



ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ

Τμήμα Μηχανικών Πληροφορικής & Υπολογιστών

Διδάσκων: ΤΡΙΓΚΑ ΜΑΡΙΑ

Διάλεξη 9-10η

Ακαδημαϊκό έτος 2024-2025

ΕΙΣΑΓΩΓΗ

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

ΕΙΣΑΓΩΓΗ

Εκμετάλλευση 3 ειδών πλεονασμού (redundancy) πληροφορίας:

Πλεονασμός κωδικοποίησης

- Αναφέρεται στη μη βέλτιστη αναπαράσταση της πληροφορίας.
- Μπορούμε να χρησιμοποιήσουμε διαφορετική αναπαράσταση για κάθε επίπεδο του γκρι που εμφανίζεται με διαφορετικές πιθανότητες (π.χ. κωδικοποίηση Huffman)
- Αντί να έχουν όλα τα επίπεδα την ίδια αναπαράσταση, μπορούμε να χρησιμοποιήσουμε τεχνικές όπως η κωδικοποίηση Huffman, όπου τα πιο συχνά εμφανιζόμενα επίπεδα αντιπροσωπεύονται με μικρότερα μήκη κωδίκων.

Παράδειγμα: Αν ένα επίπεδο εμφανίζεται συχνά, του αποδίδεται μικρότερος δυαδικός κώδικας για βέλτιστη αποθήκευση.

ΕΙΣΑΓΩΓΗ

Πλεονασμός μεταξύ των pixel (interpixel)

- Προκύπτει από τη συσχέτιση μεταξύ γειτονικών pixel, οπότε μπορούμε να εκμεταλλευτούμε τα επαναλαμβανόμενα μοτίβα ή τις μεγάλες ομαλές επιφάνειες (π.χ., κωδικοποίηση μήκους διαδρομών)
- Εικόνες συχνά περιλαμβάνουν επαναλαμβανόμενα μοτίβα ή μεγάλες περιοχές με ομαλή μετάβαση, όπου τα γειτονικά pixel έχουν παρόμοιες τιμές.
- Μπορούμε να μειώσουμε αυτόν τον πλεονασμό με τεχνικές όπως η κωδικοποίηση μήκους διαδρομών (Run-Length Encoding), όπου καταγράφεται η τιμή ενός pixel μαζί με το μήκος της ακολουθίας του.

Παράδειγμα: Σε μια μεγάλη περιοχή με την ίδια τιμή γκρι, αποθηκεύουμε την τιμή και τη συχνότητα εμφάνισής της αντί για κάθε pixel ξεχωριστά.

ΕΙΣΑΓΩΓΗ

Φυσικο-οπτικός πλεονασμός

- Προκύπτει λόγω του τρόπου με τον οποίο αντιλαμβάνεται ο άνθρωπος την οπτική πληροφορία (π.χ. κβαντισμός)
- Ο ανθρώπινος εγκέφαλος δεν αντιλαμβάνεται όλες τις λεπτομέρειες ή τα χρώματα με την ίδια ακρίβεια.
- Αυτό μας επιτρέπει να συμπίεσουμε τα δεδομένα χωρίς σημαντική απώλεια ποιότητας για τον άνθρωπο.
- Τεχνικές όπως ο κβαντισμός αξιοποιούν αυτό το φαινόμενο μειώνοντας την ακρίβεια της πληροφορίας όπου η ανθρώπινη αντίληψη είναι λιγότερο ευαίσθητη.

Παράδειγμα: Στη συμπίεση JPEG, η μείωση λεπτομερειών σε υψηλές συχνότητες δεν είναι εύκολα αντιληπτή από το ανθρώπινο μάτι.

ΚΑΤΗΓΟΡΙΕΣ ΤΕΧΝΙΚΩΝ

Μη απωλεστικές τεχνικές:

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

Απωλεστικές τεχνικές:

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

ΥΒΡΙΔΙΚΕΣ ΤΕΧΝΙΚΕΣ

- **JPEG** (Joint Photographic Experts Group)

Συνδυάζει κωδικοποίηση μετασχηματισμού (DCT) και μήκους διαδρομής (RLC)

- **H263**

Συνδυάζει κωδικοποίηση μετασχηματισμού (DCT), προβλεπτική (αντιστάθμιση κίνησης – MC: motion compensation) και μήκους διαδρομής (RLC)

- **MPEG (Moving Pictures Expert Group)**

Συνδυάζει κωδικοποίηση μετασχηματισμού (DCT), προβλεπτική (αντιστάθμιση κίνησης – MC: motion compensation), μήκους διαδρομής (RLC) και στρωματοποίησης (Subband Coding) για την συμπίεση του ήχου (MP3).

ΠΛΕΟΝΑΣΜΟΣ ΚΩΔΙΚΑ

Έστω τυχαία διακριτή μεταβλητή r_k με τιμές στο διάστημα $[0, L - 1]$ η οποία χρησιμοποιείται για να αναπαραστήσει τα επίπεδα έντασης σε μια εικόνα διάστασης $M \times N$.

Κάθε μεταβλητή εμφανίζεται με μία πιθανότητα

$$p_r(r_k) = \frac{n_k}{M \cdot N}, \text{ για } k = 0, 1, 2, 3, \dots, L - 1,$$

όπου L ο αριθμός των τιμών έντασης και n_k η συχνότητα εμφάνισής του επιπέδου k στην εικόνα.

Αν $l(r_k)$ ο αριθμός των bits που χρησιμοποιούνται για να αναπαραστήσουν κάθε τιμή του r_k , τότε ο μέσος αριθμός bits για την αναπαράσταση του κάθε pixel

$$L_{\text{avg}} = \sum_{i=1}^n P(x_i) l(x_i)$$

Ο συνολικός αριθμός bits για μία εικόνα διάστασης $M \times N$ είναι

$$L_{\text{tot}} = MNL_{\text{avg}}$$

ΠΛΕΟΝΑΣΜΟΣ ΚΩΔΙΚΑ

Ο πλεονασμός αναφέρεται στις περιττές πληροφορίες που περιέχονται στα δεδομένα και μπορούν να αφαιρεθούν χωρίς απώλεια της πληροφορίας.

Συνδέεται με την εντροπία της πηγής και το μήκος του κωδικοποιημένου μηνύματος.

Ορίζεται ως:
$$R = 1 - \frac{H(X)}{L}$$

Η απόδοση ενός συστήματος συμπίεσης εξαρτάται από το λόγο συμπίεσης, την ποιότητα του συμπιεσμένου σήματος (υποκειμενικά/αντικειμενικά κριτήρια) και την πολυπλοκότητα του συστήματος (υπολογιστικό κόστος/κόστος μνήμης).

Το πρόβλημα έγκειται στο πως θα βελτιστοποιήσουμε κάθε έναν από τους 3 παράγοντες ταυτόχρονα.

ΛΟΓΟΣ ΣΥΜΠΙΕΣΗΣ

Ο λόγος συμπίεσης (Compression Ratio) δείχνει πόσο μειώθηκε το μέγεθος των δεδομένων μετά τη συμπίεση.

Ορίζεται ως:

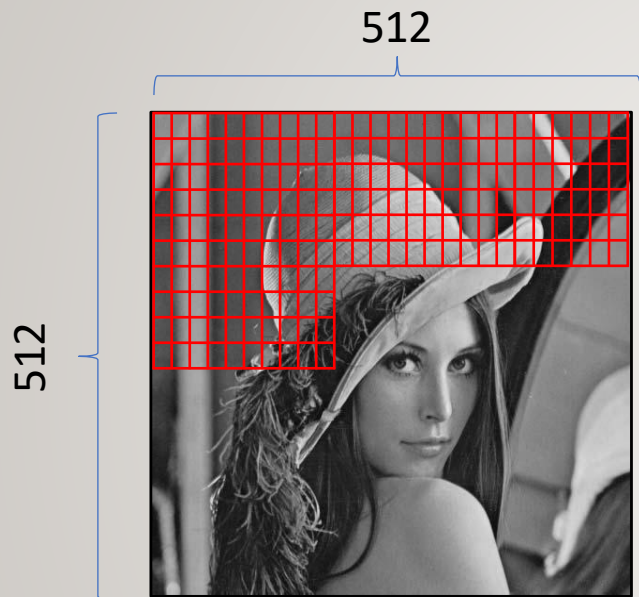
$$• R = \frac{\text{Μέγεθος σε bits αρχικών δεδομένων}}{\text{Μέγεθος σε bits συμπιεσμένων δεδομένων}}.$$

• Αν το $R > 1$ σημαίνει ότι επιτεύχθηκε συμπίεση.

• Εναλλακτικά, με ποσοστιαία μορφή:

$$\text{Ποσοστό συμπίεσης} = \left(1 - \frac{\text{Μέγεθος σε bits αρχικών δεδομένων}}{\text{Μέγεθος σε bits συμπιεσμένων δεδομένων}} \right) \times 100\%.$$

GRAYSCALE ΕΙΚΟΝΕΣ



262.144 pixels

2.097.152 bits!!

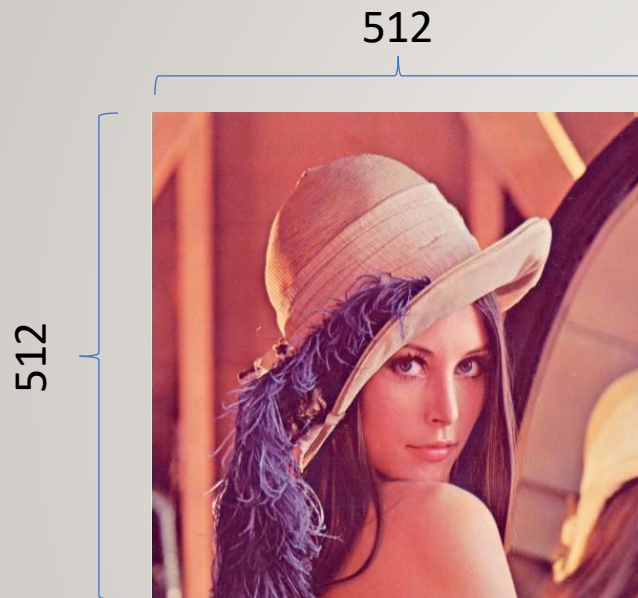
Παράδειγμα

- 256 διακριτά επίπεδα του γκρι
- Το εύρος αυτό είναι από 0 (μαύρο) και 255 (λευκό).
- Κάθε επίπεδο αναπαρίστανται από 8bits/pixel.

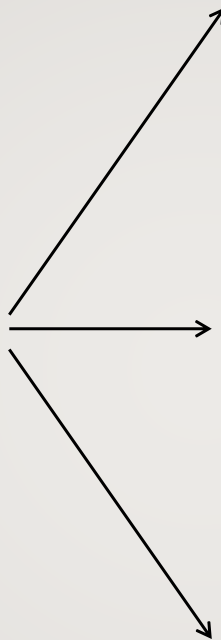
i.e. 00000101 = ?

$$0 \times 2^7 + 0 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 4 + 1 = 5$$

ΕΓΧΡΩΜΗ ΕΙΚΟΝΑ



6.291.456 bits!!



3 * 262.144 pixels

ΣΤΑΔΙΑ

Τα βασικά στάδια μιας απωλεστικής τεχνικής:

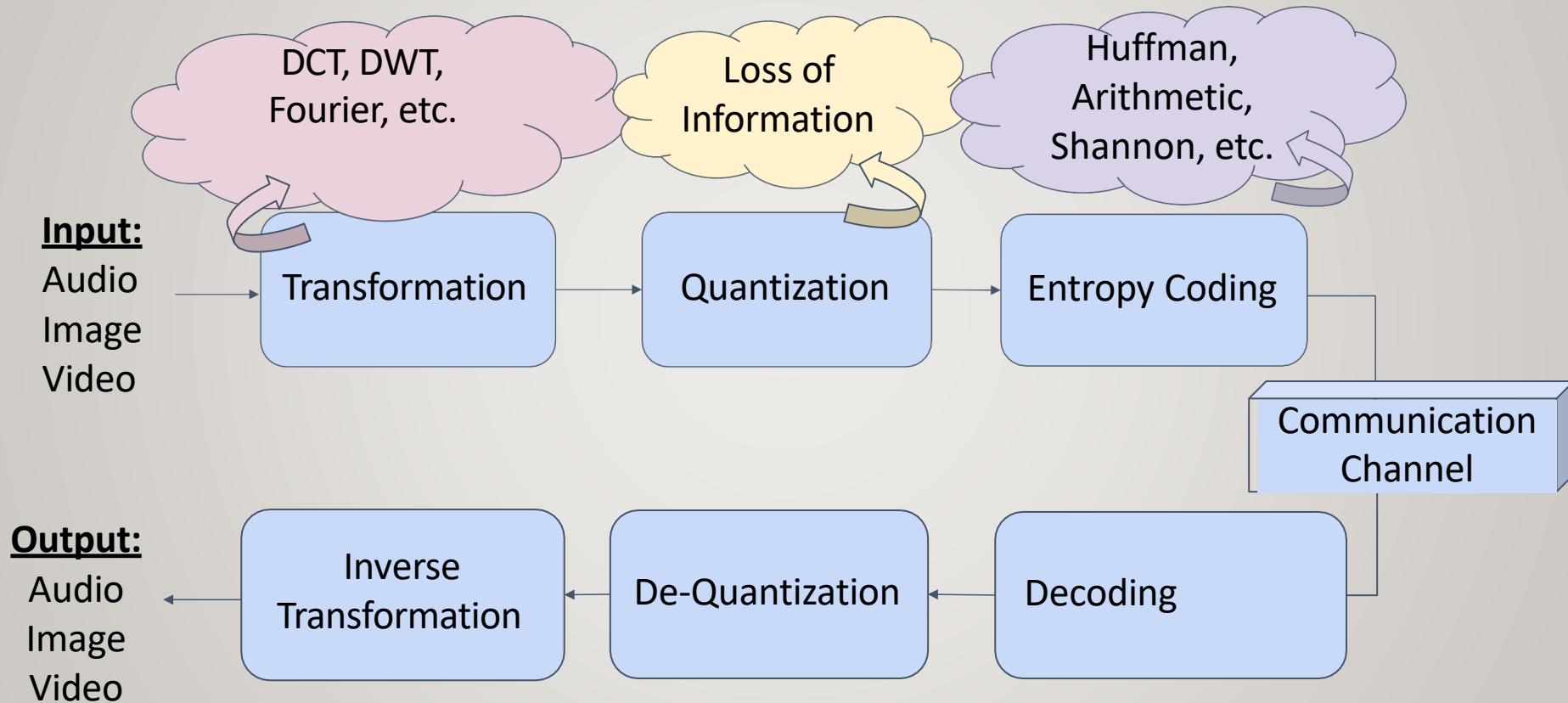
1. Μετασχηματισμός της εικόνας στο κατάλληλο πεδίο

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

ΣΤΑΔΙΑ

2. Κβαντισμός των αποτελεσμάτων του 1ου σταδίου
 - ομοιόμορφος
 - ανομοιόμορφος
 - διανυσματικός
3. Προσδιορισμός του λεξικού (codebook design) για την αναπαράσταση των εξόδων του κβαντιστή

ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΣΥΜΠΙΕΣΗΣ



ΚΒΑΝΤΟΠΟΙΗΣΗ

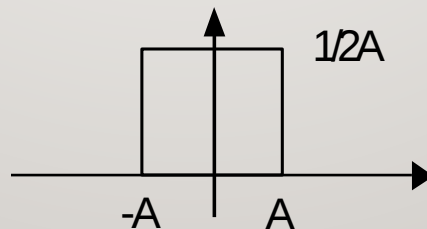
- **Κβαντοποίηση** είναι η θεμελιώδης λειτουργία ενός συστήματος συμπίεσης. Ο σκοπός της είναι να επιλέγει, για μια δεδομένη τιμή εισόδου, τον πλησιέστερο γείτονα που ανήκει σε ένα προκαθορισμένο πεπερασμένο σύνολο ψηφιακών τιμών.
- Ένας βαθμωτός κβαντιστής μεγέθους L είναι μια απεικόνιση Q από το R σε ένα πεπερασμένο σύνολο C , το οποίο ονομάζεται και λεξικό (ή βιβλίο κώδικα - codebook), που περιέχει L κβαντισμένα σύμβολα:

$$Q: R \rightarrow C = \{\hat{s}_1, \hat{s}_2, \dots, \hat{s}_L\}$$

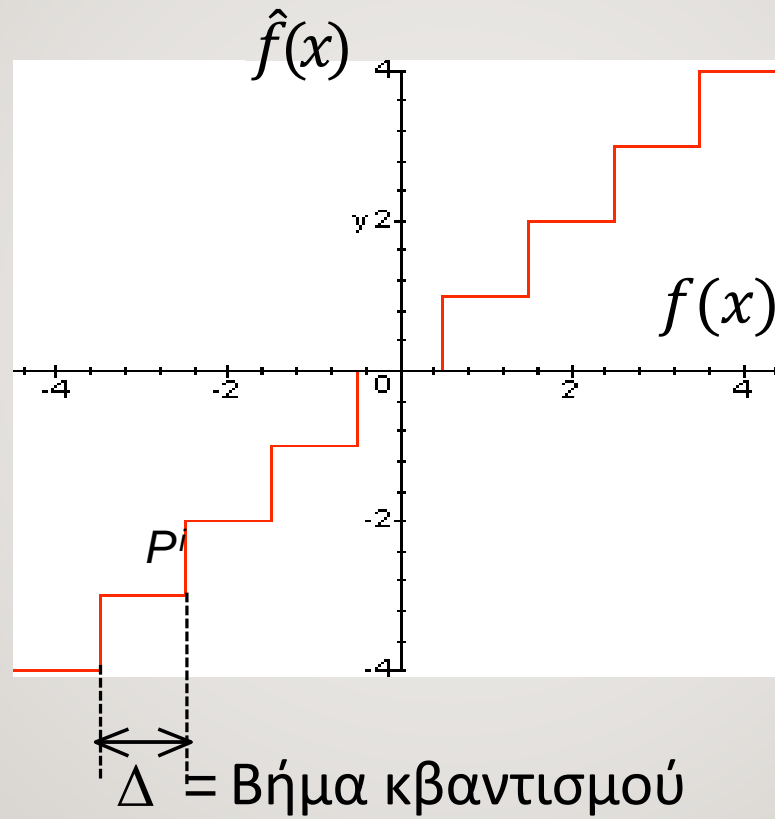
όπου $\hat{s} = Q(s)$ η κβαντισμένη εκδοχή του s

ΟΜΟΙΟΜΟΡΦΟΣ ΒΑΘΜΩΤΟΣ ΚΒΑΝΤΙΣΤΗΣ

- Θεωρούμε ένα διακριτό σήμα x με τιμές ομοιόμορφα κατανομημένες στο διάστημα $[-A, A]$, άρα συνάρτηση πυκνότητας πιθανότητα $f_k(x) = \frac{1}{2A} \quad \forall k \in [-A, A]$
- Διαιρούμε το διάστημα σε $L = 2^N$ διαφορετικά διαστήματα $\{P^1, \dots, P^L\}$, ίσου μήκους $\Delta = \frac{x_{\max} - x_{\min}}{L}$ ή $\Delta = \frac{2A}{2^N}$, που λέγονται ζώνες του κβαντιστή., όπου
N: ο αριθμός των bits που θα χρησιμοποιηθούν
max_value: η μέγιστη αποδεκτή τιμή του σήματος εισόδου
min_value: η ελάχιστη αποδεκτή τιμή του σήματος εισόδου
- Σε κάθε διάστημα ορίζεται μία αντιπροσωπευτική τιμή που λέγεται κέντρο του κβαντιστή.



ΧΑΡΑΚΤΗΡΙΣΤΙΚΗ ΣΥΝΑΡΤΗΣΗ



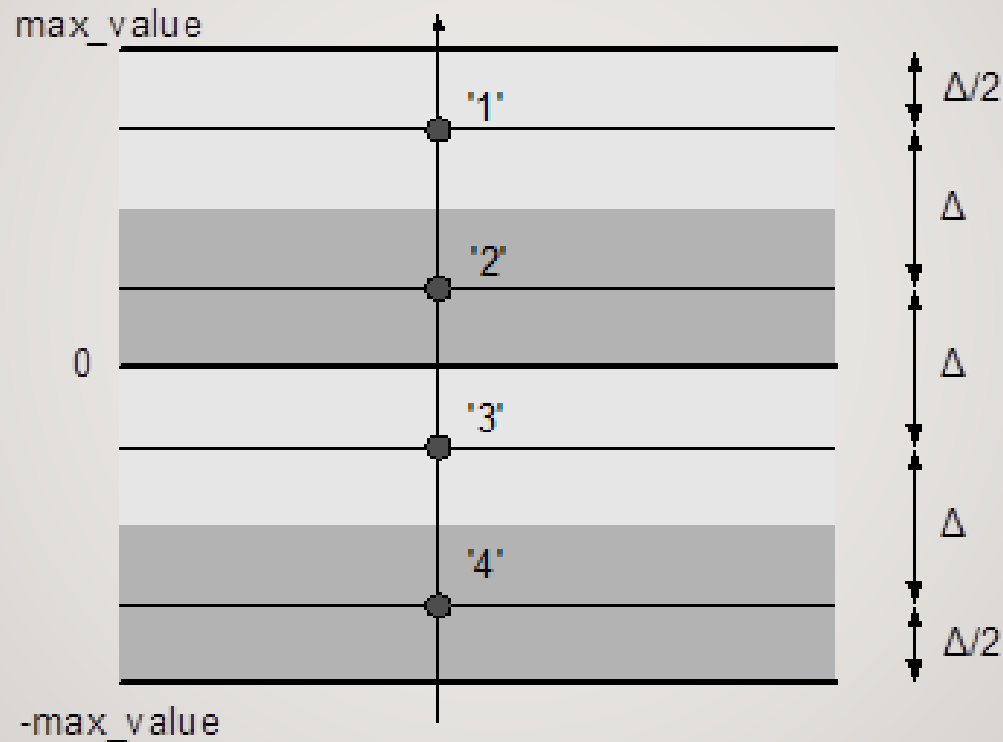
ΟΜΟΙΟΜΟΡΦΟΣ ΒΑΘΜΩΤΟΣ ΚΒΑΝΤΙΣΤΗΣ

Ειδικότερα, ο κβαντιστής ακολουθεί τα παρακάτω βήματα:

1. Περιορίζει τη δυναμική περιοχή του σήματος εισόδου στις τιμές $[\text{min_value} \text{ max_value}]$, θέτοντας τα δείγματα που βρίσκονται εκτός δυναμικής περιοχής στην αντίστοιχη ακραία αποδεκτή τιμή.
2. Στη συνέχεια υπολογίζει το βήμα κβαντισμού Δ και τα κέντρα της κάθε περιοχής.
3. Βρίσκει την περιοχή στην οποία ανήκει κάθε δείγμα του σήματος εισόδου, και βγάζει ως έξοδο το διάνυσμα
$$x_q = [1, 2, \dots, 2^R] .$$
4. Το διάνυσμα x_q μπορεί να χρησιμοποιηθεί ως δείκτης και να πάρουμε το κβαντισμένο σήμα.

ΠΑΡΑΔΕΙΓΜΑ

Παράδειγμα επιπέδων ομοιόμορφου βαθμωτού κβαντιστή για $R = 2\text{bits}$, όπου $\text{min_value} = -\text{max_value}$

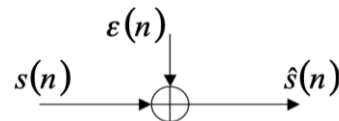


DISTORTION

Ο κβαντισμός εισάγει σφάλμα $\varepsilon(n)$ το οποίο προστίθεται στην είσοδο $s(n)$

Το distortion D ανάμεσα στο σήμα εισόδου και το σήμα εξόδου γενικά αξιολογείται μέσω του *Μέσου Τετραγωνικού Σφάλματος - MSE*

$$\begin{aligned}\varepsilon(n) &= s(n) - Q(s(n)) \\ &= s(n) - \hat{s}(n)\end{aligned}$$

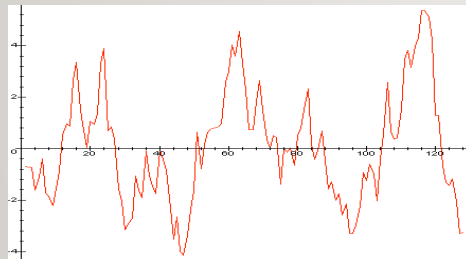


$$d = \frac{1}{N} \sum_{n=1}^N d(s(n), \hat{s}(n))$$

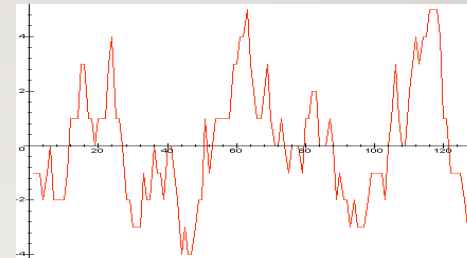
$$d(s(n), \hat{s}(n)) = |s(n) - \hat{s}(n)|^2 = \varepsilon(n)^2$$

ΠΑΡΑΔΕΙΓΜΑ

signal original

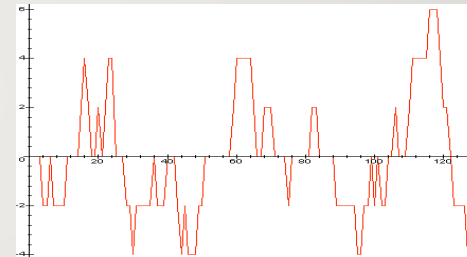


$\Delta = 1$



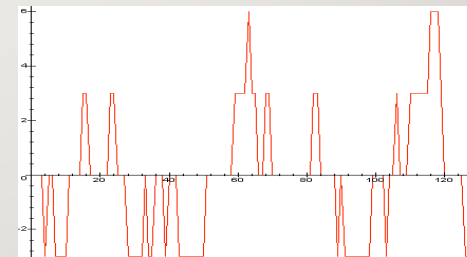
$D=0,088$

$\Delta = 2$



$D=0,348$

$\Delta = 3$



$D=0,746$

ΜΗ ΟΜΟΙΟΜΟΡΦΟΣ ΒΑΘΜΩΤΟΣ ΚΒΑΝΤΙΣΤΗΣ

- Η υπόθεση ενός ομοιόμορφα κατανεμημένου σήματος φτιαγμένο για τον ομοιόμορφο βαθμωτό κβαντιστή, φυσικά, είναι πολύ «φτωχή» όταν έχουμε να κάνουμε με σήματα ομιλίας, μουσικής, εικόνας ή βίντεο.
- Οι στατιστικές ιδιότητες εξελίσσονται συνεχώς: μόνο στις καλύτερες περιπτώσεις, ένα σήμα μπορεί να θεωρηθεί τοπικά ακίνητο.
- Οι νόμοι πιθανοτήτων δεν είναι ομοιόμορφοι

LLOYD-MAX ΑΛΓΟΡΙΘΜΟΣ

- Για τη μη ομοιόμορφη κβάντιση του διανύσματος εισόδου θα χρησιμοποιηθεί ο αλγόριθμος Lloyd-Max ο οποίος επιτρέπει την σχεδίαση βέλτιστου κβαντιστή για οποιοδήποτε αριθμό επιπέδων.
- Αρχικά επιλέγετε ένα τυχαίο σύνολο επιπέδων κβαντισμού που συνήθως είναι τα κέντρα του ομοιόμορφου κβαντιστή.

LLOYD-MAX ΑΛΓΟΡΙΘΜΟΣ

Σε κάθε επανάληψη i του Αλγόριθμου Lloyd-Max:

1. Υπολογίζουμε τα όρια των ζωνών κβαντισμού, που πρέπει να είναι στο μέσον των επιπέδων κβαντισμού, δηλαδή

$$T_k = (\tilde{x}_k^{(i)} + \tilde{x}_{k+1}^{(i)}) / 2, \quad 1 \leq k \leq M - 1$$

2. Υπολογίζουμε το κβαντισμένο σήμα με βάση τις περιοχές αυτές και μετράμε την μέση παραμόρφωση D_i με βάση το δοθέν σήμα
3. Τα νέα επίπεδα κβαντισμού είναι τα κεντροειδή των ζωνών:

$$\tilde{x}_k^{(i+1)} = E[x | T_{k-1} < x < T_k]$$

4. Επαναλαμβάνουμε τα 3 τελευταία βήματα μέχρις ότου $|D_i - D_{i-1}| < \varepsilon$

Η τιμή του ε καθορίζει και τον αριθμό των $K_{\max}$ επαναλήψεων.

LLOYD-MAX

Για να ορίσουμε τον κβαντιστή, πρέπει να βρούμε τη διαίρεση $\{P^1, \dots, P^L\}$ και τις Αντιπροσωπευτικές τιμές $\{\hat{s}_1, \hat{s}_2, \dots, \hat{s}_L\}$ που ελαχιστοποιούν το distortion

$$D = E[(S - Q(S))^2] = \sum_{i=1}^L \int_{P^i} (s - \hat{s}^i)^2 f_s(s) ds$$

ΑΝΑΓΚΑΙΕΣ ΣΥΝΘΗΚΕΣ

CONDITION 1

❖ Given a dictionary $\{s^1, \dots, s^L\}$ the **best partition** is the one which satisfies:

$$P^i = \{s \mid (s - \hat{s}^i)^2 \leq (s - \hat{s}^j)^2 \forall j \in \{1, \dots, L\}\}$$

❖ This is the **nearest neighbor** approach.

CONDITION 2

❖ Given the partition $\{P^1, \dots, P^L\}$ the **best dictionary** is the one which satisfies:

$$\hat{s}^i = \frac{\int_{x \in P^i} x f_s(x) dx}{\int_{x \in P^i} f_s(x) dx} \quad \forall i = \{1, \dots, L\}$$

❖ This is the **centroid condition**.

ΣΥΜΠΙΕΣΗ/ΑΠΟΣΥΜΠΙΕΣΗ

Η διαδικασία συμπίεσης αποφασίζει σε ποιο διάστημα ανήκει η $f(x)$, συσχετίζοντας έναν αριθμό

$$i(x) \in \{1, 2^1, \dots, 2^R\}$$

Ο αριθμός διαστήματος $i(x)$ μεταδίδεται ή αποθηκεύεται.

Η διαδικασία αποσυμπίεσης συνίσταται στη συσχέτιση του αριθμού διαστήματος $i(x)$ με την αντιπροσωπευτική τιμή

$$\hat{f}(x) = \hat{f}^{i(x)}$$

το οποίο ορίζεται από το codebook $C = \{\hat{f}^1, \hat{f}^2, \dots, \hat{f}^L\}$

ΤΕΧΝΙΚΕΣ ΚΩΔΙΚΟΠΟΙΗΣΗΣ

ΚΩΔΙΚΟΠΟΙΗΣΗ

- **Κωδικοποίηση** ονομάζουμε την αναπαράσταση των κβαντισμένων τιμών πλάτους των δειγμάτων του σήματος σε δυαδική ακολουθία (ακολουθία δυαδικών ψηφίων).
- Ο κώδικας κάθε τιμής στάθμης ονομάζεται **κωδική λέξη** (codeword).
- Το σύνολο των διαφορετικών κωδικών λέξεων που χρησιμοποιούνται για την κωδικοποίηση ονομάζεται **κώδικας**.
- Όταν το σύνολο των κωδικών λέξεων έχει το ίδιο μέγεθος (αριθμό bits) τότε ο κώδικας ονομάζεται **κώδικας σταθερού μήκους**.

ΕΝΤΡΟΠΙΑ SHANNON

- Το **Θεώρημα Εντροπίας του Shannon** είναι ένα θεμελιώδες αποτέλεσμα στη θεωρία πληροφορίας που παρέχει το όριο της συμπίεσης μιας πηγής δεδομένων χωρίς απώλειες. Διατυπώθηκε από τον Claude Shannon το 1948.
- Η **εντροπία** ενός διακριτού συνόλου πιθανών γεγονότων (ή ενός διακριτού αλφαβήτου) $X = \{x_1, x_2, \dots, x_n\}$ με πιθανότητες $P(x_i)$ ορίζεται ως

$$H(X) = - \sum_{i=1}^n P(x_i) \log_2(P(x_i))$$

όπου:

- $H(X)$: η μέση πληροφορία (εντροπία) που περιέχεται σε μια πηγή X .
 $H(X) \geq 0$, Λαμβάνει τη μέγιστη τιμή όταν τα σύμβολα της πηγής είναι ισοπίθανα.
- $P(x_i)$ η πιθανότητα εμφάνισης του i -οστού συμβόλου x_i .
- Η μονάδα της εντροπίας είναι τα **bits** αν ο λογάριθμος έχει ως βάση 2.
- Η εντροπία μετρά την αβεβαιότητα ή την πληροφορία που περιέχει η πηγή.

ΕΝΤΡΟΠΙΑ

- Όταν οι πιθανότητες είναι ίσες (π.χ., 50%-50%), η εντροπία είναι μέγιστη: Δεν μπορούμε να προβλέψουμε το αποτέλεσμα και κάθε σύμβολο μεταφέρει τη μέγιστη δυνατή πληροφορία.
- Όταν ένα σύμβολο είναι πολύ πιθανό (π.χ., 90%-10%), η εντροπία μειώνεται: Γνωρίζουμε σχεδόν σίγουρα ότι το σύμβολο θα είναι το πιο πιθανό, οπότε χρειαζόμαστε λιγότερη πληροφορία για να το περιγράψουμε.
- Μικρότερη αβεβαιότητα σημαίνει μικρότερη εντροπία: Ένα πιο προβλέψιμο σύστημα απαιτεί λιγότερα "bits" για να μεταδώσει την πληροφορία του.

Ι^ο ΘΕΩΡΗΜΑ ΤΗΣ ΕΝΤΡΟΠΙΑΣ ΤΟΥ SHANNON

- Το Θεώρημα της Εντροπίας του Shannon, γνωστό και ως Θεώρημα Κωδικοποίησης Χωρίς Θόρυβο (Noiseless Coding Theorem), διαβεβαιώνει ότι μια πηγή μπορεί να αναπαρασταθεί με τον ελάχιστο αριθμό bits.
- Μαθηματικά, περιγράφεται ως εξής: $\lim_{n \rightarrow \infty} \frac{L_{avg,n}}{n} = H$
όπου:
 - $L_{avg,n}$: το μέσο μήκος των συμβόλων σε n δείγματα.
 - H : η εντροπία της πηγής.
 - Αυτό σημαίνει ότι για μεγάλο αριθμό δεδομένων, ο μέσος αριθμός bits ανά σύμβολο πλησιάζει την εντροπία H

ΠΑΡΑΔΕΙΓΜΑ

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8		
$r_{128} = 128$	0.47	10000000	8		
$r_{186} = 186$	0.25	11000100	8		
$r_{255} = 255$	0.03	11111111	8		
r_k for $k \neq 87, 128, 186, 255$	0	—	8		

Code 1: $L_{\text{avg}} = 0.25 \times 8 + 0.47 \times 8 + 0.25 \times 8 + 0.03 \times 8 = 8$ bits /pixel

Shannon: $H = -\{0.25 \log 0.25 + 0.47 \log 0.47 + 0.25 \log 0.25 + 0.03 \log 0.03\}$
 $= 1.6614$ bits/pixel

ΠΑΡΑΔΕΙΓΜΑ

r_k	$p_r(r_k)$	Code 1	$l_1(r_k)$	Code 2	$l_2(r_k)$
$r_{87} = 87$	0.25	01010111	8	01	2
$r_{128} = 128$	0.47	10000000	8	1	1
$r_{186} = 186$	0.25	11000100	8	000	3
$r_{255} = 255$	0.03	11111111	8	001	3
r_k for $k \neq 87, 128, 186, 255$	0	—	8	—	0

Code 1: $L_{\text{avg}} = 0.25 \times 8 + 0.47 \times 8 + 0.25 \times 8 + 0.03 \times 8 = 8$ bits

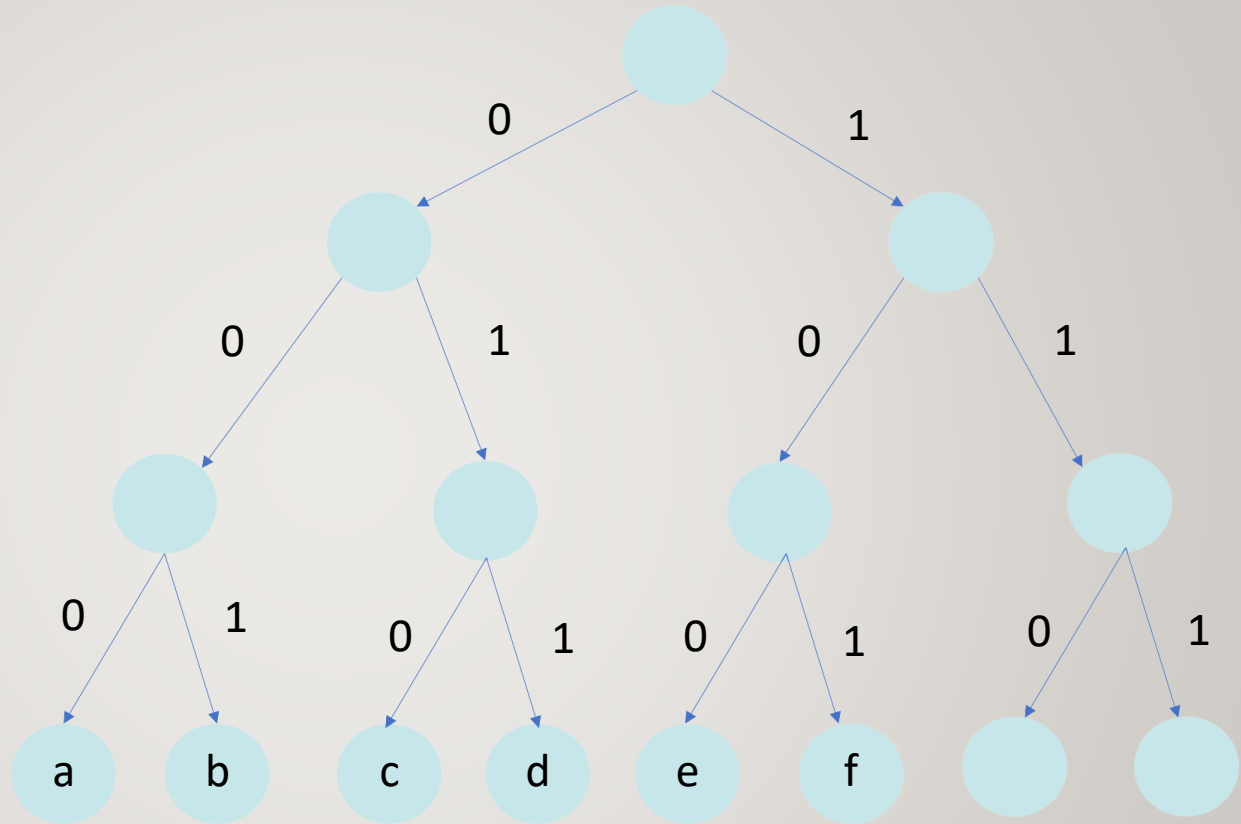
Code 2: $L_{\text{avg}} = 0.25 \times 2 + 0.47 \times 1 + 0.25 \times 3 + 0.03 \times 3 = 1.81$ bits

$$C = \frac{256 \times 256 \times 8}{118.621} = \frac{8}{1.81} \approx 4.42$$

$$R = 1 - \frac{1}{4.42} = 0.774$$

ΚΩΔΙΚΟΠΟΙΗΣΗ ΣΤΑΘΕΡΟΥ ΜΗΚΟΥΣ

Symbol	Probability	Code
a	0.4	000
b	0.3	001
c	0.1	010
d	0.1	011
e	0.06	100
f	0.04	101



$$H = 0.4 \times 3 + 0.3 \times 3 + 0.1 \times 3 + 0.1 \times 3 + 0.06 \times 3 + 0.04 \times 3 = 3\text{bits/pixel}$$

ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN

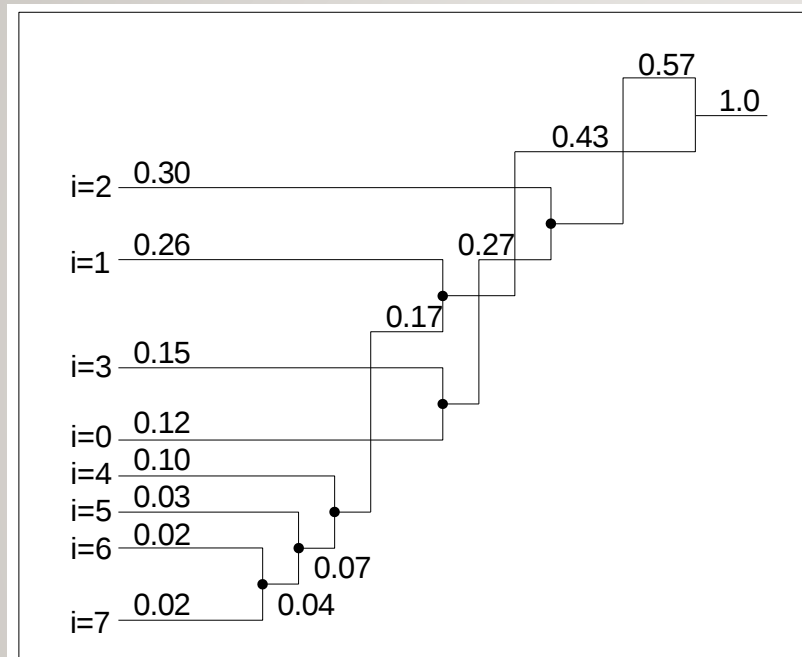
Έστω εικόνα $N \times M$ που αναπαριστάται με B bits/pixel. Μπορούμε να εκτιμήσουμε την *συνάρτηση πυκνότητας πιθανότητας* $p(i)$ κατασκευάζοντας το ιστόγραμμα της.

- **Κωδικοποίηση Εντροπίας:** αντιστοιχίζουμε μικρές κωδικές λέξεις στα επίπεδα φωτεινότητας που εμφανίζονται με μεγάλη πιθανότητα και μεγάλες κωδικές λέξεις σε αυτά που εμφανίζονται σπανιότερα.
- Τα μήκη των λέξεων επιλέγονται έτσι ώστε το μέσο μήκος να ελαχιστοποιείται. Από την Θεωρία Πληροφορίας ισχύει:

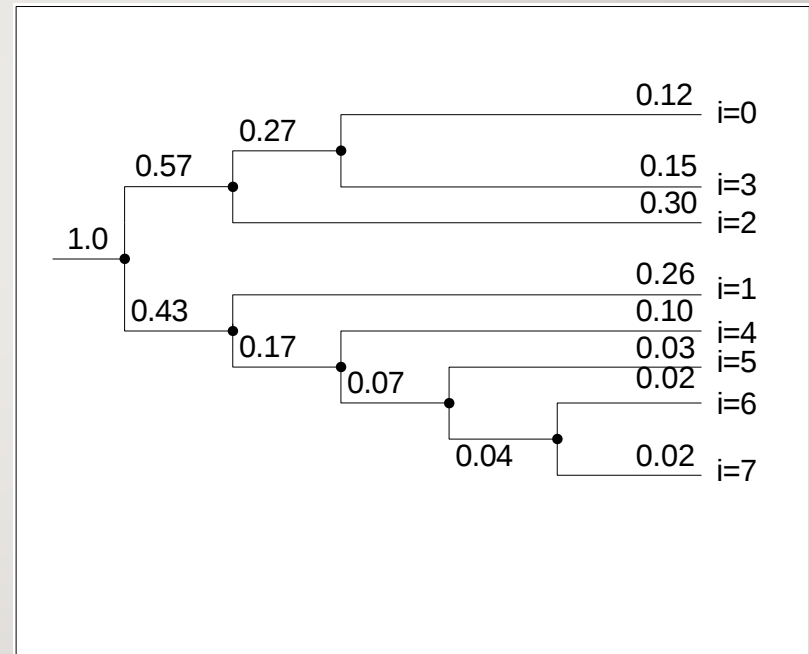
$$H(B) \leq \bar{L} \leq H(B) + 1$$

όπου $H(B)$ η εντροπία της εικόνας.

ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN



Κατασκευή του δέντρου
Huffman ←



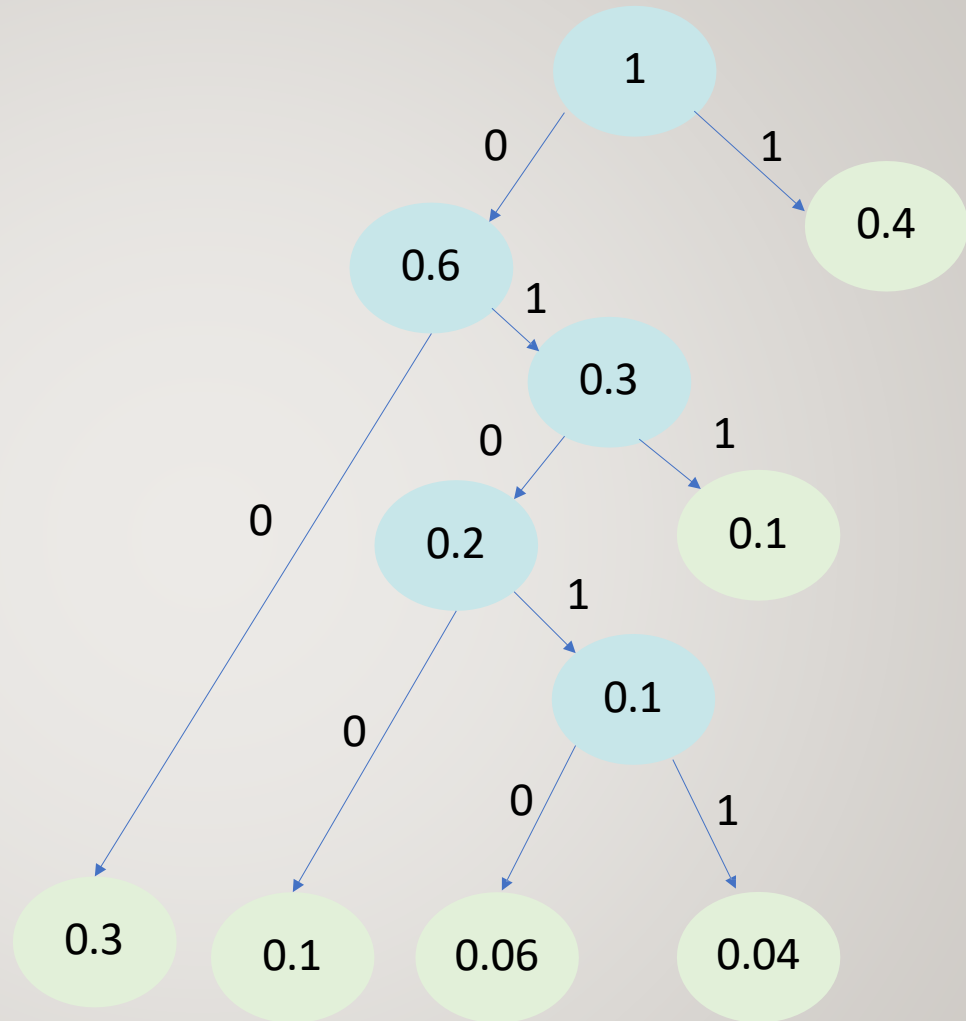
Δημιουργία των
κωδικών λέξεων →

Το δέντρο Huffman αντιστοιχεί τις κωδικές λέξεις στα σύμβολα 0-7 με τις ως
εξής: Αριστερά κλαδιά: 0, Δεξιά κλαδιά: 1

ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN

Symbol	Probability	Code
a	0.4	1
b	0.3	00
c	0.1	011
d	0.1	0100
e	0.06	01010
f	0.04	01011

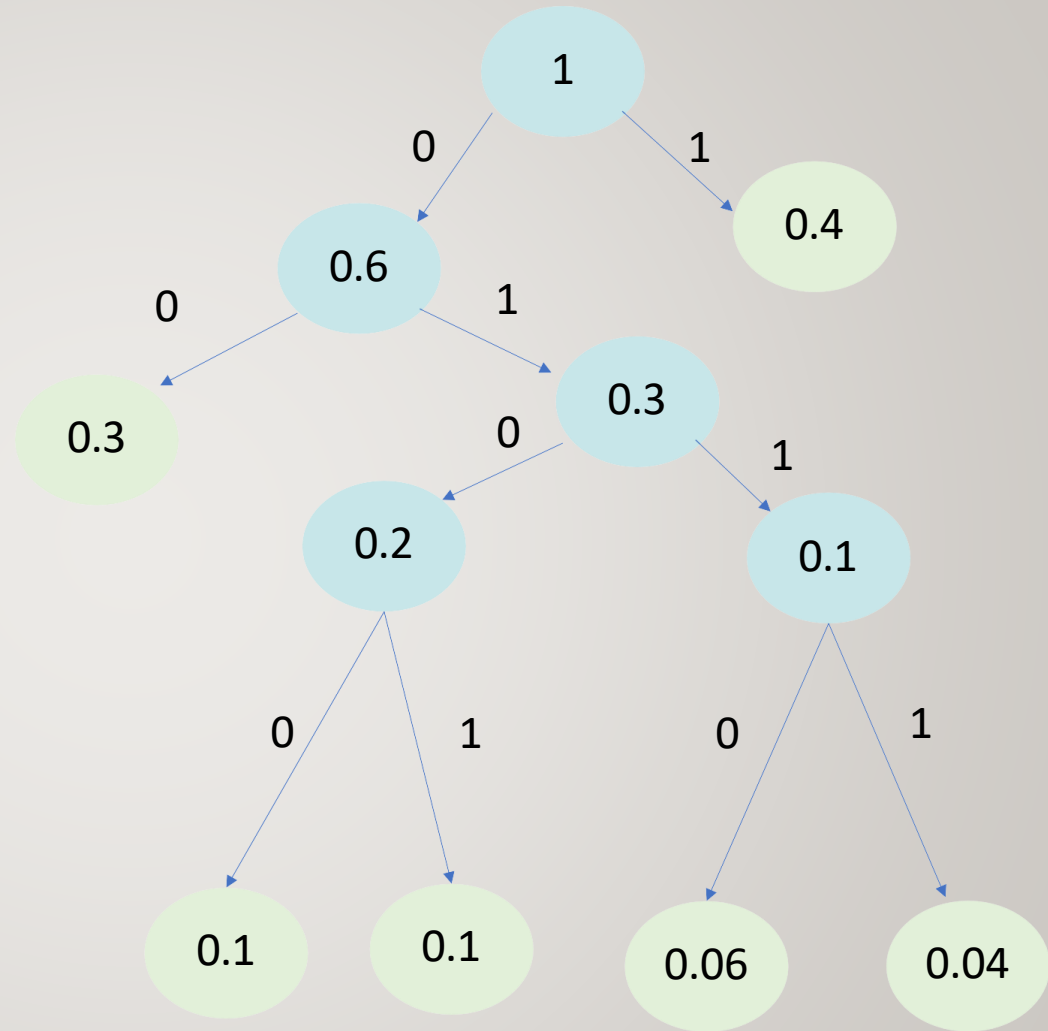
$$H = 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 3 + 0.1 \times 4 + 0.06 \times 5 + 0.04 \times 5 = 2.2 \text{ bits/pixel}$$



ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN

Symbol	Probability	Code
a	0.4	1
b	0.3	00
c	0.1	0100
d	0.1	0101
e	0.06	0110
f	0.04	0111

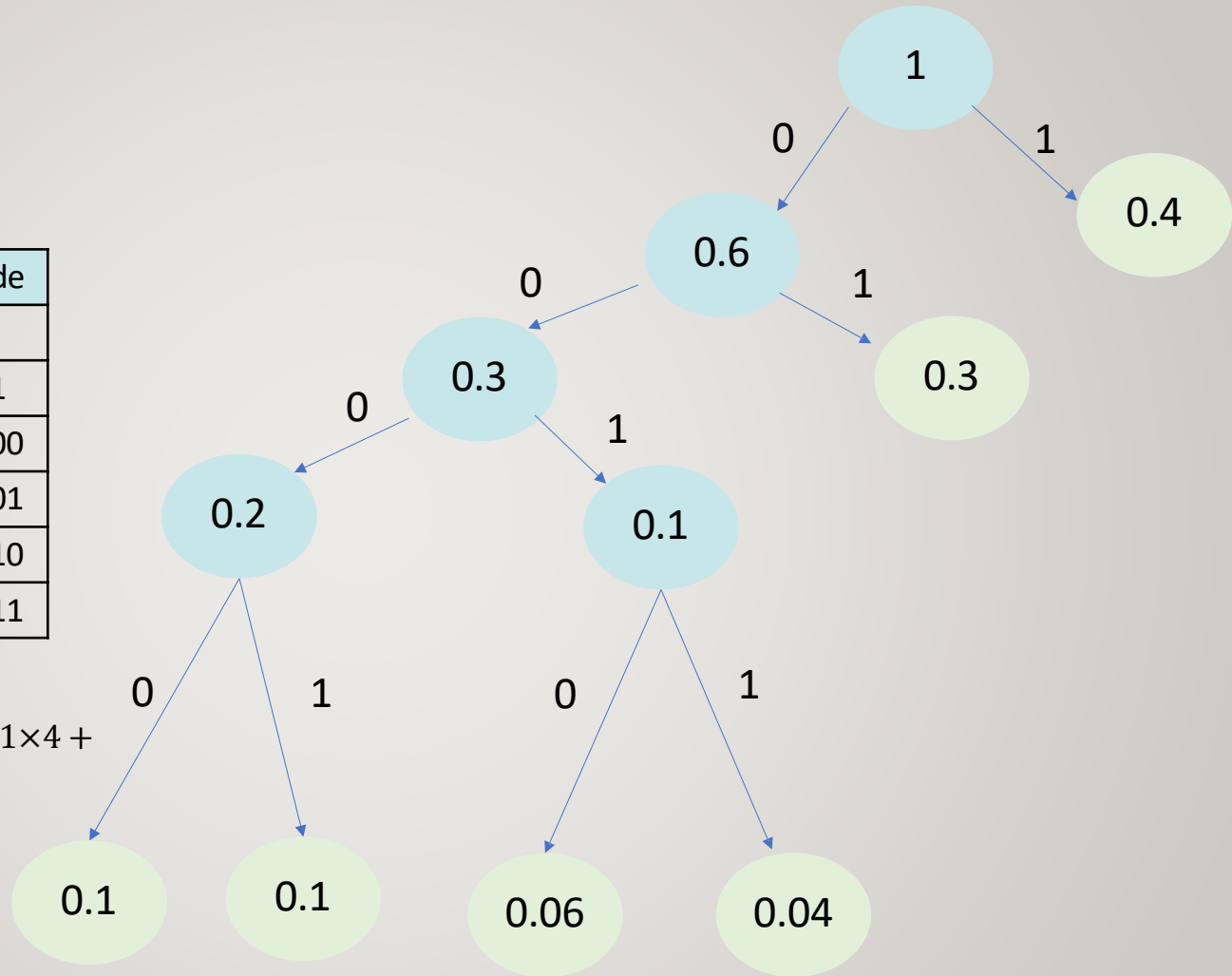
$$H = 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 4 + 0.1 \times 4 + 0.06 \times 4 + 0.04 \times 4 = 2.2 \text{ bits/pixel}$$



ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN

Symbol	Probability	Code
a	0.4	1
b	0.3	01
c	0.1	0000
d	0.1	0001
e	0.06	0010
f	0.04	0011

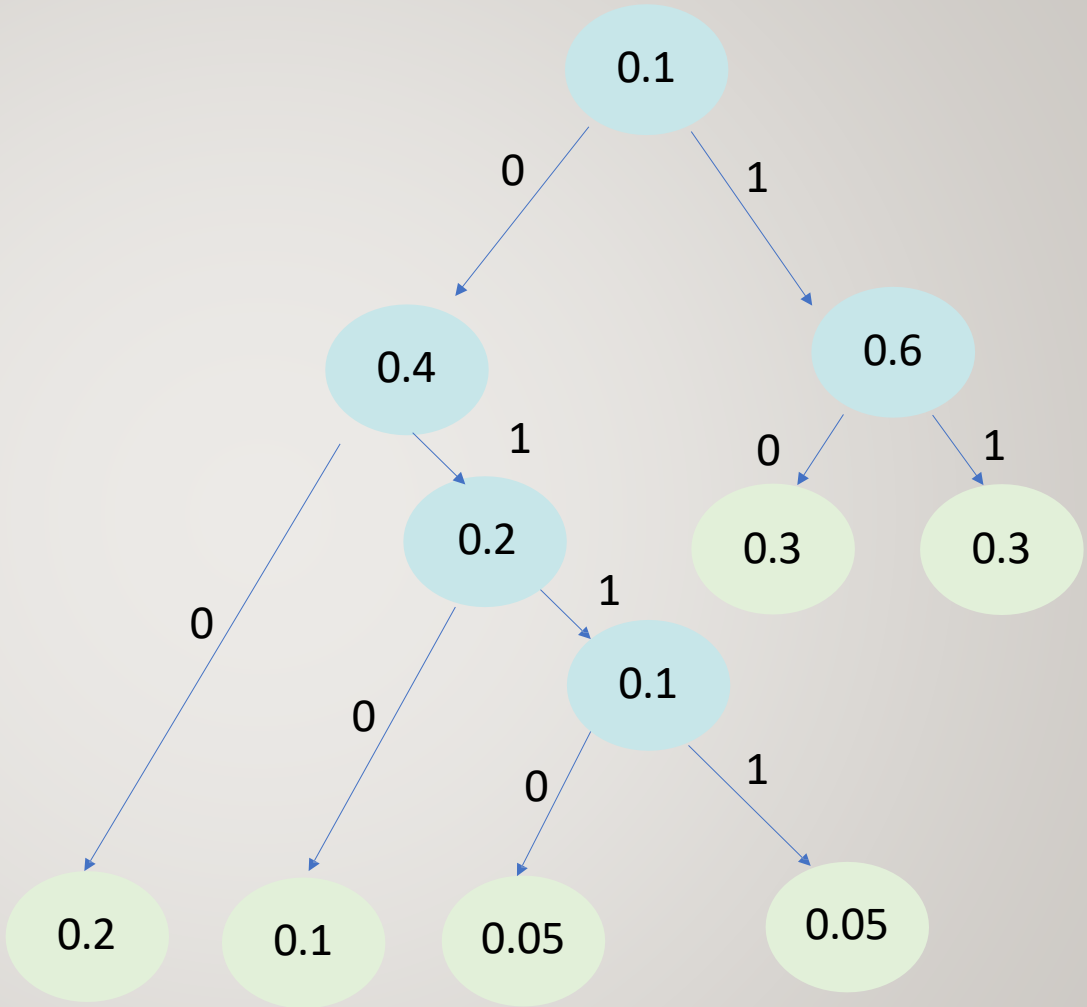
$$H = 0.4 \times 1 + 0.3 \times 2 + 0.1 \times 4 + 0.1 \times 4 + 0.06 \times 4 + 0.04 \times 4 = 2.2 \text{ bits/pixel}$$



ΚΩΔΙΚΟΠΟΙΗΣΗ HUFFMAN

Symbol	Probability	Code
f	0.3	11
r	0.2	00
d	0.05	0110
l	0.05	0111
a	0.1	010
g	0.3	10

For an input signal “frladg”
the code would be
11000111010011010= 17bits



ΑΡΙΘΜΗΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ

Η **Αριθμητική Κωδικοποίηση (Arithmetic Coding)** είναι μια τεχνική συμπίεσης δεδομένων χωρίς απώλειες που αποδίδει σε ολόκληρο το μήνυμα έναν ενιαίο αριθμό (πραγματικό αριθμό) μεταξύ 0 και 1, αντί να αντιστοιχίζει ξεχωριστούς κωδικούς σε κάθε σύμβολο, όπως γίνεται στην κωδικοποίηση Huffman.

Βασική Ιδέα:

- Αντί να κωδικοποιεί κάθε σύμβολο με έναν σταθερό κώδικα, η αριθμητική κωδικοποίηση "ζουμάρει" διαδοχικά σε ένα υποδιάστημα $[0,1)$ που αντιπροσωπεύει την πιθανότητα ολόκληρου του μηνύματος.
- Όσο περισσότερα δεδομένα κωδικοποιούνται, τόσο μικρότερο γίνεται το διάστημα.

ΑΡΙΘΜΗΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ

Βήματα Κωδικοποίησης

Ανάλυση Πιθανοτήτων:

Υπολογίζουμε τις πιθανότητες εμφάνισης κάθε συμβόλου στο μήνυμα.

Δημιουργούμε ένα εύρος $[0,1)$, όπου κάθε σύμβολο αντιστοιχεί σε ένα υποδιάστημα ανάλογο με την πιθανότητά του.

Διαδοχική Ενημέρωση Υποδιαστημάτων:

Ξεκινάμε με το διάστημα $[0,1)$.

Για κάθε σύμβολο, ενημερώνουμε το τρέχον διάστημα σύμφωνα με το υποδιάστημα που αντιστοιχεί στο σύμβολο.

Τελικό Κωδικοποιημένο Διάστημα:

Μετά την επεξεργασία όλων των συμβόλων, το τελικό διάστημα είναι πολύ μικρό.

Επιλέγουμε έναν αριθμό εντός αυτού του διαστήματος ως την τελική κωδικοποιημένη αναπαράσταση του μηνύματος.

Αποθήκευση:

Αποθηκεύουμε τον πραγματικό αριθμό (συνήθως σε δυαδική μορφή).

ΑΡΙΘΜΗΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ

Παράδειγμα

Μήνυμα: ABBA

Πιθανότητες: $A=0.6$, $B=0.4$

Αρχικό Διάστημα: $A: [0,0.6)$, $B: [0.6,1)$

Διαδοχική Κωδικοποίηση: 1ο Σύμβολο A :

Επιλέγουμε το διάστημα $[0,0.6)$.

2ο Σύμβολο B :

Ενημερώνουμε το διάστημα $[0.36,0.6)$, όπου το B είναι το 60%-100% του $[0,0.6)$.

3ο Σύμβολο B :

Ενημερώνουμε το διάστημα $[0.504,0.6)$, βασιζόμενο στο νέο διάστημα.

4ο Σύμβολο A : Τελικό διάστημα $[0.504,0.5424)$.

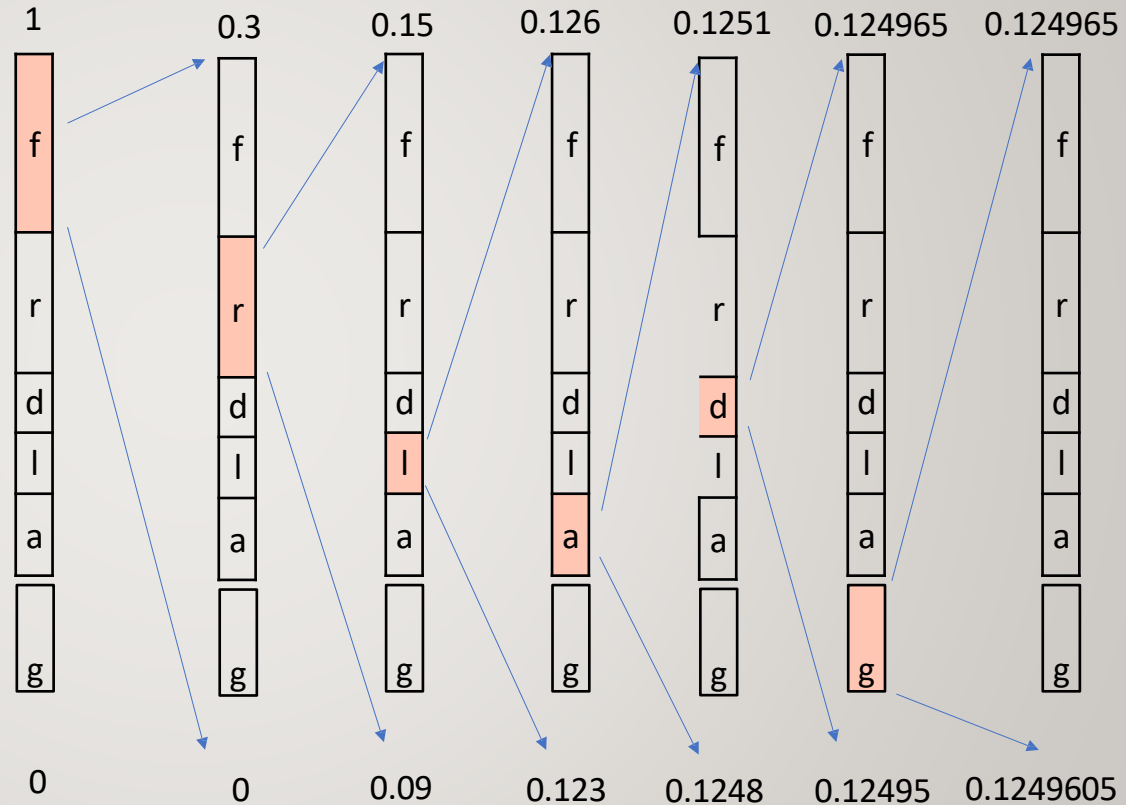
Κωδικοποιημένος Αριθμός:

Ο αριθμός 0.51 (ή οποιοσδήποτε μέσα στο διάστημα $[0.504,0.5424)$) αποθηκεύεται ως αναπαράσταση του μηνύματος.

ΑΡΙΘΜΗΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ

Symbol	Probability	subintervals
f	0.3	$[0,0.3)$
r	0.2	$[0.3,0.5)$
d	0.05	$[0.5,0.55)$
l	0.05	$[0.55,0.6)$
a	0.1	$[0.6,0.7)$
g	0.3	$[0.7,1)$

- For an input signal “frladg” the final range is 0.000045.
- To be able to encode this interval we need ≈ 16 bits.



Κωδικοποίηση Μήκους Διαδρομών

Έστω

x_1	x_2	...	x_M
-------	-------	-----	-------

 μία γραμμή εικόνας και x_i οι τιμές φωτεινότητας των pixels. Μπορεί να θεωρηθεί ότι η γραμμή αποτελείται από k τμήματα με μήκος l_i και φωτεινότητα g_i , $1 \leq i \leq k$, και να αναπαρασταθεί με ζεύγη (g_i, l_i) ως εξής:

$$(x_1, \dots, x_M) \rightarrow (g_1, l_1), (g_2, l_2), \dots, (g_k, l_k)$$

Για δυαδικές εικόνες, μόνο τα μήκη l_i είναι απαραίτητο να κωδικοποιηθούν (συνήθως με προθεματικό τρόπο)

Για επιπλέον συμπίεση, μπορεί να χρησιμοποιηθεί κωδικοποίηση Huffman για τα μήκη.

ΠΑΡΑΔΕΙΓΜΑ



Παράδειγμα γραμμής δυαδικής εικόνας και
αναπαράστασή της με ζεύγη (g_i, l_i)

$$\begin{array}{lll} k = 10 & l_1 = 3 & l_6 = 1 \\ & l_2 = 6 & l_7 = 1 \\ & l_3 = 1 & l_8 = 3 \\ & l_4 = 1 & l_9 = 4 \\ & l_5 = 1 & l_{10} = 8 \end{array}$$

Πιθανότητα εμφάνισης του 0 : p
Μέγιστο μήκος των 0 : $M=2^m$
Επιτυγχάνεται Συμπίεση C :

$$C = \frac{1 - p^M}{m(1 - p)}$$

$$(1,3),(0,6),(1,1),(0,1),(1,1),(0,1),(1,1),(0,3),(1,4),(0,8) \\ \rightarrow (3,6,1,1,1,1,1,3,4,8)$$

ΣΥΜΠΙΕΣΗ LZW (LEMPERL-ZIV-WELCH)

- Ο δημοφιλέστερος μη-απωλεστικός αλγόριθμος συμπίεσης ψηφιακής εικόνας (TIFF, GIF) που βασίζεται στη δυναμική κωδικοποίηση ακολουθιών.
- Η εικόνα αντιμετωπίζεται ως μία μονοδιάστατη ακολουθία bits.
- Δημιουργείται ένας πίνακας κωδικών λέξεων που αντιστοιχίζονται στις διάφορες ακολουθίες bits.
- Αναζητά επαναλαμβανόμενες ακολουθίες χαρακτήρων ή bytes στα δεδομένα και τις αντικαθιστά με σύντομες αναπαραστάσεις (δείκτες).
- Αντί να αποθηκεύει κάθε χαρακτήρα, ο αλγόριθμος χτίζει δυναμικά ένα λεξικό από τις εμφανιζόμενες ακολουθίες δεδομένων.
- Μειώνει το μέγεθος του αρχείου, καθώς οι ακολουθίες που εμφανίζονται συχνά κωδικοποιούνται με μικρότερα σύμβολα

ΣΥΜΠΙΕΣΗ LZW

- Οι πρώτες 256 κωδικές λέξεις αντιστοιχίζονται στις τιμές 0...255.
- Η λέξη #257 στο Clear code και η #258 στο EOI (End of Information).
- Οι υπόλοιπες λέξεις αντιστοιχίζονται σε ακολουθίες bits που εμφανίζονται στην εικόνα.
- Τα bits αναπαράστασης των κωδικών λέξεων αυξάνονται σταδιακά.

0
□
255
Clear code
EOI
string
□
string

ΑΛΛΕΣ ΜΗ-ΑΠΩΛΕΣΤΙΚΕΣ ΤΕΧΝΙΚΕΣ

Bit plane coding

Αποσύνθεση σε L δυαδικές εικόνες και διαφορετική κωδικοποίηση ανάλογα με τη «βαρύτητά» τους

Constant Area Coding

Υποδιαίρεση σε υποεικόνες και προθεματική κωδικοποίηση των υποεικόνων ανάλογα με το αν είναι Λευκές, Μαύρες ή Μικτές (0, 10, 11). Στην τελευταία περίπτωση αποστέλλεται όλη η υποεικόνα

Contour Tracing

Βρίσκεται το σύνολο των «ισοϋψών» καμπυλών της εικόνας και κωδικοποιούνται με αποδοτικό τρόπο

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

- Συνήθως τα δεδομένα της εικόνας σε μία γειτονιά είναι συσχετισμένα μεταξύ τους, οπότε είναι δυνατόν να «προβλέψουμε» την τιμή ενός pixel με βάση αυτά που βρίσκονται στην γειτονιά του.
- Για κάθε pixel (x,y) αποθηκεύουμε τη «νέα» πληροφορία που περιέχει.
- «Νέα» πληροφορία είναι η διαφορά της προβλεφθείσας από την πραγματική τιμή του pixel.

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Για κάθε pixel f_n της εικόνας, η διάταξη πρόβλεψης παράγει την αναμενόμενη τιμή βασισμένη στις προηγούμενες εισόδους (pixels).

Η τιμή αυτή στρογγυλοποιείται στον κοντινότερο ακέραιο.

Το λάθος πρόβλεψης είναι $e_n = f_n - \hat{f}_n$

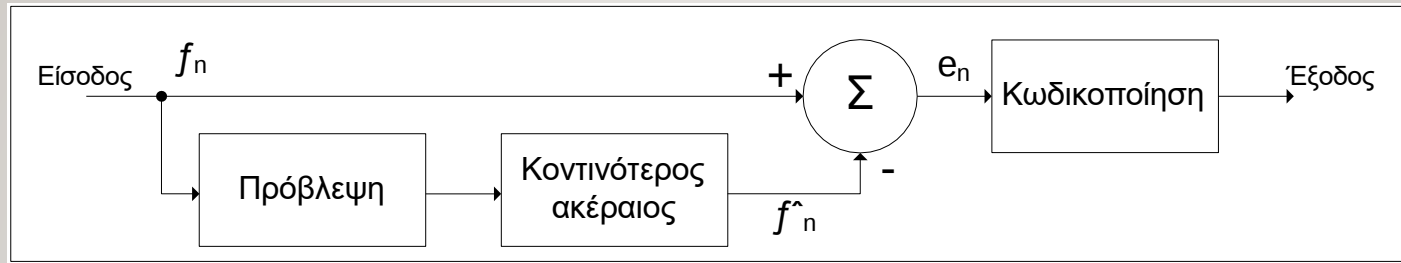
όπου $\hat{f}_n$ είναι μία συνάρτηση πρόβλεψης: $f_n = e_n + \hat{f}_n$

Η αρχική τιμή μπορεί να ανακατασκευαστεί χωρίς απώλειες ως εξής

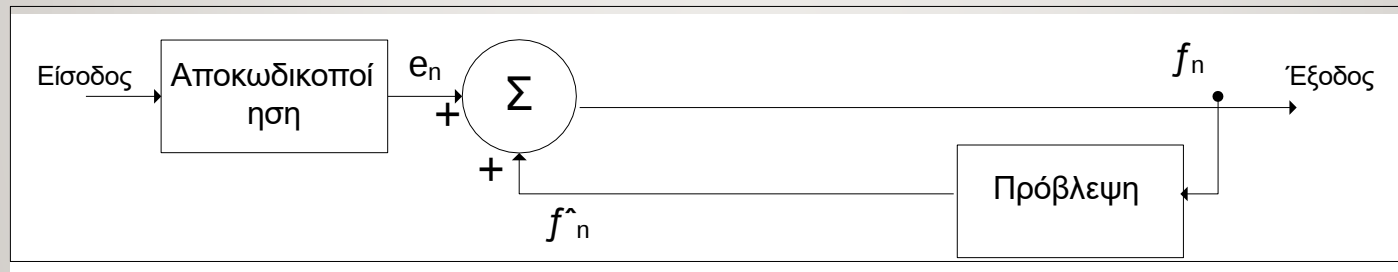
$$\hat{f}_n = \text{round} \left[\sum_{i=1}^m a_i f_{n-i} \right]$$

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Διάταξη συμπίεσης



Διάταξη ανακατασκευής



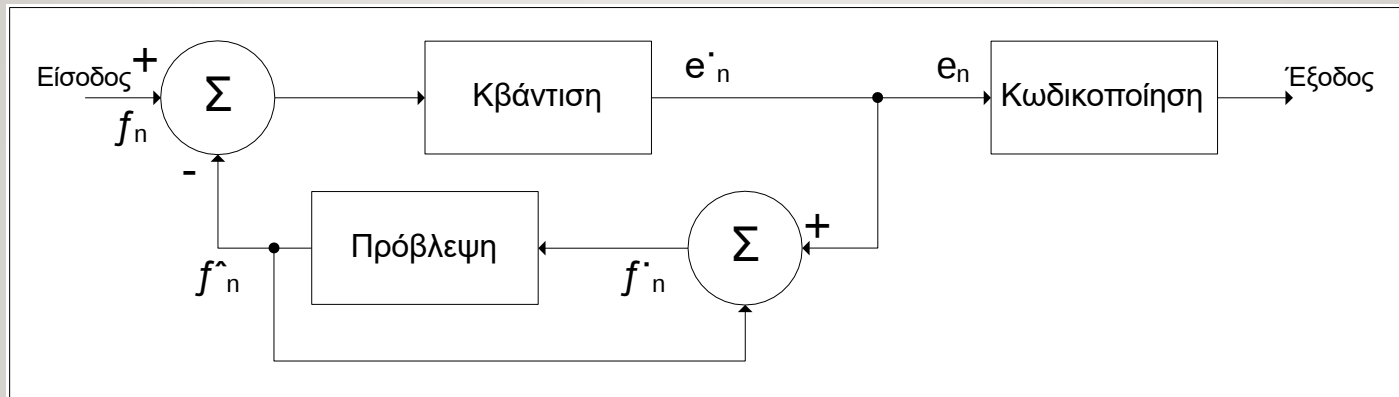
Για την ανακατασκευή απαιτείται η αποστολή των συντελεστών πρόβλεψης και του κωδικοποιημένου σφάλματος.

ΑΠΩΛΕΣΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

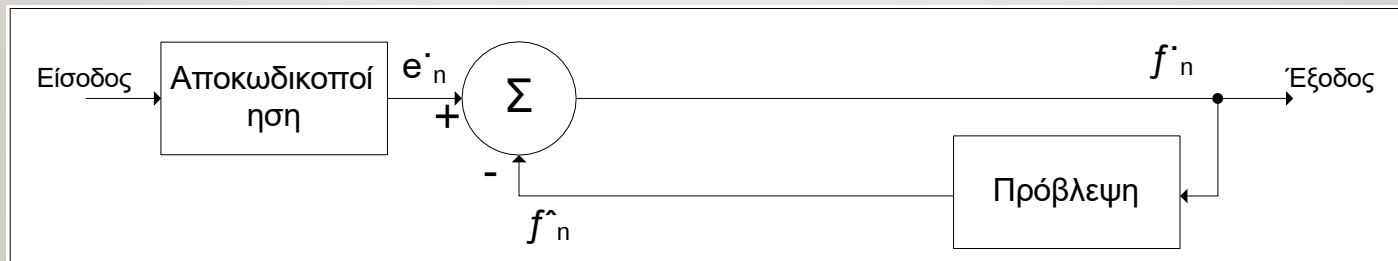
- Για καλύτερη συμπίεση εφαρμόζουμε κβαντισμό στο e_n . Σε αυτή την περίπτωση το λάθος πρόβλεψης κβαντίζεται σε ένα περιορισμένο εύρος τιμών.
- Η διαδικασία της κβάντισης είναι μη αντιστρέψιμη και απωλεστική (χάνεται πληροφορία).

ΑΠΩΛΕΣΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Διάταξη συμπίεσης



Διάταξη ανακατασκευής



ΑΠΩΛΕΣΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Ο σχεδιασμός του προβλεπτή* :

- Έστω ότι επιλέγεται μονοδιάστατη κωδικοποίηση (τύπου DPCM) σε κάθε γραμμή της εικόνας.
- Κάθε γραμμή $f(m)$ μοντελοποιείται σαν στάσιμη AR διαδικασία:

$$f(m) = \sum_{k=1}^p a(k)f(m-k) + \varepsilon(m), \quad E[\varepsilon^2(m)] = \sigma^2$$

όπου το $\varepsilon(m)$ είναι όρος λευκού Γκαουσιανού προσθετικού θορύβου, ασυσχέτιστος με το $f(m)$.

Οι παράμετροι του μοντέλου AR, μέσω του οποίου αναπαράγεται η $f(m)$, είναι ίσες με τους συντελεστές του προβλεπτή της $f(m)$.

* Παρόμοιος είναι και ο σχεδιασμός του προβλεπτή για τη μη απωλεστική διαδικασία

ΑΠΩΛΕΣΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Το κβαντισμένο σφάλμα $e_q(m)$ εκπέμπεται στον δέκτη:

$$e_q(m) = Q[e(m)] = Q[f(m) - \hat{f}(m)]$$

Η γραμμή $f(m)$ ανακατασκευάζεται ως εξής:

$$f_r(m) = \sum_{k=1}^p a(k) f_r(m-k) + e_q(m)$$

Οι συντελεστές πρόβλεψης βρίσκονται λύνοντας το σύστημα (όπου $R(k)$ η συνάρτηση αυτοσυσχέτισης):

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p-1) \\ R(1) & R(0) & \cdots & R(p-2) \\ \vdots & \vdots & \vdots & \vdots \\ R(p-1) & R(p-2) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{bmatrix} = \begin{bmatrix} R(1) \\ R(2) \\ \vdots \\ R(p) \end{bmatrix}$$

ΑΠΩΛΕΣΤΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΠΡΟΒΛΕΨΗ

Τα μονοδιάστατα μοντέλα πρόβλεψης επεκτείνονται σε δισδιάστατα της μορφής:

$$\hat{f}(n,m) = \sum_{(i,j) \in A} a(i,j) f(n-i, m-j)$$

Η εύρεση των παραμέτρων του 2D-AR μοντέλου γίνεται επεκτείνοντας στις 2 διαστάσεις την προηγούμενη διαδικασία

Γενικά:

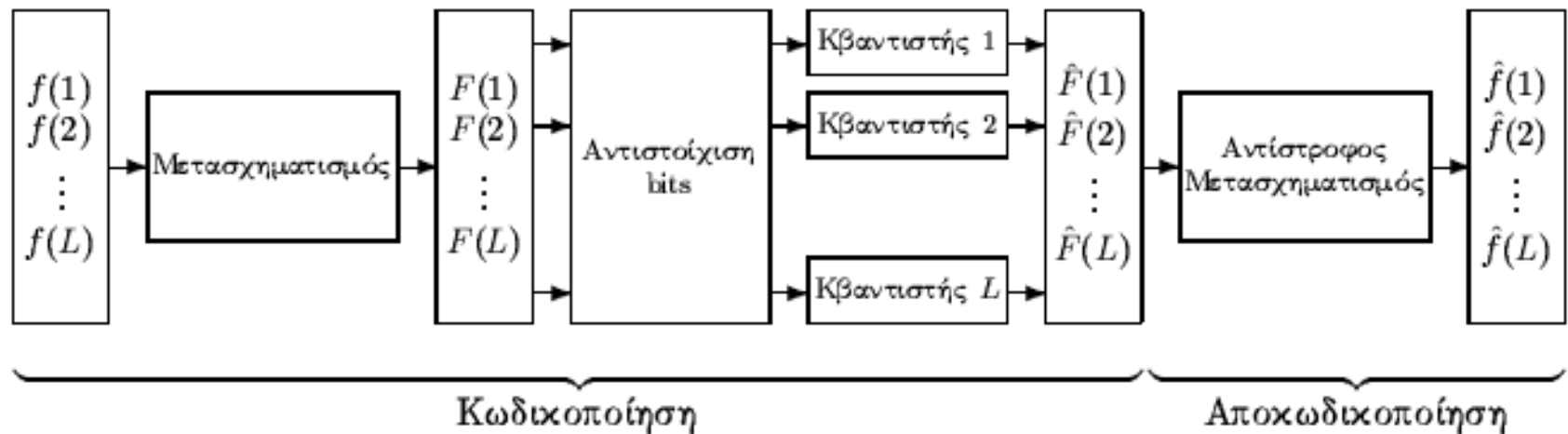
- Η τεχνική κωδικοποίησης DPCM με πρόβλεψη είναι σχετικά απλή και μπορεί εύκολα να υλοποιηθεί σε λογισμικό ή υλικό.
- Ο λόγος συμπίεσης που επιτυγχάνεται είναι μέτριος.
- Είναι ευαίσθητη στο θόρυβο καναλιού.
- Ο κρουστικός θόρυβος μπορεί να διαδοθεί σε όλο το μήκος μιας γραμμής ή και σε ολόκληρη την εικόνα.

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

- Πρόκειται για απωλεστική μέθοδο συμπίεσης με καλό λόγο συμπίεσης
- Επιλέγουμε και κωδικοποιούμε τους συντελεστές του μετασχηματισμού οι οποίοι περιέχουν την μεγάλη ενέργεια, και απορρίπτουμε τους υπόλοιπους.
- Η εικόνα ανακτάται (με απώλειες) μέσω του αντιστρόφου μετασχηματισμού.

$$F = T[f] \rightarrow \hat{f} = T^{-1}[\hat{F}]$$

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ



Διάγραμμα Βαθμίδων της Κωδικοποίησης / Αποκωδικοποίησης
μέσω Μετασχηματισμών

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

Γραμμικός Μετασχηματισμός:

Έστω $\mathbf{f}$ το διάνυσμα που αναπαριστά μία εικόνα μεγέθους $L = N \times M$, τότε το μετασχηματισμένο διάνυσμα $\mathbf{F}$ δίνεται από τη σχέση:

$$\mathbf{F} = \mathbf{A}\mathbf{f}$$

όπου $\mathbf{A}$ είναι ο πίνακας μετασχηματισμού. Ο αντίστροφος μετασχηματισμός ορίζεται ως εξής:

$$\mathbf{f} = \mathbf{A}^{-1}\mathbf{F}$$

Ορθομοναδιαίος μετασχηματισμός:

$$\mathbf{A}\mathbf{A}^{*T} = \mathbf{A}^T\mathbf{A}^* = \mathbf{I}$$

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

Ένας ορθομοναδιαίος μετασχηματισμός ικανοποιεί τη συνθήκη διατήρησης της ενέργειας:

$$\|\mathbf{f}\|^2 = \sum_{k=1}^L |f(k)|^2 = \sum_{k=1}^L |F(k)|^2 = \|\mathbf{F}\|^2$$

- Η περισσότερη ενέργεια συγκεντρώνεται στον DC όρο και σε μερικούς χαμηλής συχνότητας.
- Μεταβλητός αριθμός bits μπορεί να αντιστοιχισθεί στους άλλους συντελεστές.
- Οι συντελεστές $F(k)$, $1 \leq k \leq L$ κβαντίζονται χρησιμοποιώντας L κβαντιστές.
- Η αποκωδικοποίηση γίνεται εφαρμόζοντας τον αντίστροφο μετασχηματισμό στο διάνυσμα συντελεστών.

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

Σχεδιαστικές αποφάσεις :

- Επιλογή του συγκεκριμένου μετασχηματισμού που θα χρησιμοποιηθεί (DFT, DCT, DST κτλ.).
- Επιλογή των διαστάσεων του μπλοκ της εικόνας στο οποίο θα εφαρμοστεί ο μετασχηματισμός (8x8, 16x16)
- Καθορισμός του αριθμού των bits n_k που αντιστοιχίζονται σε κάθε συντελεστή $F(k)$, $k=1, \dots, L$. Αν ο μέσος αριθμός bits/στίγμα είναι B τότε ισχύει:

$$\frac{1}{L} \sum_{k=1}^L n_k = B$$

Οι κβαντισμένοι συντελεστές μπορούν να κωδικοποιηθούν με Huffman για επιπλέον συμπίεση

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

Η μέση παραμόρφωση εξαιτίας του κβαντισμού των συντελεστών δίνεται από τη σχέση:

$$E = \frac{1}{L} \sum_{k=1}^L E \left[|F(k) - Q[F(k)]|^2 \right] = \frac{1}{L} \sum_{k=1}^L \sigma_k^2 q(n_k)$$

όπου $Q[.]$ συμβολίζει τον κβαντισμό, σ_k^2 τη διασπορά του συντελεστή $F(k)$ και $q(n_k)$ τη συνάρτηση παραμόρφωσης του κβαντιστή (μονότονη και φθίνουσα με $q(0)=1$ και $q(\infty)=0$).

Μια δυνατή εκλογή του n_k δίνεται από τη σχέση:

$$\text{round}(n_k), \quad n_k = B + \frac{1}{2} \left\lceil \log_2 \sigma_k^2 - \frac{1}{L} \log_2 \left(\prod_{k=1}^L \sigma_k^2 \right) \right\rceil$$

ΚΩΔΙΚΟΠΟΙΗΣΗ ΜΕ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΥΣ

- Δεδομένου του n_k , μπορεί στη συνέχεια να γίνει σχεδιασμός του βέλτιστου κβαντιστή για τον κάθε συντελεστή του μετασχηματισμού βάσει της πυκνότητας πιθανότητας του κάθε συντελεστή.
- Οι κβαντισμένοι συντελεστές κωδικοποιούνται με Huffman για επιπλέον συμπίεση.
- Για την αποκωδικοποίηση γίνεται πρώτα αποκωδικοποίηση Huffman. Οι μετασχηματισμένοι συντελεστές είναι:

$$\hat{F}(k) = \begin{cases} F(k) \cdot \sigma_k^2 & 1 \leq k \leq K \\ 0 & K < k \leq L \end{cases}$$

Για να αναιρέσουμε την κανονικοποίηση που είχε γίνει κατά τη συμπίεση ώστε να έχουν διασπορά 1

Το μπλοκ της εικόνας ανακτάται εφαρμόζοντας τον αντίστροφο μετασχηματισμό:

$$\hat{\mathbf{f}} = \mathbf{A}^{-1}\hat{\mathbf{F}}$$

Ο ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ DCT

1-D DCT & IDCT

$$C(k) = \begin{cases} 2 \sum_{n=0}^{N-1} x(n) \cos \frac{k\pi}{2N} (2n+1), & 0 \leq k \leq N-1 \\ 0, & \text{αλλού} \end{cases}$$

$$x(n) = \begin{cases} \frac{1}{N} \sum_{k=0}^{N-1} w(k) C(k) \cos \frac{k\pi}{2N} (2n+1), & 0 \leq n \leq N-1 \\ 0, & \text{αλλού} \end{cases}$$

$$\text{όπου } w(k) = \begin{cases} 1/2, & k = 0 \\ 1, & 1 \leq k \leq N-1 \end{cases}$$

ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ DCT ΜΕΣΩ DFT

DCT

$$1. \quad y(n) = x(n) + x(2N - 1 - n), \quad 0 \leq n \leq 2N - 1$$

$$2. \quad Y(k) = DFT\{y(n)\}, \quad 0 \leq k \leq 2N - 1$$

$$3. \quad C(k) = \begin{cases} e^{-j \frac{k\pi}{2N}} Y(k), & 0 \leq k \leq N - 1 \\ 0, & \text{αλλού} \end{cases}$$

IDCT

$$1. \quad Y(k) = \begin{cases} e^{j \frac{k\pi}{2N}} C(k), & 0 \leq k \leq N - 1 \\ 0, & k = N \\ -e^{j \frac{k\pi}{2N}} C(2N - k), & N + 1 \leq k \leq 2N - 1 \end{cases}$$

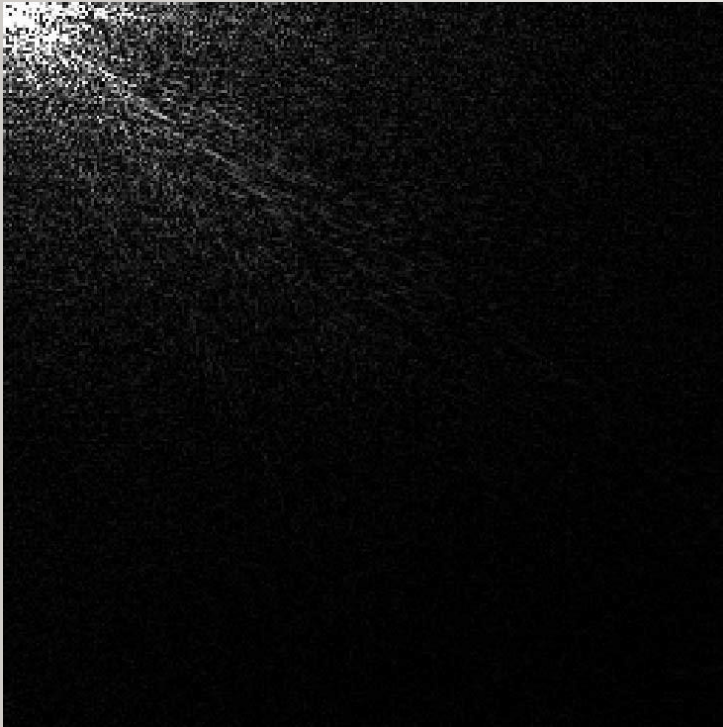
$$2. \quad y(n) = IDFT\{Y(k)\}, \quad 0 \leq n \leq 2N - 1$$

$$3. \quad x(n) = \begin{cases} y(n), & 0 \leq n \leq N - 1 \\ 0, & \text{αλλού} \end{cases}$$

■ Ο μετασχηματισμός 2-D DCT μπορεί να υπολογιστεί μέσω του 1-D DCT με τη μέθοδο γραμμών-στηλών

■ Ερμηνεία της μεγάλης ενεργειακής συγκέντρωσης

ΚΩΔΙΚΟΠΟΙΗΣΗ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ ΤΟΥ DCT



Μετασχηματισμός DCT

8	7	6	4	3	2	1	0
7	6	5	4	3	2	1	0
6	5	4	3	3	1	1	0
4	4	3	3	2	1	0	0
3	3	3	2	1	1	0	0
2	2	1	1	1	0	0	0
1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0

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

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

- Αντιστοίχιση bits κβαντισμού σε block 8x8
- ο αριθμός bits ανά συντελεστή είναι ανάλογος της μεταβλητότητάς του

ΔΙΑΔΙΚΑΣΙΑ ΣΥΜΠΙΕΣΗΣ ΜΕ DCT

Συμπίεση:

1. Διαίρεση σε block
2. Μετασχηματισμός DCT
3. Επιλογή συντελεστών (μέθοδοι: ζώνης, κατωφλίου)
4. Κβάντιση και αποθήκευση συντελεστών

Αποσυμπίεση:

1. Αντιστροφή κβάντισης με το βήμα κβαντισμού
2. Ανακατασκευή block
3. Μετασχηματισμός IDCT
4. Επανένωση των block

ΔΙΑΪΡΕΣΗ ΣΕ ΜΠΛΟΚ & ΕΦΑΡΜΟΓΗ DCT

Διαίρεση σε Μπλοκ: Η εικόνα χωρίζεται σε μικρότερα μη επικαλυπτόμενα μπλοκ, συνήθως 8x8, 16x16

1. Ο DCT εφαρμόζεται σε κάθε μπλοκ, μετατρέποντας τα δεδομένα από τον χώρο των pixel στον χώρο των συχνοτήτων.
2. Το αποτέλεσμα είναι ένας πίνακας συντελεστών DCT, όπου:
 1. Ο **DC** συντελεστής αντιπροσωπεύει τη μέση φωτεινότητα του μπλοκ.
 2. Οι **AC** συντελεστές αντιπροσωπεύουν τις μεταβολές και τις λεπτομέρειες.

ΕΠΙΛΟΓΗ ΣΥΝΤΕΛΕΣΤΩΝ DCT: ΜΕΘΟΔΟΣ ΖΩΝΗΣ

Η **σάρωση zigzag** είναι μια τεχνική που χρησιμοποιείται στη συμπίεση εικόνας, ιδιαίτερα στο πρότυπο JPEG, για τη γραμμικοποίηση των συντελεστών του **DCT** από έναν δισδιάστατο πίνακα σε ένα μονοδιάστατο διάνυσμα.

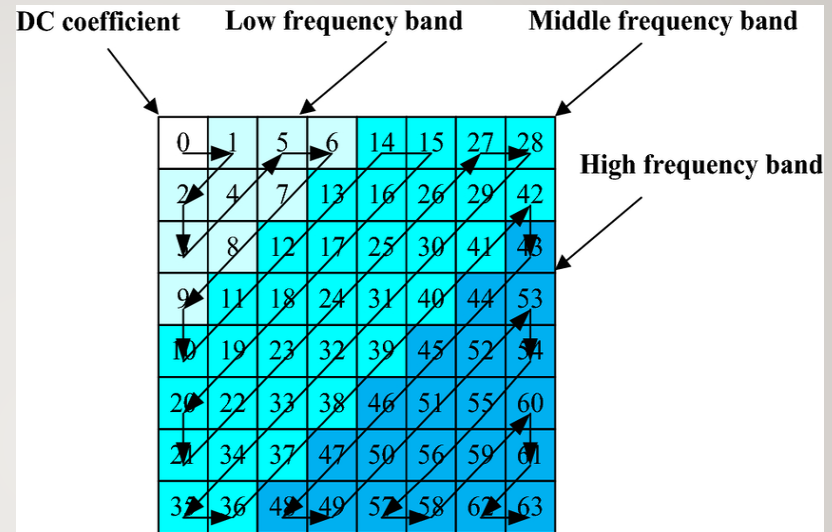
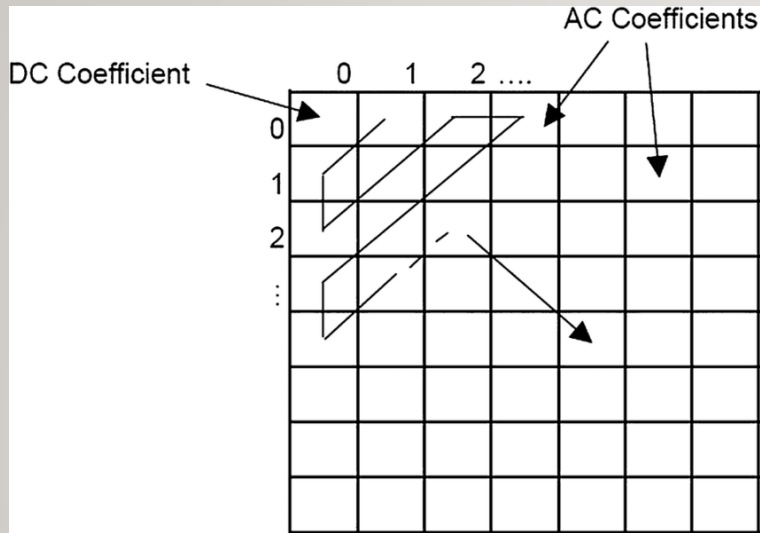
Αυτή η διαδικασία διευκολύνει την αποδοτική κωδικοποίηση των συντελεστών, εκμεταλλευόμενη την τάση των υψηλότερων συχνοτήτων να έχουν μηδενικές ή μικρές τιμές μετά την κβάντιση.

Διαδικασία Σάρωσης zigzag

Η σάρωση **zigzag** ακολουθεί μια προκαθορισμένη διαδρομή μέσω του πίνακα συντελεστών DCT, ξεκινώντας από την επάνω αριστερή γωνία (DC συντελεστής) και προχωρώντας διαγώνια μέσα από τον πίνακα.

Αυτή η διαδρομή οργανώνει τους συντελεστές κατά αύξουσα συχνότητα, τοποθετώντας πρώτα τους συντελεστές χαμηλών συχνοτήτων και ακολουθώντας με τους συντελεστές υψηλότερων συχνοτήτων.

ΕΠΙΛΟΓΗ ΣΥΝΤΕΛΕΣΤΩΝ DCT: ΜΕΘΟΔΟΣ ΖΩΝΗΣ



Αυτή η σειρά διασφαλίζει ότι οι συντελεστές χαμηλών συχνοτήτων, που περιέχουν το μεγαλύτερο μέρος της ενέργειας της εικόνας, τοποθετούνται στην αρχή του διανύσματος, ενώ ακολουθούν οι συντελεστές υψηλών συχνοτήτων, που συχνά είναι μηδενικοί μετά την κβάντιση.

ΕΠΙΛΟΓΗ ΣΥΝΤΕΛΕΣΤΩΝ DCT: ΜΕΘΟΔΟΣ ΖΩΝΗΣ

Πλεονεκτήματα της Σάρωσης zigzag

- **Αποδοτική Κωδικοποίηση:** Η οργάνωση των συντελεστών με αυτή τη σειρά επιτρέπει την εφαρμογή τεχνικών όπως η **κωδικοποίηση μήκους διαδρομής (Run-Length Encoding - RLE)**, η οποία εκμεταλλεύεται τις συνεχόμενες ακολουθίες μηδενικών για συμπίεση.

- **Βελτιωμένη Συμπίεση:** Δεδομένου ότι οι συντελεστές υψηλών συχνοτήτων τείνουν να είναι μηδενικοί μετά την κβάντιση, η σάρωση zigzag συγκεντρώνει αυτούς τους μηδενικούς συντελεστές, επιτρέποντας πιο αποδοτική συμπίεση.

ΕΠΙΛΟΓΗ ΣΥΝΤΕΛΕΣΤΩΝ DCT: ΜΕΘΟΔΟΣ ΚΑΤΩΦΛΙΟΥ

Κατωφλίωση:

1. Ορίζεται ένα κατώφλι (T), και οι συντελεστές του DCT που έχουν απόλυτη τιμή μικρότερη από αυτό το κατώφλι θεωρούνται μη σημαντικοί και μηδενίζονται.
2. Η λογική πίσω από αυτό είναι ότι οι υψηλές συχνότητες συχνά δεν είναι αντιληπτές από το ανθρώπινο μάτι ή περιέχουν θόρυβο.

Κβάντιση: Μετά την κατωφλίωση, εφαρμόζεται κβάντιση για αύξηση της συμπίεσης η οποία μειώνει περαιτέρω την ακρίβεια των συντελεστών.

Κωδικοποίηση: Οι μηδενισμένοι συντελεστές (λόγω κατωφλίωσης) και οι εναπομείναντες σημαντικοί συντελεστές κωδικοποιούνται με τεχνικές όπως Run-Length Encoding (RLE) ή Entropy Encoding (π.χ., Huffman).

ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΜΠΙΕΣΗΣ



Αρχική



126:1

15:1



51:1

8:1



28:1

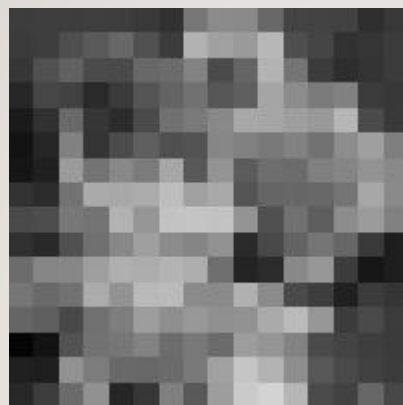
4:1



ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΜΠΙΕΣΗΣ



Αρχική



126:1

15:1



51:1

8:1

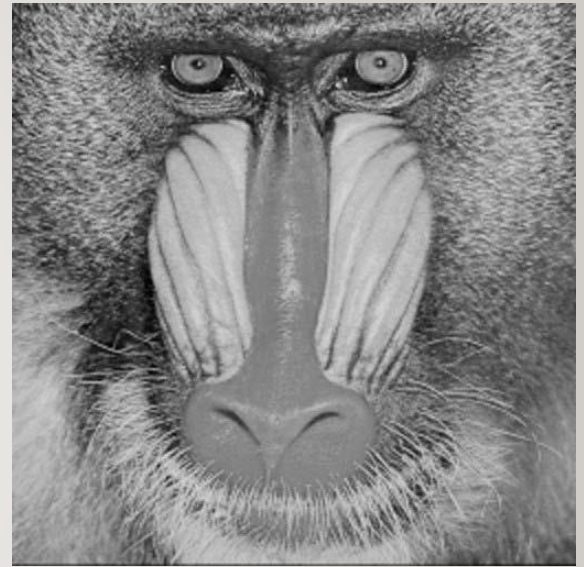
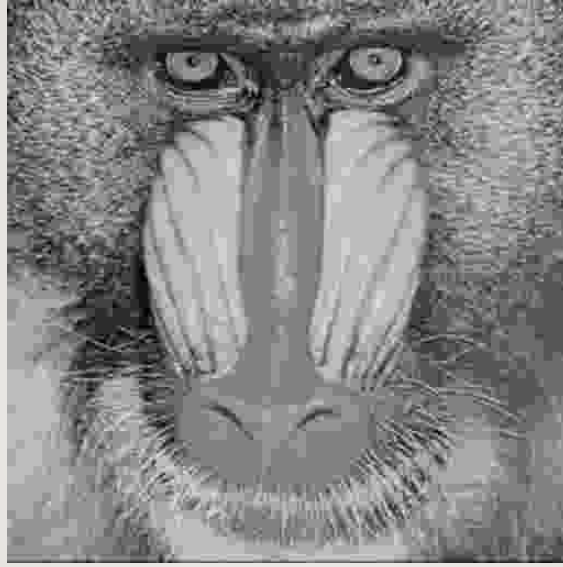
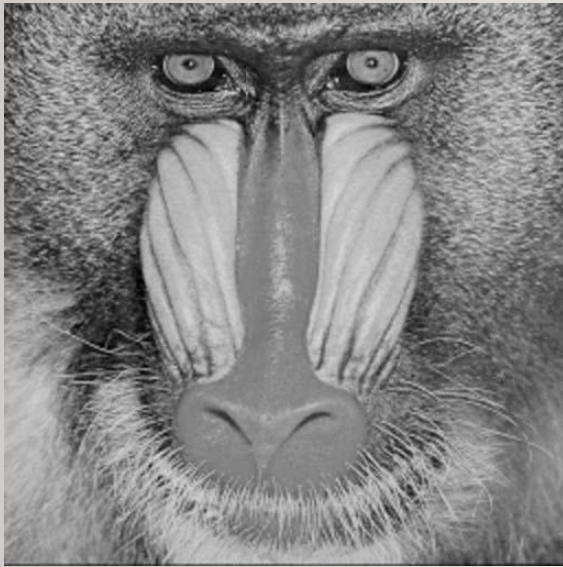


28:1

4:1



ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΥΜΠΙΕΣΗΣ



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

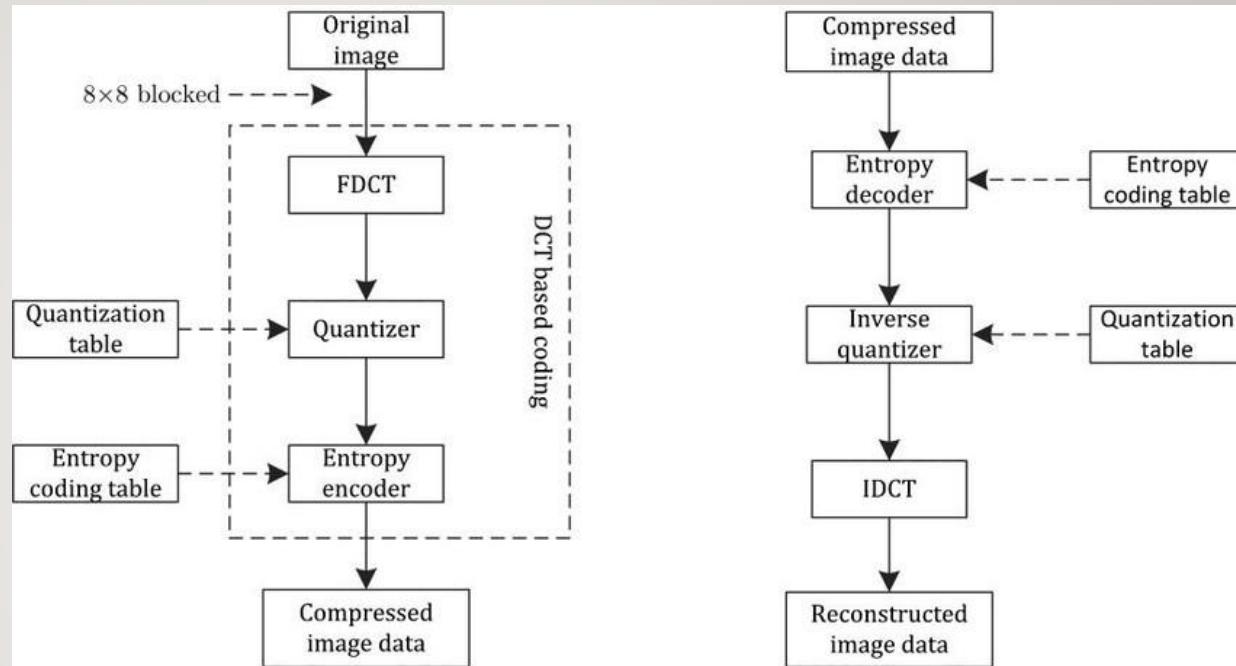
ΤΟ ΠΡΟΤΥΠΟ JPEG

- Σχεδιάστηκε από την ομάδα Joint Photographic Experts Group σε συνεργασία με την Διεθνή Ένωση Τηλεπικοινωνιών (ITU-TS)
- Μπορεί να δώσει διαφορετικό αποτέλεσμα ανάλογα με τις απαιτήσεις που έχουμε για την ποιότητα της εικόνας και το λόγο συμπίεσης:
 - 10:1 έως 20:1 – υψηλή ποιότητα
 - 30:1 έως 50:1 – μέτρια ποιότητα
 - 60:1 έως 100:1 – κακή ποιότητα
- Είναι μορφή κωδικοποίησης με βάση την αντίληψη και βασίζεται στο γεγονός ότι:
 - Το μάτι είναι λιγότερο ευαίσθητο στις χρωματικές συνιστώσες (μικρότερη διακριτική ικανότητα) μιας εικόνα από ότι στη φωτεινότητα
 - Έχει πεπερασμένη διακριτική ικανότητα σε διαδοχικές αλλαγές της φωτεινότητας
- Αποτελεί συνδυασμό διαφόρων τεχνικών συμπίεσης:
 - Μετασχηματισμού (DCT)
 - Εντροπίας (Huffman)
 - Μήκους διαδρομής (RLE)

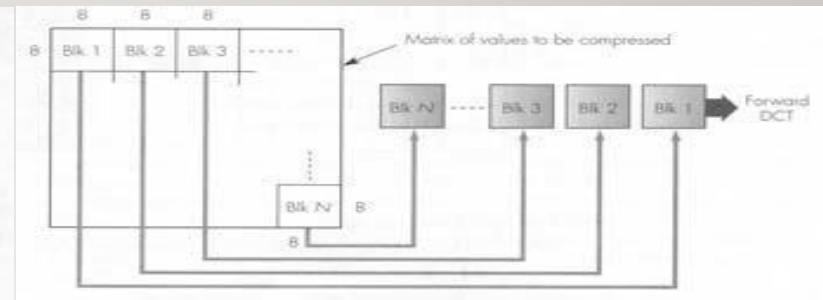
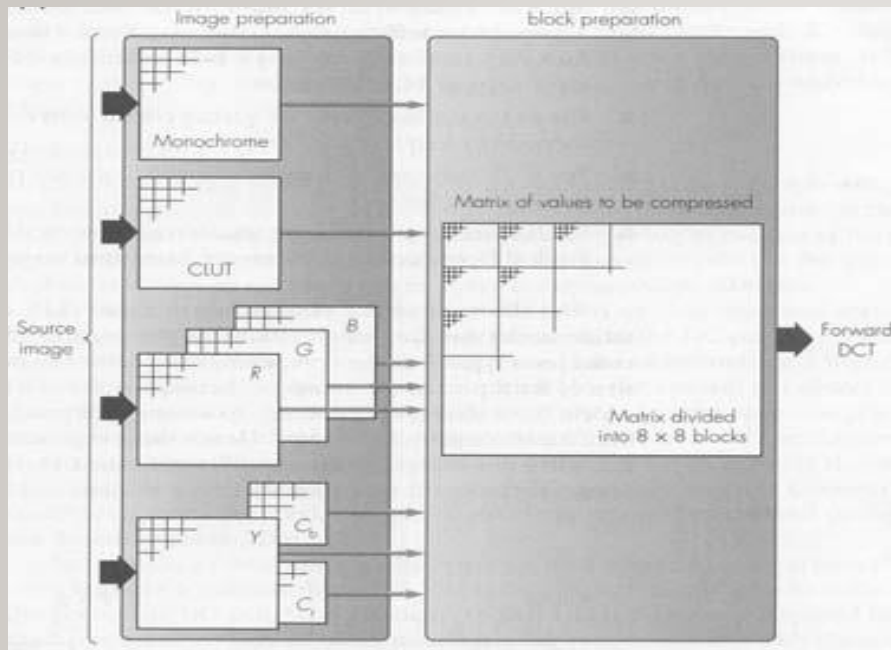
ΤΟ ΠΡΟΤΥΠΟ JPEG

- Πέντε βασικά στάδια:

- Προετοιμασία εικόνας/μπλοκ
- Ευθύς μετασχηματισμός DCT
- Κβαντισμός
- Κωδικοποίηση εντροπίας
- Δημιουργία Frame (πλαισίου)



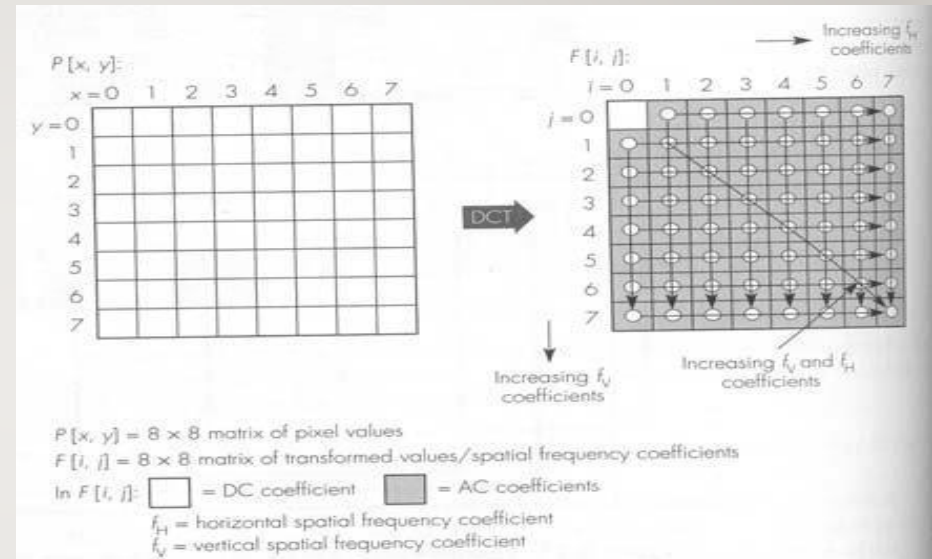
ΠΡΟΕΤΟΙΜΑΣΙΑ ΕΙΚΟΝΑΣ/ΜΠΛΟΚ



- Προετοιμασία εικόνας:
 - Τα διάφορα χρωματικά κανάλια (π.χ. R,G,B ή Y, Cr, Cb) κωδικοποιούνται ως διαφορετικές εικόνες
- Προετοιμασία μπλοκ
 - Η εικόνα χωρίζεται σε μπλοκ 8x8 pixels
 - Τα μπλοκ μεταδίδονται με πρώτο το πάνω αριστερά και τελευταίο το κάτω δεξιά

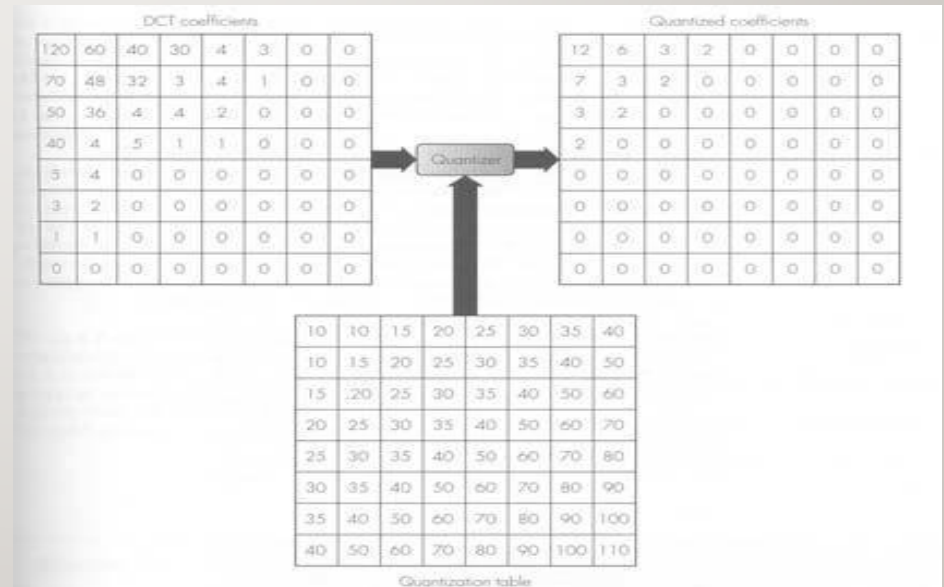
ΕΥΘΥΣ ΜΕΤΑΣΧΗΜΑΤΙΣΜΌΣ DCT

- Σε κάθε μπλοκ εφαρμόζεται μετασχηματισμός DCT:
 - Πριν την εφαρμογή του DCT από κάθε pixel του μπλοκ αφαιρείται η τιμή 128 (οπότε τελικά οι τιμές του μπλόκ θα λαμβάνουν τιμές στο διάστημα $[-128 \ 127]$ αντί στο διάστημα $[0 \ 255]$)
 - Ο συντελεστής $F(0,0)$ ονομάζεται DC.
 - Οι υπόλοιποι συντελεστές ονομάζονται AC.



ΚΒΑΝΤΙΣΜΌΣ

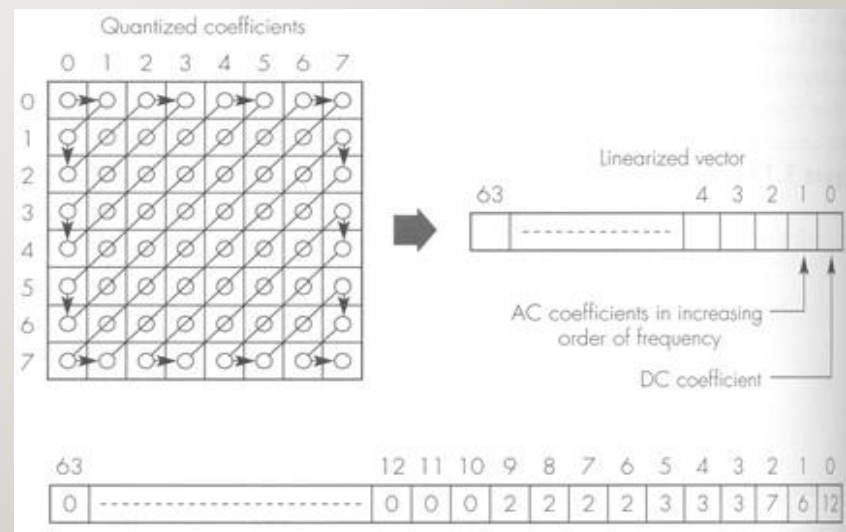
- Το μάτι έχει διαφορετική ευαισθησία και διακριτική ικανότητα όσον αφορά τις χωρικές μεταβολές της φωτεινότητας (χωρικές συχνότητες):
 - Οι πίνακες κβαντισμού αντιπροσωπεύουν την ευαισθησία αυτή.
 - Όσο μικρότερος είναι ο αντίστοιχος συντελεστής στον πίνακα κβαντισμού τόσο μεγαλύτερη θεωρείται η διακριτική ικανότητα του ματιού στη συγκεκριμένη συχνότητα.



ΚΩΔΙΚΟΠΟΙΗΣΗ

περιλαμβάνει τα εξής στάδια:

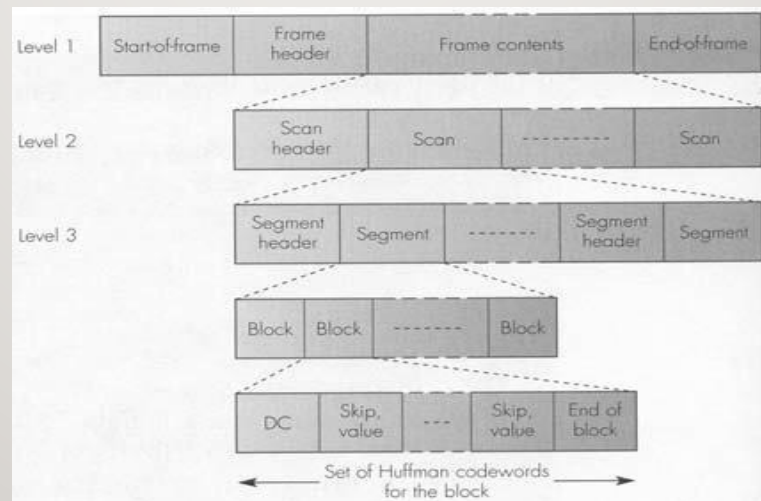
- Διανυσματοποίηση του μπλοκ από τους κβαντισμένους συντελεστές με zig-zag scanning
- Διαφορική κωδικοποίηση των συντελεστών DC γειτονικών μπλοκ – Κωδικοποίηση μήκους διαδρομής για τους AC συντελεστές κάθε μπλοκ
- Huffman coding επί των συμβόλων που προέκυψαν από τη διαφορική κωδικοποίηση και την κωδικοποίηση μήκους διαδρομής



ΔΗΜΙΟΥΡΓΙΑ JPEG FRAME

Δημιουργία frame:

- Προσθήκη ένδειξης αρχής και τέλους frame
- Επικεφαλίδες (χρωματικά κανάλια, διαστάσεις εικόνας, κλπ)
- Περιεχόμενα εικόνας (πολυπλεξία μπλοκ από διαφορετικά χρωματικά κανάλια ώστε να επιτυγχάνεται σταδιακή αποκωδικοποίηση)
- Δεδομένα των μπλοκ



ΠΑΡΑΔΕΙΓΜΑ JPEG CODING(ENCODER)

Original 8×8 block

52	55	61	66	70	61	64	73
63	59	66	90	109	85	69	72
62	59	68	113	144	104	66	73
63	58	71	122	154	106	70	69
67	61	68	104	126	88	68	70
79	65	60	70	77	68	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94

DCT

The DCT outcome

-415	-29	-62	25	55	-20	-1	3
7	-21	-62	9	11	-7	-6	6
-46	8	77	-25	-30	10	7	-5
-50	13	35	-15	-9	6	0	3
11	-8	-13	-2	-1	1	-4	1
-10	1	3	-3	-1	0	2	-1
-4	-1	2	-1	2	-3	1	-2
-1	-1	-1	-2	-1	-1	0	-1

ΠΑΡΑΔΕΙΓΜΑ JPEG CODING(ENCODER)

$-415/16 = -26$ Πίνακας κβαντισμού

-415	-29	-62	25	55	-20	-1	3	16	11	10	16	24	40	51	61
7	-21	-62	9	11	-7	-6	6	12	12	14	19	26	58	60	55
-46	8	77	-25	-30	10	7	-5	14	13	16	24	40	57	69	56
-50	13	35	-15	-9	6	0	3	14	17	22	29	51	87	80	62
11	-8	-13	-2	-1	1	-4	1	18	22	37	56	68	109	103	77
-10	1	3	-3	-1	0	2	-1	24	35	55	64	81	104	113	92
-4	-1	2	-1	2	-3	1	-2	49	64	78	87	103	121	120	101
-1	-1	-1	-2	-1	-1	0	-1	72	92	95	98	112	100	103	99

Αν πολλαπλασιάσω τον πίνακα με 2, 4 και 8. Χάνω πληροφορία, δηλαδή συμπιέζω επιπλέον την εικόνα.

ΠΑΡΑΔΕΙΓΜΑ JPEG CODING(ENCODER)

-415	-29	-62	25	55	-20	-1	3	-26	-3	-6	2	2	0	0	0
7	-21	-62	9	11	-7	-6	6	1	-2	-4	0	0	0	0	0
-46	8	77	-25	-30	10	7	-5	-3	1	5	-1	-1	0	0	0
-50	13	35	-15	-9	6	0	3	-4	1	2	-1	0	0	0	0
11	-8	-13	-2	-1	1	-4	1	1	0	0	0	0	0	0	0
-10	1	3	-3	-1	0	2	-1	0	0	0	0	0	0	0	0
-4	-1	2	-1	2	-3	1	-2	0	0	0	0	0	0	0	0
-1	-1	-1	-2	-1	-1	0	-1	0	0	0	0	0	0	0	0

Σημαντική παρατήρηση: η πληροφορία μετά την κβάντιση έχει μαζευτεί πάνω και αριστερά.

Με άλλα λόγια: Η κωδικοποίηση πρέπει να γίνεται εστιάζοντας στα πάνω αριστερά σημεία.

ΕΠΕΚΤΑΣΕΙΣ ΤΟΥ ΒΑΣΙΚΟΥ JPEG

- **Προοδευτική κωδικοποίηση (Progressive mode)**
 - Δημιουργήθηκε για την μετάδοσης εικόνων σε πραγματικό χρόνο (streaming).
 - Οι συντελεστές DCT των μπλοκ μεταδίδονται σε πολλαπλά «περάσματα» της εικόνας αρχίζοντας από τους DC συντελεστές.
 - Με κάθε πέρασμα ο αποκωδικοποιητής μπορεί να παράγει μια υψηλότερης ποιότητας εκδοχή της εικόνας, οπότε μια προεκδοχή της μπορεί να σταλθεί γρήγορα και να αποφασίσει ο χρήστης που την λαμβάνει εάν θέλει να την αφήσει να ολοκληρωθεί και να βελτιωθεί σε ποιότητα.

ΕΠΕΚΤΑΣΕΙΣ ΤΟΥ ΒΑΣΙΚΟΥ JPEG

- **Ιεραρχική κωδικοποίηση (Hierarchical mode)**

- Η εικόνα αναπαρίσταται σε διαφορετικές αναλύσεις. Έτσι, μια ανάλυση της εικόνας που θέλουμε να επεξεργαστούμε θα μπορούσε να είναι 128×128 , μια άλλη 256×256 και μια άλλη 512×512 .
- Η κάθε ανάλυση κωδικοποιείται ως το σύνολο των διαφορών που έχει από την αμέσως χαμηλότερη της ανάλυση και επομένως δεν έχουμε άσκοπη επανάληψη δεδομένων που ήδη έχουμε λάβει με την προηγούμενη εκδοχή της εικόνας.

ΕΠΕΚΤΑΣΕΙΣ ΤΟΥ ΒΑΣΙΚΟΥ JPEG: ΙΕΡΑΡΧΙΚΗ ΚΩΔΙΚΟΠΟΙΗΣΗ



Ανάλυση 128x128



Ανάλυση 256x256



Ανάλυση 512x512