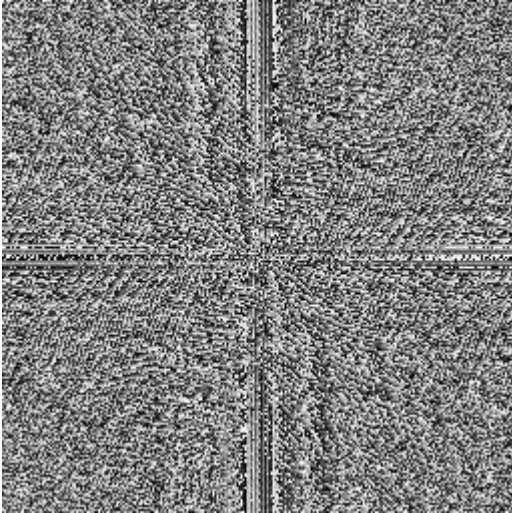
$\log(1 + \text{abs}(DFT))$



Τεράστιο δυναμικό εύρος
οπτικά οι χαμηλού πλάτους μεγέθους
συχνότητες θα είναι δυσδιάκριτες
Μόνο DC

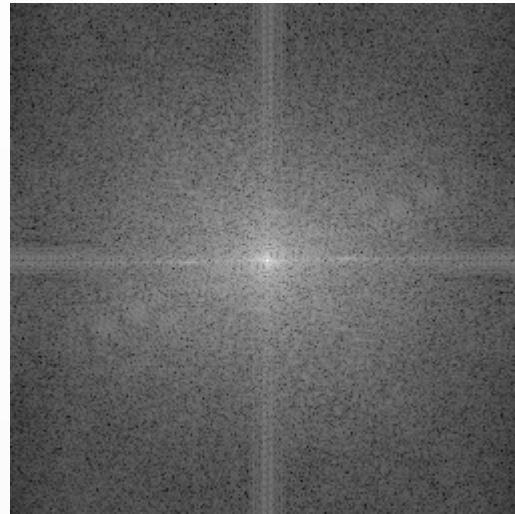
- Οι χαμηλές συχνότητες έχουν περισσότερη πληροφορία για την εικόνα από τις υψηλές
 - Στο μετασχηματισμό βλέπουμε ότι υπάρχουν δύο κυρίαρχες κατευθύνσεις στην εικόνα που διέρχονται οριζόντια και κάθετα από το κέντρο της εικόνας, αυτές προέρχονται από τα επαναλαμβανόμενα μοτίβα στο υπόβαθρο(background) της εικόνας.
- Πλήθος συχνοτήτων με μειούμενο πλάτος όσο αυξάνεται η συχνότητα
 - Συμμετρικό ως προς το κέντρο

Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier



Φάση DFT

Η τιμή σε κάθε σημείο προσδιορίζει τη φάση της αντίστοιχης συχνότητας
Επίσης και εδώ βλέπουμε τις δύο κατευθύνσεις όπως και στο πλάτος αλλά όχι πολύ άλλη πληροφορία, οπότε μπορούμε να την αγνοήσουμε για ορισμένες εφαρμογές.

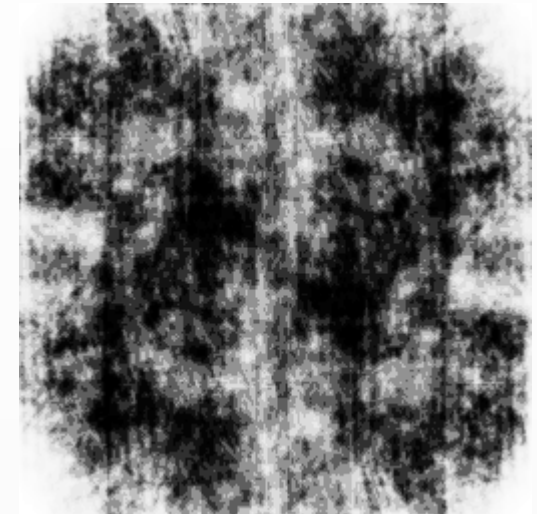


Λογαριθμικός μετ/σμός
πλάτους DFT

Inverse DFT, χωρίς
πληροφορία φάσης

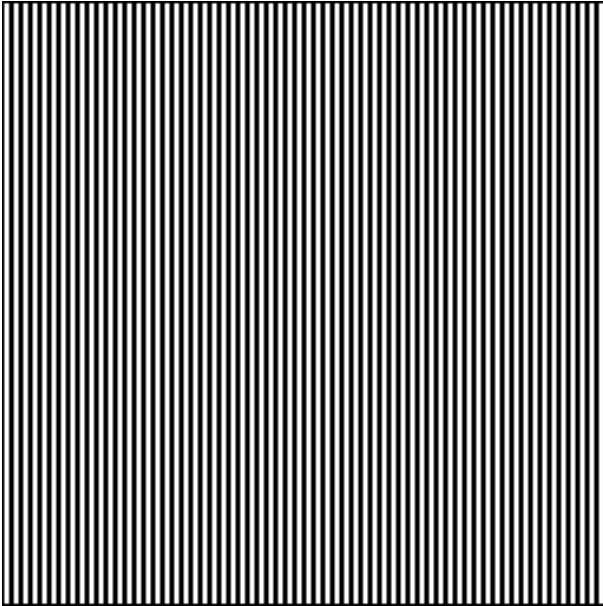


Εξίσωση
Ιστογράμματος μετά
τον IDFT

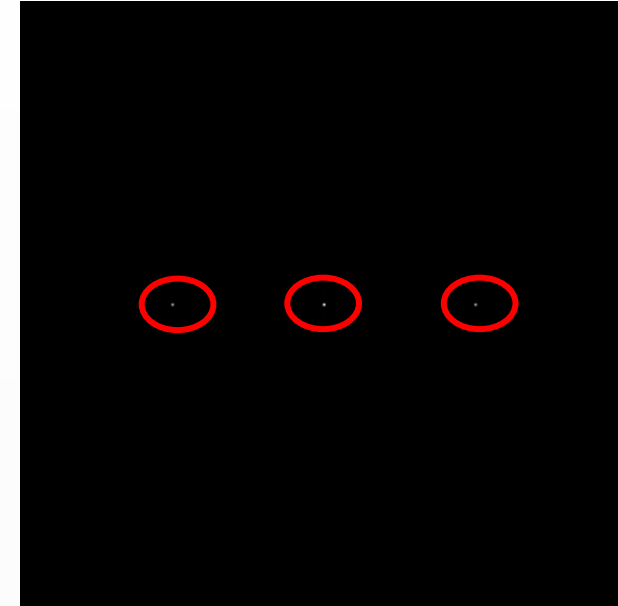


Η εικόνα δεν επιστρέφει την αρχική, δηλαδή η πληροφορία της φάσης είναι αναγκαία για να επιστρεψουμε στο spatial domain (χώρο εικόνας)

Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier

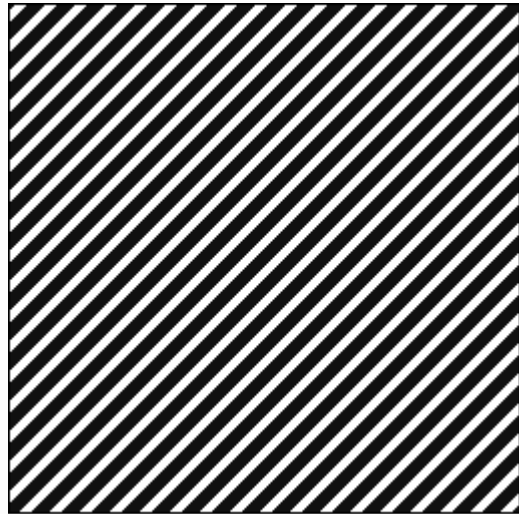


Πλάτος DFT, χωρίς
πληροφορία φάσης

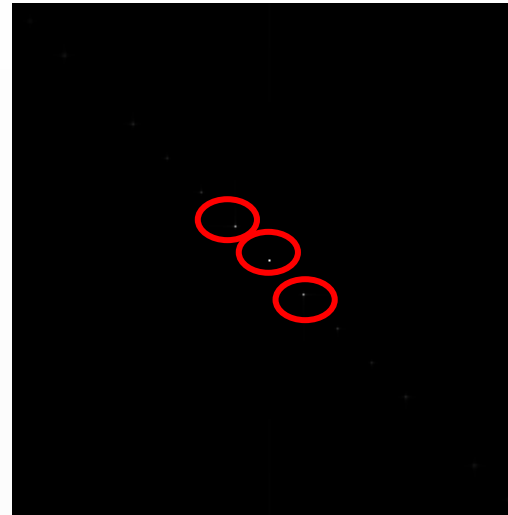


- 3 τιμές, η μέγιστη στο DC και 1 τιμή συχνότητας (και η συμμετρική της) λόγω των λωρίδων.
- Τα 2 σημεία βρίσκονται στην οριζόντια γραμμή που διέρχεται από το κέντρο επειδή η φωτεινότητα αλλάζει στο χώρο της εικόνας εάν μετακινηθούμε οριζόντια

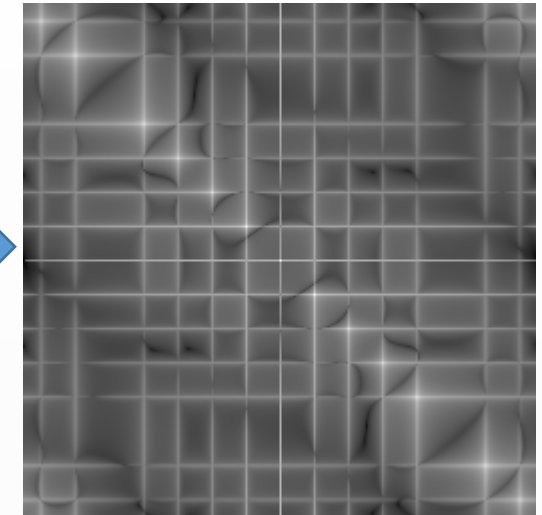
Εικόνα στο χώρο της συχνότητας – Διακριτός Μετασχηματισμός Fourier



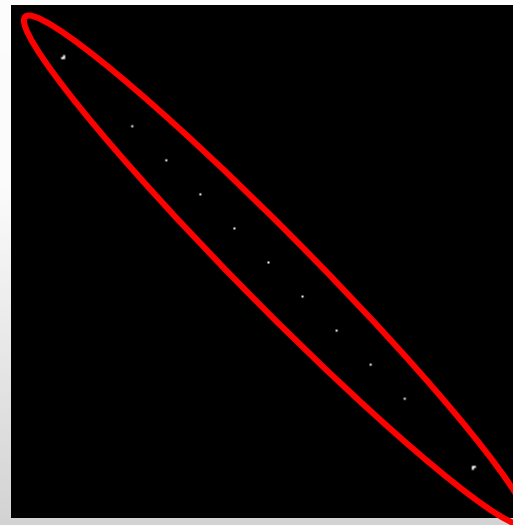
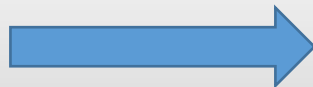
Πλάτος DFT, χωρίς
πληροφορία φάσης



Λογαριθμικός
μετ/σμός



Threshold >13



Συμπίεση Εικόνας

Οι τεχνικές συμπίεσης βασίζονται στην απόρριψη της πλεονάζουσας πληροφορίας

- Ανάγκες που καλύπτονται
 - Εξοικονόμηση μνήμης
 - Ελάττωση χρόνου και εύρους ζώνης μετάδοσης
- Ιδιαίτερα μεγάλη η σημασία της συμπίεσης στις 2-D και 3-D εφαρμογές
- Κατηγορίες τεχνικών συμπίεσης
 - Απωλεστικές
 - Μη απωλεστικές

Συμπίεση Εικόνας

- Εκμετάλλευση 3 ειδών πλεονασμού πληροφορίας:
 1. Στατιστικός **Πλεονασμός (per pixel)**
 - Μπορούμε να χρησιμοποιήσουμε διαφορετική αναπαράσταση για gray levels που εμφανίζονται με διαφορετικές πιθανότητες (π.χ. κωδικοποίηση Huffman)
 2. **Πλεονασμός μεταξύ των pixel (interpixel)**
 - Μπορούμε να εκμεταλλευτούμε τα επαναλαμβανόμενα μοτίβα ή τις μεγάλες ομαλές επιφάνειες (πχ κωδικοποίηση μήκους διαδρομών)
 3. **Φυσικό-οπτικός πλεονασμός**
 - Αξιοποίηση του τρόπου με τον οποίο αντιλαμβάνεται ο άνθρωπος την οπτική πληροφορία (π.χ. ακμές)

Συμπίεση Εικόνας

Τα βασικά στάδια μιας απωλεστικής τεχνικής:

1. Μετασχηματισμός της εικόνας στο κατάλληλο πεδίο

- Πεδίο εικονοστοιχείων (π.χ., ADPCM)
- Πεδίο συχνοτήτων (π.χ., DCT, Wavelet)
- Πεδίο παραμέτρων μοντέλου (π.χ., στοχαστικά μοντέλα, fractals κλπ)

2. Κβαντισμός των αποτελεσμάτων του 1ου σταδίου

- ομοιόμορφος
- ανομοιόμορφος
- διανυσματικός

3. Προσδιορισμός του λεξικού (codebook design) για την αναπαράσταση των εξόδων του κβαντιστή

Συμπίεση Εικόνας

Παραδείγματα μη απωλεστικών τεχνικών:

- Κωδικοποίηση Huffman/Εντροπίας
- Bit plane coding
- Constant Area Coding
- Contour Tracing
- Κωδικοποίηση Μήκους Διαδρομών
- Συμπίεση LZW
- Κωδικοποίηση με Πρόβλεψη
- ... και άλλες

Παραδείγματα απωλεστικών τεχνικών:

- Κωδικοποίηση με Πρόβλεψη (και κβαντισμό)
- Κωδικοποίηση με Μετασχηματισμούς
- ... και άλλες

Συμπίεση Εικόνας

Εκμετάλλευση χωρικού πλεονασμού

- Ο ανθρώπινος οφθαλμός παρουσιάζει μειωμένη ευαισθησία τις απότομες διαφοροποιήσεις σε μια εικόνα (μικρός αριθμός pixels - υψηλές χωρικές συχνότητες)
- Έτσι, 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 στο μέγεθος της εικόνας
- εκμετάλλευση του χωρικού πλεονασμού εντός εικόνας με το πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)
- Ο βασικός αλγόριθμος που χρησιμοποιείται στο JPEG είναι ο **Διακριτός Μετασχηματισμός Συνημιτόνου** (Discrete Cosine Transform, DCT)
- αυξημένη ευαισθησία του ανθρώπινου οφθαλμού στη φωτεινότητα σε σχέση με την ευαισθησία του στα χρώματα.

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

- Βήματα (αρχικά)
- Μετατροπή εικόνας από το χώρο RGB σε χώρο YCbCr
- Υπόδειγματοληψία χρώματος (λόγω αυξημένης ευαισθησία του ανθρώπινου οφθαλμού στη φωτεινότητα σε σχέση με την ευαισθησία του στα χρώματα)
- Κατάτμηση σε μπλοκ Minimum Coded Unit (MCU) συνήθως 8x8 τιμών ή 16x16

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

Μετατροπή από RGB σε YCbCr

Η περιγραφή γίνεται μέσω της φωτεινότητας (luminance, **Y**) και των χρωμοδιαφορών **Cb** (μπλε), **Cr** (κόκκινο),

Για τον υπολογισμό της φωτεινότητας (**Y**) επιμερίζουμε το ποσοστό με το οποίο συνεισφέρουν τα (τρία) βασικά χρώματα δηλαδή τους συντελεστές με τους οποίους κάθε βασικό χρώμα συνεισφέρει στη φωτεινότητα.

Οι συντελεστές αυτοί έχουν προκύψει από πειραματικές μετρήσεις της ευαισθησίας του ανθρώπινου οφθαλμού στα συγκεκριμένα χρώματα και έχουν προτυποποιηθεί.

Η φωτεινότητα δίνεται από την σχέση:

$$Y = 0,299R + 0,587G + 0,114B \text{ ή } Y = 0,3R + 0,59G + 0,11B$$

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

Μετατροπή από RGB σε YCbCr

Η φωτεινότητα μεταβάλλεται από το 0 (μαύρο) ως το 1 (λευκό)

Ως **χρωμοδιαφορές** ορίζονται η διαφορά της φωτεινότητας από την τιμή της συνιστώσας του εκάστοτε χρώματος (συγκεκριμένα του μπλε και του κόκκινου).

Η χρωμοδιαφορά του κόκκινου προκύπτει ως η διαφορά της φωτεινότητας από την αναλογία του κόκκινου:

$$Cr = R - Y = R - (0,299R + 0,587G + 0,114B) = -0,299R - 0,587G + 0,886B$$

με R, G, B τις αναλογίες συμμετοχής στα τρία βασικά χρώματα, ως R, G, B,

Η χρωμοδιαφορά του μπλε προκύπτει ως η διαφορά της φωτεινότητας από την αναλογία του μπλέ:

$$Cb = B - Y = B - (0,299R + 0,587G + 0,114B) = 0,701R - 0,587G - 0,114B$$

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

- Παραδείγματα Υπολογισμού φωτεινότητας και χρωμοδιαφορών (κατά προσέγγιση στο 2^ο δεκαδικό) :

Χρώμα	R	G	B	Y	C_r	C_g	C_b
Μαύρο	0	0	0	0	0	0	0
Μπλε	0	0	1	0,11	-0,11	-0,11	0,89
Πράσινο	0	1	0	0,59	-0,59	0,41	-0,59
Γαλανό	0	1	1	0,70	-0,70	0,30	0,30
Κόκκινο	1	0	0	0,30	0,70	-0,30	-0,30
Πορφυρό	1	0	1	0,41	0,59	-0,41	0,59
Κίτρινο	1	1	0	0,89	0,11	0,11	-0,89
Λευκό	1	1	1	1	0	0	0

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

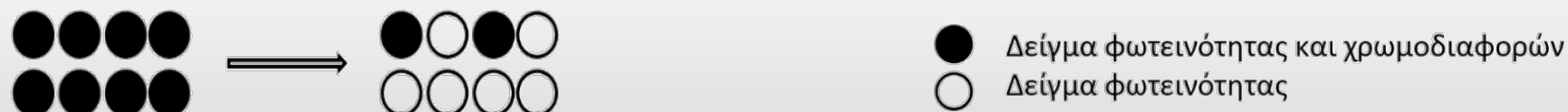
Υπόδειγματοληψία χρώματος

- λόγω αυξημένης ευαισθησία του ανθρώπινου οφθαλμού στη φωτεινότητα σε σχέση με την ευαισθησία του στα χρώματα μπορούμε να λαμβάνουμε λιγότερα δείγματα χρωματικής πληροφορίας σε σχέση με τα δείγματα τα οποία αφορούν τη φωτεινότητα.

→ Σήμα 4:2:2



- Ο ανθρώπινος οφθαλμός έχει χαμηλή ευαισθησία στην ανάλυση χρώματος στην οριζόντια και κατακόρυφη διεύθυνση → λαμβάνονται δείγματα χρωμοδιαφορών ανά δύο γραμμές → Σήμα 4:2:0



4:4:4 → 4:2:0 (DVB, HDTV)

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG (Joint Pictures Expert Group)

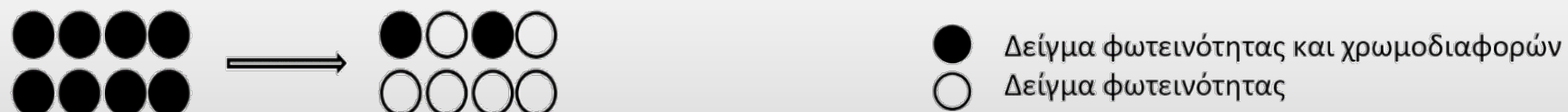
Υπόδειγματοληψία χρώματος

- λόγω αυξημένης ευαισθησίας του ανθρώπινου οφθαλμού στη φωτεινότητα σε σχέση με την ευαισθησία του στα χρώματα μπορούμε να λαμβάνουμε λιγότερα δείγματα χρωματικής πληροφορίας σε σχέση με τα δείγματα τα οποία αφορούν τη φωτεινότητα.

→ Σήμα 4:2:2



- Ο ανθρώπινος οφθαλμός έχει χαμηλή ευαισθησία στην ανάλυση χρώματος στην οριζόντια και κατακόρυφη διεύθυνση → λαμβάνονται δείγματα χρωμοδιαφορών ανά δύο γραμμές → Σήμα 4:2:0



4:4:4 → 4:2:0 (DVB, HDTV)

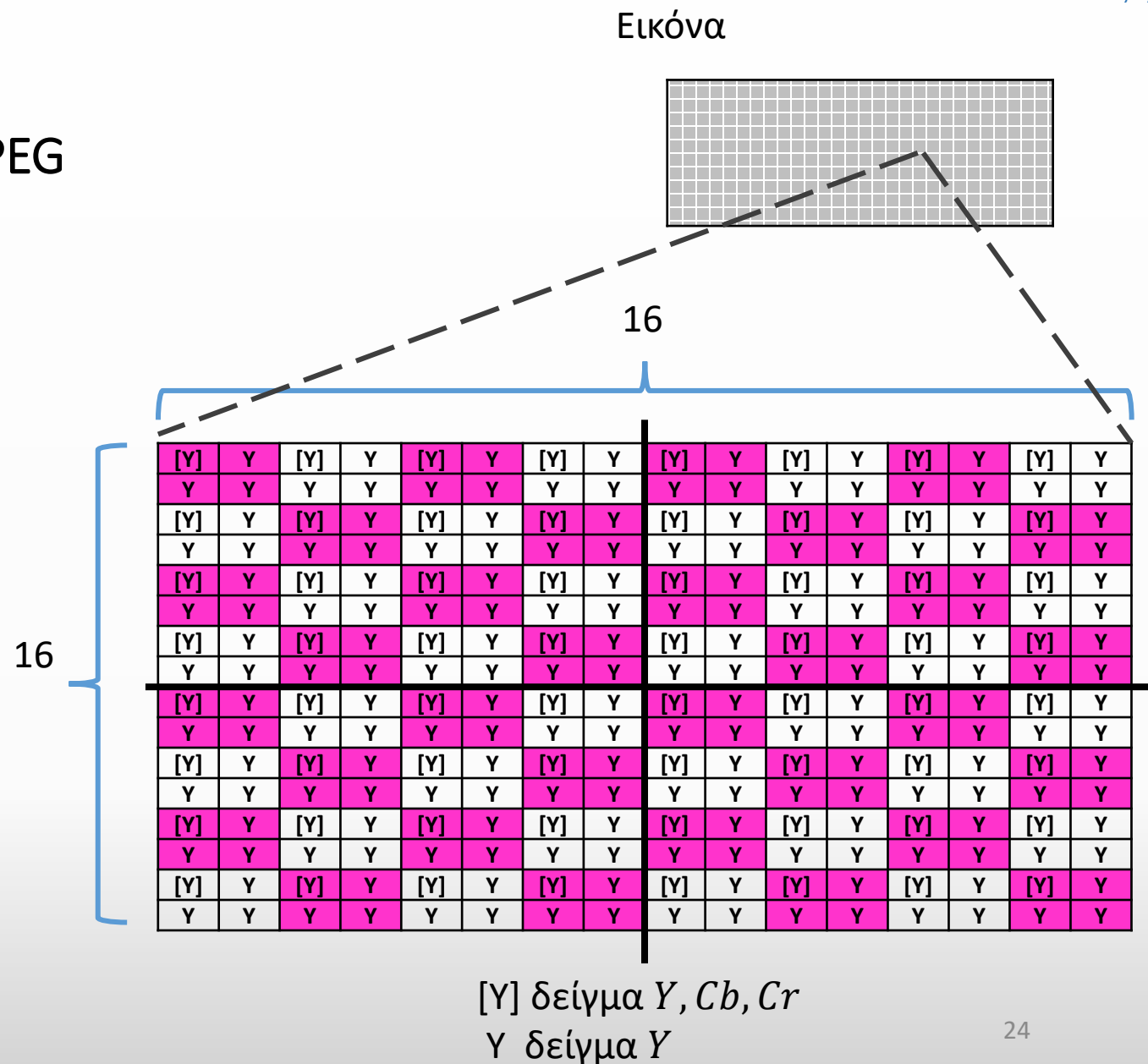
Συμπύεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG

- μπλοκ των 16x16 (pixels).


$$4:4:4 \rightarrow 4:2:0$$

- μπλοκ των 16x16 pixel φωτεινότητας Y ,
→ ή 4 μπλοκ 8x8 pixel Y
- μπλοκ των 8x8 pixel Cb
- μπλοκ των 8x8 pixel Cr



Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG

Διακριτός Μετασχηματισμός Συνημιτόνου (DCT) : ειδική μορφή του Discrete Fourier Transform (DFT).

- Με χρήση του μετασχηματισμού αυτού, ένα πλήθος τιμών μετασχηματίζεται σε ίδιου πλήθους τιμές στο πεδίο των συχνοτήτων.
- στις δύο διαστάσεις οι ανεξάρτητες μεταβλητές του DCT είναι οι συντεταγμένες του pixel (p, q)

$$y(p, q) = a(p) \cdot a(q) \cdot \sum_{m=1}^M \sum_{n=1}^N x(m, n) \cos \frac{\pi(2m-1)(p-1)}{2M} \cos \frac{\pi(2n-1)(q-1)}{2N},$$

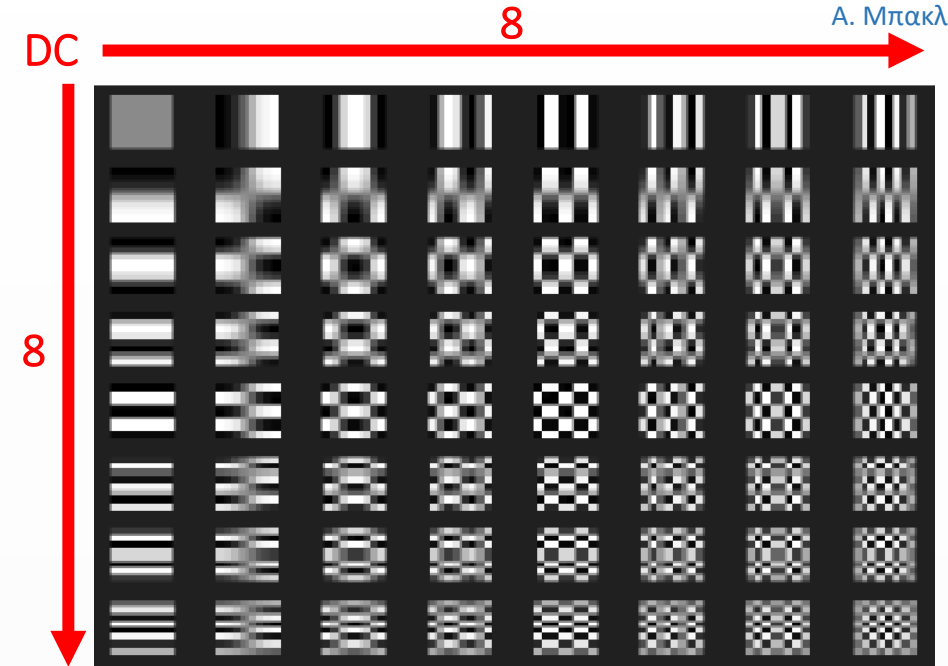
$$p = 1, 2, \dots, M, q = 1, 2, \dots, N, \quad a(p) = \begin{cases} \frac{1}{\sqrt{M}}, & p = 1 \\ \sqrt{\frac{2}{M}}, & 2 \leq p \leq M \end{cases} \quad a(q) = \begin{cases} \frac{1}{\sqrt{N}}, & q = 1 \\ \sqrt{\frac{2}{N}}, & 2 \leq q \leq N \end{cases}$$

Greyscale εικόνες
 $x(m, n) \in [0, 255]$

Συμπίεση Εικόνας

Διακριτός Μετασχηματισμός Συνημιτόνου (DCT)

- Μετασχηματισμός κατά DCT → συντελεστές χωρικών συχνοτήτων
- Οι συντελεστές αυτοί δηλώνουν τη σχετική βαρύτητα κάθε συχνότητας που συμβάλει στη σύνθεση του διακριτού σήματος της εικόνας και έχουν πραγματικές ακέραιες τιμές
- το στοιχείο (1,1) (πάνω-αριστερά) είναι η DC συνιστώσα (μέση φωτεινότητα της αρχικής εικόνας),
- οι συντελεστές στην πρώτη γραμμή από αριστερά προς δεξιά δείχνουν την συνεισφορά των «τραχειών» (χαμηλόσυχνων) δομών της εικόνας έως των πιο «λεπτών»(υψίσυχνων) στην οριζόντια διεύθυνση
- οι συντελεστές στην πρώτη στήλη από πάνω προς τα κάτω δείχνουν την συνεισφορά των «τραχειών» (χαμηλόσυχνων) δομών της εικόνας έως των πιο «λεπτών»(υψίσυχνων) στην κατακόρυφη διεύθυνση
- οι συντελεστές της διαγώνιου ξεκινώντας από το (1,1) δείχνουν την συνεισφορά των «τραχειών» (χαμηλόσυχνων) δομών της εικόνας έως των πιο «λεπτών»(υψίσυχνων) στην διαγώνιο
- πάνω και αριστερά θέσεις αντιστοιχούν στις πιο σημαντικές συχνότητες

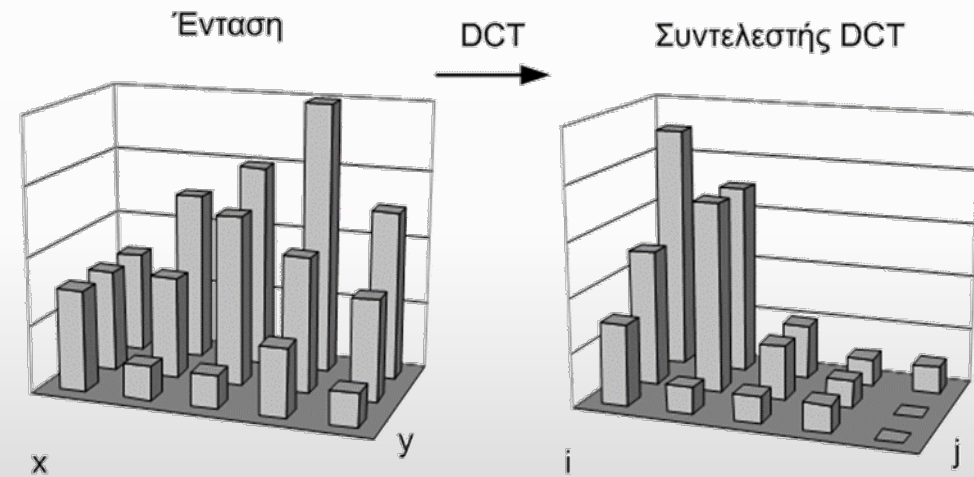


Συμπίεση Εικόνας

Διακριτός Μετασχηματισμός Συνημιτόνου (DCT)

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

Συγκέντρωση υψηλών τιμών κοντά στο DC
Υψηλές τιμές στις χαμηλές χωρικές συχνότητες



Συμπίεση Εικόνας

Γιατί Διακριτός Μετασχηματισμός Συνημιτόνου (DCT);

- Φυσικές εικόνες
 - Μεγάλες περιοχές με παρόμοια χρώματα
 - Πολλοί συντελεστές AC σχεδόν μηδενικοί
 - Δείχνουν την απόσταση των εναλλαγών
 - Όταν δεν υπάρχουν, οι συντελεστές είναι μηδέν
- Διαφοροποίηση σε σχέση με γραφικά
 - Δεν υπάρχουν απότομες αλλαγές χρωμάτων οι οποίες παράγουν μη μηδενικούς συντελεστές AC
- Ο DCT επιδεικνύει ανθεκτικότητα στην παράλειψη ενός ποσοστού από τους συντελεστές που προκύπτουν αφού εφαρμοστεί σε κάποια δεδομένα και οδηγεί σε σημαντικό βαθμό συμπίεσης

Συμπίεση Εικόνας

Διακριτός Μετασχηματισμός Συνημιτόνου (DCT)

- Βήματα:
 - Η εικόνα τεμαχίζεται σε μπλοκ (συνήθως) 8x8 pixel (άρα $\mathbf{M} = \mathbf{N} = \mathbf{8}$)*
 - αφαίρεση μέσης τιμής (στην περίπτωση κβαντισμού 8 bits το 128 άρα από $[0,255] \rightarrow [-128,127]$)
 - μετασχηματισμός DCT σε κάθε μπλοκ*
 - προσαρμογή συντελεστών με βάση την ευαισθησία του ανθρώπινου οφθαλμού αλλά και του επιθυμητού επιπέδου ποιότητας Q
 - Zig-zag σάρωση συντελεστών με ομαδοποίηση τιμών

* Τα συνημίτονα δεν εξαρτώνται από $x(m, n)$
→ υπολογίζονται προκαταβολικά

$$* a(p) = \begin{cases} \frac{1}{\sqrt{8}}, & p = 1 \\ \sqrt{\frac{2}{8}} = \sqrt{\frac{1}{4}} = \frac{1}{2}, & 2 \leq p \leq M \end{cases} \quad a(q) = \begin{cases} \frac{1}{\sqrt{8}}, & q = 1 \\ \sqrt{\frac{2}{8}} = \sqrt{\frac{1}{4}} = \frac{1}{2}, & 2 \leq q \leq N \end{cases}$$

Συμπίεση Εικόνας

Διακριτός Μετασχηματισμός Συνημιτόνου (DCT)

- Επεξεργασία συντελεστών DCT → οι συντελεστές DCT διαιρούνται με συγκεκριμένους συντελεστές κβάντισης (ευαισθησία του ανθρώπινου οφθαλμού).
- Όσο μεγαλύτερες είναι οι τιμές των συντελεστών κβάντισης Q (δηλαδή του διαιρέτη) τόσο μειώνονται οι τιμές των συντελεστών προς κβάντιση.
- Σε αρκετές περιπτώσεις οι συντελεστές αυτοί μηδενίζονται (ειδικά καθώς αυξάνουν οι συχνότητες) → Οι μηδενικοί συντελεστές κωδικοποιούνται με ελάχιστες απαιτήσεις → η συνεισφορά των υψηλών συχνοτήτων αποκόπτεται
- **Απωλεστική συμπίεση**
- Οι συντελεστές κβάντισης υπάρχουν σε έτοιμους πίνακες, προτυποποιημένους κατά το JPEG.

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG

Κβαντοποίηση - Πίνακες κλιμάκωσης συντελεστών DCT

8x8 πίνακας κλιμάκωσης συντελεστών DCT σήματος φωτεινότητας (**luma**) ποιότητας 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
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

8x8 πίνακας κλιμάκωσης συντελεστών DCT σήματος χρωμικότητας (**chrominance**) ποιότητας 50% (Q_{50})

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG Κβαντοποίηση - Πίνακες κλιμάκωσης συντελεστών DCT

8x8 πίνακας κλιμάκωσης συντελεστών DCT σήματος φωτεινότητας ποιότητας 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
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

Υπολογισμός πίνακα κλιμάκωσης συντελεστών DCT ποιότητας $a\%$ (Q_a) :

$$Q_a = \begin{cases} Q_{50} \cdot \frac{100 - a}{50}, & a > 50 \\ Q_{50} \cdot \frac{50}{a}, & a < 50 \end{cases}$$

8x8 πίνακας κλιμάκωσης συντελεστών DCT ποιότητας 90% (Q_{90})

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

8x8 πίνακας κλιμάκωσης συντελεστών DCT ποιότητας 10% (Q_{10})

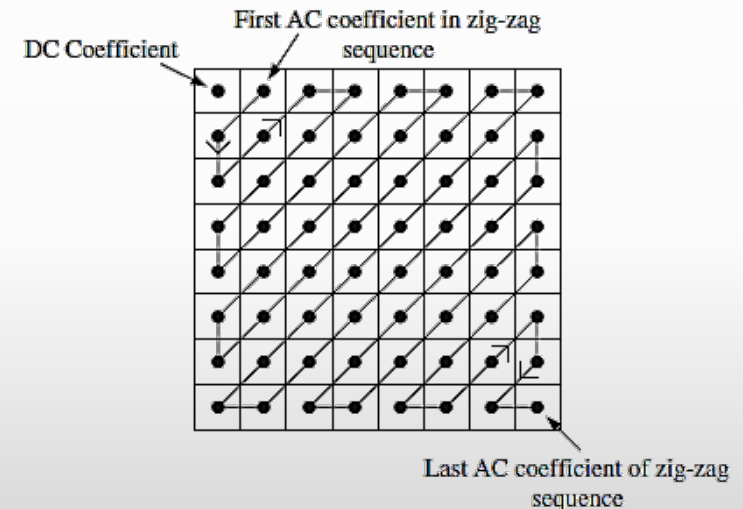
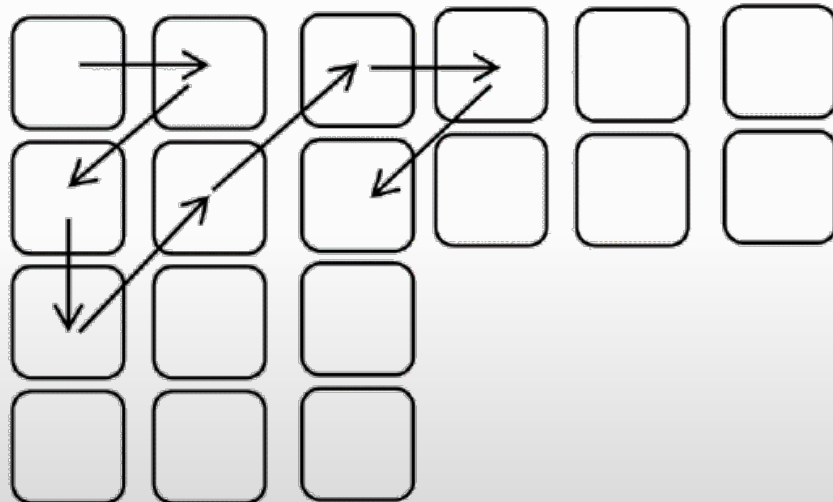
80	60	50	80	120	200	255	255
55	60	70	95	130	255	255	255
70	65	80	120	200	255	255	255
70	85	110	145	255	255	255	255
90	110	185	255	255	255	255	255
120	175	255	255	255	255	255	255
245	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG

Κωδικοποίηση Μήκους Διαδρομής - Zig-zag σάρωση συντελεστών με ομαδοποίηση τιμών :

- αναμένεται συγκέντρωση των μεγαλύτερων συντελεστών στο πάνω αριστερά μέρος και των μηδενικών κάτω δεξιά. → συγκέντρωση μέγιστου πλήθους μηδενικών
- με σκοπό την ομαδοποίηση των τιμών και κατά το δυνατόν συγκέντρωση των μηδενικών τιμών, η σάρωση γίνεται κατά zig-zag, ξεκινώντας από το άνω αριστερά τμήμα και προχωρώντας προς τα δεξιά και κάτω.



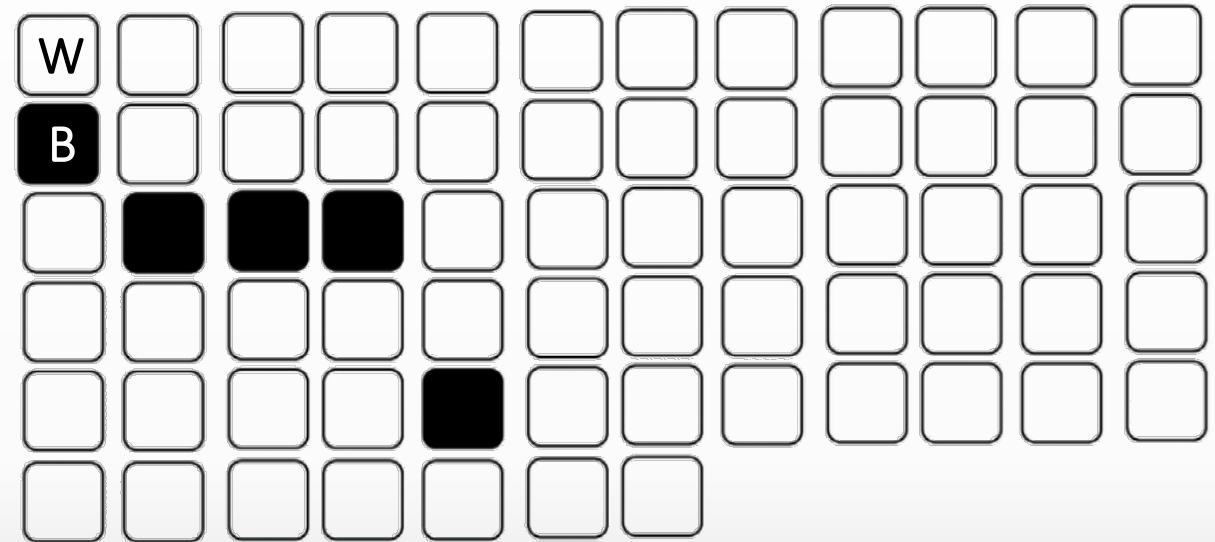
Συμπύεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας JPEG

Κωδικοποίηση Μήκους Διαδρομής (Run Length Encoding, RLE) :

- Περιγραφή δεδομένων που προσφέρει αποτελεσματική συμπίεση **μόνο** όταν επαναλαμβάνονται τιμές (π.χ. 0)

Ο όγκος των δεδομένων διπλασιάζεται
για συνεχείς εναλλαγές τιμών ☹️



$\text{WWWWWWWWWWWWWBWWWWWWWWWWWWBBBWW}$

$\text{WWWWWWWWWW} = (\text{W}, 12), (\text{B}, 1), (\text{W}, 12), (\text{B}, 3), (\text{W}, 24), (\text{B}, 1), (\text{W}, 14)$

Συμπίεση Εικόνας

Πρότυπο συμπίεσης στατικής εικόνας (JPEG -Joint Pictures Expert Group)

- Βήματα (συνέχεια):
 - Η εικόνα τεμαχίζεται σε μπλοκ (συνήθως) 8x8 pixel ή 16x16
 - αφαίρεση μέσης τιμής (στην περίπτωση κβαντισμού 8 bits το 128)
 - μετασχηματισμός DCT σε κάθε μπλοκ
 - προσαρμογή συντελεστών με βάση την ευαισθησία του ανθρώπινου οφθαλμού αλλά και του επιθυμητού επιπέδου ποιότητας Q
 - Zig-zag σάρωση συντελεστών με ομαδοποίηση τιμών
 - Κωδικοποίηση εντροπίας (Huffman)
 - Αποθήκευση συντελεστών

Συμπίεση Εικόνας

Παράδειγμα Συμπίεσης

8x8 block εικόνας με τιμές φωτεινότητας 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

-128

26	-5	-5	-5	-5	-5	-5	8
64	52	8	26	26	26	8	-18
126	70	26	26	52	26	-5	-5
111	52	8	52	52	38	-5	-5
52	26	8	39	38	21	8	8
0	8	-5	8	26	52	70	26
-5	-23	-18	21	8	8	52	38
-18	8	-5	-5	-5	8	26	8

DCT (dct2)

162	41	20	72	30	12	-20	-11
30	108	10	32	28	-16	18	-2
-94	-60	12	-43	-31	6	-3	7
-39	-83	-5	-22	-14	15	-1	4
-31	18	-6	-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

κλιμάκωση συντελεστών
για την επιθυμητή ποιότητα 50%

Συμπύεση Εικόνας

Παράδειγμα Συμπίεσης

8x8 πίνακας κλιμάκωσης συντελεστών DCT ποιότητας 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
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

κλιμάκωση 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
6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Οι μη μηδενικές συνιστώσες μαζεύονται στο πάνω αριστερά μέρος του πίνακα.

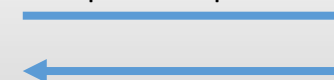
Με τη μέθοδο σάρωσης zig-zag προκύπτουν οι εξής τιμές:

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 (64 χαρακτήρες)

10, 4, 3, -7, 9, 2, 5, 1, -5, -3, -2, -5, 1, 2, 1, 0, 1, -2, 0, 1, 4*(0), 2*(-1), 38*(0) (46 χαρακτήρες)

Run Length Coding

~28.1% Συμπίεση



Συμπύεση Εικόνας

Παράδειγμα Συμπίεσης

8x8 πίνακας κλιμάκωσης συντελεστών DCT ποιότητας 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
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

κλιμάκωση 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
6	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Οι μη μηδενικές συνιστώσες μαζεύονται στο πάνω αριστερά μέρος του πίνακα.

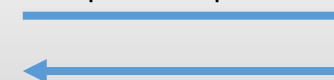
Με τη μέθοδο σάρωσης zig-zag προκύπτουν οι εξής τιμές:

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 (64 χαρακτήρες)

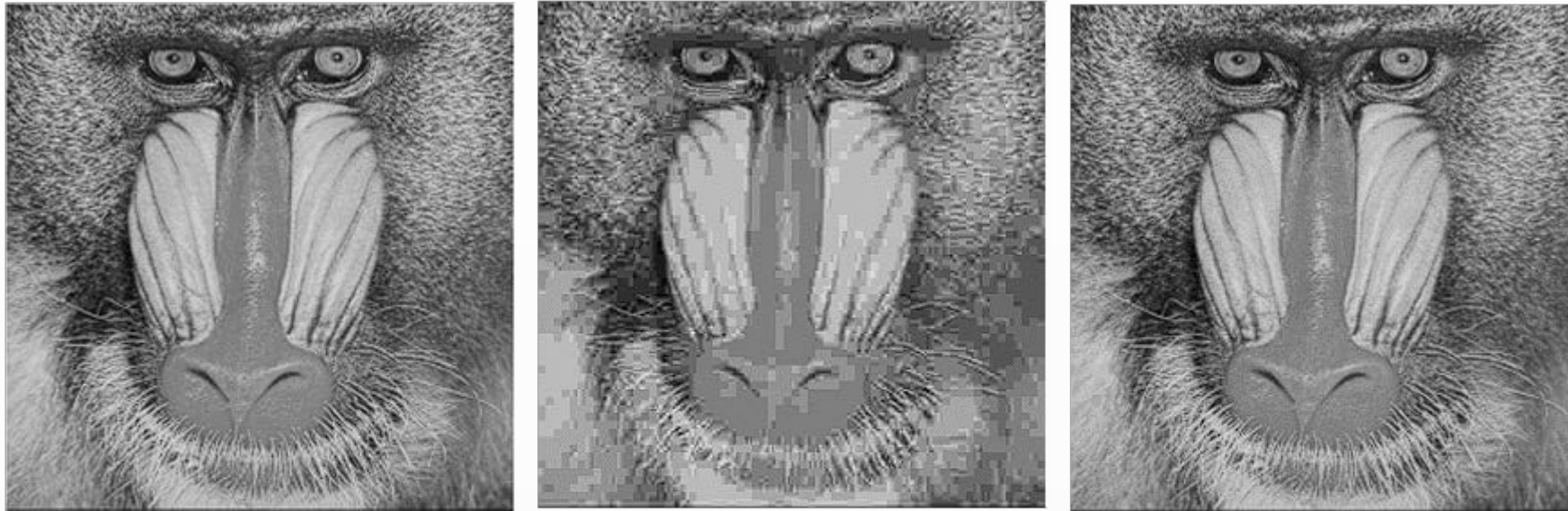
10, 4, 3, -7, 9, 2, 5, 1, -5, -3, -2, -5, 1, 2, 1, 0, 1, -2, 0, 1, 4*(0), 2*(-1), 38*(0) (46 χαρακτήρες)

Run Length Coding

~28.1% Συμπίεση



Συμπίεση Εικόνας

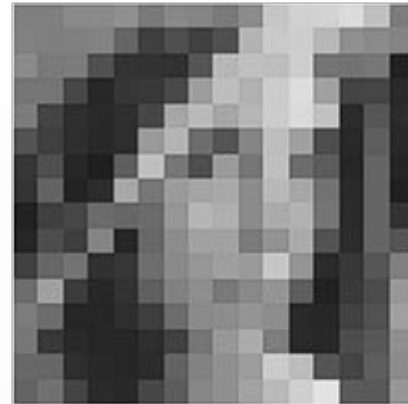


(α) Αρχική εικόνα. (β) Εικόνα κωδικοποιημένη με 50% ποιότητα. (γ) Εικόνα κωδικοποιημένη με 90% ποιότητα.

Συμπίεση Εικόνας



Αρχική



126:1

15:1



51:1

8:1



28:1

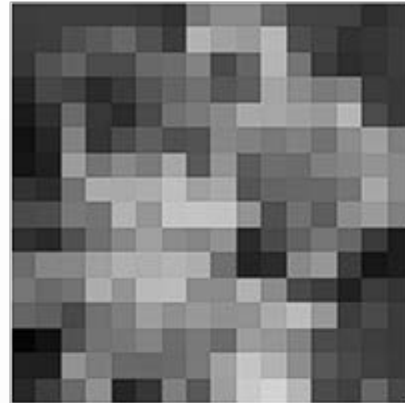
4:1



Συμπίεση Εικόνας

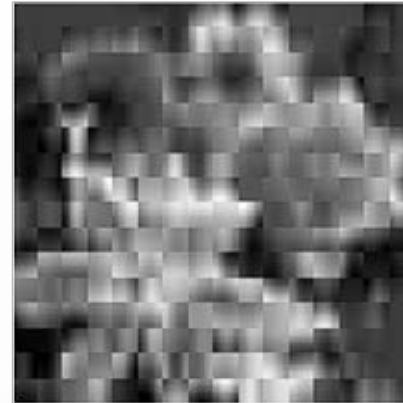


Αρχική



126:1

15:1



51:1

8:1



28:1

4:1



Συμπίεση Εικόνας

Κωδικοποίηση Huffman (Εντροπίας)

Έστω εικόνα $N \times M$ που αναπαριστάται με B bits/pixel.

Μπορούμε να εκτιμήσουμε την συνάρτηση πυκνότητας πιθανότητας $p(i)$ κατασκευάζοντας το ιστόγραμμα της. (Συνολικό μέγεθος $N \times M \times B$ bits)

Κωδικοποίηση Εντροπίας: αντιστοιχίζουμε μικρές κωδικές λέξεις (σύμβολα) στα επίπεδα φωτεινότητας που εμφανίζονται με μεγάλη πιθανότητα (συχνά) και μεγάλες κωδικές λέξεις σε αυτά που εμφανίζονται σπανιότερα.

Τα μήκη των λέξεων επιλέγονται έτσι ώστε το μέσο μήκος να ελαχιστοποιείται

– συμπίεση σε σχέση με κωδικοποίηση με σταθερού (ίδιου) μήκους λέξεις (ομοιόμορφη κωδικοποίηση)

Συμπίεση Εικόνας

Ομοιόμορφη κωδικοποίηση

υπολογίζεται το πλήθος των bits με βάση το πλήθος των συμβόλων προς κωδικοποίηση από τη σχέση υπολογισμού

$$K \leq 2^N$$

K 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 :

$K = 6$ σύμβολα,

πλήθος των bits: $N = 3$ ($2^2 < 6$ και $2^3 > 6$)

Απόδοση Κωδικοποίησης = $\sum_{i=1}^K w_i n_i = 3$ (bits/σύμβολο) για την ομοιόμορφη

n_i πλήθος bit αναπαράστασης συμβόλου i

w_i πιθανότητα εμφάνισης συμβόλου i

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

Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

- Δημιουργούνται οι αρχικοί κόμβοι (φύλλα) από τα προς κωδικοποίηση σύμβολα και σε κάθε ένα αναγράφεται το βάρος του (πιθανότητα εμφάνισης). Τα σύμβολα γράφονται με αύξουσα ή φθίνουσα σειρά (όσον αφορά τα βάρη τους).
- Ξεκινώντας τους δύο κόμβους με τα μικρότερα βάρη, δημιουργείται ένα νέος εσωτερικός κόμβος, ο οποίος τους συνδέει. Το βάρος του νέου κόμβου προκύπτει από το άθροισμα των βαρών των δύο κόμβων.
- Επανάληψη μέχρι να εξαντληθούν όλοι οι κόμβοι (να δημιουργηθεί η ρίζα του δένδρου - βάρος μονάδα).
- 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.
- Ακολουθώντας τις διαδρομές του προηγούμενου βήματος υπολογίζεται η κωδικοποίηση που αντιστοιχεί σε κάθε φύλλο (αρχικό σύμβολο), ως συνδυασμός των 0 και 1 που αντιστοιχούν στους ενδιάμεσους κλάδους που συνθέτουν τη διαδρομή.

→ Κάθε σύμβολο αντιστοιχεί σε μία μοναδική κωδικοποίηση.

→ Η κωδικοποίηση οποιουδήποτε συμβόλου δεν είναι πρόθεμα της κωδικοποίησης άλλου συμβόλου.

Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

- Καταγράφονται τα σύμβολα (τιμές φωτεινότητας) με φθίνουσα πιθανότητα εμφάνισης.
- Αν δεν είναι γνωστές οι πιθανότητες εμφάνισης, υπολογίζονται

Σύμβολα	Πλήθος εμφανίσεων	Πιθανότητα εμφάνισης
A	6	
B	4	
C	3	
D	2	
E	4	
F	1	
Σύνολο	20	

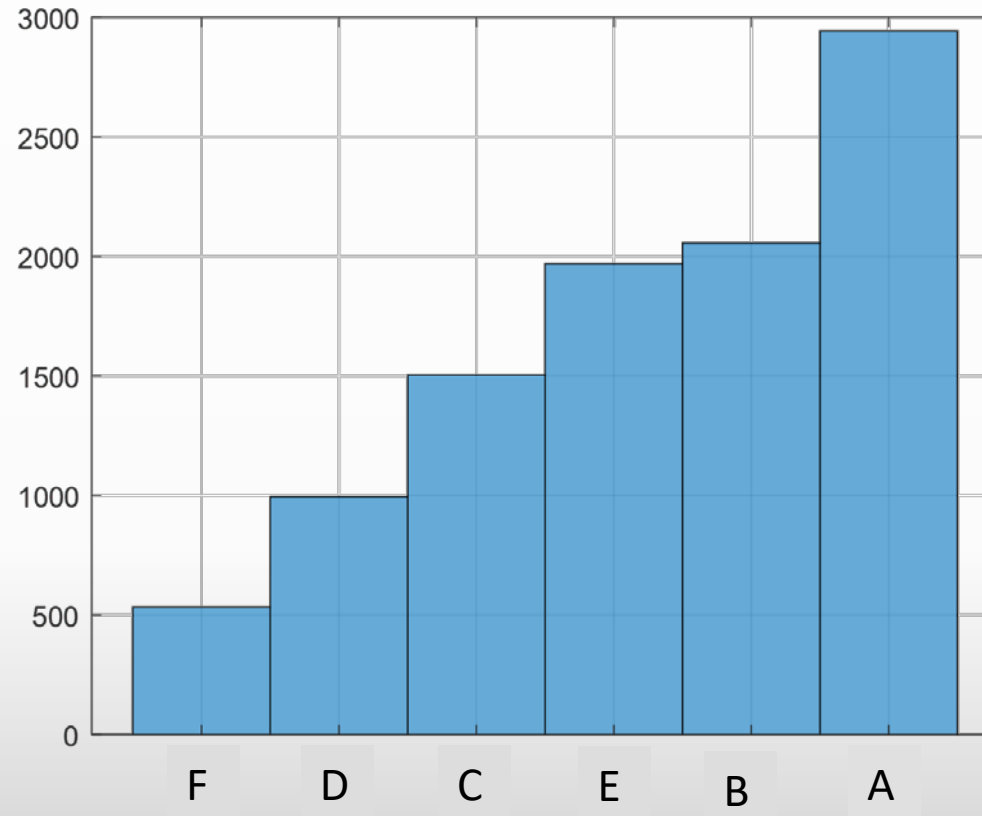
$$p_i = \frac{n_i}{\sum n_i}$$

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

Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

Ενδεικτικό Ιστόγραμμα εικόνας (10k pixels) με 6 σύμβολα – τιμές φωτεινότητας



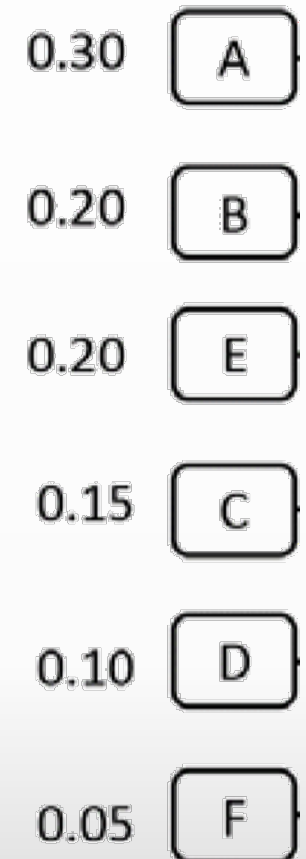
Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

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.

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

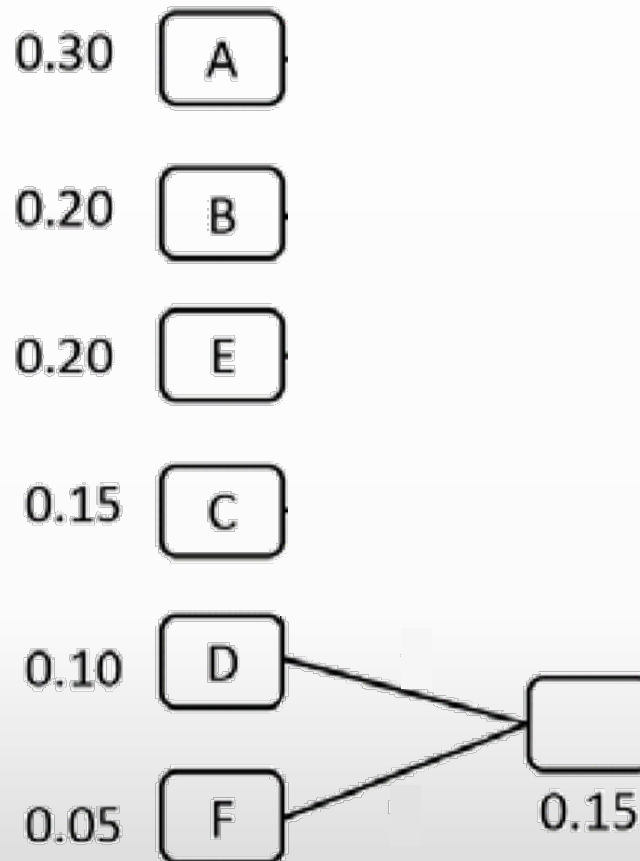


Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

Βήμα 2:

- Επιλέγονται οι δύο κόμβους με τα μικρότερα βάρη
- Δημιουργείται ένα νέος εσωτερικός σύνθετος 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.

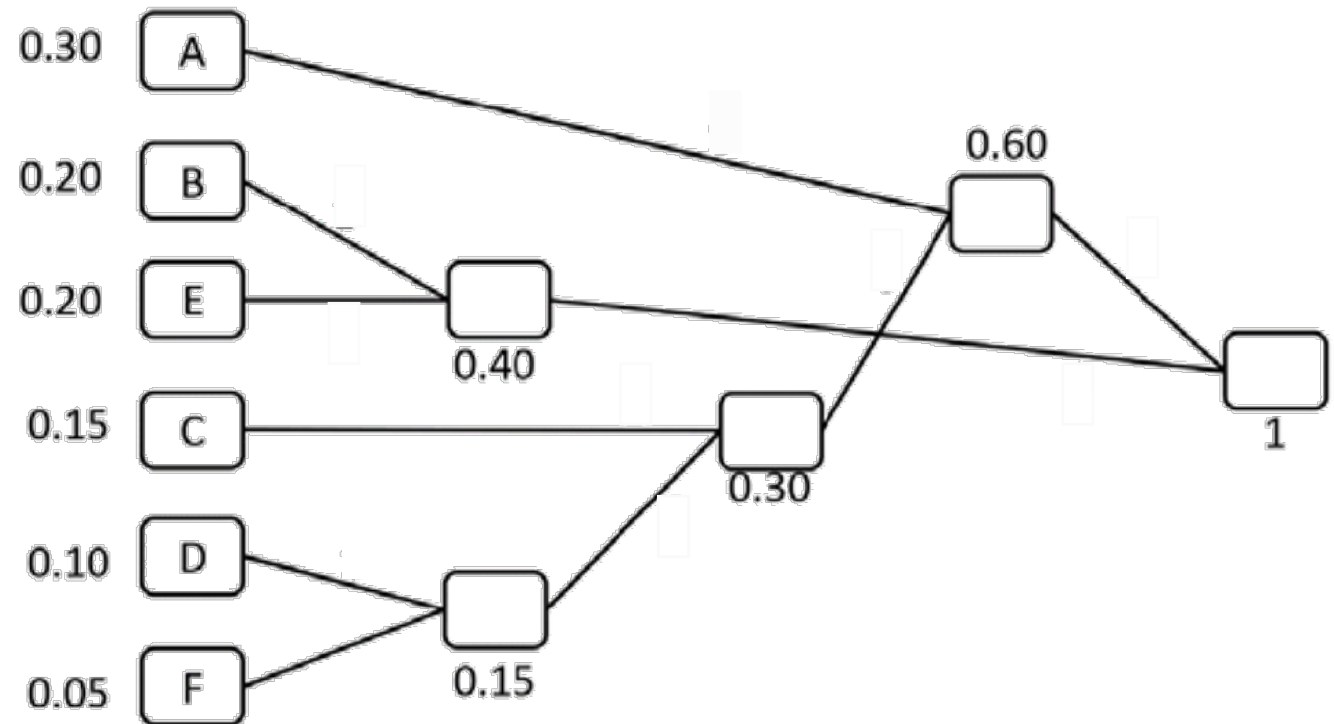


Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

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.

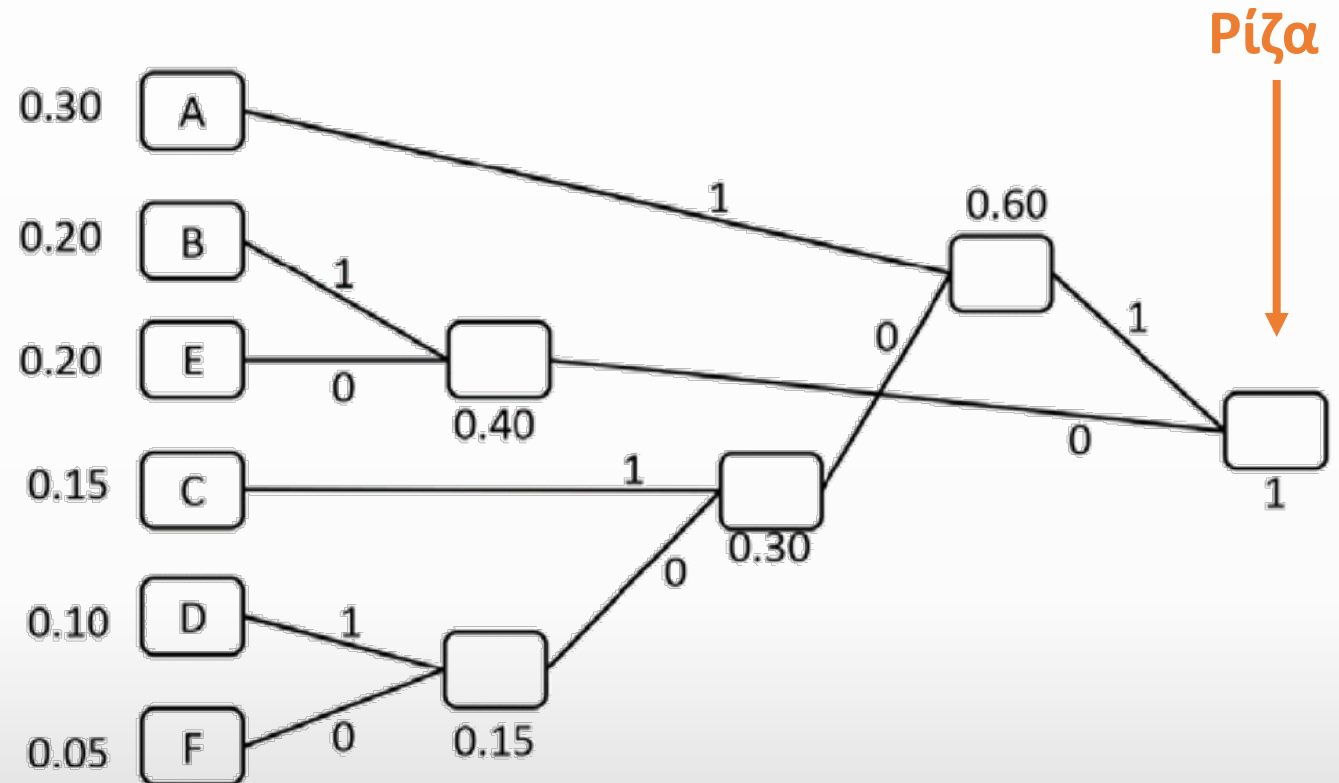


Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

Βήμα 4:

- Ξεκινώντας από τη ρίζα του δένδρου βρίσκουμε τις διαδρομές προς τα φύλλα και σε κάθε κλάδο τοποθετούμε το 0 (χαμηλή πιθανότητα) ή το 1 (υψηλή πιθανότητα)



Συμπίεση Εικόνας

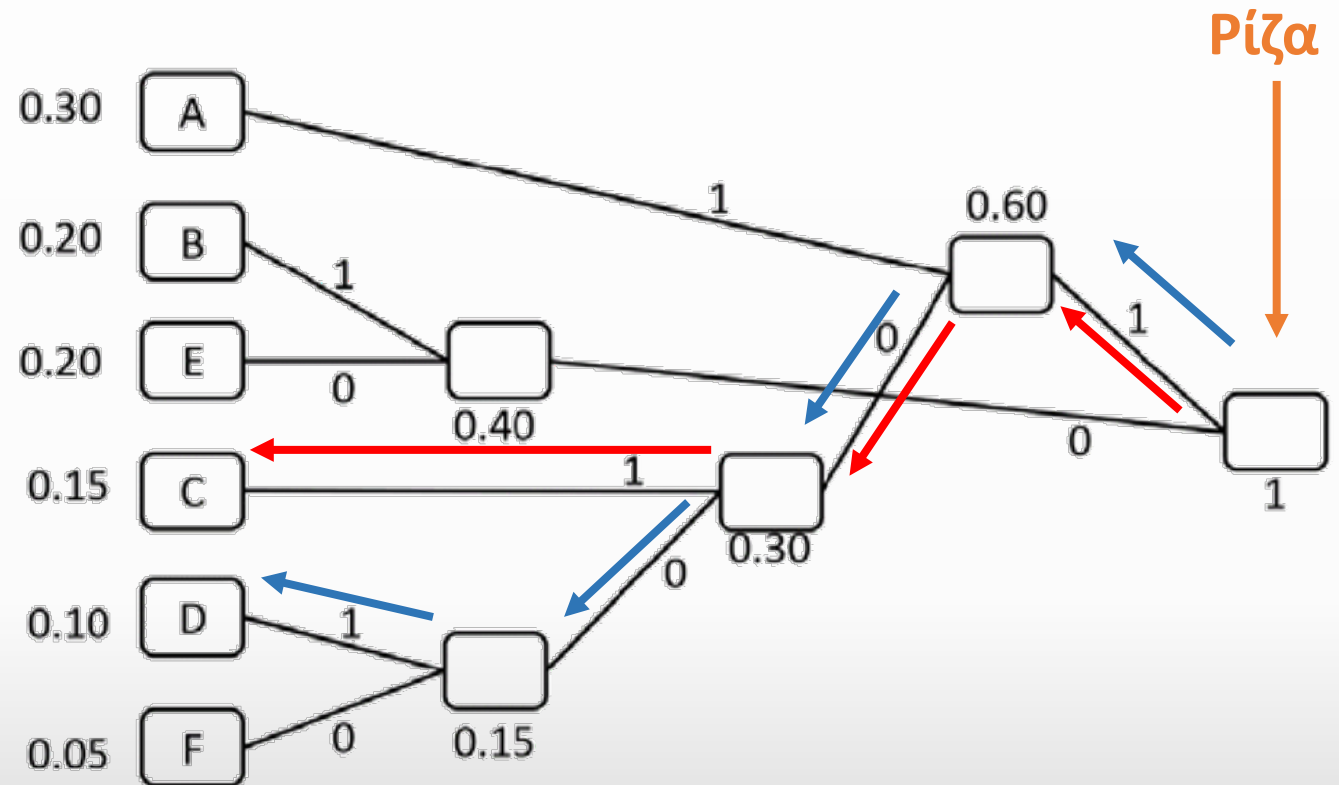
Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

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.

D: 1001

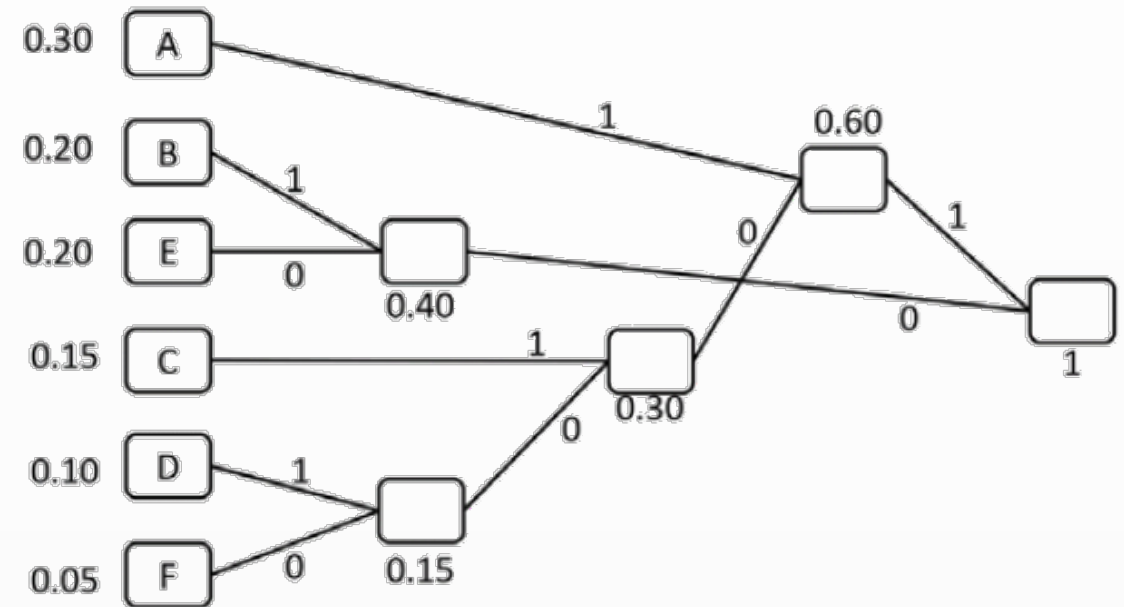
C: 101



Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

Σύμβολα	Πλήθος εμφανίσεων	Πιθανότητα εμφάνισης	Κωδικοποίηση
A	6	0,3	11
B	4	0,2	01
C	3	0,15	101
D	2	0,1	1001
E	4	0,2	00
F	1	0,05	1000
	20	1	



Χαρακτηριστικά :

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

Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

- Στο παράδειγμα,
 - 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:

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

- κάθε σύμβολο κωδικοποιείται με 2.45 bits, ενώ στην ομοιόμορφη με 3 bits ($2^2 = 4 < 6$ σύμβολα, $2^3 = 8 > 6$ σύμβολα)
- η κωδικοποίηση Huffman είναι πιο αποδοτική περίπου 18% από την ομοιόμορφη κωδικοποίηση.

Συμπίεση Εικόνας

Αλγόριθμος Huffman βέλτιστης κωδικοποίησης (δυναμικό δένδρο Huffman) :

	10									
10	5	3	6	4	3	5	4	3	6	4
	6	2	4	4	4	6	3	6	6	5
	2	2	5	3	6	4	4	4	4	5
	5	2	2	4	6	2	1	5	6	3
	6	5	5	4	2	6	6	4	6	6
	3	5	2	5	5	6	6	4	6	6
	2	5	4	5	5	3	6	5	5	6
	2	3	5	5	6	4	6	6	5	5
	6	3	6	6	6	2	5	6	6	2
	5	6	6	3	4	2	2	4	6	4

Με ομοιόμορφη κωδικοποίηση
Συνολικό μέγεθος:
 $N \times M \times B = 10 \times 10 \times 3 = 300 \text{bits}$



Σύμβολα	Πιθανότητα εμφάνισης	Κωδικοποίηση	Πλήθος εμφανίσεων στην εικόνα
A (6)	0,3	11	32
B (5)	0,2	01	23
C (3)	0,15	101	11
D (2)	0,1	1001	14
E (4)	0,2	00	19
F (1)	0,05	1000	1
	1		100

Με κωδικοποίηση Huffman
Συνολικό μέγεθος:
 $32 \times 2 + 23 \times 2 + 11 \times 3 + 14 \times 4 + 19 \times 2 + 1 \times 4 = 241 \text{bits}$

Ψηφιακή Επεξεργασία Σήματος & Εικόνας

Ελληνικό Μεσογειακό Πανεπιστήμιο
Τμήμα Ηλεκτρονικών Μηχανικών

Α. Μπακλέζος

abaklezos@hmu.gr