

Ψηφιακή Επεξεργασία Εικόνας

Συμπίεση εικόνας

Τι είναι; Γιατί χρειάζεται;

Συμπίεση εικόνας: μείωση του ποσού δεδομένων που απαιτείται για την αναπαράσταση μιας εικόνας

Γιατί - μείωση αποθηκευτικού χώρου

Έστω video 2 ωρών, ανάλυσης 720×480.

Για κάθε frame απαιτούνται 720×480×24 bits (ή 720×480×3 bytes).

Το frame rate είναι γύρω στα 30fps (30 frames per second).

$30 \text{ frames/sec} \times (720 \times 480) \text{ pixels/frame} \times 3 \text{ bytes/pixel} = 31,104,000 \text{ bytes/sec}$

$31,104,000 \text{ bytes/sec} \times (60^2) \text{ sec/hr} \times 2 \text{ hrs} \approx 2.24 \times 10^{11} \text{ bytes (224 GB) - 27 DVD (8.5 GB)}$

Για αποθήκευση σε 1 DVD (8.5 GB), συμπίεση κατά 26.3

Γιατί χρειάζεται;

Γιατί - μείωση χρόνου μετάδοσης

Έστω δύο συνδέσεις Internet με ταχύτητες 56 Kbps (kilobits per second) και 12 Mbps (megabits per second).

Για τη μετάδοση μιας έγχρωμης εικόνας ($128 \times 128 \times 24$ bits) χρειάζονται περίπου 7 sec και 0.03 sec, αντίστοιχα.

Η συμπίεση μπορεί να μειώσει το χρόνο μετάδοσης κατά παράγοντα 2 ως 10 ή περισσότερο.

Συμπίεση δεδομένων (data compression)

Data compression: Η διαδικασία μείωσης του ποσού των δεδομένων που απαιτείται για την αναπαράσταση μιας δεδομένης ποσότητας πληροφορίας.

Δεδομένα και πληροφορία δεν είναι το ίδιο πράγμα.

Δεδομένα: το μέσο μέσω του οποίου μεταφέρεται η πληροφορία

Τεράστια ποσά δεδομένων μπορούν να αναπαριστούν την ίδια πληροφορία με μικρότερα ποσά δεδομένων.

Αναπαραστάσεις που περιέχουν επαναλαμβανόμενη ή «άσχετη» πληροφορία λέγεται ότι περιέχουν πλεονάζοντα δεδομένα (redundant data).

Συμπίεση δεδομένων (data compression)

Έστω b , b' ο αριθμός bits που χρησιμοποιείται σε δύο αναπαραστάσεις της ίδιας πληροφορίας.

Ο σχετικός πλεονασμός δεδομένων (relative data redundancy) R της αναπαράστασης με τα b bits είναι:

$$R = 1 - \frac{1}{C}$$

όπου C είναι ο λόγος συμπίεσης (compression ratio), ο οποίος ορίζεται σαν:

$$C = \frac{b}{b'}$$

Αν $C = 10$ (10:1), η μεγαλύτερη αναπαράσταση έχει 10 bits δεδομένων για κάθε 1 bit της μικρότερης αναπαράστασης.

Το αντίστοιχο R της μεγαλύτερης αναπαράστασης είναι ίσο με 0.9, που σημαίνει ότι το 90% των δεδομένων της είναι πλεονάζοντα.

Συμπίεση εικόνας

b: αριθμός bits που χρησιμοποιείται για την αναπαράσταση της εικόνας σαν ένα 2D-array τιμών έντασης

2D-array τιμών έντασης: κατάλληλο για απεικόνιση, όχι όμως για συμπυκνωμένη αναπαράσταση της εικόνας

3 τύποι πλεονασμού δεδομένων (data redundancy):

Coding redundancy

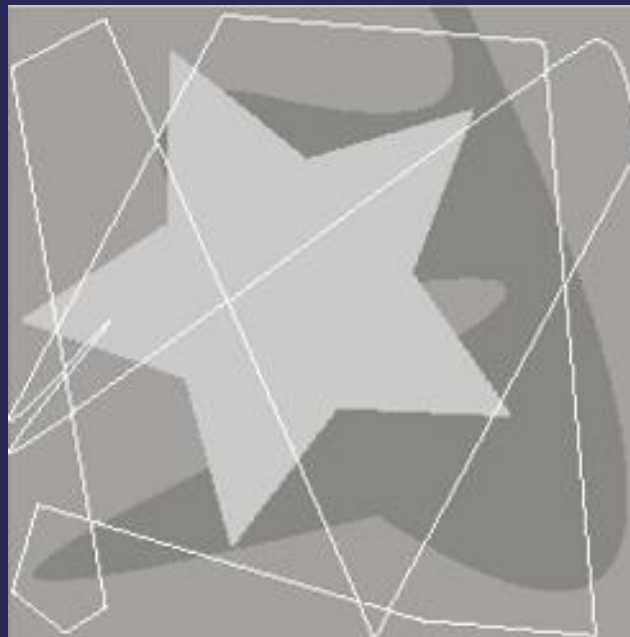
Ένας κώδικας (code) είναι ένα σύστημα συμβόλων (γράμματα, αριθμοί, bits) που χρησιμοποιείται για την αναπαράσταση πληροφορίας.

Κάθε τμήμα πληροφορίας εκχωρείται σε μια ακολουθία από σύμβολα του κώδικα (code symbols), που λέγεται κωδική λέξη (code word).

Συμπίεση εικόνας

Ο αριθμός συμβόλων σε κάθε κωδική λέξη ορίζουν το **μήκος της (length)**.

Οι 8-bit κώδικες που χρησιμοποιούνται για την αναπαράσταση των τιμών της έντασης σε ένα 2D array περιέχουν περισσότερα bits από όσα πραγματικά χρειάζονται.



4 τιμές έντασης

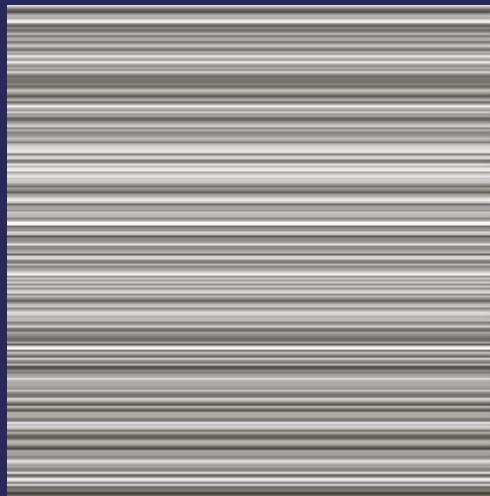
Συμπίεση εικόνας

Spatial and temporal redundancy

Σε ένα 2D array τα pixels συσχετίζονται χωρικά (κάθε pixel είναι παρόμοιο με τα γειτονικά του pixels ή εξαρτάται από αυτά).

Στις αναπαραστάσεις συσχετισμένων pixels επαναλαμβάνεται πληροφορία.

Σε μια ακολουθία video υπάρχουν χρονικά συσχετισμένα pixels, για τα οποία επίσης υπάρχει επανάληψη πληροφορίας.



Spatial redundancy: σε κάθε γραμμή η τιμή της έντασης είναι η ίδια για όλα τα pixels της γραμμής

Συμπίεση εικόνας

Irrelevant information

Τα περισσότερα 2D arrays περιέχουν πληροφορία η οποία αγνοείται από το ανθρώπινο σύστημα όρασης (human visual system - HVS).

Η πληροφορία αυτή είναι πλεονάζουσα, με την έννοια ότι δεν χρησιμεύει σε κάτι.



Η ένταση φαίνεται σταθερή (αλλά δεν είναι).

Coding redundancy

Έστω r_k τ.μ. στο διάστημα $[0, L-1]$, η οποία χρησιμοποιείται για την αναπαράσταση των εντάσεων μιας $M \times N$ εικόνας.

Επίσης, κάθε r_k έχει πιθανότητα εμφάνισης $p_r(r_k)$

$$p_r(r_k) = \frac{n_k}{MN}, \quad k = 0, 1, 2, \dots, L-1$$

όπου L ο αριθμός των τιμών της έντασης και n_k ο αριθμός των εμφανίσεων της έντασης k στην εικόνα.

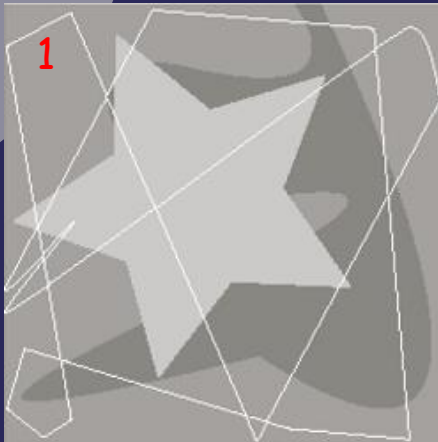
Αν ο αριθμός των bits που χρησιμοποιείται για την αναπαράσταση κάθε τιμής r_k είναι $l(r_k)$, ο μέσος αριθμός bits που απαιτείται για την αναπαράσταση κάθε pixel (το μέσο μήκος κώδικα) είναι:

$$L_{avg} = \sum_{k=0}^{L-1} l_k(r_k) p_r(r_k)$$

Ο συνολικός αριθμός bits για την αναπαράσταση μιας $M \times N$ εικόνας είναι MNL_{avg} .

Coding redundancy

Natural m -bit fixed-length code: τότε $L_{\text{avg}}=m$, γιατί $l(r_k)=m$ για όλα τα r_k (και το άθροισμα των πιθανοτήτων $p_r(r_k)$ είναι προφανώς ίσο με 1).



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

natural 8-bit code

variable-length coding

$$L_{\text{avg}} = 0.25(2) + 0.47(1) + 0.25(3) + 0.03(3) = 1.81 \text{ bits}$$

$$\text{MNL}_{\text{avg}} = 256 \times 256 \times 1.81 = 118,621$$

$$C = \frac{256 \times 256 \times 8}{118,621} = \frac{8}{1.81} \approx 4.42$$

$$R = 1 - \frac{1}{4.42} = 0.774 \quad 77.4\% \text{ redundant data}$$

Coding redundancy

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

variable-length coding

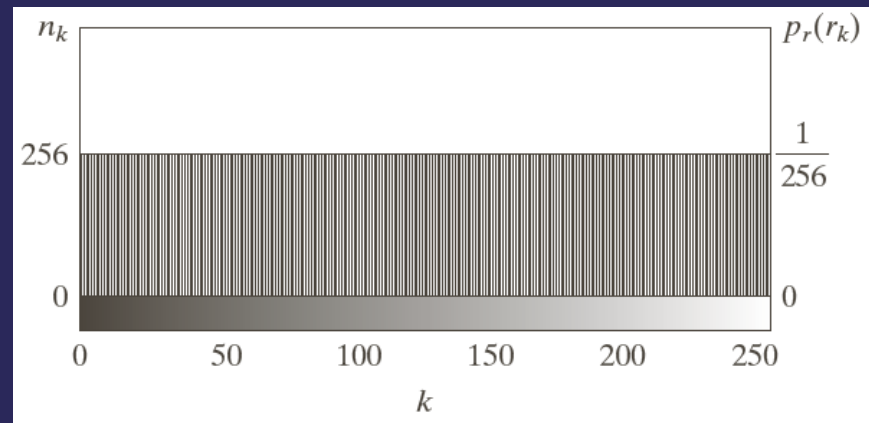
Η συμπίεση (Code 2) προκύπτει από την ανάθεση λιγότερων bits στις πιο πιθανές τιμές έντασης.

$$l_2(r_{128}) = 1, \quad l_2(r_{255}) = 3$$

Ο καλύτερος fixed-length code που θα μπορούσε να χρησιμοποιηθεί είναι ο 2-bit {00, 01, 10, 11}.

Η συμπίεση σε αυτή την περίπτωση θα ήταν 8/2 (4:1), δηλαδή 10% λιγότερο από την 4.42:1 συμπίεση που πετυχαίνουμε με τον variable-length code.

Spatial and Temporal redundancy



1. Οι 256 εντάσεις εμφανίζονται με την ίδια πιθανότητα (ιστόγραμμα με ομοιόμορφή κατανομή).
2. Σε κάθε γραμμή, η τιμή της έντασης επιλέχτηκε τυχαία, άρα τα pixels της εικόνας είναι ανεξάρτητα στην κάθετη διεύθυνση.
3. Κατά μήκος μιας γραμμής, τα pixels έχουν την ίδια τιμή, είναι δηλαδή πλήρως συσχετισμένα στην οριζόντια διεύθυνση.

Spatial and Temporal redundancy

1. Η συγκεκριμένη εικόνα, όταν αναπαρίσταται σαν 8-bit array, δεν μπορεί να συμπιεστεί από έναν variable-length κώδικα μόνο. Οι τιμές της έντασης είναι ισοπίθανες και ένας 8-bit fixed-length κώδικας ελαχιστοποιεί το L_{avg} .

2, 3. Υπάρχει σημαντικός χωρικός (spatial) πλεονασμός, ο οποίος μπορεί να εξαλειφθεί αναπαριστώντας την εικόνα με run-length ζευγάρια.

Run-length pair: καθορίζει την αρχή μιας τιμής έντασης και τον αριθμό των διαδοχικών pixels που έχουν αυτή την ένταση.

Συμπίεση: $(256 \times 256 \times 8) / [(256 + 256) \times 8]$ ή 128:1

Κάθε γραμμή (256 pixels) της αρχικής αναπαράστασης αντικαθίσταται στη run-length αναπαράσταση με μία 8-bit τιμή έντασης και μήκος ίσο με 256.

Spatial and Temporal redundancy

Στις περισσότερες εικόνες υπάρχει χωρική συσχέτιση (κατά x και κατά y) και χρονική συσχέτιση αν η εικόνα είναι τμήμα μιας ακολουθίας video.

Οι περισσότερες τιμές έντασης των pixels μπορούν να προβλεφθούν από τις εντάσεις της γειτονιάς.

Η πληροφορία που φέρει ένα απλό pixel είναι μικρή (αρκετή από τη συνεισφορά του στην απεικόνιση είναι πλεονάζουσα, αφού μπορεί να εξαχθεί από τους γείτονές του).

Μείωση της συγκεκριμένης μορφής πλεονασμού με αναπαραστάσεις μη απεικονιστικές (π.χ. run-lengths, διαφορές ανάμεσα σε γειτονικά pixels): mappings

Reversible mapping: αν τα pixels της αρχικής εικόνας μπορούν να αναπαραχθούν χωρίς σφάλματα

Irrelevant information

Απλός τρόπος συμπίεσης δεδομένων: απομάκρυνση περιττών δεδομένων (π.χ. στην εικόνα, πληροφορία που δεν είναι αντιληπτή από το μάτι μπορεί να παραλειφθεί).

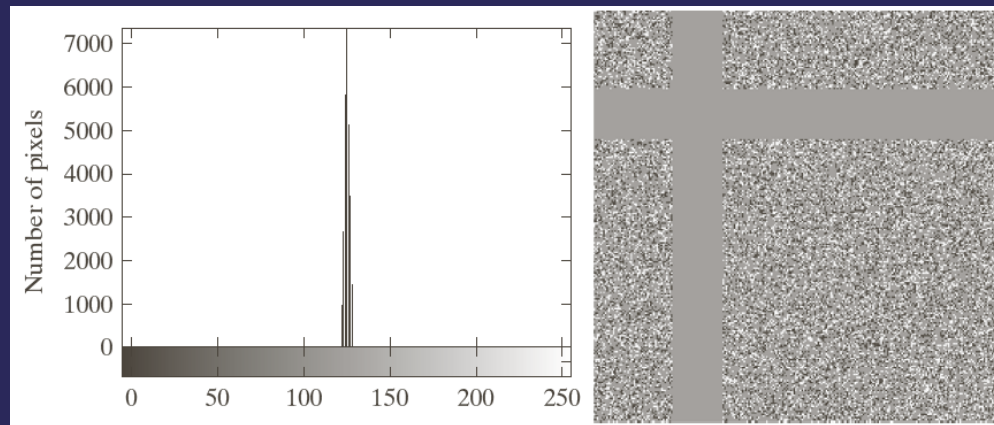


Αρχική εικόνα: φαίνεται ομοιογενής, αναπαρίσταται από τη μέση τιμή της έντασης - μία 8-bit τιμή.

Συμπίεση $(256 \times 256 \times 8) / 8$ ή 65,536:1.

Η εικόνα πρέπει να ανακατασκευαστεί για να απεικονιστεί μετά τη συμπίεση.

Irrelevant information



Αρχική εικόνα φαίνεται ομοιογενής, αλλά υπάρχουν αρκετές τιμές έντασης (125 ως 131).

Αν η εικόνα αναπαρασταθεί από τη μέση τιμή της, η πληροφορία αυτή (που φαίνεται στην εξισωμένη εικόνα) θα χαθεί.

Εξαρτάται από την εφαρμογή αν θέλουμε να διατηρήσουμε αυτή την πληροφορία (η οποία υπό normal συνθήκες δεν είναι ορατή) ή όχι.

Irrelevant information

Πλεονασμός διαφορετικός από τους άλλους.

Οφείλεται στο γεγονός ότι η πληροφορία δεν είναι απαραίτητη σε μια συνηθισμένη διαδικασία απεικόνισης.

Απώλεια πληροφορίας -> κβαντισμός

Κβαντισμός: αντιστοίχιση μεγάλου εύρους τιμών έντασης σε έναν περιορισμένο αριθμό τιμών

Κβαντισμός: μη αντιστρέψιμη διαδικασία (απώλεια πληροφορίας)

Μέτρηση πληροφορίας εικόνας

Υπάρχει κάποιο ελάχιστο ποσό δεδομένων, ικανό να περιγράψει μια εικόνα χωρίς απώλεια πληροφορίας;

Η θεωρία πληροφορίας παρέχει το μαθηματικό υπόβαθρο για την απάντηση σε αυτή την ερώτηση.

Ένα τυχαίο γεγονός E με πιθανότητα $P(E)$, περιέχει

$$I(E) = \log \frac{1}{P(E)} = -\log P(E)$$

μονάδες πληροφορίας.

Αν $P(E) = 1$ (το γεγονός συμβαίνει πάντα), $I(E) = 0$ (δεν φέρει καμία πληροφορία).

Η βάση του λογαρίθμου καθορίζει τη μονάδα που χρησιμοποιείται για τη μέτρηση της πληροφορίας.

Για βάση ίση με 2, η μονάδα μέτρησης της πληροφορίας είναι το bit.

Μέτρηση πληροφορίας εικόνας

Αν $P(E) = \frac{1}{2}$, $I(E) = -\log_2 \frac{1}{2} = 1 \text{ bit}$.

Ένα bit πληροφορίας παρέχεται, όταν συμβαίνει ένα από δύο ισοδύναμα γεγονότα (π.χ. ρίχνω ένα νόμισμα και ανακοινώνω το αποτέλεσμα).

Δεδομένης μιας πηγής η οποία παράγει ανεξάρτητα τυχαία γεγονότα από ένα σύνολο πιθανών γεγονότων $\{a_1, a_2, \dots, a_J\}$ με πιθανότητες εμφάνισης $\{P(a_1), P(a_2), \dots, P(a_J)\}$, η μέση πληροφορία που παράγει η πηγή ανά έξοδό της λέγεται **εντροπία** και είναι:

$$H = -\sum_{j=1}^J P(a_j) \log P(a_j)$$

Τα a_j λέγονται σύμβολα πηγής, και επειδή είναι στατιστικά ανεξάρτητα η πηγή λέγεται μηδενικής μνήμης.

Μέτρηση πληροφορίας εικόνας

Αν θεωρήσουμε την εικόνα έξοδο μιας πηγής εντάσεων, μπορούμε να χρησιμοποιήσουμε το ιστόγραμμά της για να υπολογίσουμε τις πιθανότητες των συμβόλων (εντάσεων) της πηγής.

Τότε, η εντροπία της πηγής εντάσεων είναι:

$$\tilde{H} = -\sum_{k=1}^{L-1} p_r(r_k) \log_2 p_r(r_k)$$

Δεν είναι δυνατό να κωδικοποιήσουμε τις τιμές της έντασης μιας εικόνας με λιγότερα από \tilde{H} bits/pixel.

Το ποσό της εντροπίας (και άρα της πληροφορίας) μιας εικόνας δεν μπορούμε να το καταλάβουμε διαισθητικά.

Μέτρηση πληροφορίας εικόνας

r_k	$p_r(r_k)$
$r_{87} = 87$	0.25
$r_{128} = 128$	0.47
$r_{186} = 186$	0.25
$r_{255} = 255$	0.03
r_k for $k \neq 87, 128, 186, 255$	0

$$\begin{aligned}\tilde{H} &= -[0.25 \log_2 0.25 + 0.47 \log_2 0.47 + 0.25 \log_2 0.25 + 0.03 \log_2 0.03] \\ &\approx -[0.25(-2) + 0.47(-1.09) + 0.25(-2) + 0.03(-5.06)] \\ &\approx 1.6614 \text{ bits/pixel}\end{aligned}$$

Παρομοίως, οι εντροπίες των άλλων δύο εικόνων είναι 8 bits/pixel και 1.566 bits/pixel.

Η πρώτη εικόνα έχει την περισσότερη πληροφορία από άποψη απεικόνισης, αλλά έχει την μικρότερη εντροπία (1.6614 bits/pixel).

Η δεύτερη εικόνα, με εντροπία περίπου πέντε φορές μεγαλύτερη, φαίνεται να έχει την ίδια ή λιγότερη «απεικονιστική» πληροφορία σε σχέση με την πρώτη εικόνα.

Η τρίτη εικόνα, με εντροπία περίπου ίση με την πρώτη εικόνα, φαίνεται να φέρει μικρή ή καθόλου πληροφορία.

Μέτρηση πληροφορίας εικόνας

Χρησιμοποιώντας variable-length κώδικα για την πρώτη εικόνα
-> αναπαράσταση τιμών έντασης με 1.81 bits/pixel.

Υπολογίσαμε όμως εντροπία ίση με 1.6614 bits/pixel.

$$0 \leq H \leq \log_2 K$$

όπου K το πλήθος των συμβόλων της πηγής

$H = 0$, αν η πιθανότητα εμφάνισης ενός συμβόλου είναι 1.

$H = \log_2 K$, αν οι πιθανότητες εμφάνισης των K συμβόλων είναι ίσες ($1/K$).

Μέτρηση πληροφορίας εικόνας

1^ο Θεώρημα Shannon (Θεώρημα κωδικοποίησης πηγής)

Για δεδομένη διακριτή πηγή χωρίς μνήμη με εντροπία H , το μέσο μήκος κώδικα L_{avg} οποιουδήποτε κωδικοποιητή πηγής φράσσεται από την ανισότητα:

$$L_{avg} \geq H$$

Άρα $(L_{avg})_{min} = H$ (το ελάχιστο μέσο μήκος κώδικα είναι ίσο με την εντροπία της πηγής).

Η αποδοτικότητα ενός κωδικοποιητή πηγής είναι:

$$\eta = \frac{(L_{avg})_{min}}{L_{avg}} = \frac{H}{L_{avg}}$$

Μέτρηση πληροφορίας εικόνας

Παρατηρήσεις

Εντροπία: κάτω όριο στη συμπίεση που μπορεί να επιτευχθεί στην κωδικοποίηση στατιστικά ανεξάρτητων σύμβολων.

Δεν ισχύει στην περίπτωση συσχετισμένων συμβόλων, όπου ομάδες συμβόλων μπορούν να κωδικοποιηθούν με λιγότερα bits (π.χ. run-length).

Όταν η έξοδος μιας πηγής πληροφορίας εξαρτάται από ένα πεπερασμένο αριθμό προηγούμενων εξόδων, τότε η πηγή λέγεται Markov ή πεπερασμένης μνήμης.

Κριτήρια πιστότητας (fidelity)

Απώλεια πληροφορίας από εικόνα \rightarrow μείωση ποιότητας εικόνας

Αποτίμηση απώλειας μέσω:

- 1) Αντικειμενικών κριτηρίων πιστότητας (μαθηματικές συναρτήσεις)
- 2) Υποκειμενικών κριτηρίων πιστότητας (όπως την αντιλαμβάνεται ο άνθρωπος)

Αντικειμενικά κριτήρια:

$$e(x, y) = \hat{f}(x, y) - f(x, y)$$

$$e_{total} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]$$

$$e_{rms} = \left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2 \right]^{1/2}$$

Κριτήρια πιστότητας (fidelity)

$$SNR = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x,y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x,y) - f(x,y)]^2}$$

Υποκειμενικά κριτήρια

Άνθρωποι αποτιμούν την ποιότητα μιας εικόνας

π.χ. $\{-3, -2, -1, 0, 1, 2, 3\}$

ή

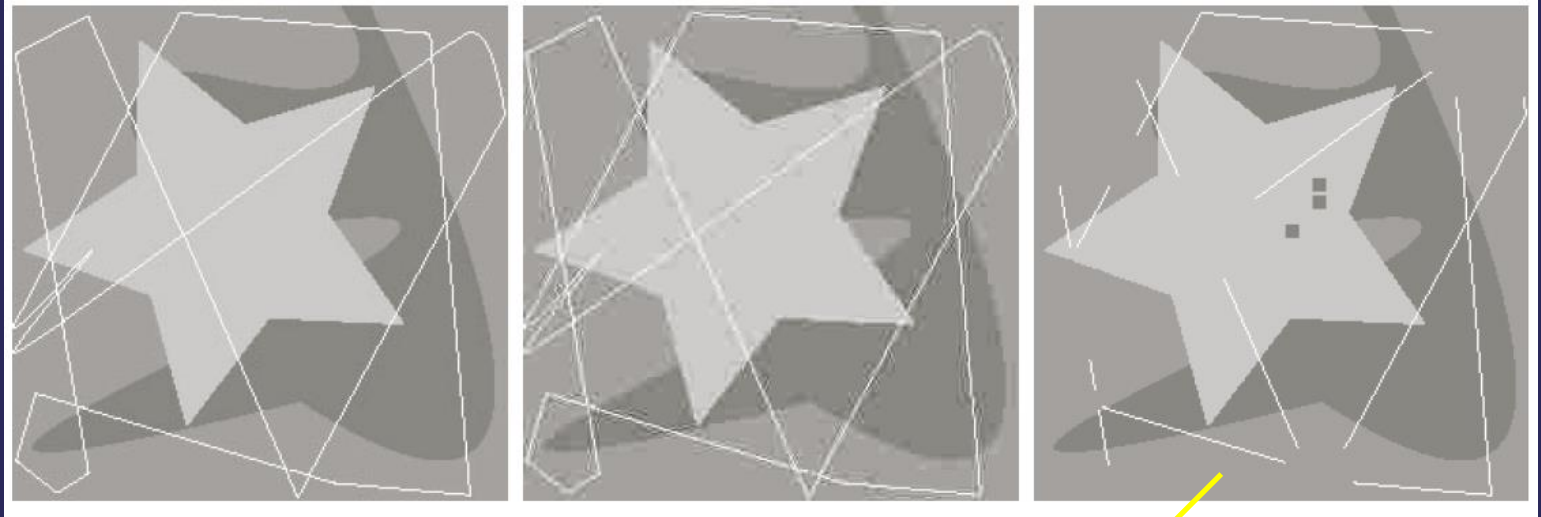
$\{\text{πολύ χειρότερη, χειρότερη, λίγο χειρότερη, ίδια, λίγο καλύτερη, καλύτερη, πολύ καλύτερη}\}$

Κριτήρια πιστότητας (fidelity)

Υποκειμενικά κριτήρια

Value	Rating	Description
1	Excellent	An image of extremely high quality, as good as you could desire.
2	Fine	An image of high quality, providing enjoyable viewing. Interference is not objectionable.
3	Passable	An image of acceptable quality. Interference is not objectionable.
4	Marginal	An image of poor quality; you wish you could improve it. Interference is somewhat objectionable.
5	Inferior	A very poor image, but you could watch it. Objectionable interference is definitely present.
6	Unusable	An image so bad that you could not watch it.

Κριτήρια πιστότητας (fidelity)

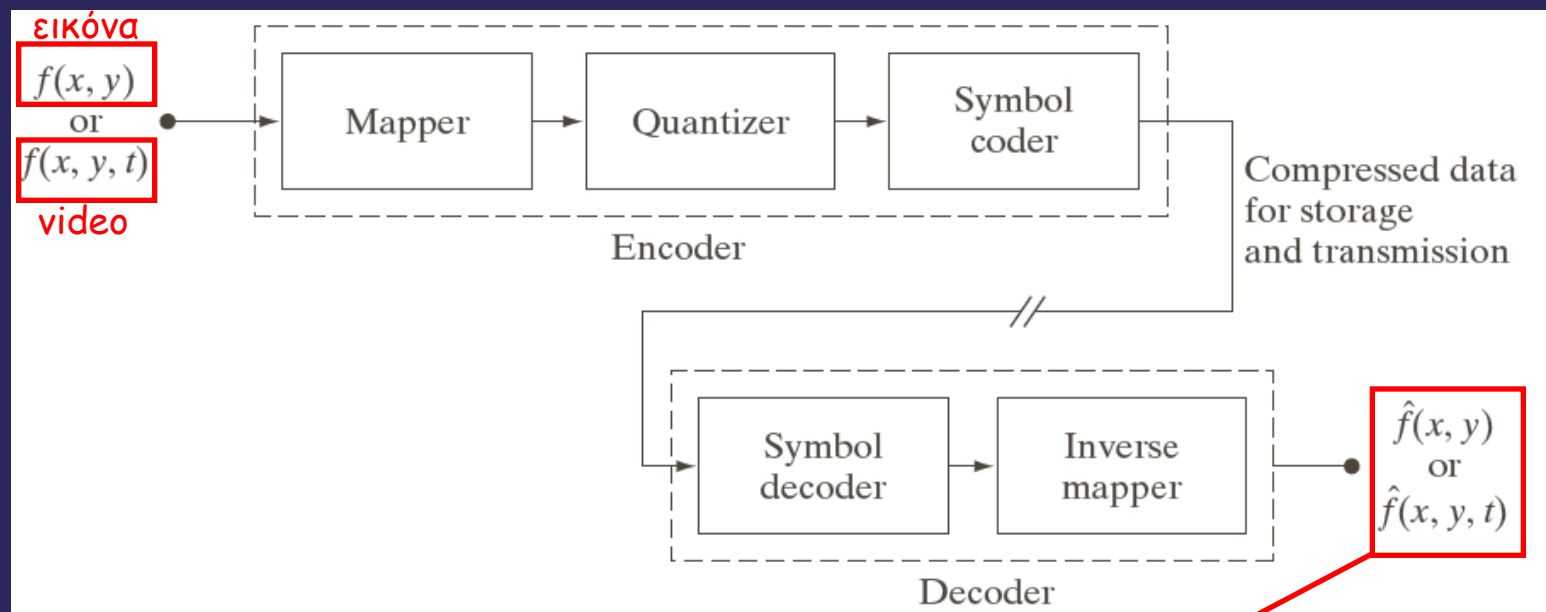


Περιορισμοί αντικειμενικών κριτηρίων

e_{rms} : 5.17, 15.67, 14.17

excellent, passable or marginal, inferior or unusable

Μοντέλα συμπίεσης εικόνας



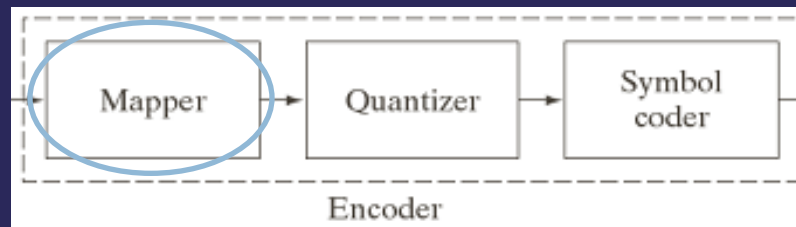
Ίδια με τα αρχικά: **lossless compression**
Αλλιώς: **lossy compression**

Encoder (κωδικοποιητής): συμπίεση

Decoder (αποκωδικοποιητής): αποσυμπίεση

Codec: συσκευή ή πρόγραμμα που κάνει και τις δύο διαδικασίες (κωδικοποίηση και αποκωδικοποίηση)

Μοντέλα συμπίεσης εικόνας



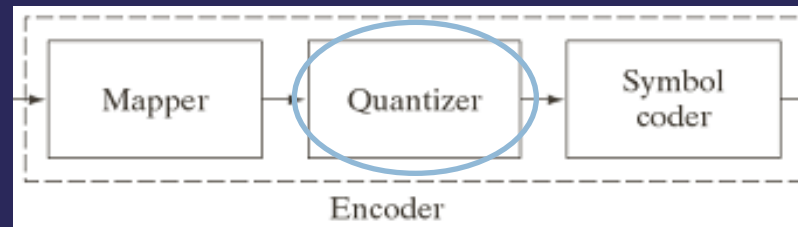
Mapper: μετατρέπει τα δεδομένα σε κάποιο format κατάλληλο για μείωση της πλεονάζουσας πληροφορίας

Διαδικασία γενικά αναστρέψιμη, η οποία μπορεί (μπορεί και όχι) να μειώνει άμεσα το ποσό των δεδομένων που χρειάζονται για την αναπαράσταση της εικόνας.

Run-length coding: mapping που αποφέρει συμπίεση στο πρώτο στάδιο της κωδικοποίησης

Αντίθετο παράδειγμα: η αντιστοίχιση της εικόνας σε ένα μετασχηματισμό (οι συντελεστές του μετασχηματισμού είναι λιγότερο συσχετισμένοι από τα pixels της εικόνας)

Μοντέλα συμπίεσης εικόνας



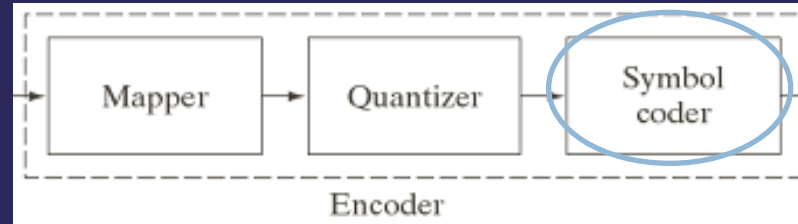
Quantizer: μειώνει την ακρίβεια της εξόδου του mapper με βάση κάποιο προκαθορισμένο κριτήριο πιστότητας

Ο σκοπός είναι να κρατήσουμε την «άσχετη» πληροφορία έξω από τη συμπιεσμένη αναπαράσταση.

Μη αναστρέψιμη διαδικασία. Παραλείπεται στη lossless συμπίεση.

Video: το bit rate της κωδικοποιημένης εξόδου χρησιμοποιείται για τον καθορισμό της λειτουργίας του κβαντιστή

Μοντέλα συμπίεσης εικόνας



Symbol coder: παράγει ένα *fixed-length* ή *variable-length* κώδικα για την αναπαράσταση της εξόδου του κβαντιστή

Στην περίπτωση *variable-length* κώδικα, οι μικρότερες κωδικές λέξεις χρησιμοποιούνται για τις εξόδους του κβαντιστή με την πιο συχνή εμφάνιση.

Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman

Prefix-free κώδικας: καμία κωδική λέξη δεν αποτελεί πρόθεμα κάποιας άλλης

Αποκωδικοποιήσιμος

Βέλτιστος: για δεδομένη διακριτή πηγή χωρίς μνήμη, πετυχαίνει το μικρότερο μέσο μήκος

Το μέσο μήκος του πλησιάζει την εντροπία της πηγής που κωδικοποιεί. $L_{avg} \geq H$ όπου:

$$H = -\sum_{j=1}^J P(a_j) \log P(a_j)$$

Στην περίπτωση της εικόνας, σαν σύμβολα πηγής (a_j) μπορούν να θεωρηθούν οι εντάσεις της εικόνας ή η έξοδος μιας διαδικασίας αντιστοίχισης των εντάσεων (π.χ. διαφορές pixels, run lengths).

Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman

- 1: Τα σύμβολα της πηγής ταξινομούνται από το περισσότερο πιθανό στο λιγότερο πιθανό.
- 2: Τα δύο λιγότερα πιθανά σύμβολα συγχωνεύονται σε ένα, με πιθανότητα ίση με το άθροισμα των πιθανοτήτων των συμβολών αυτών.
- 3: Το νέο σύμβολο ταξινομείται στη σωστή θέση, σύμφωνα με την πιθανότητα που έχει. Στην περίπτωση που το νέο σύμβολο έχει πιθανότητα ίση με κάποιο υπάρχων σύμβολο, συνηθίζεται στη διαδικασία της ταξινόμησης να προηγείται του υπάρχοντος συμβόλου.
- 4: Η διαδικασία επαναλαμβάνεται μέχρι να φτάσουμε στη ρίζα του δυαδικού δέντρου.
- 5: Αναθέτουμε δυαδικά σύμβολα (0 και 1) στα κλαδιά του δέντρου και διαβάζοντας προς τα κάτω σχηματίζονται οι κωδικές λέξεις.

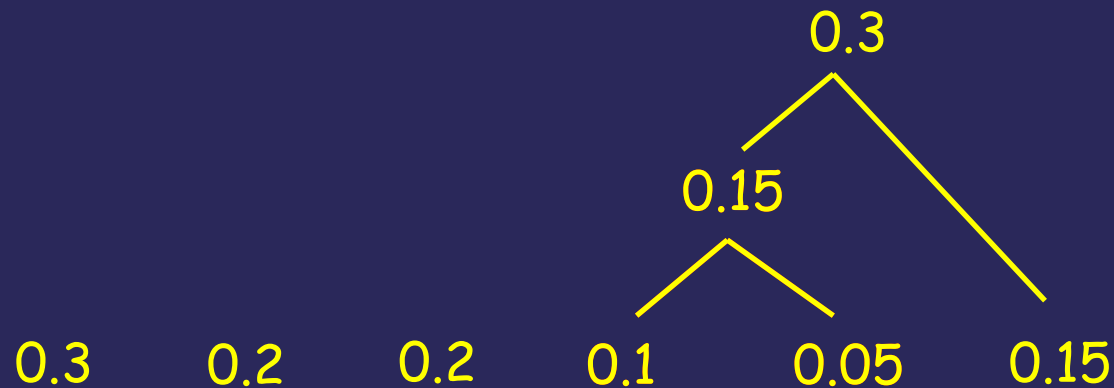
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



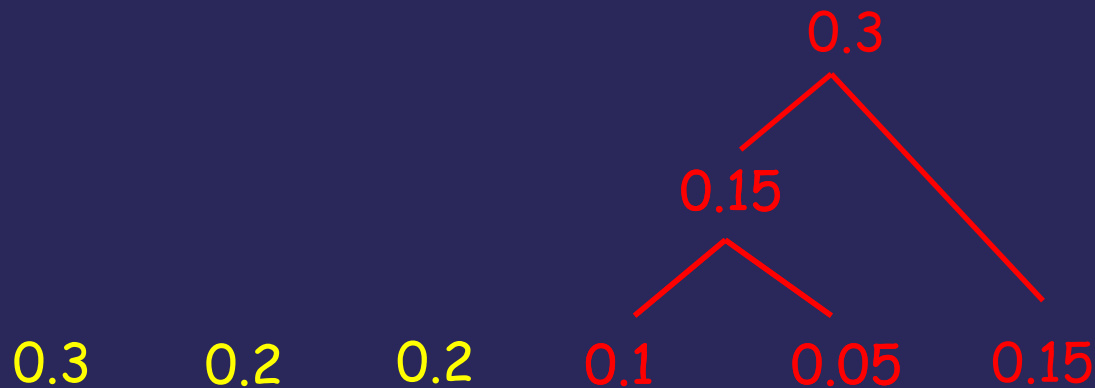
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



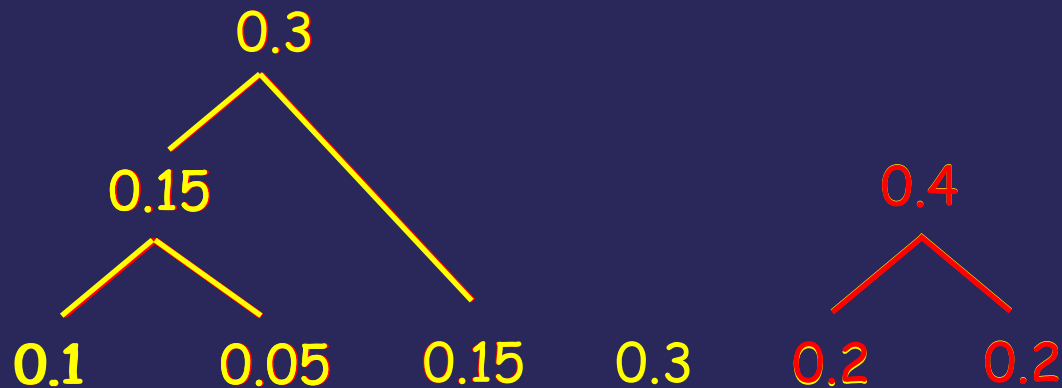
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



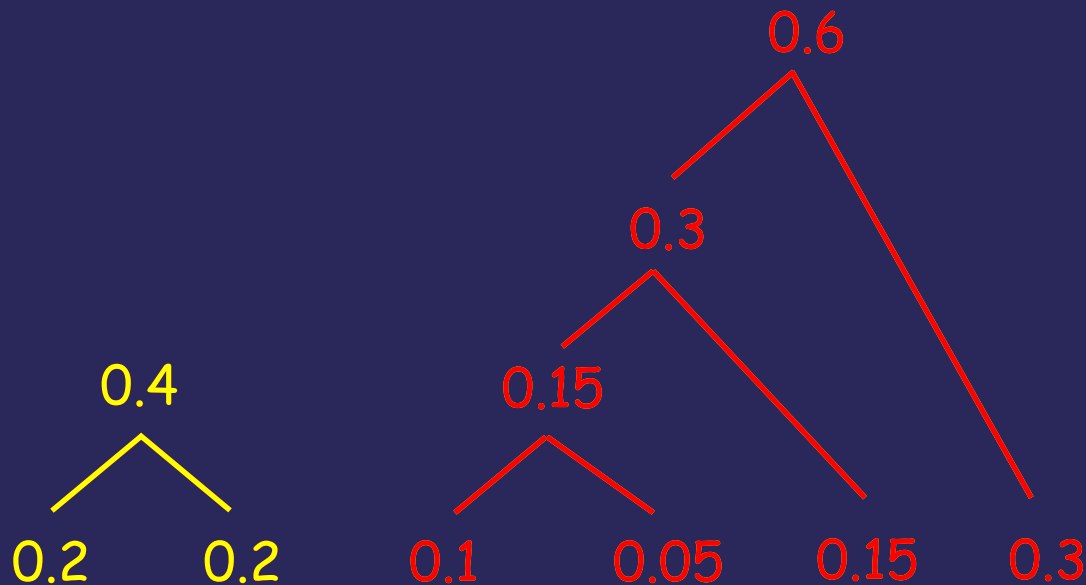
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



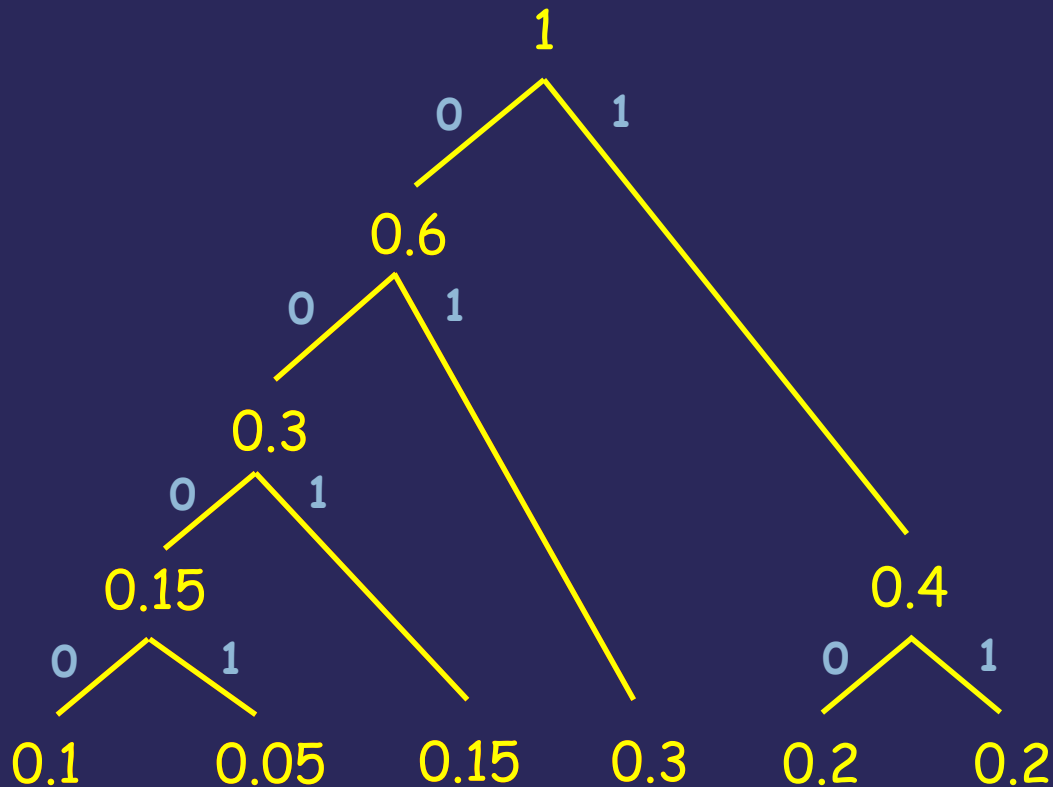
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



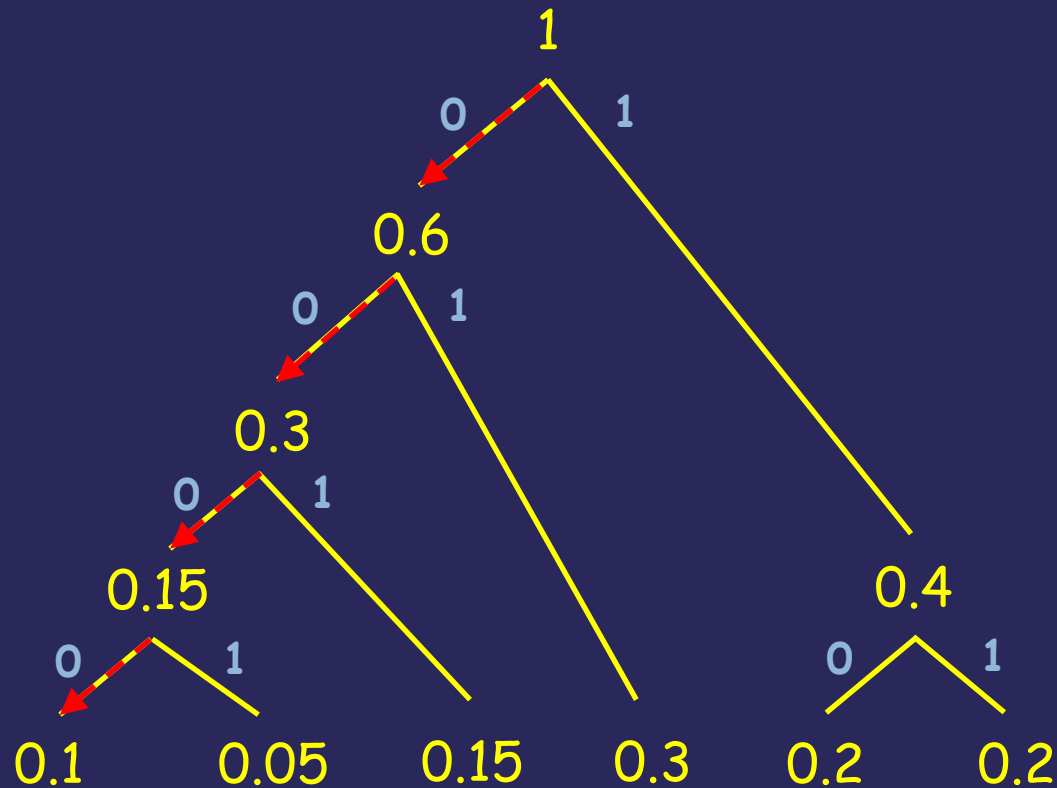
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



Lossless αλγόριθμοι συμπίεσης

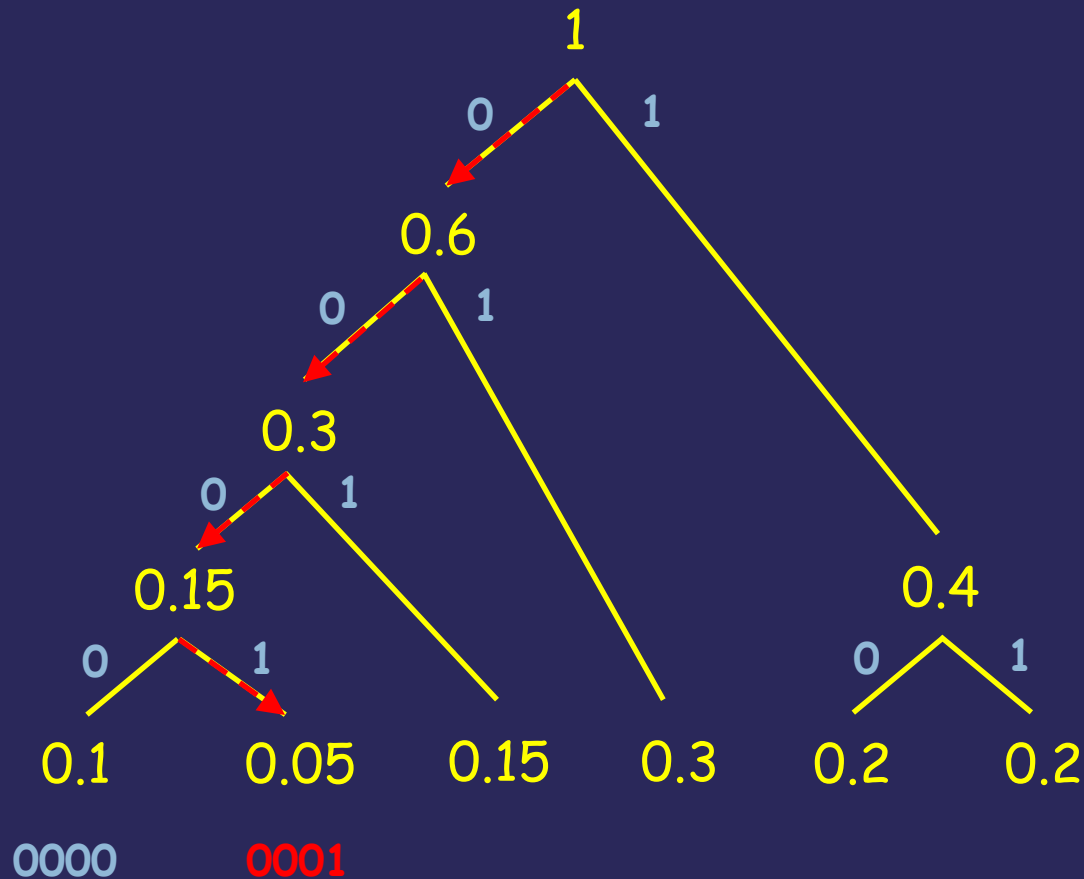
Κωδικοποίηση Huffman



0000

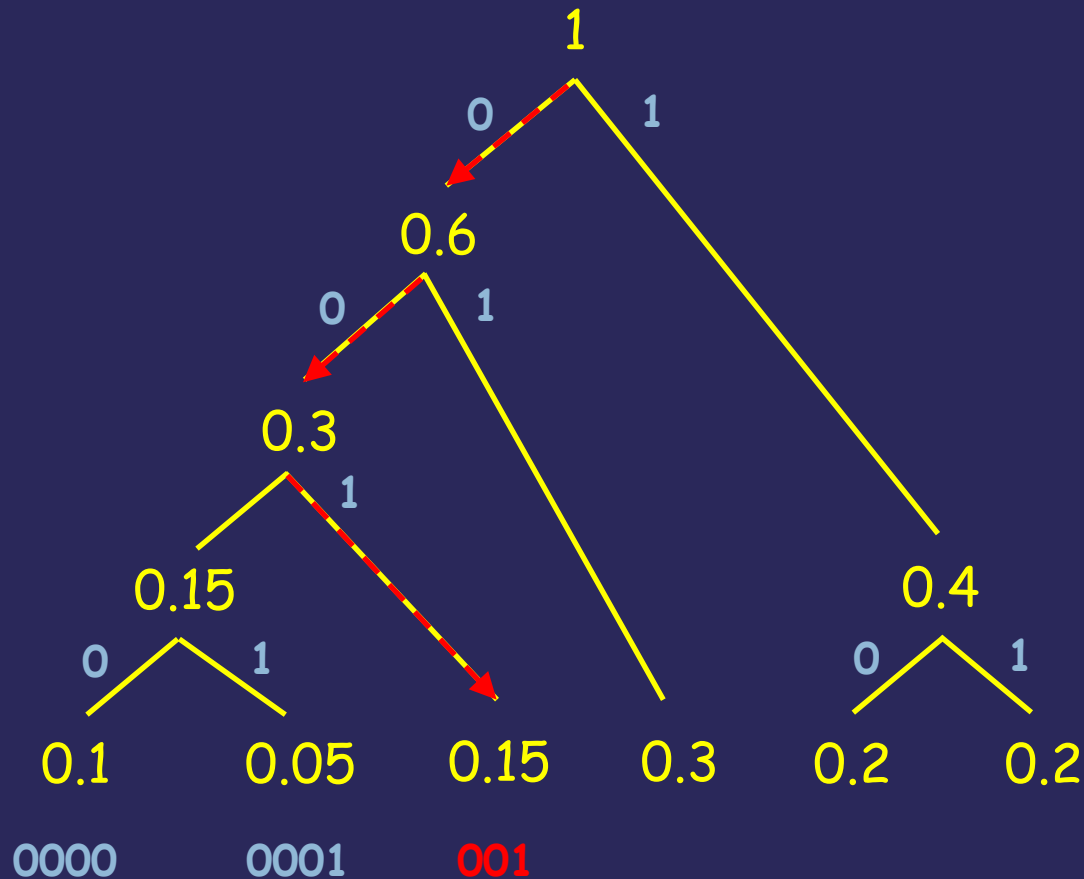
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



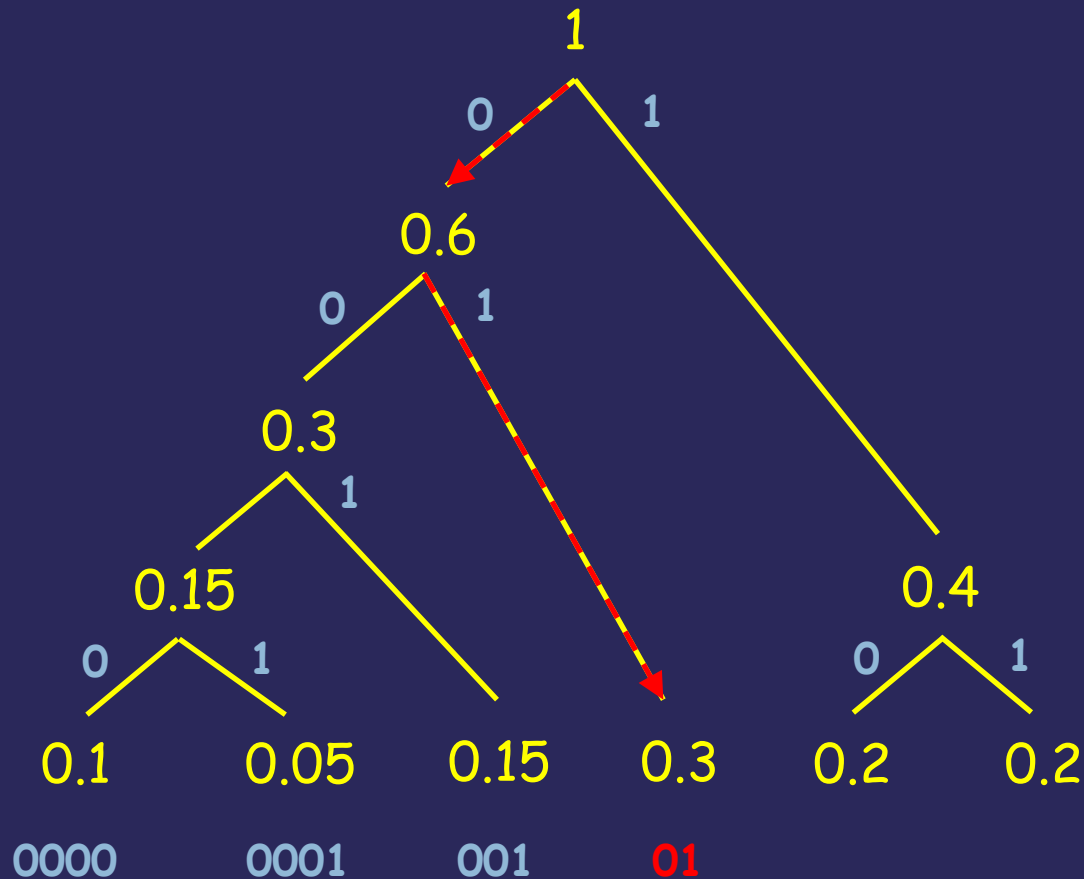
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



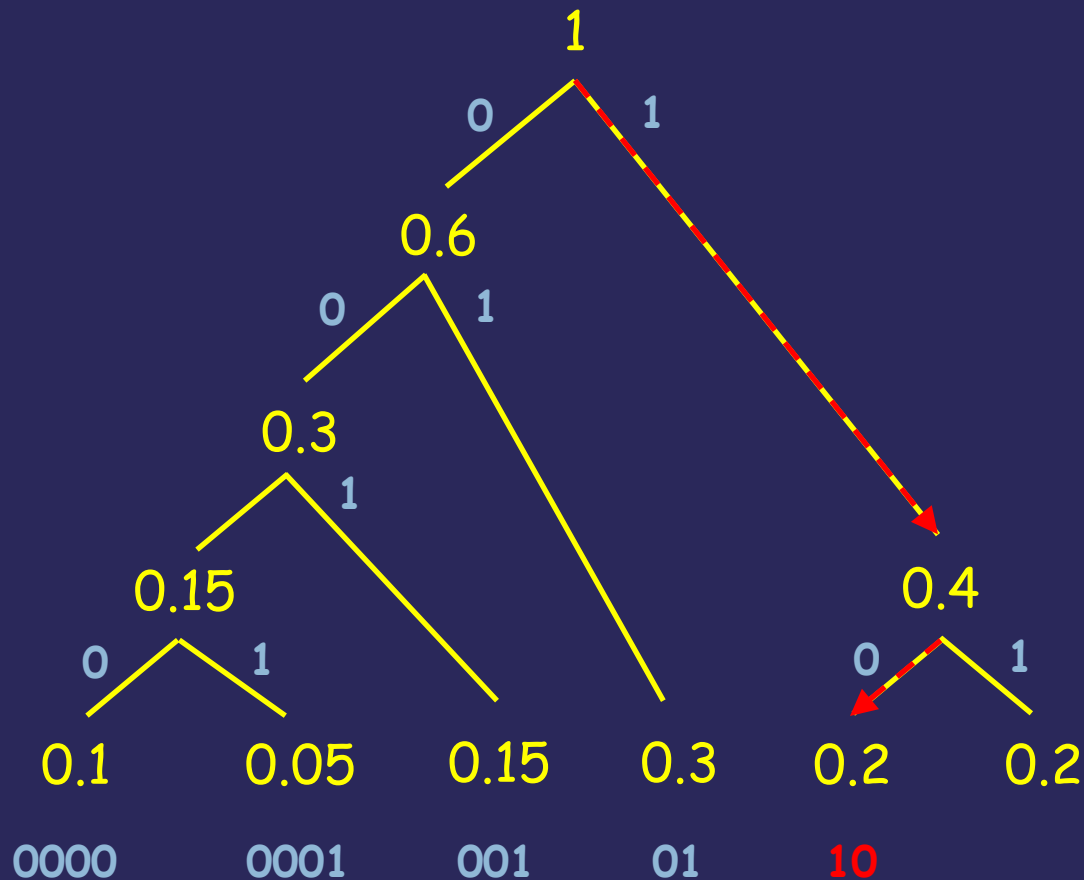
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



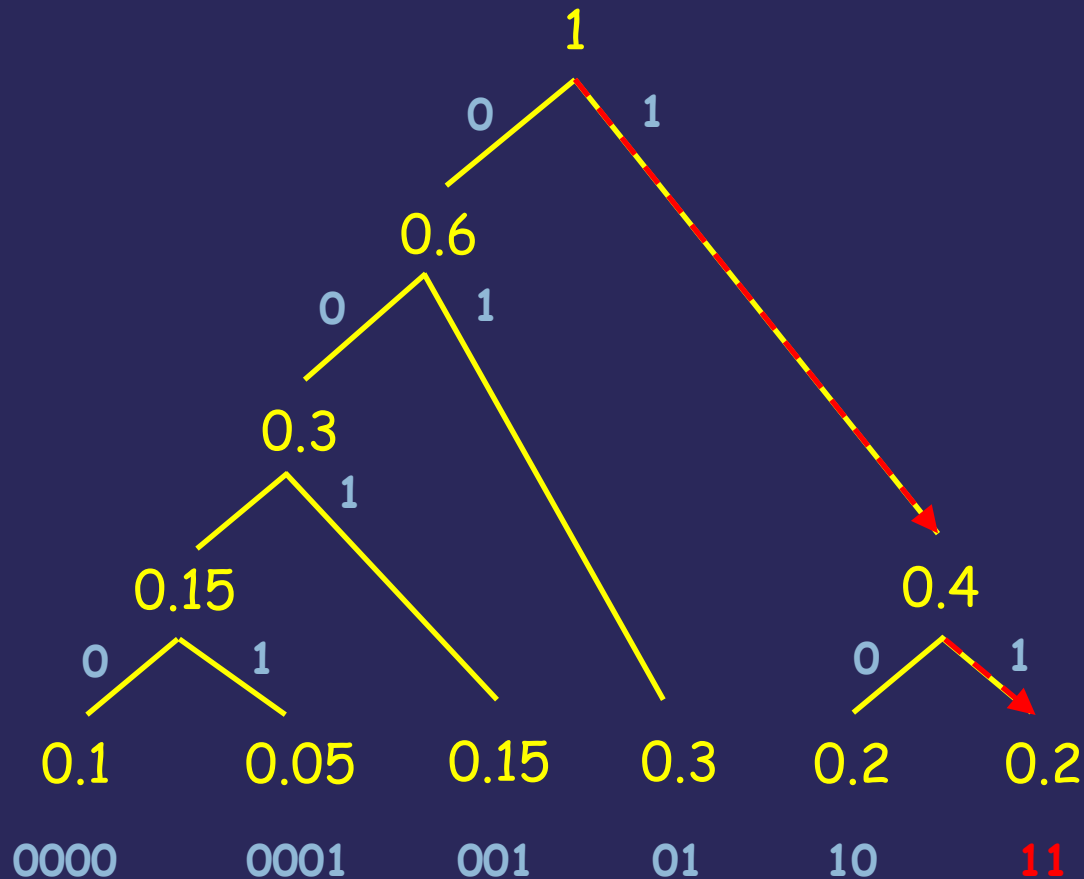
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



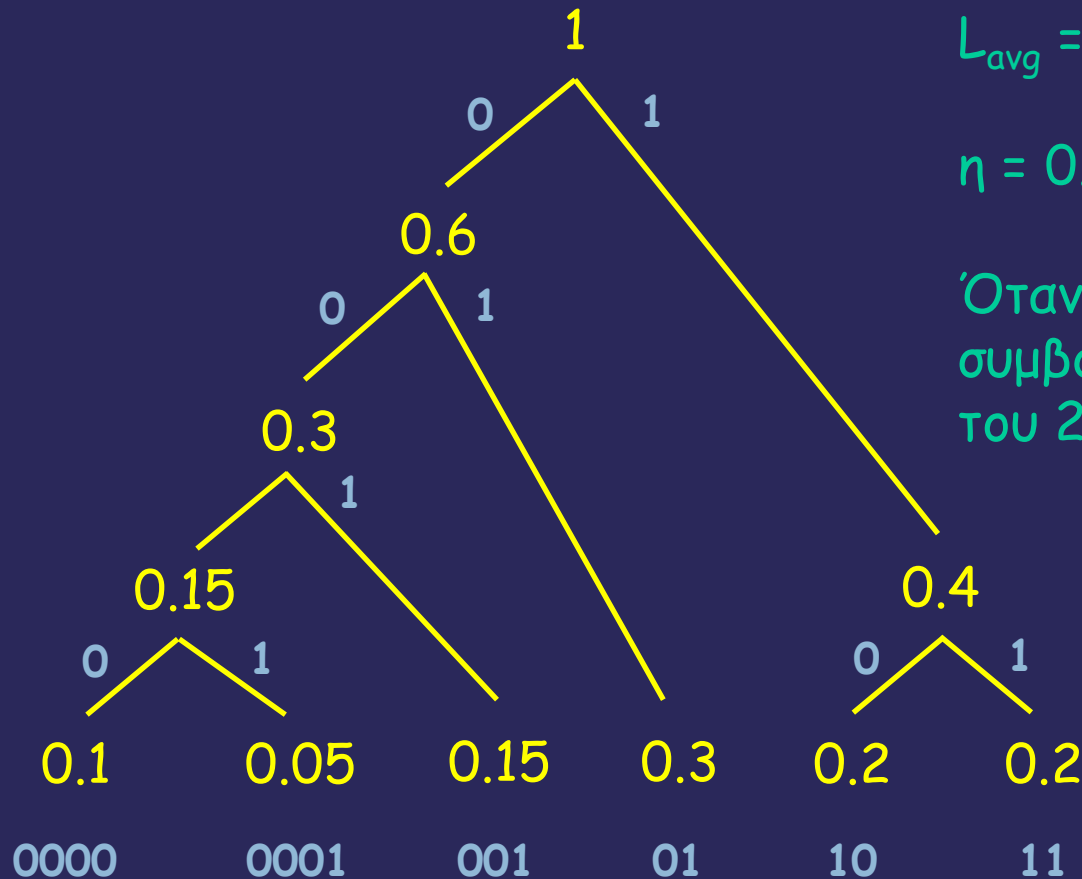
Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Huffman



$$H = 2.40869$$

$$L_{avg} = 2.45$$

$$\eta = 0.98314$$

Όταν οι πιθανότητες των συμβόλων είναι δυνάμεις του 2, $\eta = 1$ ($L_{avg} = H$).

Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση LZW (Lempel-Ziv-Welch)

Δεν απαιτεί εκ των προτέρων γνώση της πιθανότητας εμφάνισης των συμβόλων που κωδικοποιεί.

Αρχικά, κατασκευάζεται ένα codebook (dictionary) το οποίο περιέχει τα σύμβολα που πρόκειται να κωδικοποιηθούν (π.χ. για 8-bit εικόνες, οι πρώτες 256 λέξεις του codebook είναι οι τιμές της έντασης της εικόνας, 0, 1, 2, ..., 255).

Στη συνέχεια, καθώς ο κωδικοποιητής εξετάζει σειριακά τα pixels της εικόνας, τοποθετούνται στο dictionary ακολουθίες από τιμές έντασης.

Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση LZW (Lempel-Ziv-Welch)

Π.χ. αν τα δύο πρώτα pixels μιας εικόνας είναι άσπρα, στη θέση 256 του dictionary θα μπει το «256-256». Την επόμενη φορά που θα συναντήσουμε 2 λευκά pixels, θα χρησιμοποιήσουμε την κωδική λέξη 256.

Σημαντική παράμετρος: μέγεθος dictionary

Μικρό -> μικρή πιθανότητα εύρεσης ακολουθιών από τιμές έντασης

Μεγάλο -> το μέγεθος των κωδικών λέξεων θα έχει επίπτωση στην απόδοση της συμπίεσης

Lossless αλγόριθμοι συμπίεσης

Κωδικοποίηση Run-Length

Run-length pair: καθορίζει την αρχή μιας νέας τιμής έντασης και τον αριθμό των διαδοχικών pixels που έχουν αυτή την τιμή

Όταν δεν υπάρχουν ίδια pixels (spatial redundancy), η run-length κωδικοποίηση έχει σαν αποτέλεσμα αύξηση των δεδομένων.

Καλή απόδοση για δυαδικές εικόνες.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding

Η εικόνα διαιρείται σε μικρά μη επικαλυπτόμενα blocks ίσου μεγέθους (π.χ. 8×8).

Η διαδικασία της συμπίεσης γίνεται σε κάθε block ξεχωριστά με τη χρήση κάποιου 2D μετασχηματισμού.

Ένας αντιστρέψιμος, γραμμικός μετασχηματισμός (όπως ο DFT) χρησιμοποιείται για να αντιστοιχίσει κάθε υποεικόνα σε ένα σύνολο συντελεστών του μετασχηματισμού, οι οποίοι κβαντίζονται και κωδικοποιούνται.

Για τις περισσότερες εικόνες, οι περισσότεροι συντελεστές έχουν μικρές τιμές και μπορούν να κβαντιστούν ή και να αφαιρεθούν χωρίς να προκαλέσουν ιδιαίτερο πρόβλημα στην ποιότητα της εικόνας.

Lossy αλγόριθμοι συμπίεσης

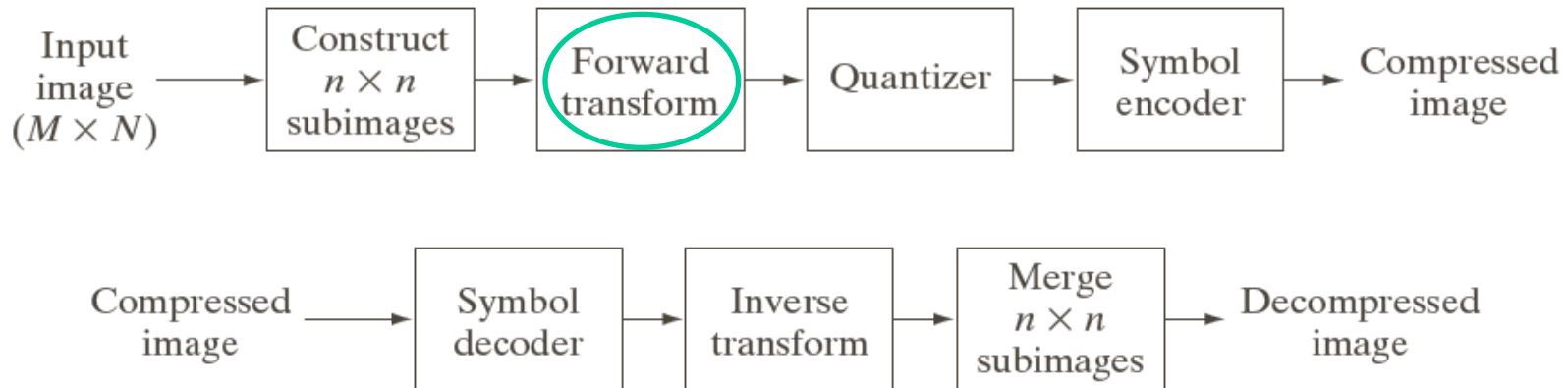
Block Transform Coding

a
b

FIGURE 8.21

A block transform coding system:

(a) encoder;
(b) decoder.



Αποσυσχέτιση των pixels της κάθε υποεικόνας

Συσσώρευση όσο το δυνατό περισσότερης πληροφορίας σε λίγους συντελεστές του μετασχηματισμού

Lossy αλγόριθμοι συμπίεσης

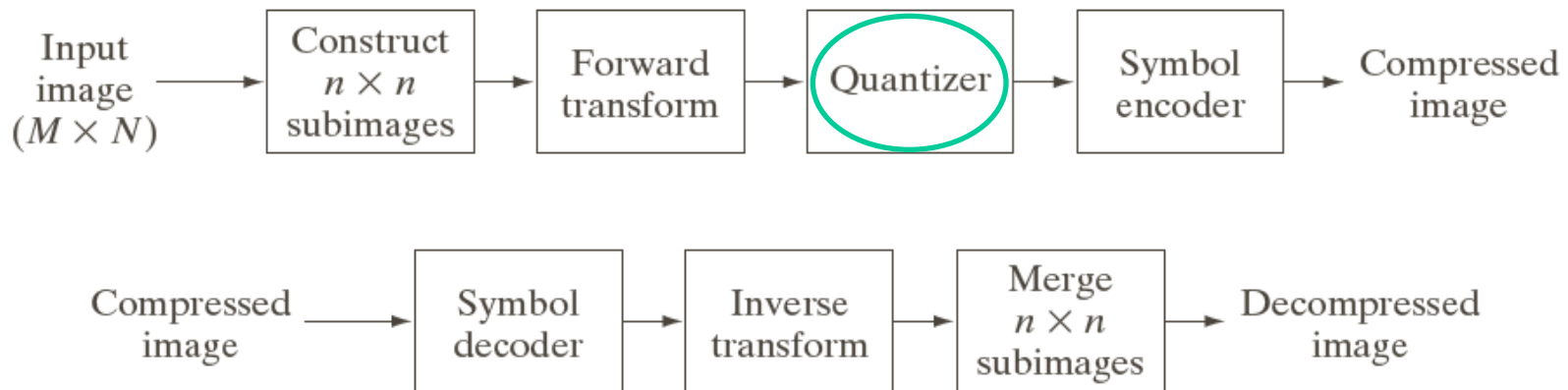
Block Transform Coding

a
b

FIGURE 8.21

A block transform coding system:

(a) encoder;
(b) decoder.



Επιλεκτική διαγραφή ή κβαντισμός των συντελεστών που φέρουν τη λιγότερη πληροφορία

Οι συντελεστές αυτοί έχουν τη μικρότερη επίδραση στην ποιότητα της ανακατασκευασμένης εικόνας.

Lossy αλγόριθμοι συμπίεσης

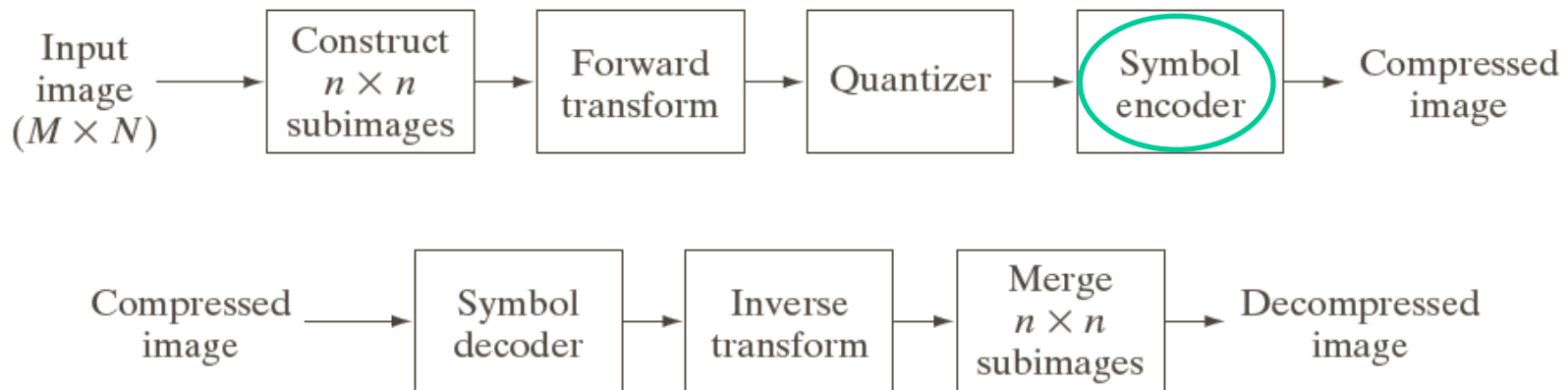
Block Transform Coding

a
b

FIGURE 8.21

A block transform coding system:

(a) encoder;
(b) decoder.



Κωδικοποίηση κβαντισμένων συντελεστών, συνήθως με κώδικες μεταβλητού μήκους

Κάθε ένα ή όλα τα βήματα της κωδικοποίησης μπορεί να προσαρμόζονται στα τοπικά χαρακτηριστικά της εικόνας ή όχι.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

Εξαρτάται:

- από το σφάλμα ανακατασκευής που μπορεί να γίνει ανεκτό από κάποια εφαρμογή
- από τους διαθέσιμους υπολογιστικούς πόρους

Η συμπίεση επιτυγχάνεται με τον κβαντισμό των συντελεστών του μετασχηματισμού (και όχι με το μετασχηματισμό της εικόνας).

Δυνατοί μετασχηματισμοί: FFT, WHT (Walsh-Hadamard Transform), DCT (Discrete Cosine Transform) -> JPEG, KLT (Karhunen-Loève Transform), Wavelets -> JPEG2000.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

Κρατάμε το 50% των συντελεστών με βάση το μέγεθός τους

Σφάλμα στις τιμές της έντασης

a	b	c
d	e	f

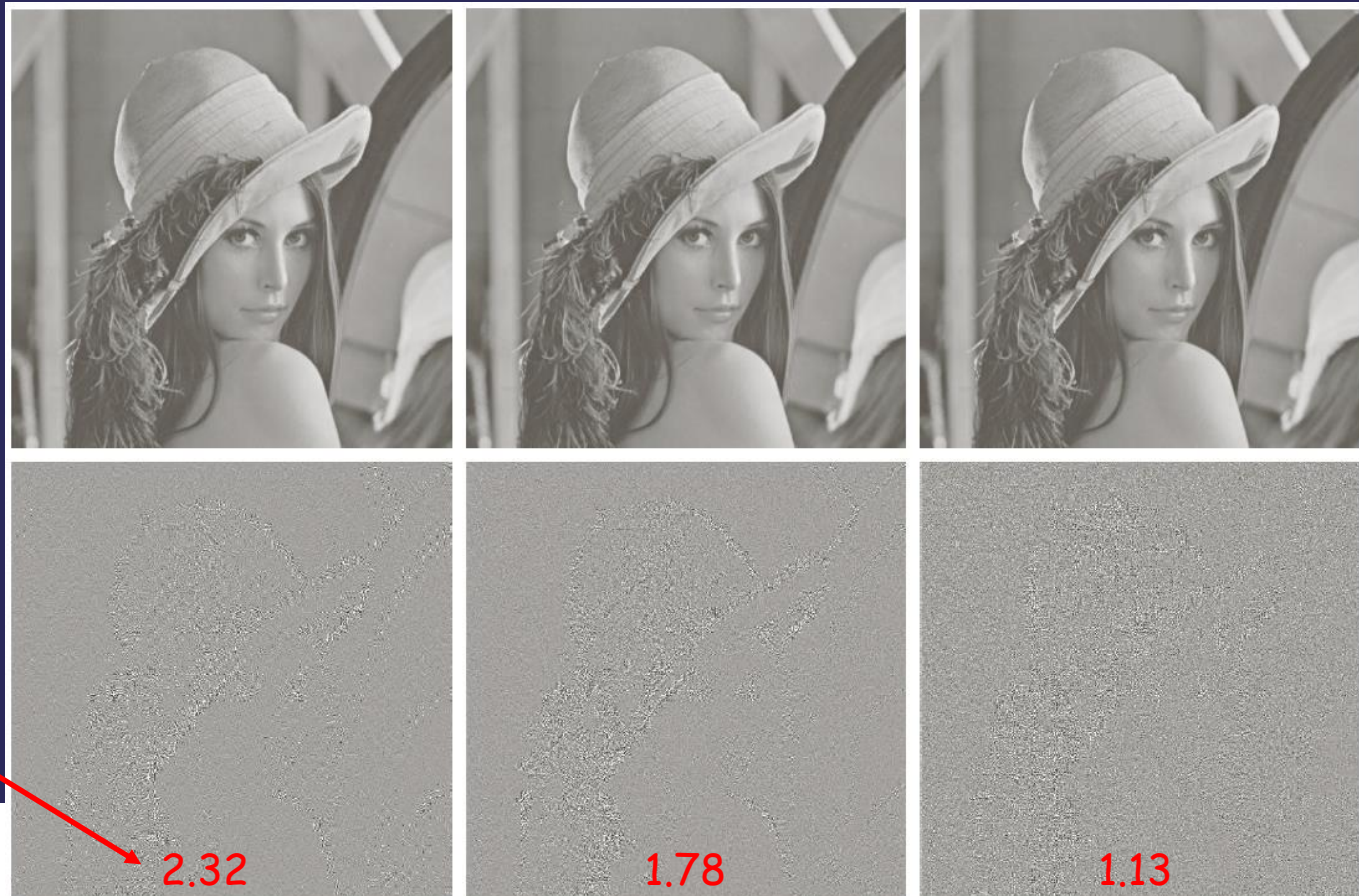


FIGURE 8.24 Approximations of Fig. 8.9(a) using the (a) Fourier, (b) Walsh-Hadamard, and (c) cosine transforms, together with the corresponding scaled error images in (d)–(f).

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

Μικρό σφάλμα (DCT) -> μεγάλη συγκέντρωση ενέργειας σε λίγους συντελεστές του μετασχηματισμού

KLT: ο βέλτιστος μετασχηματισμός με αυτή την ιδιότητα (ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα ανάμεσα στην αρχική και την αποσυμπιεσμένη εικόνα), μεγάλη υπολογιστική πολυπλοκότητα

Επιλέγουμε το DCT (συμπίεση κατά JPEG)

DCT (Discrete Cosine Transform): ένα διακριτό σήμα μπορεί να προσεγγιστεί με το άθροισμα συνημιτονοειδών σημάτων με διαφορετικά πλάτη

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

Ορισμός DCT

$$F(p, q) = a(p)a(q) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cos \left[\frac{(2m+1)\pi p}{2M} \right] \cos \left[\frac{(2n+1)\pi q}{2N} \right]$$

$$0 \leq p \leq M-1$$

$$0 \leq q \leq N-1$$

όπου:

$$a(p) = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases}$$

$$a(q) = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases}$$

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

DCT

Μετασχηματισμός από το πεδίο του χρόνου (ή του χώρου) στο πεδίο των συχνοτήτων

Η πληροφορία των εικονοστοιχείων της εικόνας μετασχηματίζεται σε πληροφορία που αφορά τις συχνότητες μεταβολής της εικόνας.

Συντελεστές DCT: δείχνουν τη σχετική βαρύτητα που έχει η κάθε συχνότητα μεταβολής της εικόνας στη σύνθεσή της

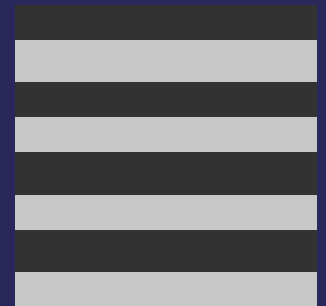
Συχνότητα: γρήγορες ή αργές μεταβολές στην πληροφορία των pixels της εικόνας σε διάφορες περιοχές της

Lossy αλγόριθμοι συμπίεσης

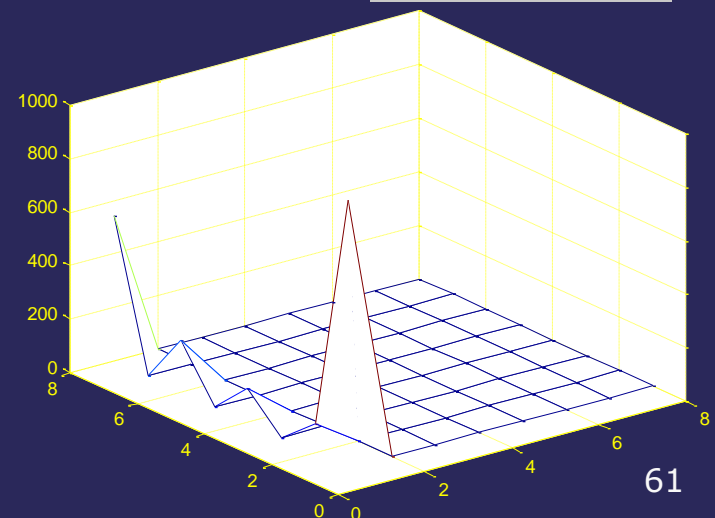
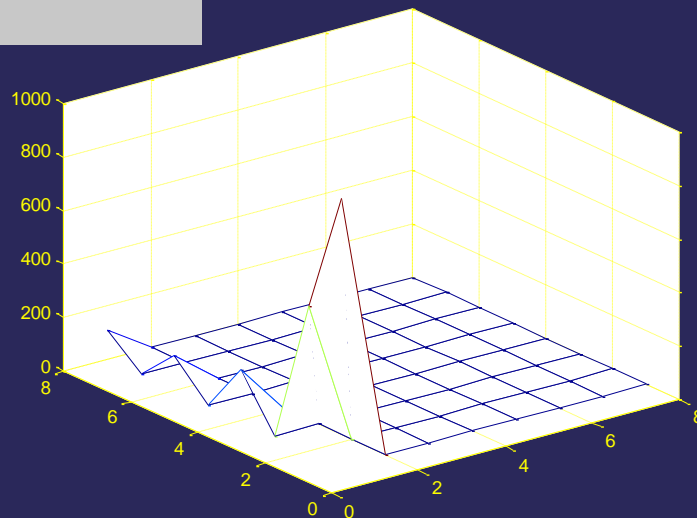
Block Transform Coding - Επιλογή μετασχηματισμού

DCT - συχνότητες μεταβολής της εικόνας

Όσο πιο γρήγορα μεταβάλλεται η ένταση (φωτεινότητα) της εικόνας, τόσο ενισχύονται οι υψηλές συχνότητες



$\text{abs}(\text{dct2}(\text{image}))$

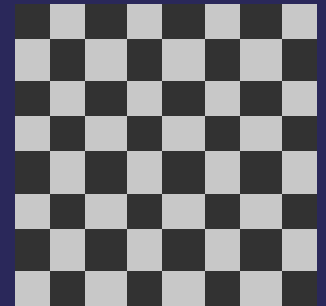
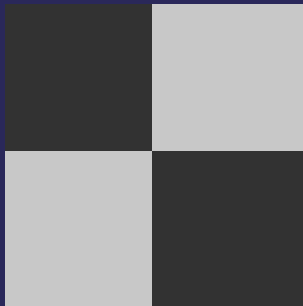


Lossy αλγόριθμοι συμπίεσης

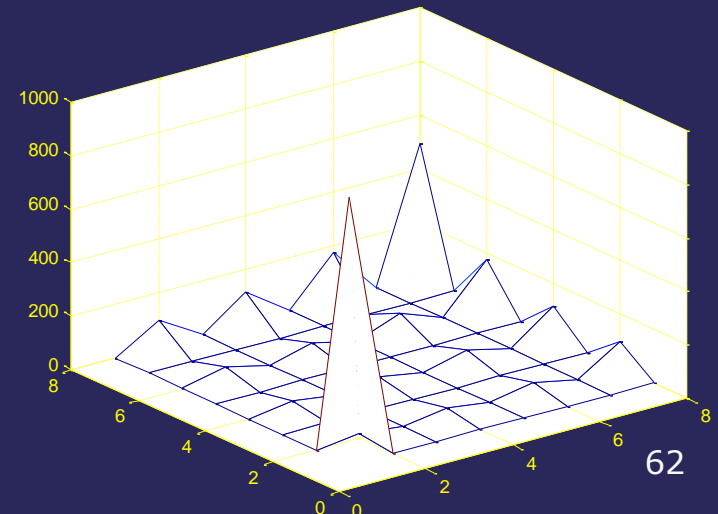
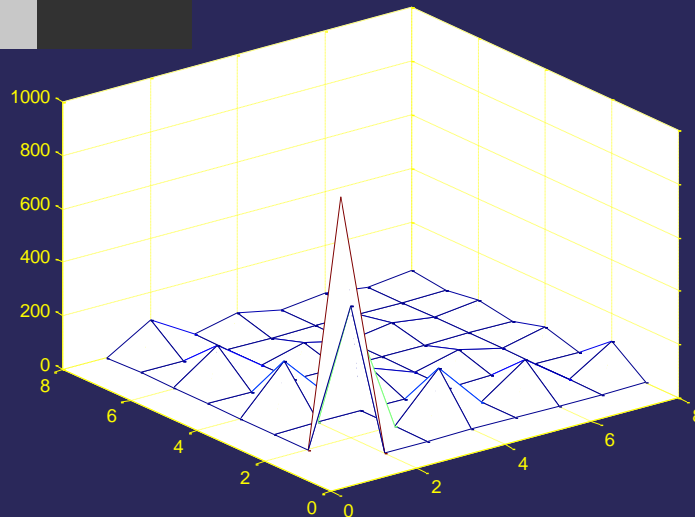
Block Transform Coding - Επιλογή μετασχηματισμού

DCT - συχνότητες μεταβολής της εικόνας

Όσο πιο γρήγορα μεταβάλλεται η ένταση (φωτεινότητα) της εικόνας, τόσο ενισχύονται οι υψηλές συχνότητες



$\text{abs}(\text{dct2}(\text{image}))$

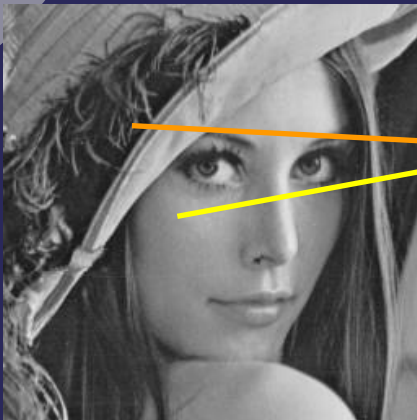


Lossy αλγόριθμοι συμπίεσης

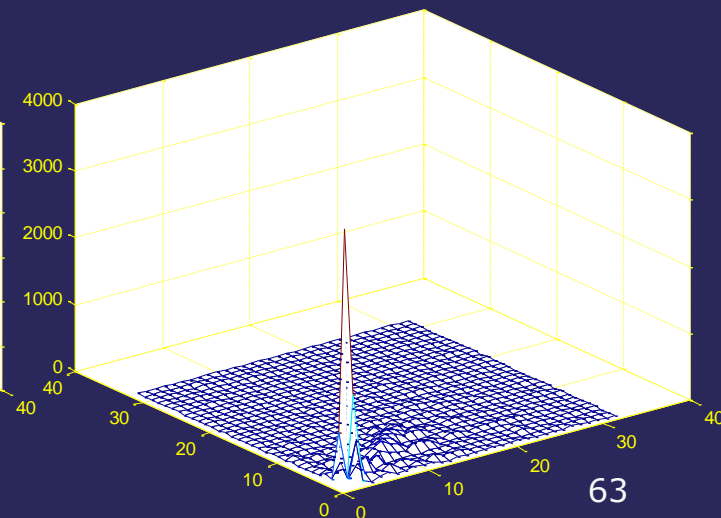
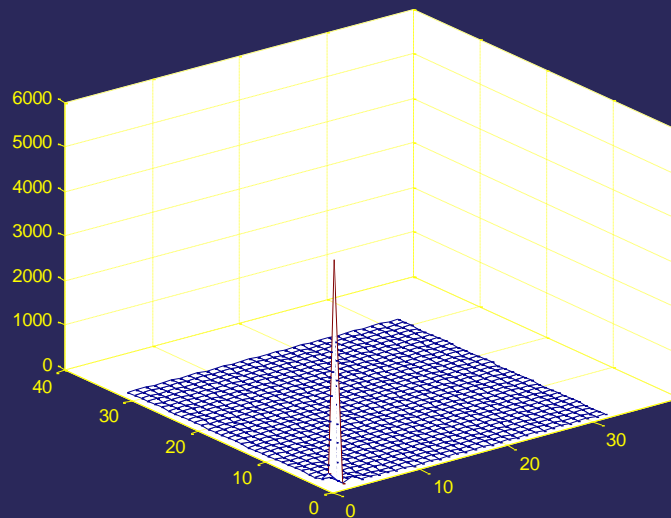
Block Transform Coding - Επιλογή μετασχηματισμού

DCT - συχνότητες μεταβολής της εικόνας

Όσο πιο γρήγορα μεταβάλλεται η ένταση (φωτεινότητα) της εικόνας, τόσο ενισχύονται οι υψηλές συχνότητες



`abs(dct2(image))`



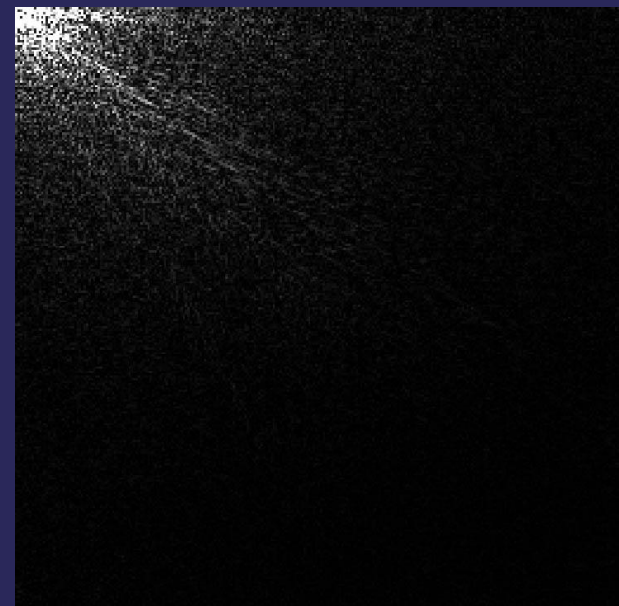
Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μετασχηματισμού

DCT

Συγκέντρωση ενέργειας στις χαμηλές συχνότητες -> η εικόνα δεν παρουσιάζει συχνές απότομες αλλαγές

Υψηλές συχνότητες -> σημαντικές αλλαγές στην πληροφορία του χρώματος μεταξύ γειτονικών pixels



Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μεγέθους υποεικόνας

Η εικόνα υποδιαιρείται με σκοπό τη μείωση της συσχέτισης μεταξύ γειτονικών υποεικόνων.

Συνηθισμένα μεγέθη υποεικόνων (δυνάμεις του 2): 8×8 , 16×16

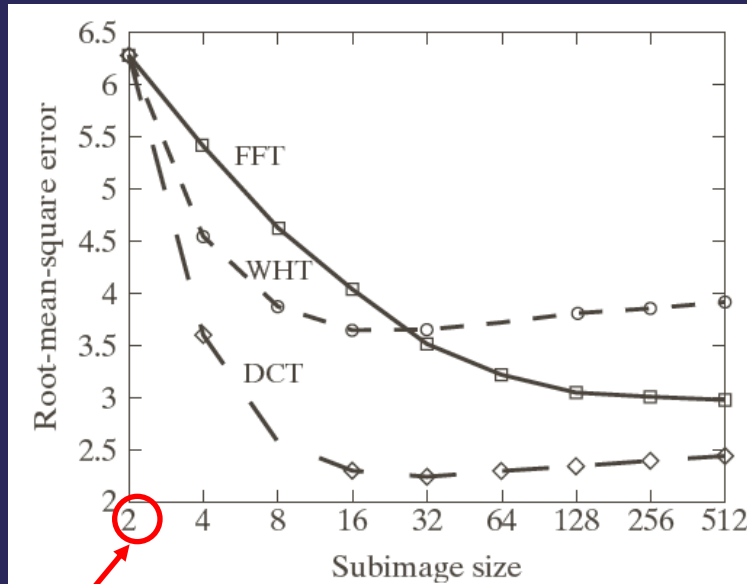


FIGURE 8.26
Reconstruction
error versus
subimage size.

Μόνο ο DC όρος διατηρείται

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Επιλογή μεγέθους υποεικόνας

Η εικόνα υποδιαιρείται με σκοπό τη μείωση της συσχέτισης μεταξύ γειτονικών υποεικόνων.

Συνηθισμένα μεγέθη υποεικόνων (δυνάμεις του 2): 8×8 , 16×16

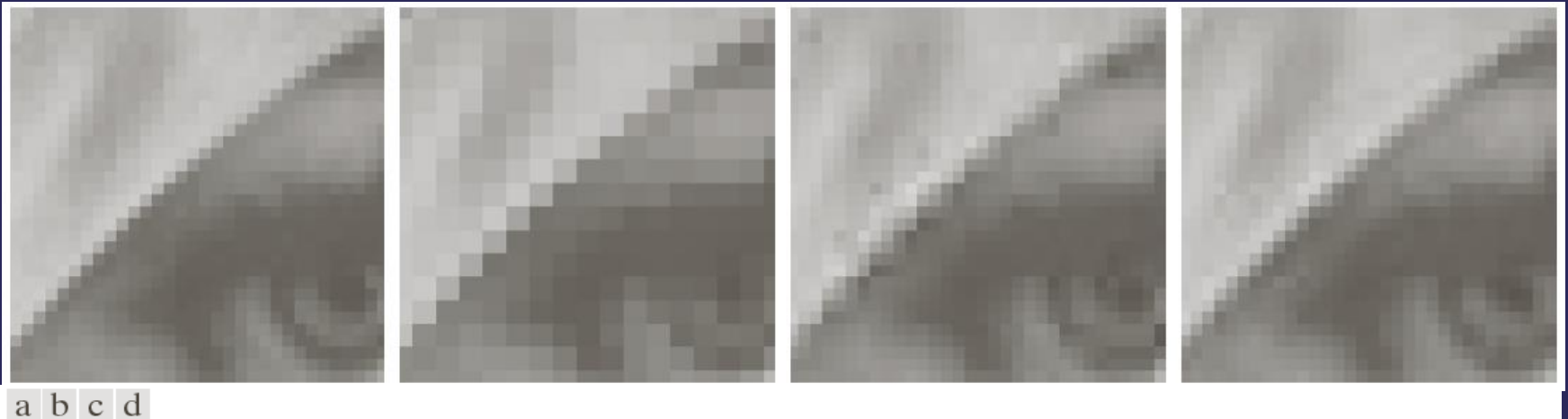


FIGURE 8.27 Approximations of Fig. 8.27(a) using 25% of the DCT coefficients and (b) 2×2 subimages, (c) 4×4 subimages, and (d) 8×8 subimages. The original image in (a) is a zoomed section of Fig. 8.9(a).

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Bit allocation

Το σφάλμα ανακατασκευής της εικόνας εξαρτάται:

- από τον αριθμό και τη σημαντικότητα των συντελεστών που απορρίπτονται
- από την ακρίβεια που χρησιμοποιείται για την αναπαράσταση των συντελεστών που διατηρούνται

Η επιλογή των συντελεστών γίνεται με βάση τη μέγιστη διασπορά (zonal coding) ή τη μέγιστη (απόλυτη) τιμή (threshold coding).

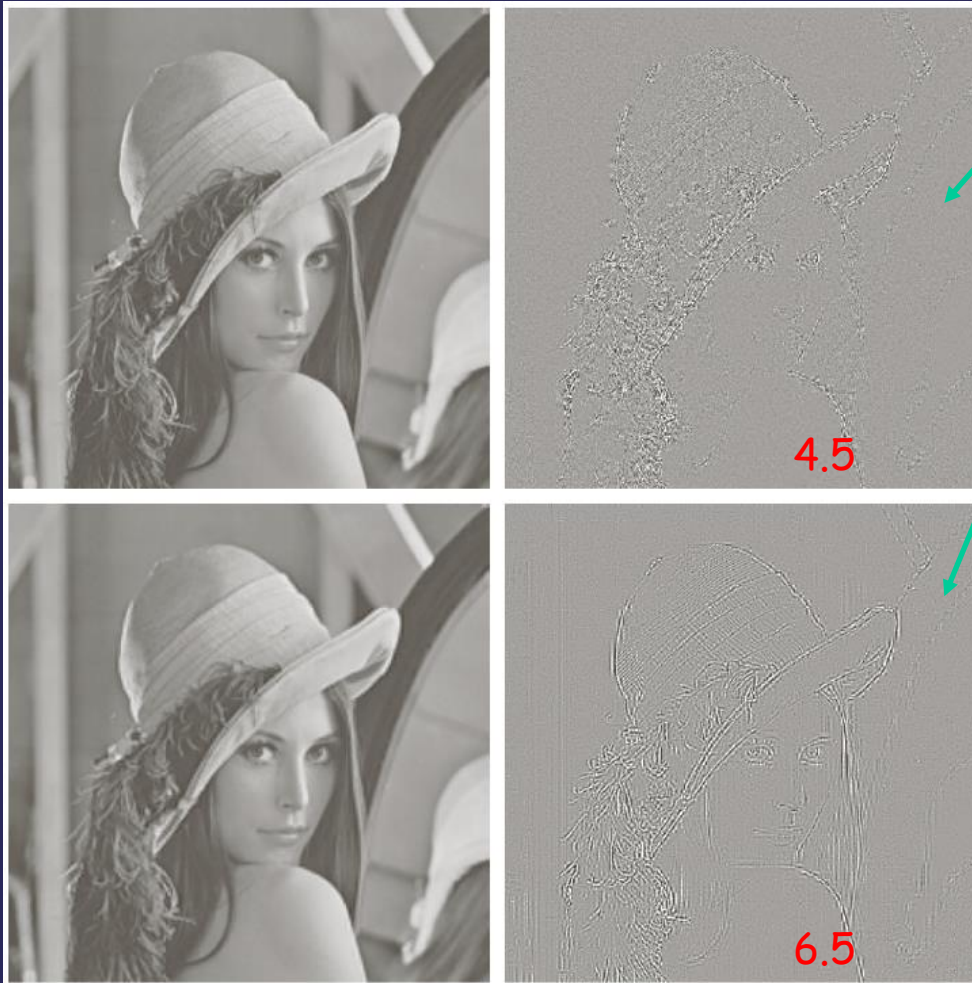
Όλη η διαδικασία του κοψίματος, της κβάντισης και της κωδικοποίησης των συντελεστών λέγεται bit allocation.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Bit allocation

a b
c d

FIGURE 8.28
Approximations
of Fig. 8.9(a) using
12.5% of the
 8×8 DCT
coefficients:
(a)–(b) threshold
coding results;
(c)–(d) zonal
coding results. The
difference images
are scaled by 4.



Διατήρηση των 8
μεγαλύτερων
συντελεστών

Κάθε συντελεστής
θεωρείται τυχαία
μεταβλητή, της οποίας
η κατανομή
υπολογίζεται με βάση
τις τιμές τους σε όλες
τις υποεικόνες.
Οι 8 κατανομές με τη
μεγαλύτερη διασπορά
χρησιμοποιήθηκαν για
την επιλογή των
συντελεστών

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση ζώνης

Βασίζεται στη θεωρία πληροφορίας, η οποία αντιμετωπίζει την πληροφορία σαν αβεβαιότητα.

Έτσι, οι συντελεστές με τη μεγαλύτερη μεταβλητότητα φέρουν την περισσότερη πληροφορία της εικόνας και πρέπει να διατηρηθούν στη διαδικασία κωδικοποίησης.

Οι μεταβλητότητες (διασπορές) μπορούν να υπολογιστούν άμεσα από όλες τις υποεικόνες ή με βάση κάποιο υποθετικό μοντέλο εικόνας.

Σε κάθε περίπτωση, μπορούμε να δούμε την επιλογή συντελεστών με τη μέθοδο ζώνης σαν τον πολλαπλασιασμό των συντελεστών με μία κατάλληλη μάσκα (1 για τους συντελεστές που θέλουμε να διατηρήσουμε, 0 για τους υπόλοιπους συντελεστές).

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση ζώνης

Οι συντελεστές με τη μεγαλύτερη μεταβλητότητα βρίσκονται συνήθως γύρω από τον DC όρο.

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

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση ζώνης

Οι συντελεστές που διατηρούνται πρέπει στη συνέχεια να κβαντιστούν και να κωδικοποιηθούν.

Έτσι, κάποιες φορές οι μάσκες ζώνης απεικονίζονται με τον αριθμό των bits που χρησιμοποιείται για την κωδικοποίηση του κάθε συντελεστή.

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

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση ζώνης

Στις περισσότερες περιπτώσεις, χρησιμοποιείται ο ίδιος αριθμός bits για όλους τους συντελεστές. Οι συντελεστές κανονικοποιούνται από τις τυπικές αποκλίσεις τους και κβαντίζονται ομοιόμορφα.

Σε άλλες περιπτώσεις, κάποιος συγκεκριμένος αριθμός από bits διανέμεται στους συντελεστές με ανομοιόμορφο τρόπο. Για κάθε συντελεστή πρέπει να σχεδιαστεί ένας κβαντιστής.

Για την κατασκευή των απαιτούμενων κβαντιστών, οι DC και μηδενικοί συντελεστές μοντελοποιούνται από Rayleigh κατανομή, ενώ οι λοιποί συντελεστές μοντελοποιούνται από Laplacian ή Gaussian κατανομή.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

Η κωδικοποίηση κατωφλίου είναι κατά κάποιο τρόπο προσαρμοστική. Οι συντελεστές που διατηρούνται διαφέρουν από υποεικόνα σε υποεικόνα.

Η μέθοδος κατωφλίου χρησιμοποιείται πιο συχνά στην πράξη, επειδή είναι πιο απλή υπολογιστικά.

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	0	0	0	0	0
0	0	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

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

Όταν η μάσκα εφαρμοστεί στην υποεικόνα για την οποία έχει παραχθεί, το $n \times n$ array που θα προκύψει μπορεί να αναδιαταχθεί σε ένα διάνυσμα με n^2 στοιχεία, με βάση το παρακάτω σχήμα (zigzag ordering).

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Για τις ακολουθίες από 0 που προκύπτουν μπορεί να γίνει run-length κωδικοποίηση.

Για τους υπόλοιπους συντελεστές μπορούν να χρησιμοποιηθούν κώδικες διαφορετικού μήκους.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

Υπάρχουν τρεις βασικοί τρόποι επιλογής κατωφλίου:

- 1) Ένα κατώφλι για όλες τις υποεικόνες. Ο βαθμός συμπίεσης διαφέρει από υποεικόνα σε υποεικόνα (εξαρτάται από τον αριθμό των συντελεστών που είναι μεγαλύτεροι από το κατώφλι).
- 2) Διαφορετικό κατώφλι για κάθε υποεικόνα (N-largest coding). Διατηρείται ο ίδιος αριθμός συντελεστών για κάθε υποεικόνα.
- 3) Κατώφλι σαν συνάρτηση της θέσης του κάθε συντελεστή σε μια υποεικόνα. Διαφορετικός αριθμός συντελεστών. Η κατωφλίωση και ο κβαντισμός μπορούν να συνδυαστούν.

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

Η κατωφλίωση και ο κβαντισμός μπορούν να συνδυαστούν.

Στη θέση του γινομένου της μάσκας με τους συντελεστές του μετασχηματισμού, μπορούμε να χρησιμοποιήσουμε:

$$\hat{T}(u,v) = \text{round} \left[\frac{T(u,v)}{Z(u,v)} \right]$$

όπου το $Z(u,v)$ είναι ένα στοιχείο του πίνακα κανονικοποίησης του μετασχηματισμού:

$$\mathbf{Z} = \begin{bmatrix} Z(0,0) & Z(0,1) & \dots & Z(0,n-1) \\ Z(1,0) & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ Z(n-1,0) & Z(n-1,1) & \dots & Z(n-1,n-1) \end{bmatrix}$$

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

Τα στοιχεία του $Z(u,v)$ μπορούν να προσαρμοστούν ώστε να επιτευχθούν διάφορα επίπεδα συμπίεσης.

Ένας τυπικός πίνακας κανονικοποίησης, ο οποίος έχει χρησιμοποιηθεί στο JPEG, φαίνεται στη συνέχεια. Τα βάρη του έχουν καθοριστεί με βάση το πόσο αντιληπτός είναι ο κάθε συντελεστής.

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

Lossy αλγόριθμοι συμπίεσης

Block Transform Coding - Κωδικοποίηση κατωφλίου

FIGURE 8.31 Approximations of Fig. 8.9(a) using the DCT and normalization array of Fig. 8.30(b): (a) Z , (b) $2Z$, (c) $4Z$, (d) $8Z$, (e) $16Z$, and (f) $32Z$.

Βαθμοί συμπίεσης: 12, 19, 30, 49, 85 και 182 προς 1

Σφάλματα στις τιμές της έντασης: 3.83, 4.93, 6.62, 9.35, 13.94 και 22.46



Συμπίεση εικόνας - JPEG

Σε κάθε 8×8 block:

- Ολίσθηση τιμών -128 (2^7)
- 2D DCT
- Κανονικοποίηση και κβαντισμός
- Διάταξη συντελεστών (zigzag) σε ένα διάνυσμα
- Κωδικοποίηση Huffman, Run-Length

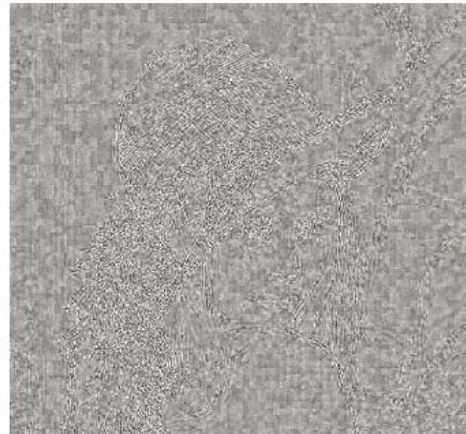
Ανακατασκευή εικόνας με αντίστροφη διαδικασία

Κβαντισμός: μη αναστρέψιμη διαδικασία -> lossy συμπίεση

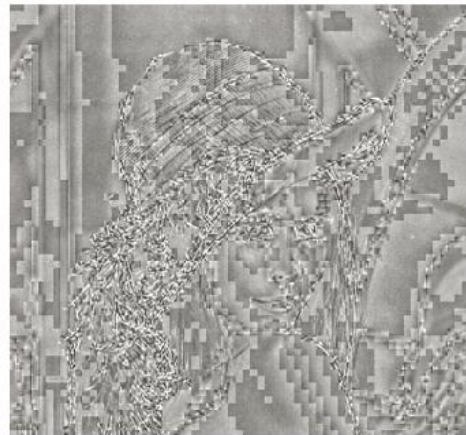
Σφάλμα: διαφορές ανάμεσα στην αρχική εικόνα και την αποσυμπιεσμένη

Συμπίεση εικόνας - JPEG

25:1



50:1



a	b	c
d	e	f

FIGURE 8.32 Two JPEG approximations of Fig. 8.9(a). Each row contains a result after compression and reconstruction, the scaled difference between the result and the original image, and a zoomed portion of the reconstructed image.