

# Συμπύεση Εικόνας

## Διάλεξη 7 - Ψηφιακή Επεξεργασία Εικόνας

TEL750 – ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ

Δρ. Α. Κούτρας, Αναπληρωτής Καθηγητής  
koutras@uop.gr

# Περιεχόμενα

- Εισαγωγή – Συμπύεση
- Θεωρητικό υπόβαθρο
- Κωδικοποίηση Huffman
- Κωδικοποίηση Golomb
- Αριθμητική κωδικοποίηση
- Κωδικοποίηση LZW
- Κωδικοποίηση μήκους διαδρομής
- Κωδικοποίηση επιπέδων bits
- Κωδικοποίηση μετασχηματισμού ομάδων pixel
- Προβλεπτική κωδικοποίηση

# Συμπίεση – Εισαγωγή

- Μια ψηφιακή ταινία τυπικής ευκρίνειας διάρκειας 2 ωρών με 720X480 pixels frame απαιτεί με 30fps:

224GB δεδομένων ή 27 DVD-DL (8.5GB)

- Για να μπορέσει να χωρέσει σε ένα DVD θα πρέπει να συμπιεστεί κατά έναν παράγοντα 26.3

# Χρήσεις συμπίεσης

- Εικόνες ιστοσελίδων και διαδικτύου
- Ψηφιακή φωτογραφική μηχανή
- Τηλεδιάσκεψη, τηλεεπισκόπηση
- Ιατρική απεικόνιση
- Αποστολή και λήψη FAX

Πληροφορία και Πλεονασμός

# Θεωρητικό υπόβαθρο

- *Συμπύεση*: Μείωση της ποσότητας των **δεδομένων** που απαιτούνται για την αναπαράσταση μιας δεδομένης ποσότητας **πληροφορίας**.
- *Δεδομένα*: Το μέσο με το οποίο μεταφέρεται η πληροφορία.
- Μπορούν να χρησιμοποιηθούν πολλές διαφορετικές ποσότητες δεδομένων για την αναπαράσταση της ίδιας πληροφορίας.
- Οι αναπαραστάσεις που περιέχουν άσχετη ή επαναλαμβανόμενη πληροφορία λέγεται πως περιέχουν πλεονάζοντα δεδομένα.

# Παράδειγμα

- Αν συμβολίσουμε με  $b$  και  $b'$  το πλήθος των bits σε δύο αναπαραστάσεις της ίδιας πληροφορίας ο σχετικός πλεονασμός δεδομένων  $R$  της αναπαράστασης με τα  $b$  bits είναι:

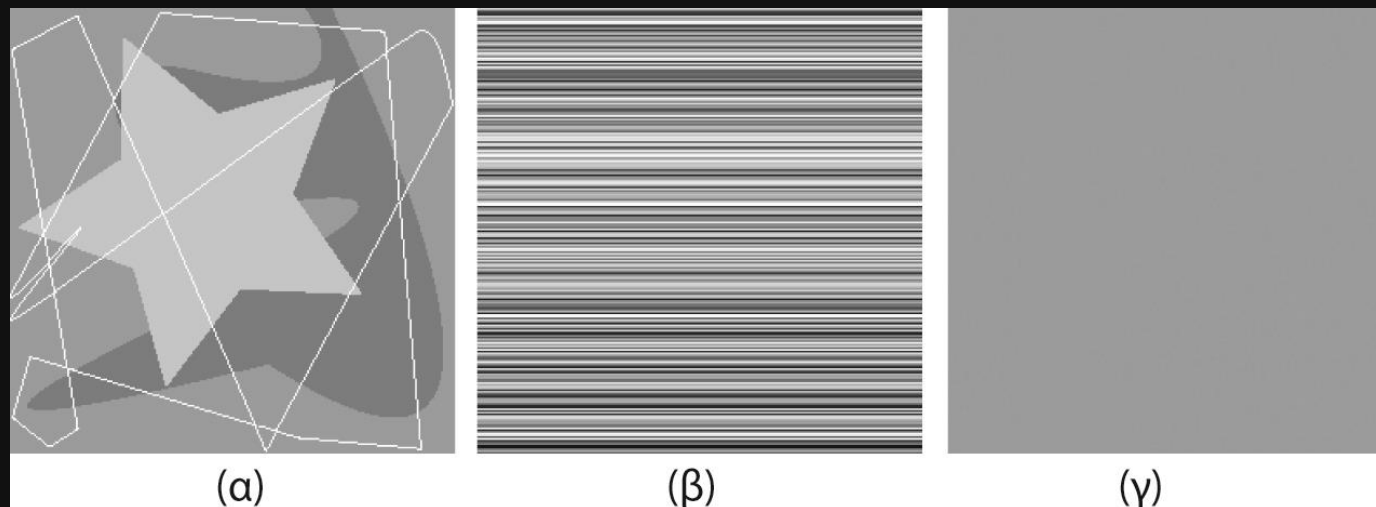
$$R = 1 - \frac{1}{C}$$

όπου  $C$  είναι ο λόγος συμπίεσης

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

# Τύποι πλεονασμού δεδομένων

- Πλεονασμός κωδικοποίησης
- Χρονικός και χωρικός πλεονασμός
- Άσχετη πληροφορία



εικόνες που έχουν δημιουργηθεί από ηλεκτρονικό υπολογιστή και χαρακτηρίζονται από: (α) πλεονασμό κωδικοποίησης (β) χωρικό πλεονασμό (γ) άσχετη πληροφορία



# Πλεονασμός κωδικοποίησης

- Έστω το ιστόγραμμα μιας εικόνας

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

- Εάν το πλήθος των bits για κάθε τιμή του  $r_k$  είναι  $l(r_k)$ , τότε το μέσο πλήθος bits που απαιτούνται για την αναπαράσταση του κάθε pixel είναι

$$L_{\text{μέσο}} = \sum_{k=0}^{L-1} l(r_k) p_r(r_k)$$

# Πλεονασμός κωδικοποίησης

- Προκύπτει όταν οι κώδικες που εκχωρούνται στις τιμές έντασης δεν εκμεταλεύονται στο έπακρο τις πιθανότητες των συμβάντων.
- Εμφανίζεται σχεδόν πάντοτε όταν οι εντάσεις μιας εικόνας αναπαρίστανται μέσω ενός φυσικού κώδικα.

# Παράδειγμα

$r_k$	$p_r(r_k)$	Κώδικας 1	$l_1(r_k)$	Κώδικας 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
$r_k$ για $k \neq 87, 128, 186, 255$	0	—	8	—	0

Παράδειγμα κωδικοποίησης μεταβλητού μήκους

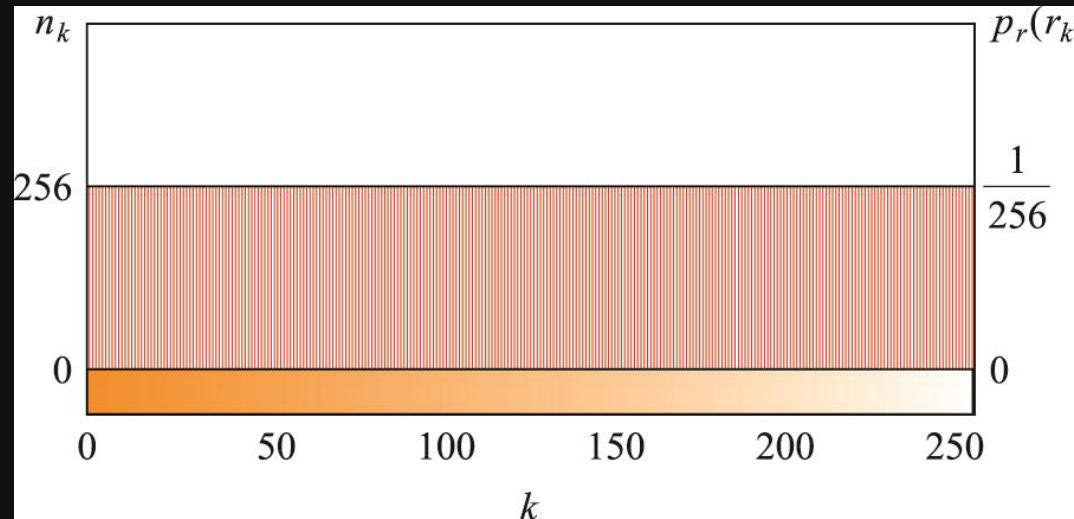
# Χωρικός και χρονικός πλεονασμός

- Όλες οι **256** διαφορετικές τιμές έντασης χαρακτηρίζονται από την ίδια πιθανότητα εμφάνισης (ομοιόμορφο ιστόγραμμα)
- Επειδή η ένταση της κάθε γραμμής έχει επιλεγεί με τυχαίο τρόπο, τα εικονοστοιχεία είναι ανεξάρτητα στην κατακόρυφη διεύθυνση
- Επειδή τα εικονοστοιχεία κατά μήκος της κάθε γραμμής είναι ταυτόσημα, συσχετίζονται μεταξύ τους στον μέγιστο βαθμό στην οριζόντια διεύθυνση.

# Χωρικός και χρονικός πλεονασμός

## □ Παρατήρηση 1

- Αν η εικόνα αναπαρίσταται ως πίνακας εντάσεων 8 bits δεν μπορεί να συμπιεστεί χρησιμοποιώντας μόνο κωδικοποίηση μεταβλητού μήκους. Αρκεί η αναπαράσταση σταθερού μήκους 8 bits



Το ιστόγραμμα της δοκιμαστικής εικόνας ( $\beta$ )

# Χωρικός και χρονικός πλεονασμός

## □ Παρατήρηση 2 & 3

- Υπάρχει σημαντικός χωρικός πλεονασμός ο οποίος μπορεί να απαλειφθεί αν αναπαρασταθεί η εικόνα ως μια ακολουθία από ζεύγη μήκους διαδρομής
- Το κάθε ζεύγος προσδιορίζει την αρχή μιας νέας τιμής έντασης και το πλήθος των διαδοχικών εικονοστοιχείων που χαρακτηρίζονται από αυτή την τιμή έντασης.
- Η αναπαράσταση της εικόνας συμπιέζει τον αρχικό διδιάστατο πίνακα εντάσεων 8 bits κατά έναν λόγο της τάξεως του  $(256 \times 256 \times 8) / [(256 + 256) \times 8]$  ή ισοδύναμα **128: 1**

# Χωρικός και χρονικός πλεονασμός

- Αντί να κωδικοποιηθούν οι εντάσεις των pixels που δεν έχουν μεγάλη πληροφορία, ελατώνουμε τον πλεονασμό που παρατηρείται με μετασχηματισμό που δημιουργεί μια πιο αποτελεσματική αλλά όχι παραστατική αναπαράσταση.
- Για παράδειγμα μπορούμε να χρησιμοποιήσουμε μήκη διαδρομών που κωδικοποιούν την διαφορά εντάσεων ανάμεσα σε διαφορετικά εικονοστοιχεία.
- Οι μετασχηματισμοί αυτοί ονομάζονται **απεικονίσεις**.
- **Αντιστρέψιμες**: τα εικονοστοιχεία του αρχικού πίνακα εντάσεων μπορούν να ανακατασκευαστούν χωρίς σφάλματα από το σύνολο των μετασχηματισμένων δεδομένων
- **Μη-αντιστρέψιμες**: το αντίθετο

# Άσχετη πληροφορία

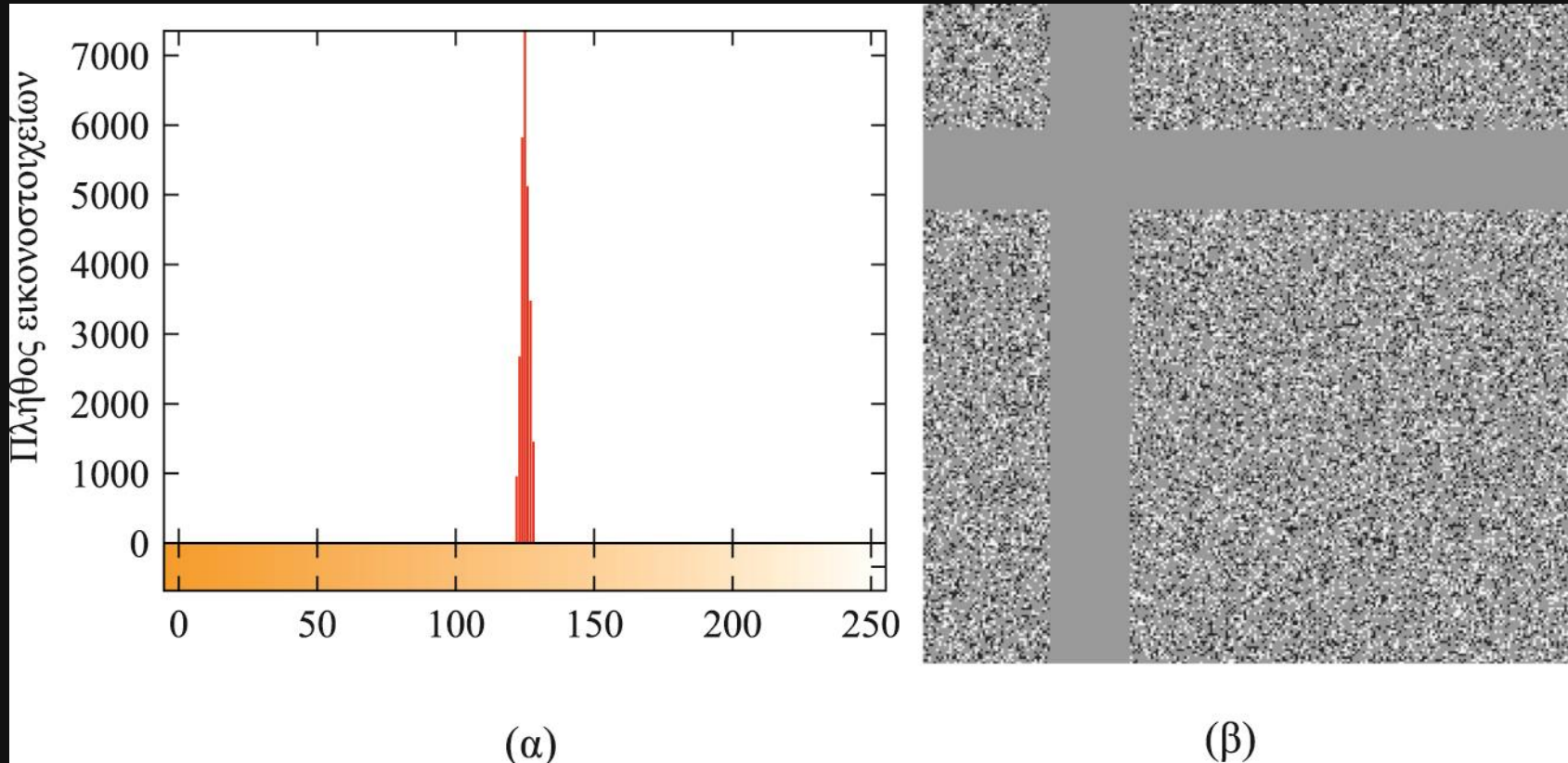
- Συμπίεση μέσω απομάκρυνσης περιττών δεδομένων.
- Αντικατάσταση με μέση φωτεινότητα δηλαδή μια απλή τιμή των 8 bits.
- Οι πληροφορίες που αγνοούνται από το ανθρώπινο οπτικό σύστημα ή είναι άσχετες με τον επιδιωκόμενο τρόπο χρήσης, είναι υποψήφιος για απομάκρυνση.



# Άσχετη πληροφορία

- Η τρίτη εικόνα του παραδείγματος εμφανίζεται ως μια ομογενής έκταση γκρι χρώματος
- Για την αναπαράσταση της θα μπορούσαμε να χρησιμοποιήσουμε μια μέση τιμή έντασης, ή μια απλή τιμή των 8 bits.
- Ο αρχικός πίνακας τιμών ανάγεται σε 1 byte και η συμπίεση που προκύπτει χαρακτηρίζεται από λόγο με τιμή  $(256 \times 256 \times 8)/8 = 65536:1$

# Άσχετη πληροφορία



(α) Το ιστόγραμμα της προηγούμενης δοκιμαστικής εικόνας (γ) (β) μια έκδοση εικόνας που έχει προκύψει έπειτα από εξίσωση ιστογράμματος

# Άσχετη πληροφορία

- Από το προηγούμενο ιστόγραμμα παρατηρούμε ότι υπάρχουν φωτεινότητες μεταξύ **125** και **131**.
- Το ανθρώπινο μάτι τις αγνοεί και αντιλαμβάνεται μόνο την μέση φωτεινότητα
- Η εικόνα μετά την ισοστάθμιση ιστογράμματος δείχνει αυτές τις μεταβολές έντασης και ταυτόχρονα αποκαλύπτει δύο περιοχές σταθερής έντασης της εικόνας που δεν ήταν αντιληπτές πριν.
- Αν συμπίεσουμε την εικόνα αντικαθιστώντας όλες τις φωτεινότητες με μια μέση τιμή φωτεινότητας, οι δύο αυτές περιοχές θα χαθούν.
- Μας ενδιαφέρει; Εξαρτάται από την εφαρμογή (π.χ. σε μια ιατρική εφαρμογή μας ενδιαφέρει)

# Άσχετη πληροφορία

- Ο τελευταίος τύπος πλεονασμού είναι διαφορετικός από τους προηγούμενους και η αφαίρεση του είναι δυνατή γιατί η πληροφορία δεν είναι ουσιαστική για την επεξεργασία της εικόνας.
- Επειδή προκαλεί απώλεια ποσοτικής πληροφορίας, ονομάζεται κβαντισμός
- Επειδή παρουσιάζεται απώλεια πληροφορίας, οι επιδράσεις του κβαντισμού είναι μη αντιστρέψιμες.

# Καταμετρώντας την πληροφορία μιας εικόνας

- Με ποιον τρόπο είναι δυνατό, ένα μικρό πλήθος δυαδικών στοιχείων να αναπαραστήσουν την πληροφορία που περιλαμβάνεται σε μια εικόνα;
- Υπάρχει μια ελάχιστη ποσότητα δεδομένων η οποία είναι αρκετή για να περιγράψει μια εικόνα χωρίς την απώλεια πληροφορίας;

# Καταμετρώντας την πληροφορία μιας εικόνας

## □ Θεωρία της πληροφορίας

- Η δημιουργία της πληροφορίας μπορεί να μοντελοποιηθεί ως μια πιθανοτική διαδικασία που μπορεί να μετρηθεί με ένα τρόπο ο οποίος συμφωνεί με την διαίσθηση μας.

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

- Ένα τυχαίο συμβάν  $E$  με πιθανότητα  $P(E)$  περιέχει  $l(E)$  μονάδες πληροφορίας.
- Αν η πιθανότητα είναι ίση με  $1$ , τότε η πληροφορία είναι ίση με μηδέν.

# Εντροπία πηγής

- Αν θεωρήσουμε μια πηγή στατιστικώς ανεξάρτητων τυχαίων συμβάντων που ανακτώνται μέσα από ένα διακριτό σύνολο πιθανών συμβάντων  $a_j$  με αντίστοιχες πιθανότητες  $P(a_j)$ , η εντροπία της πηγής είναι

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

- Οι παράμετροι είναι γνωστές ως σύμβολα της πηγής και επειδή είναι στατιστικώς ανεξάρτητες μεταξύ τους, η πηγή είναι μηδενικής μνήμης

# Εντροπία εικόνας

- Αν θεωρήσουμε εικόνα ως έξοδο μιας φανταστικής πηγής έντασης μηδενικής μνήμης χρησιμοποιούμε το ιστόγραμμα της εικόνας για να εκτιμήσουμε τις πιθανότητες των συμβόλων της πηγής:

$$\tilde{H} = - \sum_{k=0}^{L-1} p_r(r_k) \log_2 p_r(r_k)$$

- Είναι η ελάχιστη τιμή πληροφορίας ανά εικονοστοιχείο της εικόνας σε bits.



# Παράδειγμα

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

*Να βρεθεί η εντροπία των εικόνων (α), (β), (γ)*

# Κριτήρια πιστότητας

- Τρόποι εκτίμησης της φύσης της απώλειας της πληροφορίας που έχει χαθεί:
- Υπάρχουν δύο τύποι κριτηρίων:
  - ▣ Αντικειμενικά κριτήρια πιστότητας
  - ▣ Υποκειμενικά κριτήρια πιστότητας

# Αντικειμενικά κριτήρια πιστότητας

- Η πληροφορία μπορεί να διατυπωθεί ως μια μαθηματική συνάρτηση της εισόδου και της εξόδου κάποιας διαδικασίας συμπίεσης

Σφάλμα ανάμεσα στις εικόνες:  $e(x, y) = \hat{f}(x, y) - f(x, y)$

συνολικό σφάλμα ανάμεσα σε δύο εικόνες:  $\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]$

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

μέσος τετραγωνικός λόγος σήματος προς θόρυβο:

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

# Υποκειμενικά κριτήρια πιστότητας

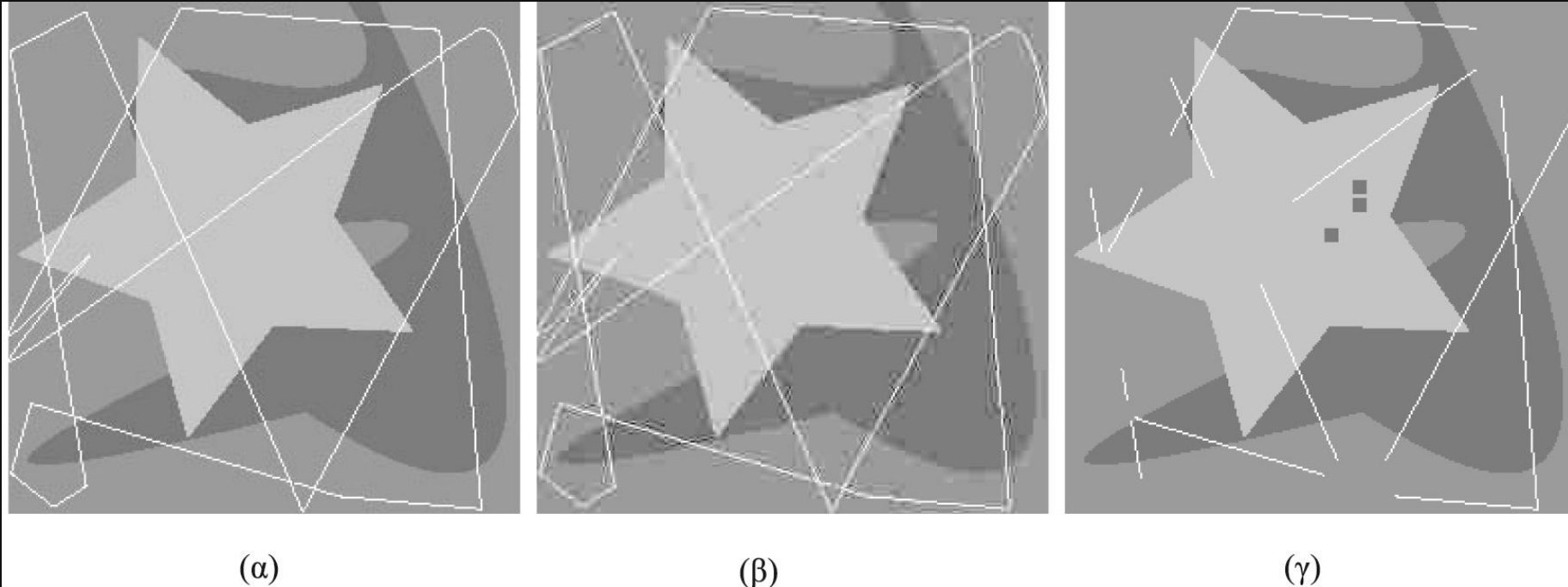
- Η καταμέτρηση της ποιότητας της εικόνας γίνεται με τη βοήθεια υποκειμενικής αξιολόγησης από παρατηρητές
- Παρουσίαση συμπιεσμένων εικόνων σε παρατηρητές και καταγραφή μέσου όρου των συμπερασμάτων τους.
- Η αξιολόγηση γίνεται με χρήση απόλυτης κλίμακας κατάταξης και σύγκρισης των εικόνων τοποθετώντας αυτές τη μία δίπλα στην άλλη.

# Κλίμακα κατάταξης του TASO

Τιμή	Κατάταξη	Περιγραφή
1	Άριστη	Μία εικόνα εξαιρετικά υψηλής ποιότητας, τόσο υψηλής όσο θα θέλατε
2	Καλή	Μία εικόνα υψηλής ποιότητας που παρατηρείται ευχάριστα. Η παρεμβολή δεν είναι απαράδεκτη.
3	Υποφερτή	Μία εικόνα αποδεκτής ποιότητας με παρεμβολή η οποία δεν είναι απαράδεκτη.
4	Οριακή	Μία εικόνα χαμηλής ποιότητας την οποία θα ευχόσαστε να μπορούσατε να βελτιώσετε. Η παρεμβολή είναι απαράδεκτη σε κάποιο βαθμό.
5	Κατώτερη	Μία πολύ φτωχή εικόνα την οποία ωστόσο μπορείτε να καταλάβετε. Η ύπαρξη της παρεμβολής είναι αναμφισβήτητη και μάλιστα σε απαράδεκτο βαθμό.
6	Ακατάλληλη προς χρήση	Η εικόνα είναι τόσο κακή που δεν μπορείτε να καταλάβετε τι απεικονίζει.

Κλίμακα κατάταξης TASO (television allocations study organization)

# Παράδειγμα



ΜΤΣ = 5.17

ΜΤΣ = 15.67

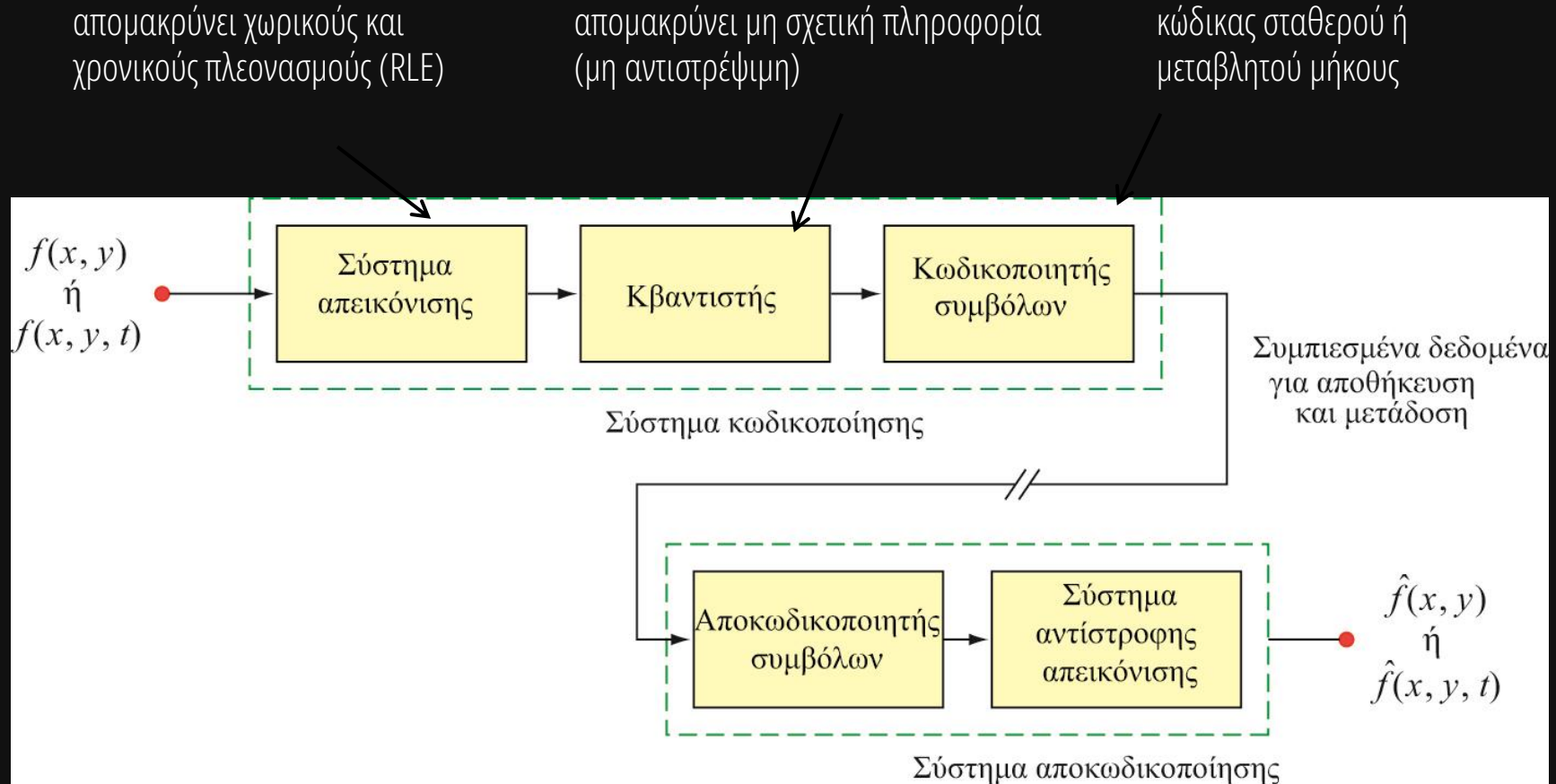
ΜΤΣ = 14.17

Οι τρεις προσεγγίσεις της δοκιμαστικής εικόνας (α). Για κάθε εικόνα αναγράφεται και το Μέσο Τετραγωνικό Σφάλμα (ΜΤΣ).

Η υποκειμενική αξιολόγηση χαρακτήρισε άριστη την (α), οριακή την (β) και κατώτερη ή ακατάλληλη προς χρήση την (γ).

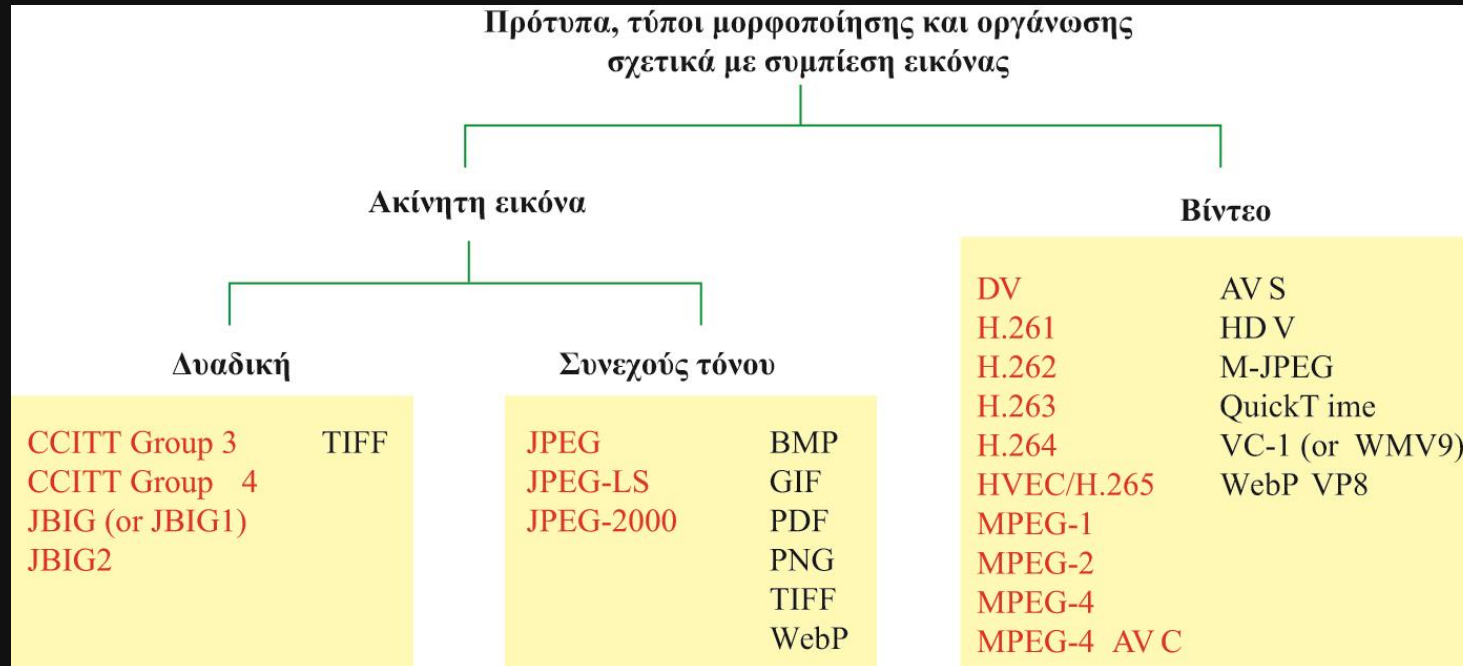
Με υποκειμενικά κριτήρια, η εικόνα (β) θα ταξινομηθεί σε καλύτερη θέση σε σχέση με την (γ).

# Μοντέλα συμπίεσης εικόνας



Σχηματικό διάγραμμα ενός γενικού συστήματος συμπίεσης εικόνας

# Τύποι αρχείων εικόνας



Δημοφιλή πρότυπα συμπίεσης εικόνας μαζί με τύπους αρχείων και υποδοχείς εικόνας.



# Πρότυπα συμπίεσης

Όνομα	Οργανισμός	Περιγραφή
<i>Ασπρόμαυρες ακίνητες εικόνες</i>		
CCITT Group 3	ITU-T	Σχεδιάστηκε ως μέθοδος αποστολής ψηφιακών κειμένων δια μέσου τηλεφωνικών γραμμών με FAX. Υποστηρίζει μονοδιάστατα / διδιάστατα μήκη διαδρομής [8.6] και κωδικοποίηση Huffman [8.2].
CCITT Group 4	ITU-T	Μία απλοποιημένη και βελτιστοποιημένη έκδοση του προτύπου της CCITT Group 3 που υποστηρίζει μόνο διδιάστατη κωδικοποίηση μήκους διαδρομής.
JBIG ή JBIG1	ISO/IEC/ ITU-T	Ένα πρότυπο της ομάδας JBIG για την προοδευτική συμπίεση ασπρόμαυρων εικόνων. Οι εικόνες συνεχούς τόνου χρωματικού βάθους έως 6 bits ανά εικονοστοιχείο κωδικοποιούνται με βάση την έννοια των επιπέδων bits [8.8]. Το πρότυπο στηρίζεται στην αριθμητική κωδικοποίηση που εξαρτάται από το περιβάλλον [8.4] και μια έκδοση της εικόνας που αρχικά είναι χαμηλής ανάλυσης βελτιώνεται σταδιακά με πρόσθετα συμπεσμένα δεδομένα.
JBIG2	ISO/IEC/ ITU-T	Ένα επακόλουθο του προτύπου JBIG1 για ασπρόμαυρες εικόνες στην επιφάνεια εργασίας, το διαδίκτυο και τις εφαρμογές FAX. Η μέθοδος συμπίεσης που χρησιμοποιείται, εξαρτάται από το περιεχόμενο με μεθόδους που στηρίζονται στη χρήση λεξικού [8.7] για ημιτονισμένες περιοχές και περιοχές κειμένου και αριθμητική κωδικοποίηση [8.4] ή κωδικοποίηση Huffman [8.2] για περιεχόμενο εικόνας άλλου τύπου. Η μέθοδος μπορεί να χαρακτηρίζεται ή όχι από απώλειες.
<i>Ακίνητες εικόνες συνεχούς τόνου</i>		
JPEG	ISO/IEC/ ITU-T	Ένα πρότυπο της ομάδας Joint Photographic Expert Group κατάλληλο για εικόνες φωτογραφικής ποιότητας. Το σύστημα κωδικοποίησης που χρησιμοποιεί, οδηγεί σε απώλεια πληροφορίας και (τουλάχιστον στις πιο πολλές υλοποιήσεις του) χρησιμοποιεί κβαντισμένους διακριτούς μετασχηματισμούς συνημιτόνου που εφαρμόζονται σε τμήματα εικόνας [8.9], κωδικοποίηση Huffman [8.2] και κωδικοποίηση μήκους διαδρομής [8.6]. Αυτό το πρότυπο αποτελεί μία από τις πιο δημοφιλείς μεθόδους για τη συμπίεση εικόνων στο παγκόσμιο διαδίκτυο.
JPEG-LS	ISO/IEC/ ITU-T	Ένα μη απωλεστικό / σχεδόν μη απωλεστικό πρότυπο κατάλληλο για εικόνες συνεχούς τόνου, που βασίζεται στην προσαρμοζόμενη πρόβλεψη [8.10], στη μοντελοποίηση περιεχομένου [8.4] και στην κωδικοποίηση Golomb [8.3].
JPEG-2000	ISO/IEC/ ITU-T	Ένα επακόλουθο του προτύπου JPEG για ακόμη μεγαλύτερη συμπίεση των φωτογραφικών εικόνων ποιότητας, που στηρίζεται στη χρήση αριθμητικής κωδικοποίησης [8.4] και κβαντισμένων διακριτών μετασχηματισμών κυματιδίων [8.11]. Η συμπίεση που επιφέρει, μπορεί να είναι απωλεστική ή μη απωλεστική.

Όνομα	Οργανισμός	Περιγραφή
<i>Βίντεο</i>		
DV	IEC	Ψηφιακό βίντεο. Ένα πρότυπο βίντεο, προσαρμοσμένο σε οικιακές και ημιαγγελματικές εφαρμογές παραγωγής βίντεο (καθώς και στον αντίστοιχο εξοπλισμό) όπως είναι η συλλογή ηλεκτρονικών νέων και οι βιντεοκάμερες. Τα πλαίσια συμπίεζονται ανεξάρτητα για απλή επεξεργασία χρησιμοποιώντας μία προσέγγιση που στηρίζεται στο μετασχηματισμό DCT [8.9] παρόμοια με το JPEG.
H.261	ITU-T	Ένα πρότυπο αμφίδρομης βιντεοδιάσκεψης για γραμμές ISDN (Integrated Services Digital Network). Υποστηρίζει όχι πεπλεγμένες οθόνες διαστάσεων 352 x 288 και 176 x 144 οι οποίες είναι γνωστές ως CIF (Common Intermediate Format) και QCIF (Quarter CIF) αντίστοιχα. Σε αυτό το πρότυπο, χρησιμοποιείται μία προσέγγιση συμπίεσης που στηρίζεται στη χρήση του μετασχηματισμού DCT [8.9] παρόμοιας με το JPEG με πρόβλεψη διαφορών από καρέ σε καρέ [8.10] για την ελάττωση του χρονικού πλεονασμού. Προκειμένου να αντισταθμιστεί η κίνηση ανάμεσα στα πλαίσια, χρησιμοποιείται μία τεχνική τύπου block.
H.262	ITU-T	Δείτε το πρότυπο MPEG-2 που ακολουθεί.
H.263	ITU-T	Μία βελτιωμένη έκδοση του προτύπου H.261 που έχει σχεδιαστεί για χρήση με τα συνηθισμένα τηλεφωνικά modems (με ταχύτητες 28.8 Kb/s) με πρόσθετες αναλύσεις: SQCIF (Sub Quarter CIF 128 x 96), 4CIF (704 x 576) και 16 CIF (1408 x 512).
H.264	ITU-T	Μία επέκταση των προτύπων H.261 - H.263 για εφαρμογές βιντεοδιάσκεψης, μεταδόσεων μέσω Internet και τηλεοπτικής εκπομπής. Υποστηρίζει διαφορές πρόβλεψης μέσα στο πλαίσιο [8.10], μετασχηματισμούς ακεραίων μεταβλητού μεγέθους ομάδας (αντί για το μετασχηματισμό DCT) και αριθμητική κωδικοποίηση προσαρμοσμένη στο περιεχόμενο [8.4].
H.265 MPEG-H HEVC	ISO/IEC ITU-T	Κωδικοποίηση Video υψηλής απόδοσης (HVEC). Αποτελεί μία επέκταση του H.264 που περιλαμβάνει υποστήριξη για macroblocks με μεγέθη μέχρι και 64x64 και πρόσθετες δυνατότητες πρόβλεψης για ανεξάρτητα πλαίσια, τα οποία αμφότερα είναι χρήσιμα σε εφαρμογές 4K video.
MPEG-1	ISO/IEC	Ένα πρότυπο της ομάδας Motion Pictures Expert Group για εφαρμογές CD-ROM με συνεχές φάσμα μέχρι τα 1.5 Mb/s. Αυτό το πρότυπο είναι παρόμοιο με το H.261, με τη διαφορά πως οι προβλέψεις των πλαισίων μπορούν να στηριχθούν στο προηγούμενο πλαίσιο, ή σε μία παρεμβολή αυτών των δύο πλαισίων. Υποστηρίζεται από όλους σχεδόν τους υπολογιστές και τις συσκευές αναπαραγωγής DVD.
MPEG-2	ISO/IEC	Μία επέκταση του προτύπου MPEG-1 σχεδιασμένη για DVDs με ρυθμούς μεταφοράς δεδομένων έως και 15 Mb/s. Υποστηρίζει πεπλεγμένο βίντεο και τηλεόραση υψηλής ευκρίνειας και αποτελεί το πιο επιτυχημένο πρότυπο βίντεο των ημερών μας.
MPEG-4	ISO/IEC	Μία επέκταση του προτύπου MPEG-2 που υποστηρίζει μεταβλητά μεγέθη ομάδας και προβλέψεις διαφορών [8.10] ανάμεσα στα πλαίσια.
MPEG-4 AVC	ISO/IEC	MPEG-4 Part 10 Advanced Video Coding (AVC). Είναι ταυτόσημο με το πρότυπο H.264 που περιγράψαμε παραπάνω.

Όνομα	Οργανισμός	Περιγραφή
<i>Ακίνητες εικόνες συνεχούς τόνου</i>		
BMP	Microsoft	Εικόνα Bitmap των Microsoft Windows. Πρόκειται για έναν τύπο αρχείου που χρησιμοποιείται κυρίως για απλές ασυμπίεστες εικόνες.
GIF	CompuServe	Graphic Interchange Format. Πρόκειται για έναν τύπο αρχείου που χρησιμοποιεί τη μη απωλεστική μέθοδο κωδικοποίησης LZW [8.5] η οποία χρησιμοποιείται για εικόνες με χρωματικό βάθος από 1 έως 8 bits. Χρησιμοποιείται συχνά για την κατασκευή μικρών κινούμενων σχεδίων.
PDF	Adobe Systems	Portable Document Format. Ένας τύπος αρχείου για την αναπαράσταση διδιάστατων εγγράφων με έναν τρόπο που να είναι ανεξάρτητος από το σύστημα και την ανάλυση της απεικόνισης. Μπορεί να λειτουργήσει ως αποθηκευτής για εικόνες τύπου JPEG, JPEG2000, CCITT καθώς και πολλών άλλων τύπων συμπίεσμένων εικόνων.
PNG	<i>World Wide Web Consortium (W3C)</i>	Portable Network Graphics. Ένας τύπος αρχείου για τη συμπίεση χωρίς απώλειες εικόνων πλήρους χρώματος με διαφάνεια μέχρι 48 bits / εικονοστοιχείο που στηρίζεται στην κωδικοποίηση της διαφοράς ανάμεσα στην τιμή του κάθε εικονοστοιχείου και σε μία τιμή η οποία να είναι σε θέση να προβλεφθεί με βάση τις τιμές των προηγούμενων εικονοστοιχείων [8.2.9].
TIFF	Aldus	Tagged Image File Format. Εύκαμπος τύπος αρχείου που υποστηρίζει ποικιλία πρότυπων συμπίεσης εικόνας, συμπεριλαμβανομένων των JPEG, JPEG-LS JPEG-2000, JBIG2 και πολλά άλλα.
WebP	Google	<i>Το WebP υποστηρίζει απωλεστική συμπίεση μέσω της συμπίεσης video Web VP8 (δείτε πιο κάτω) και μη απωλεστική συμπίεση χρησιμοποιώντας χωρική πρόβλεψη [8.10] και μία παραλλαγή της αναφοράς προς τα πίσω του LZW [8.5] και κωδικοποίηση εντροπίας κατά Huffman [8.2]. Επίσης υποστηρίζεται το χαρακτηριστικό της διαφάνειας.</i>
<i>Βίντεο</i>		
AVS	MII	Audio Video Standard. Είναι παρόμοιο με το H.264 αλλά στηρίζεται στη χρήση εκθετικής κωδικοποίησης Golomb. Αναπτύχθηκε στην Κίνα.
HDV	Συνεργασία εταιρειών	High Definition Video. Αποτελεί μία επέκταση του προτύπου DV για τηλεόραση υψηλής ευκρίνειας που χρησιμοποιεί συμπίεση τύπου MPEG-2, η οποία περιλαμβάνει απομάκρυνση χρονικού πλεονασμού με τη βοήθεια της πρόβλεψης διαφορών [8.2.9].
M-JPEG	Διάφορες εταιρείες	Ένας τύπος συμπίεσης στον οποίο το κάθε πλαίσιο συμπιέζεται ανεξάρτητα από τα υπόλοιπα, χρησιμοποιώντας τον αλγόριθμο JPEG.
Quick-Time	Apple Computer	Ένας υποδοχέας αρχείων που υποστηρίζει DV, H.261, H.262, H.264, MPEG-1, MPEG-2, MPEG-4 καθώς και πολλά άλλα πρότυπα συμπίεσης βίντεο.
VC-1	SMPTE	Ο πιο ευρέως χρησιμοποιούμενος τύπος βίντεο στο παγκόσμιο διαδίκτυο που έχει υιοθετηθεί για HD και Blue-ray HD DVDs. Είναι παρόμοιος με το πρότυπο H.264 / AVC και χρησιμοποιεί έναν ακέραιο μετασχηματισμό DCT με μεταβλητά μήκη ομάδων [8.2.8 και 8.2.9] καθώς και πίνακες κωδικών μεταβλητού μήκους ανεξάρτητων περιεχομένου [8.2.1] αλλά όχι τεχνικές πρόβλεψης ανάμεσα στα πλαίσια.
WMV9	Microsoft	
WebP VP8	Google	Ένας τύπος αρχείου που στηρίζεται στην κωδικοποίηση μετασχηματισμού ομάδων εικονοστοιχείων [5.9] των διαφορών πρόβλεψης μέσα στα πλαίσια και ανάμεσα στα πλαίσια [8.10]. Η κωδικοποίηση των διαφορών στηρίζεται στην εντροπία και σε έναν κωδικοποιητή προσαρμοστικής κωδικοποίησης.

# Κωδικοποίηση Huffman

# Κωδικοποίηση Huffman

- Τεχνική απομάκρυνσης πλεονασμού κωδικοποίησης
- Οδηγεί στο μικρότερο δυνατό πλήθος κωδικοσυμβόλων ανά σύμβολο πηγής.
- Είναι βέλτιστος για μια δεδομένη τιμή του  $n$  με τον περιορισμό πως η κωδικοποίηση της πηγής πραγματοποιείται ένα σύμβολο κάθε φορά.
- Τα σύμβολα μπορεί να είναι ή οι φωτεινότητες της εικόνας ή η έξοδος μιας διαδικασίας απεικόνισης έντασης (π.χ. Διαφορές έντασης εικονοστοιχείων, μήκη διαδρομών κλπ)

# Κωδικοποίηση Huffman

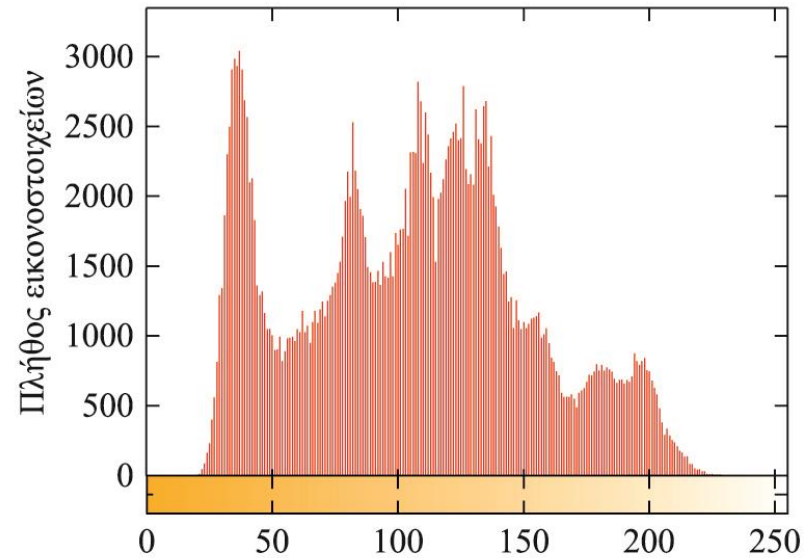
Αρχική πηγή		Μείωση της πηγής			
Σύμβολο	Πιθανότητα	1	2	3	4
$a_2$	0.4	0.4	0.4	0.4	0.6 0.4
$a_6$	0.3	0.3	0.3	0.3	
$a_1$	0.1	0.1	0.2	0.3	
$a_4$	0.1	0.1	0.1		
$a_3$	0.06	0.1			
$a_5$	0.04				

Αρχική πηγή			Μείωση της πηγής							
Σύμβολο	Πιθανότητα	Κώδικας	1	2	3	4				
$a_2$	0.4	1	0.4	1	0.4	1	0.4	1	0.6	0
$a_6$	0.3	00	0.3	00	0.3	00	0.3	00	0.4	1
$a_1$	0.1	011	0.1	011	0.2	010	0.3	01		
$a_4$	0.1	0100	0.1	0100	0.1	011				
$a_3$	0.06	01010	0.1	0101						
$a_5$	0.04	01011								

# Παράδειγμα κωδικοποίησης Huffman



(α)



(β)

Η **εντροπία** της αρχικής εικόνας: **7.3838** bits/pixel

Κωδικοποίηση κατά **Huffman**: **7.428** bits/pixel

Υπερβαίνει κατά **11.587** bits περίπου **0.6%**

$$C = 8/7.428 = 1.077$$

$$R = 1 - (1/1.077) = 0.0715$$

Το **7.15%** της αρχικής αναπαράστασης της έντασης με **8** bits απομακρύνθηκε χαρακτηριζόμενο ως **πλεονασμός κωδικοποίησης**.

# Κωδικοποίηση Golomb



# Κωδικοποίηση Golomb

- Κωδικοποιούμε μη-αρνητικούς ακέραιους αριθμούς με εκθετικά ελαττούμενες κατανομές πιθανότητας.
- Οι είσοδοι κωδικοποιούνται στο βέλτιστο βαθμό με πιο απλούς υπολογιστικά κωδικούς σε σχέση με τους κώδικες Huffman.

# Κωδικοποίηση Golomb

- Για κάποιον μη-αρνητικό ακέραιο  $n$  και ενός ακέραιου διαιρέτη  $m$ , ο κώδικας Golomb του  $n$  ως προς  $m$ ,  $G_m(n)$  είναι συνδυασμός του μοναδιαίου κώδικα του πηλίκου της διαίρεσης  $\lfloor m/n \rfloor$  και της διαδικής αναπαράστασης του υπόλοιπου  $n \bmod m$ 
  - Κατασκευάζουμε τον μοναδιαίο κώδικα του πηλίκου  $\lfloor m/n \rfloor$
  - Μοναδιαίος κώδικας ενός ακέραιου αριθμού  $q$  είναι μια ακολουθία  $q$  δυαδικών ψηφίων με τιμή 1 ακολουθούμενα από ένα ψηφίο με τιμή 0.
  - Θέτουμε  $k = \lceil \log_2 m \rceil$ ,  $c = 2^k - m$ ,  $r = n \bmod m$  και υπολογίζουμε με αποκοπή το υπόλοιπο  $r'$

$$r' = \begin{cases} r, \text{ με αποκοπή σε } k - 1 \text{ bits} & 0 \leq r \leq c \\ r + c, \text{ με αποκοπή σε } k \text{ bits} & \text{διαφορετικά} \end{cases}$$

- Ενώνουμε τα αποτελέσματα των δύο προηγούμενων βημάτων

# Κώδικες Golomb

$n$	$G_1(n)$	$G_2(n)$	$G_4(n)$	$G_{\text{exp}}^0(n)$
0	0	00	000	0
1	10	01	001	100
2	110	100	010	101
3	1110	101	011	11000
4	11110	1100	1000	11001
5	111110	1101	1001	11010
6	1111110	11100	1010	11011
7	11111110	11101	1011	1110000
8	111111110	111100	11000	1110001
9	1111111110	111101	11001	1110010

Ορισμένοι κώδικες Golomb για τους ακεραίους 0-9

# Κώδικες Golomb

- Οι κώδικες αυτοί μπορούν να χρησιμοποιηθούν για αναπαράσταση μη αρνητικών αριθμών
- Υπάρχουν πολλοί κώδικες από τους οποίους μπορούμε να διαλέξουμε
- Το πιο σημαντικό βήμα για την αποδοτική εφαρμογή τους είναι η επιλογή του διαιρέτη  $m$ .
- Αν οι ακέραιοι κατανέμονται γεωμετρικά με συνάρτηση πυκνότητας πιθανότητας

$$P(n) = (1 - \rho)\rho^n$$

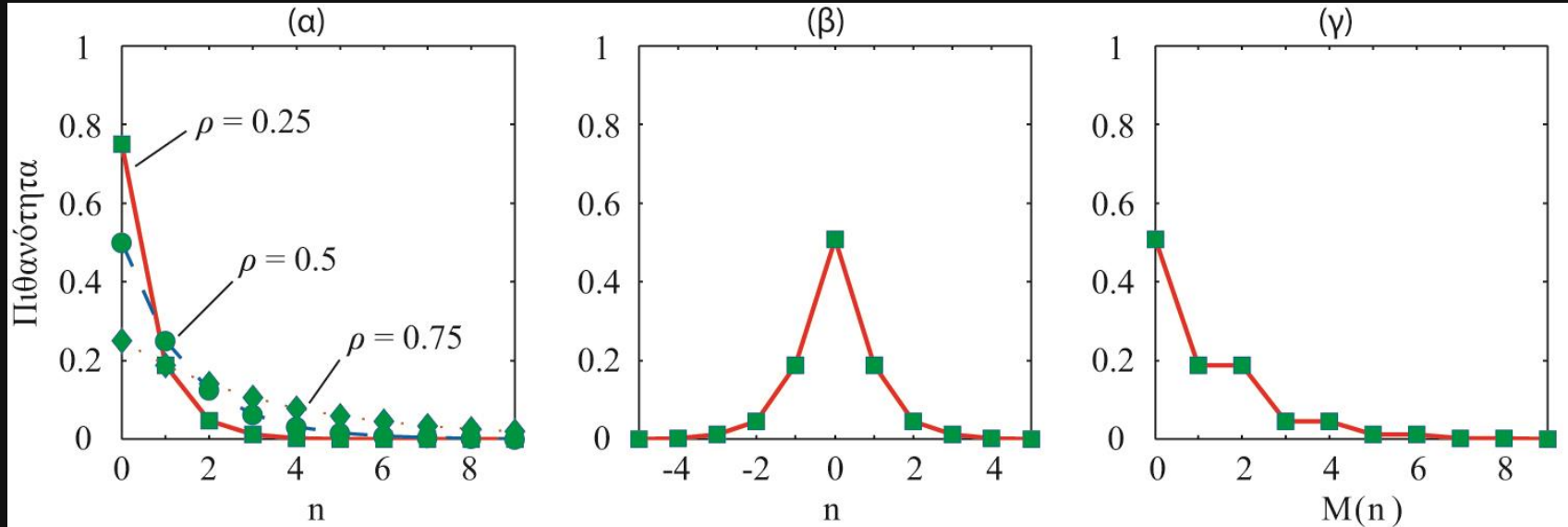
η βέλτιστη λύση δίνεται για  $m = \left\lceil \frac{\log_2(1+\rho)}{\log_2(\frac{1}{\rho})} \right\rceil$

# Κώδικες Golomb

- Γενικά οι πιθανότητες των εντάσεων μιας εικόνας είναι απίθανό να ταιριάζουν με τις πιθανότητες που προκύπτουν από την εξίσωση της κατανομής.
- Οι κώδικες Golomb χρησιμοποιούνται σπάνια για κωδικοποίηση της έντασης
- Αν όμως κωδικοποιήσουμε τις διαφορές εντάσεων, οι πιθανότητες αυτές μοιάζουν με αυτές που περιγράφονται στην προηγούμενη εξίσωση.
- Για να αντιμετωπίσουμε τις αρνητικές διαφορές στην κωδικοποίηση, χρησιμοποιούμε την απεικόνιση της μορφής:

$$M(n) = \begin{cases} 2n, & n \geq 0 \\ 2|n| - 1, & n < 0 \end{cases}$$

# Παράδειγμα κωδικοποίησης Golomb



- (α) τρεις γεωμετρικές κατανομές θετικών τιμών που προκύπτουν από την εξίσωση κατανομών
- (β) μια εκθετικώς φθίνουσα κατανομή θετικών και αρνητικών τιμών
- (γ) η επαναδιατεταγμένη έκδοση της κατανομής της εικόνας (β) που κατασκευάστηκε χρησιμοποιώντας την εξίσωση μετασχηματισμού.

# Αριθμητική Κωδικοποίηση

# Αριθμητική κωδικοποίηση

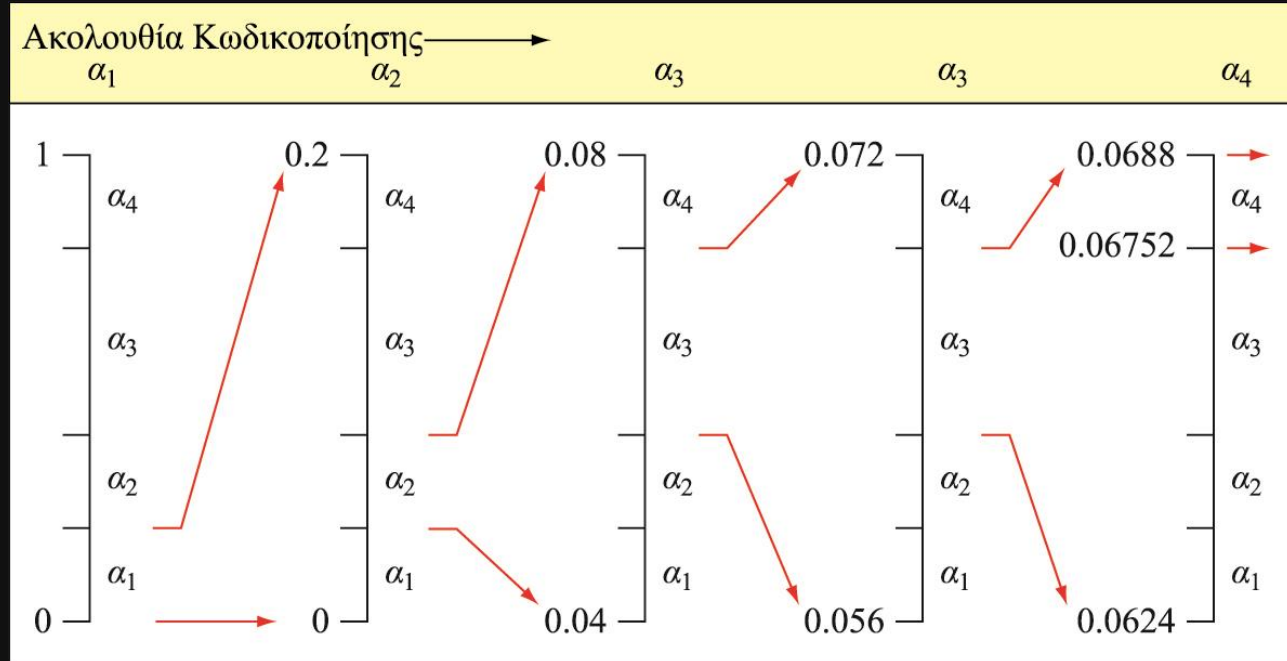
- Σε αντίθεση με τους κώδικες μεταβλητού μήκους, οι αριθμητική κωδικοποίηση δημιουργεί non-block κώδικες.
- Δεν υπάρχει μια προς μια αντιστοιχία ανάμεσα στα σύμβολα της πηγής και τις κωδικολέξεις
- Μια απλή κωδικολέξη εκχωρείται σε μια ολόκληρη ακολουθία συμβόλων πηγής (μήνυμα)
- Η κωδικολέξη ορίζει μια περιοχή αριθμών από μηδέν μέχρι το 1.



# Αριθμητική κωδικοποίηση

- Είναι διαφορετική από τους κώδικες μεταβλητού μήκους γιατί δημιουργεί nonblock κώδικες.
- Δεν δίνουν μία-προς-μία αντιστοιχία ανάμεσα στα σύμβολα της πηγής και τις κωδικολέξεις που τα περιγράφουν
- Μια αριθμητική κωδικολέξη εκχωρείται σε μια ολόκληρη ακολουθία συμβόλων πηγής (μήνυμα).
- Η κάθε κωδικολέξη από μόνη της ορίζει μια περιοχή πραγματικών αριθμών με τιμές ανάμεσα στο 0 και το 1.
- Όσο μεγαλώνει ο αριθμός των συμβόλων του μηνύματος, το διάστημα μικραίνει, ενώ τα bits που απαιτούνται για την αναπαράσταση της πληροφορίας γίνονται περισσότερα.
- Το κάθε επιπλέον σύμβολο της πηγής, ελαττώνει το διάστημα ανάλογα με την πιθανότητα εμφάνισης του.
- Η τεχνική αυτή δεν απαιτεί το κάθε σύμβολο να μετατρέπεται σε σειρά ακέραιου πλήθους κωδικοσυμβόλων, και έτσι προσεγγίζει (θεωρητικά) το όριο του Shannon

# Παράδειγμα αριθμητικής κωδικοποίησης



Σύμβολο πηγής	Πιθανότητα	Αρχικό υποδιάστημα
$a_1$	0.2	[0.0, 0.2)
$a_2$	0.2	[0.2, 0.4)
$a_3$	0.4	[0.4, 0.8)
$a_4$	0.2	[0.8, 1.0)

# Αριθμητική κωδικοποίηση

- Όσο μεγαλώνει το μήκος του μηνύματος, τόσο το διάστημα μικραίνει και ο κώδικας προσεγγίζει το όριο που υπαγορεύεται από το πρώτο θεώρημα Shannon.
- Δεν μπορεί να το φτάσει όμως γιατί:
  - ▣ Υπάρχει προσθήκη συμβόλου που δείχνει το τέλος του μηνύματος και είναι απαραίτητο για να ξεχωρίζει το ένα μήνυμα από το άλλο
  - ▣ Η χρήση αριθμητικής πεπερασμένης ακρίβειας.

# Παράδειγμα αριθμητικής κωδικοποίησης

- Στο προηγούμενο παράδειγμα το μήνυμα των 5 συμβόλων κωδικοποιείται με 3 δεκαδικά ψηφία ή 0.6 δεκαδικά ψηφία ανά σύμβολο.
- Δεν μπορεί να γίνει ίσο με αυτό που υπαγορεύει το πρώτο θεώρημα του Shannon γιατί:
  - προσθέτουμε σύμβολο που δηλώνει το τέλος του μηνύματος και είναι απαραίτητο να ξεχωρίζει το ένα από το άλλο
  - χρησιμοποιούμε αριθμητική πεπερασμένης ακρίβειας.

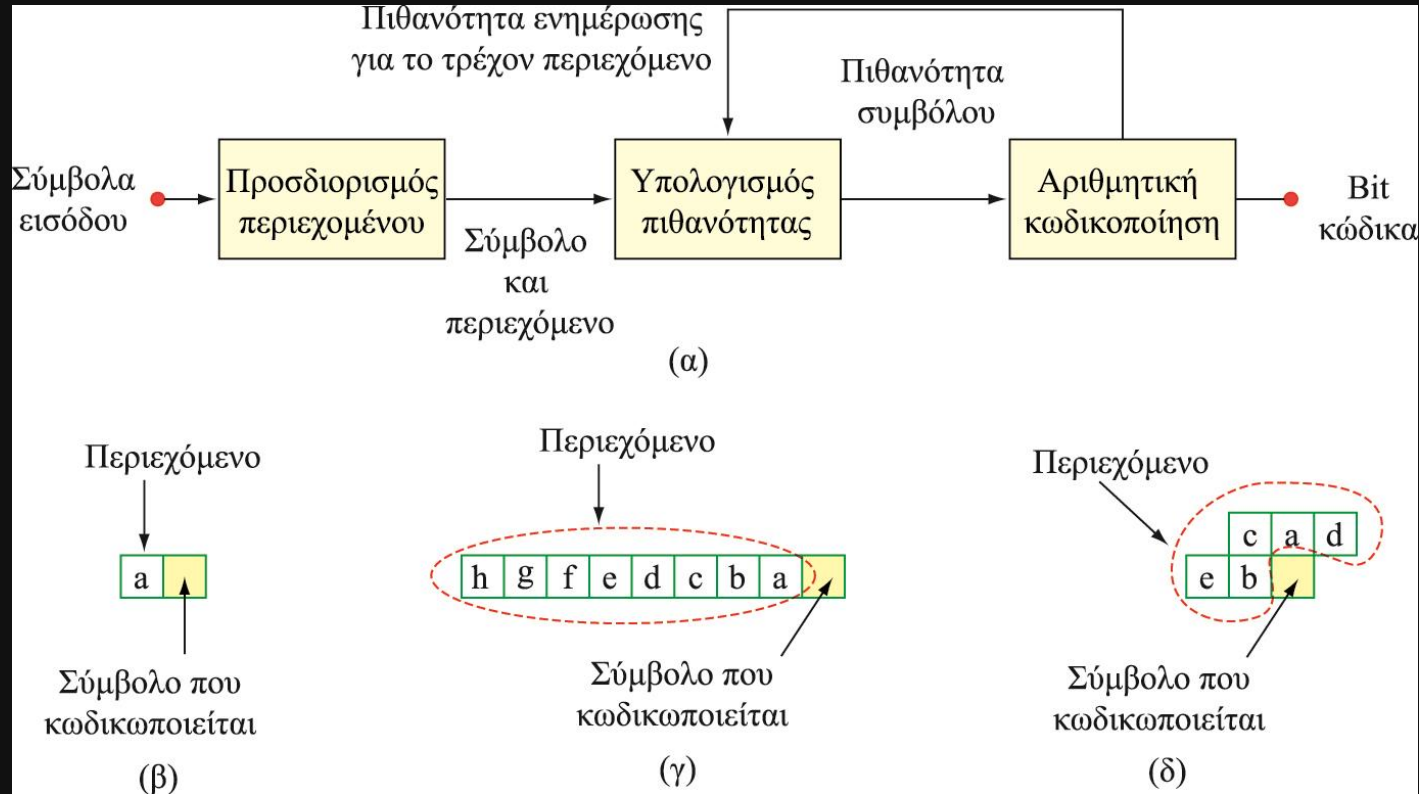
# Προσαρμοζόμενες εξαρτώμενες από το πλαίσιο εκτιμήσεις

- Οι αριθμητικοί κωδικοποιητές είναι βέλτιστοι στην απόδοση τους όταν έχουμε υπολογισμό πραγματικών πιθανοτήτων των συμβόλων που κωδικοποιούνται.
- Αν χρησιμοποιούνται όμως ανακριβή μοντέλα εκτίμησης πιθανοτήτων, τα αποτελέσματα είναι μη βέλτιστα.
- Για την λύση του προβλήματος αυτού, χρησιμοποιούμε μοντέλο που εξαρτάται από το υφιστάμενο σε κάθε περίπτωση περιβάλλον.
- Τα **προσαρμοζόμενα** πιθανοτικά μοντέλα, ενημερώνουν τις πιθανότητες των συμβόλων ταυτόχρονα με την κωδικοποίησή τους
- Τα μοντέλα τα **εξαρτώμενα** από το εννοιολογικό πλαίσιο υπολογίζουν τιμές πιθανότητας που στηρίζονται στη χρήση μιας προκαθορισμένης γειτονιάς εικονοστοιχείων, γύρω από τα σύμβολα που κωδικοποιούνται.

# Προσαρμοζόμενες εξαρτώμενες από το πλαίσιο εκτιμήσεις

- Η τεχνική **Q-coder** αλλά και **MQ-coder** έχουν ενσωματωθεί στα πρότυπα **JBIG**, **JPEG-2000** και χρησιμοποιούν πιθανοτικά μοντέλα τα οποία είναι προσαρμοζόμενα αλλά και εξαρτώμενα από το πλαίσιο.
- Η τεχνική Q-coder ενημερώνει με δυναμικό τρόπο τις πιθανότητες των συμβόλων κατά τη διάρκεια των επανακανονικοποιήσεων των διαστημάτων στην αριθμητική κωδικοποίηση
- Τα μοντέλα αυτά χρησιμοποιούνται και στην κωδικοποίηση Golomb στο πρότυπο συμπίεσης JPEG-LS

# Προσαρμοζόμενες εξαρτώμενες από το πλαίσιο εκτιμήσεις



(α) μια προσέγγιση προσαρμοστικής και εξαρτώμενης από το πλαίσιο αριθμητικής κωδικοποίησης (β) - (δ) τρία πιθανά μοντέλα περιεχομένου

# Κωδικοποίηση LZW



# Κωδικοποίηση LZW

- Οι προηγούμενες τεχνικές έχουν στόχο την απομάκρυνση πλεονασμού κωδικοποίησης
- Η κωδικοποίηση Lempel-Ziv-Welch (LZW) εκχωρεί κωδικολέξεις πεπερασμένου μήκους σε ακολουθίες συμβόλων πηγής μεταβλητού μήκους.
- Δεν απαιτεί γνώση εκ των προτέρων των πιθανοτήτων εμφάνισης των συμβόλων προς κωδικοποίηση
- Έχει ενσωματωθεί σε διάφορα αρχεία όπως GIF, TIFF, PDF
- Ο τύπος PNG έχει γίνει για να αρθούν οι περιορισμοί ευρεσιτεχνίας.

# Κωδικοποίηση LZW

- Κατά την έναρξη της διαδικασίας συμπίεσης, δημιουργείται ένα βιβλίο κώδικα ή λεξικό που περιέχει τα σύμβολα της πηγής που θέλουμε να κωδικοποιήσουμε
- Για μονόχρωμες εικόνες 8bit εκχωρούνται οι φωτεινότητες **0, 1, ..., 255**.
- Καθώς ο κωδικοποιητής εξετάζει τα εικονοστοιχεία της εικόνας το ένα μετά το άλλο, οι ακολουθίες τιμών έντασης που δεν ανήκουν στο λεξικό τοποθετούνται σε θέσεις, οι τιμές των οποίων υπολογίζονται από τον αλγόριθμο (πρώτη μη κατειλημμένη θέση που θα βρεθεί).

# Παράδειγμα κωδικοποίησης LZW

→

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

Τρέχουσα Αναγνωρισθείσα Ακολουθία	Εικονοστοιχείο προς επεξεργασία	Κωδικοποιημένη Έξοδος	Θέση στο Λεξικό (Κωδικολέξη)	Είσοδος Λεξικού
	39			
39	39	39	256	39-39
39	126	39	257	39-126
126	126	126	258	126-126
39	39	126	259	126-39
39	39			
39-39	126	256	260	39-39-126
126	126			
126-126	39	258	261	126-126-39
39	39			
39-39	126			
39-39-126	126	260	262	39-39-126-126
126	39			
126-39	39	259	263	126-39-39
39	126			
39-126	126	257	264	39-126-126
126		126		

Η αρχική εικόνα των 128bits μετατρέπεται σε εικόνα των 90bits με λόγο συμπίεσης 1.42:1

Κωδικοποίηση μήκους διαδρομής

# Κωδικοποίηση μήκους διαδρομής

- Για εικόνες που χαρακτηρίζονται από επαναλαμβανόμενες τιμές έντασης κατά μήκος των γραμμών (ή στηλών), μπορούμε να τις συμπίεσουμε με αναπαράσταση ομάδων διαδοχικών ταυτόσημων εντάσεων ως ζεύγη της μορφής ομάδα-μήκος.
- Κάθε ζεύγος καθορίζει την έναρξη μιας νέας τιμής έντασης και το πλήθος των διαδοχικών εικονοστοιχείων αυτής της έντασης.
- Η τεχνική αυτή ονομάζεται Running-Length Encoding (RLE).
- Η συμπίεση επιτυγχάνεται δια της απαλοιφής ενός απλού τύπου χωρικού πλεονασμού (ομάδα ταυτόσημων εντάσεων).
- Όταν έχουμε μικρό αριθμό ταυτόσημων εικονοστοιχείων, αυτός ο τύπος κωδικοποίησης οδηγεί σε αύξηση του όγκου των δεδομένων

# Η τεχνική RLE

- Είναι αποδοτική σε περιπτώσεις ασπρόμαυρων εικόνων.
- Επειδή έχουν μόνο 2 φωτεινότητες μπορεί μια ολόκληρη γραμμή να κωδικοποιηθεί με μια ακολουθία μηκών αντί για ζεύγη της μορφής μήκος-τιμή έντασης.
- Σαρώνουμε γραμμή από αριστερά προς τα δεξιά και κωδικοποιούμε κάθε ομάδα συνεχόμενων εικονοστοιχείων (δηλαδή την κάθε διαδρομή) που συναντάται (και τα οποία έχουν τιμή 0 ή 1) χρησιμοποιώντας το μήκος της καθιερώνοντας μια σύμβαση για τον καθορισμό της τιμής της διαδρομής.

# Συμβάσεις RLE

- Οι πιο διαδεδομένες συμβάσεις καθορισμού τιμής διαδρομής:
  - ▣ καθορισμός της τιμής της πρώτης ομάδας εικονοστοιχείων σε κάθε γραμμή της εικόνας
  - ▣ υπόθεση πως η γραμμή ξεκινά με ένα σύνολο λευκών εικονοστοιχείων, το πλήθος των οποίων μπορεί να είναι ίσο με μηδέν.

# Κωδικοποίηση μήκους διαδρομής

- Μπορεί να επιτευχθεί περισσότερη συμπίεση αν κωδικοποιήσουμε ξεχωριστά τα μαύρα και τα λευκά εικονοστοιχεία χρησιμοποιώντας διαφορετικούς κώδικες μεταβλητού μήκους.
- Τα μήκη των λευκών και μαύρων ομάδων εικονοστοιχείων κωδικοποιούνται χωριστά με κώδικες μεταβλητού μήκους προσαρμοσμένοι στη δική τους στατιστική.
- Έστω το σύμβολο  $\alpha$  για την κωδικοποίηση ομάδας διαδοχικών εικονοστοιχείων μαύρων μήκους  $j$ .
- Η πιθανότητα εκπομπής του συμβόλου  $\alpha_j$  από μια υποθετική μαύρη πηγή μήκους διαδρομής, διαιρώντας το πλήθος των μαύρων μηκών διαδρομής μήκους  $j$  ολόκληρης της εικόνας με τον συνολικό αριθμό των μαύρων μηκών.
- Από αυτά μπορούμε να υπολογίσουμε την εντροπία  $H_0$  της πηγής μαύρων μηκών διαδρομής
- Το ίδιο μπορεί να γίνει και με την πηγή λευκών μηκών διαδρομής για την εντροπία  $H_1$
- Η κατά προσέγγιση εντροπία μήκους διαδρομής της εικόνας, υπολογίζεται από την εξίσωση

$$H_{RL} = \frac{H_0 + H_1}{L_0 + L_1}$$

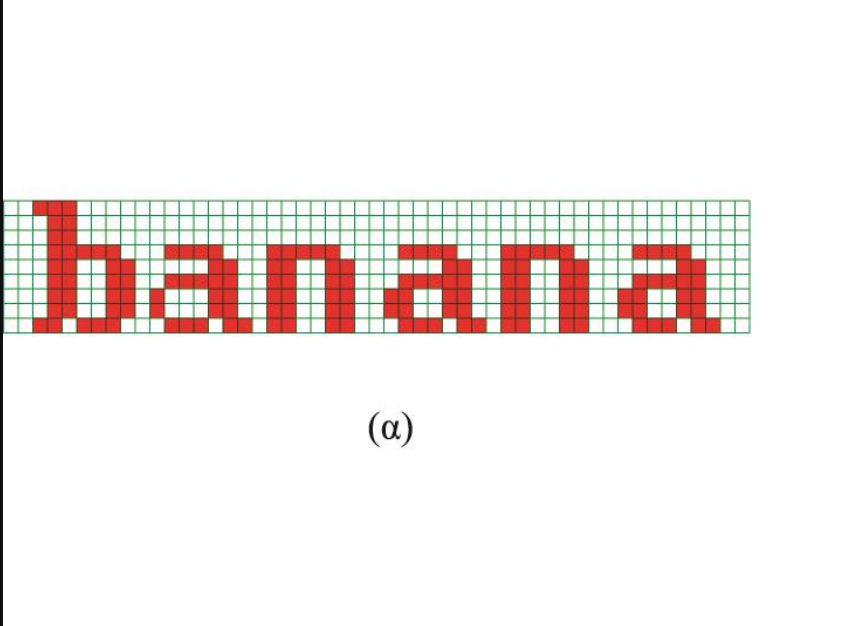





Κωδικοποίηση βασισμένη σε σύμβολα

# Κωδικοποίηση βασισμένη σε σύμβολα

- Κάθε εικόνα αναπαρίσταται ως μια συλλογή από συχνά εμφανιζόμενα σύμβολα (υπο-εικόνες).
- Κάθε σύμβολο αποθηκεύεται σε λεξικό συμβόλων και η εικόνα κωδικοποιείται.
- Η κωδικοποίηση περιλαμβάνει μια τριπλέτα  $(x_i, y_i, t_i)$  όπου  $x_i, y_i$  είναι η θέση του στην εικόνα και  $t_i$  είναι η διεύθυνση του στο λεξικό.
- Η κάθε τριπλέτα αναπαριστά ένα στιγμιότυπο του συμβόλου του λεξικού στην εικόνα
- Η αποθήκευση του μόνο μια φορά, προκαλεί την συμπίεση τη εικόνας σε σημαντικό βαθμό (ιδιαίτερα για εικόνες τύπου bitmap).

# Κωδικοποίηση βασισμένη σε σύμβολα - Παράδειγμα

 <p>(α)</p>	Token	Σύμβολο	Τριπλέτα
	0		(0, 2, 0) (3, 10, 1) (3, 18, 2)
	1		(3, 26, 1) (3, 34, 2) (3, 42, 1)
2			

(α) έγγραφο δύο επιπέδων (β) το λεξικό των συμβόλων (γ) οι τριπλέτες που χρησιμοποιούνται για τον εντοπισμό των συμβόλων στο έγγραφο

Αρχική εικόνα:  $9 \times 51 \times 1 = 459$  bits

Αν κάθε τριπλέτα αποτελείται από 3 bytes, τότε η αναπαράσταση θα έχει:

$(6 \text{ τριπλέτες} \times 3 \text{ αριθμούς} \times 8 \text{ bits}) + (9 \times 7 + 6 \times 7 + 6 \times 6) = 285$  bits

$$C = 1.61$$

Κωδικοποίηση επιπέδων bits

# Κωδικοποίηση επιπέδων bits

- Οι προηγούμενες τεχνικές μπορούν να εφαρμοστούν και σε εικόνες με περισσότερες τιμές έντασης από δύο με επεξεργασία καθενός από τα επίπεδα bits που περιέχουν.
- Η τεχνική αυτή ονομάζεται κωδικοποίηση επιπέδων bits.
- Στηρίζεται στην αποσύνθεση της εικόνας σε σύνολο δυαδικών εικόνων.
- Για την ανάλυση της εικόνας σε ένα σύνολο δυαδικών εικόνων, διαχωρίζουμε τους  $m$  συντελεστές του πολυωνύμου σε  $m$  επίπεδα bits του ενός bit.

$$a_{m-1}2^{m-1} + a_{m-2}2^{m-2} + \dots + a_12^1 + a_02^0$$

# Κωδικοποίηση επιπέδων bits

- Το επίπεδο του χαμηλότερου στην τάξη bit αντιστοιχεί μετά την καταγραφή των  $a_0$  bits
- Το επίπεδο του υψηλότερου στην τάξη bit αντιστοιχεί μετά την καταγραφή των  $a_{m-1}$  bits
- Μειονέκτημα: Μικρές αλλαγές στην φωτεινότητα επιρρεάζουν την πολυπλοκότητα των επιπέδων bits. (πχ εικονοστοιχείο με ένταση **127 01111111** και εικονοστοιχείο με ένταση **128 10000000**).
- Προκύπτουν απότομες μεταβάσεις από το **0** στο **1**.
- Εναλλακτικά για την ελάττωση των επιδράσεων των μικρών μεταβολών έντασης, μπορεί να χρησιμοποιηθεί αρχικά ο κώδικας Gray μήκους  $m$  bits.
- Ο κώδικας αυτός μήκους  $m$  bits  $g_{m-1} \dots g_2 g_1 g_0$  που αντιστοιχεί στο προηγούμενο πολυώνυμο, μπορεί να υπολογιστεί από

$$g_i = a_i \oplus a_{i+1}, 0 \leq i \leq m - 2$$

$$g_{m-1} = a_{m-1}$$

Με το σύμβολο  $\oplus$  να δηλώνει την πράξη αποκλειστικού OR (XOR)

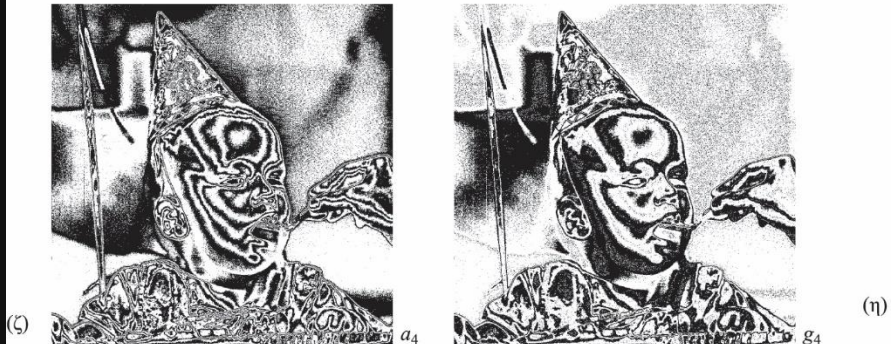
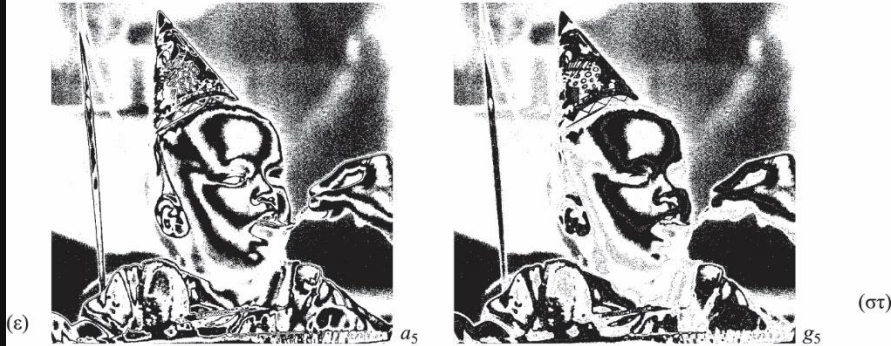
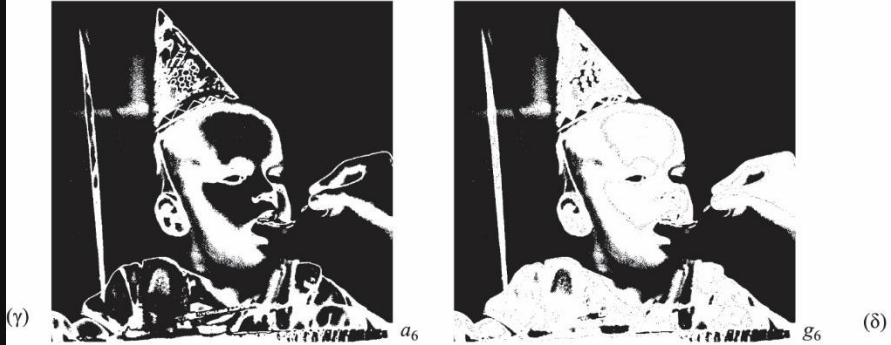
# Κωδικοποίηση επιπέδων bits

- Με χρήση κώδικα Gray, δεν έχουμε μεγάλες μεταβολές στα bits κωδικολέξεων, παρά μόνο σε μια μία θέση bit.
- Για παράδειγμα όταν τα επίπεδα έντασης με τιμές **127** και **128** αναφέρονται σε συνεχόμενα εικονοστοιχεία της εικόνας, μόνο το επίπεδο των πιο σημαντικών bit θα περιέχει μετάπτωση από το **0** στο **1**, αφού οι κώδικες Gray είναι:

127: 01000000

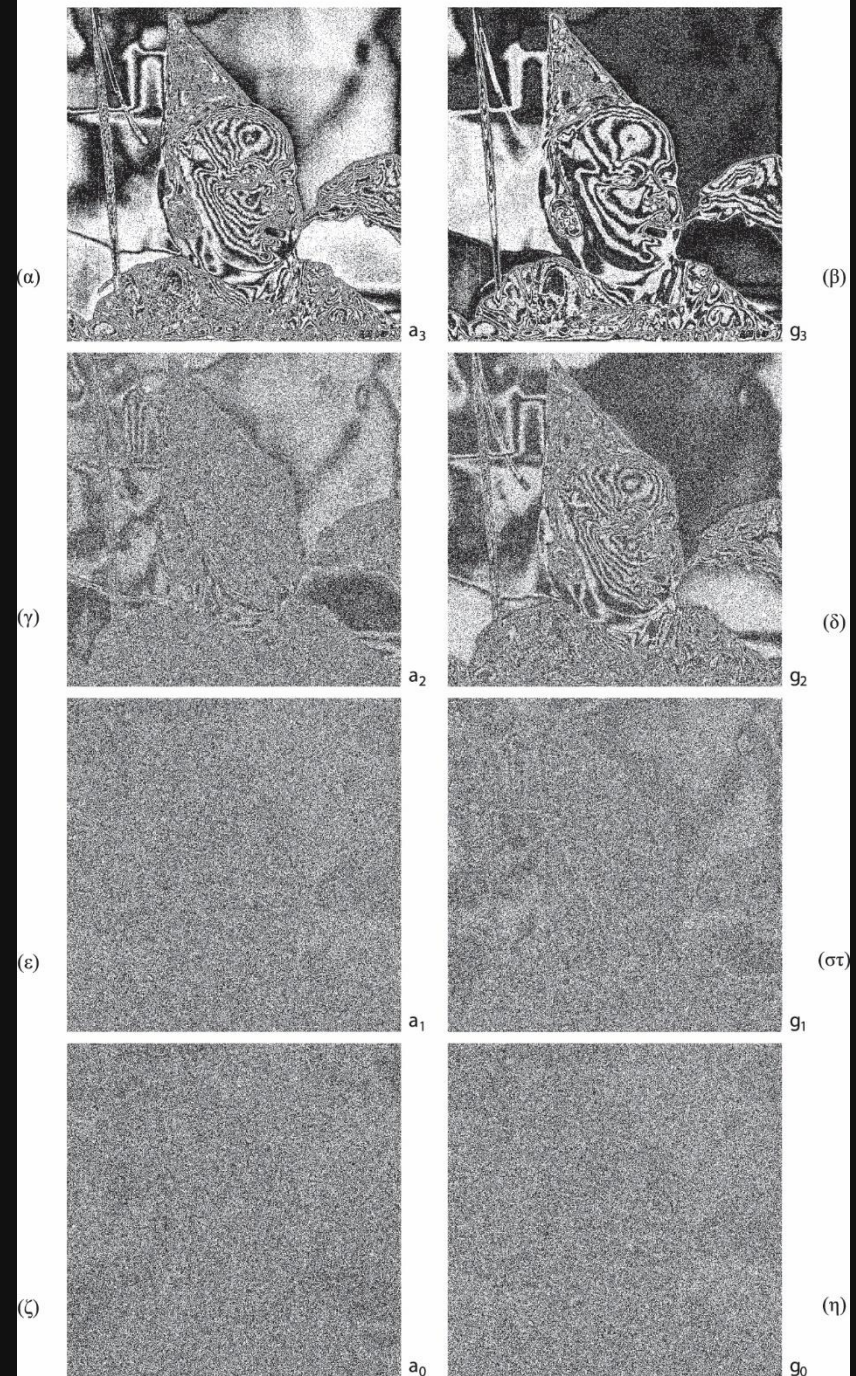
128: 11000000

(α) μια μονόχρωμη εικόνα των 256 bits  
(β)-(η) τα τέσσερα πιο σημαντικά επίπεδα bits  
της εικόνας (α) τα οποία είναι κωδικοποιημένα  
δυναδικά (αριστερή στήλη) και σύμφωνα με τον  
κώδικα Gray (δεξιά στήλη).





(α) – (η) τα τέσσερα λιγότερο σηµατικά επίπεδα bits της εικόνας (α) τα οποία είναι δυαδικά (αριστερά) και κωδικοποιηµένα σύµφωνα µε τον κώδικα Gray (δεξιά στήλη).



# Παράδειγμα κωδικοποίησης επιπέδων γκρι

- Τα επίπεδα που αντιστοιχούν στο περισσότερο σημαντικό bit είναι λιγότερο πολύπλοκα σε σχέση με αυτά που αντιστοιχούν στο λιγότερο σημαντικό bit.
- Περιέχουν μεγάλες ομοιόμορφες περιοχές με σημαντικά μικρότερο όγκο πληροφορίας, βαθμού επεξεργασίας και τυχειότητας.
- Το ίδιο ισχύει και για τα επίπεδα που είναι κωδικοποιημένα με κώδικα Gray.

# Παράδειγμα κωδικοποίησης επιπέδων γκρι

Συντελεστής <i>m</i>	Δυαδικός κώδικας (PDF bits)	Κώδικας Gray (PDF bits)	Λόγος Συμπίεσης
7	6,999	6,999	1.00
6	12,791	11,024	1.16
5	40,104	36,914	1.09
4	55,911	47,415	1.18
3	78,915	67,787	1.16
2	101,535	92,630	1.10
1	107,909	105,286	1.03
0	99,753	107,909	0.92

Αποτελέσματα της χωρίς σφάλμα κωδικοποίησης JBIG2 για τα δυαδικά και τα κωδικοποιημένα κατά Gray επίπεδα bits της προηγούμενης εικόνας.

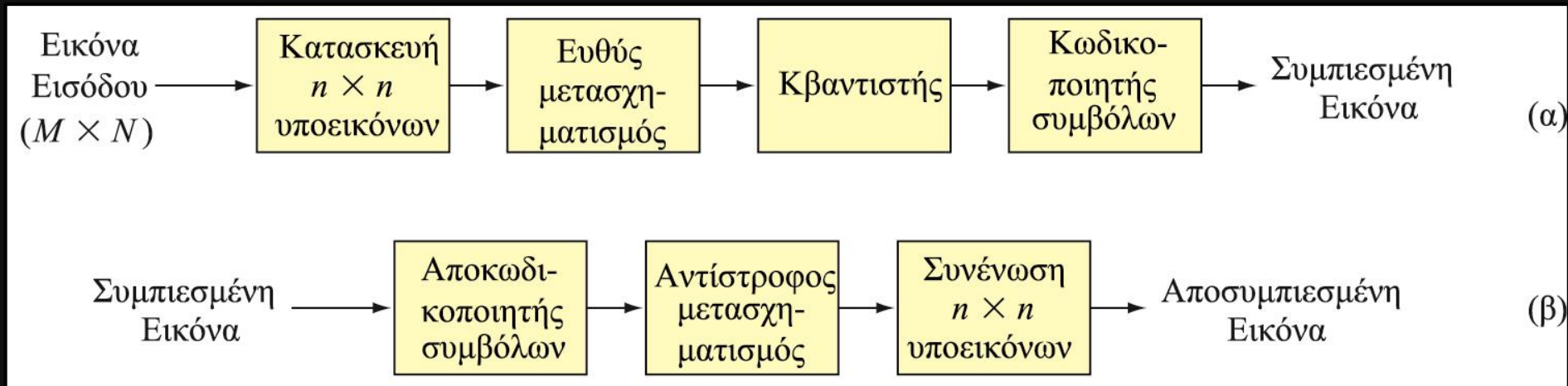
Η αρχική μονόχρωμη εικόνα συμπιέζεται κατά  $678676/475964 = 1.43:1$  για την κωδικοποίηση Gray και κατά  $678676/503916 = 1.35:1$  χωρίς την κωδικοποίηση Gray

Κωδικοποίηση μετασχηματισμού ομάδων εικονοστοιχείων

# Κωδικοποίηση μετασχηματισμού ομάδων εικονοστοιχείων

- Υποδιαιρεί την εικόνα σε μικρές μη επικαλυπτόμενες μεταξύ τους περιοχές εικονοστοιχείων ίσου μεγέθους (π.χ.  $8 \times 8$ )
- επεξεργάζεται κάθε μια ομάδα ξεχωριστά και ανεξάρτητα από τις άλλες χρησιμοποιώντας δισδιάστατο μετασχηματισμό.
- Μετά υφίσταται κβαντισμό και κωδικοποίηση.
- πραγματοποιείται με αντιστρέψιμο γραμμικό μετασχηματισμό όπως ο μετασχηματισμός Fourier
- Για τις περισσότερες εικόνες υπάρχει ένας αριθμός συντελεστών με μικρές τιμές που μπορούν να υποστούν μια διαδικασία κβαντισμού (ή και να αγνοηθούν παντελώς) με μικρή παραμόρφωση της εικόνας.

# Κωδικοποίηση μετασχηματισμού ομάδων εικονοστοιχείων

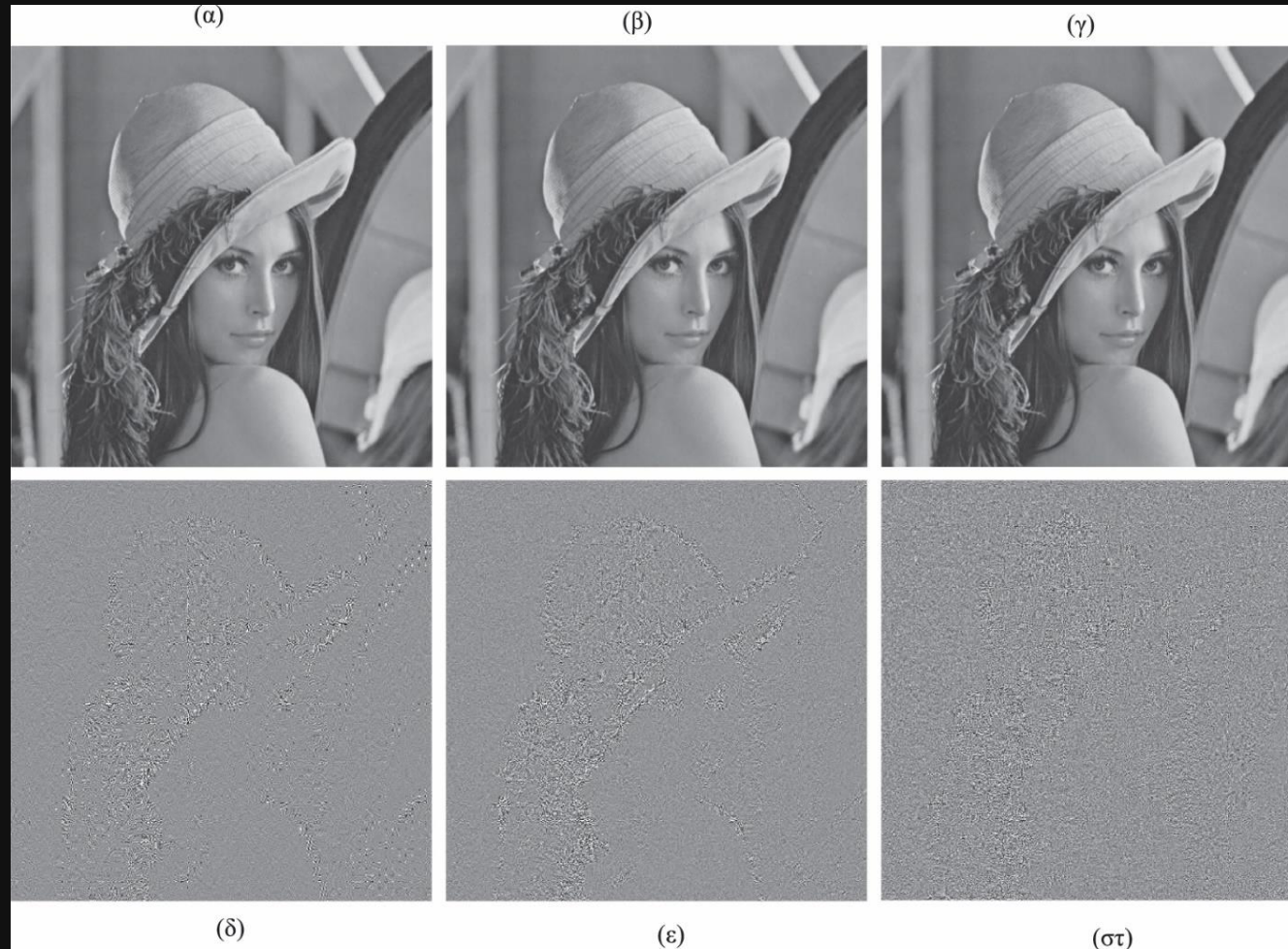


- ο αποκωδικοποιητής εκτελεί την αντίστροφη διαδικασία των βημάτων του κωδικοποιητή, εκτός του κβαντισμού (που δεν έχει αντιστροφή)
- Η εφαρμογή του μετασχηματισμού **αποσυσχετίζει** τα εικονοστοιχεία της κάθε υποεικόνας ή και να συσκευάσει την μεγαλύτερη πληροφορία σε μικρότερο αριθμό συντελεστών μετασχηματισμού
- Ο κβαντισμός **εξαλείφει** επιλεκτικά ή **κβαντίζει** κατά προσέγγιση τους συντελεστές με την ελάχιστη ποσότητα πληροφορίας
- Οι συντελεστές που απομένουν, κωδικοποιούνται με **κωδικοποίηση μεταβλητού μήκους**
- Η παραπάνω διαδικασία μπορεί να γίνει (α) με **προσαρμογή στο τοπικό περιεχόμενο** της εικόνας (προσαρμοσμένη κωδικοποίηση μετασχηματισμού) (β) **σταθερή** για όλες τις υποεικόνες (μη προσαρμοσμένη κωδικοποίηση μετασχηματισμού)

# Επιλογή μετασχηματισμού

- Υπάρχει μια πληθώρα μετασχηματισμών που μπορούν να χρησιμοποιηθούν
- Η επιλογή τους εξαρτάται από το μέγεθος του σφάλματος ανακατασκευής που μπορεί να γίνει ανεκτό και από τους διαθέσιμους υπολογιστικούς πόρους.
- Η συμπίεση επιτυγχάνεται κατά τη διάρκεια του κβαντισμού των συντελεστών του μετασχηματισμού (και όχι κατά το στάδιο εφαρμογής του μετασχηματισμού).

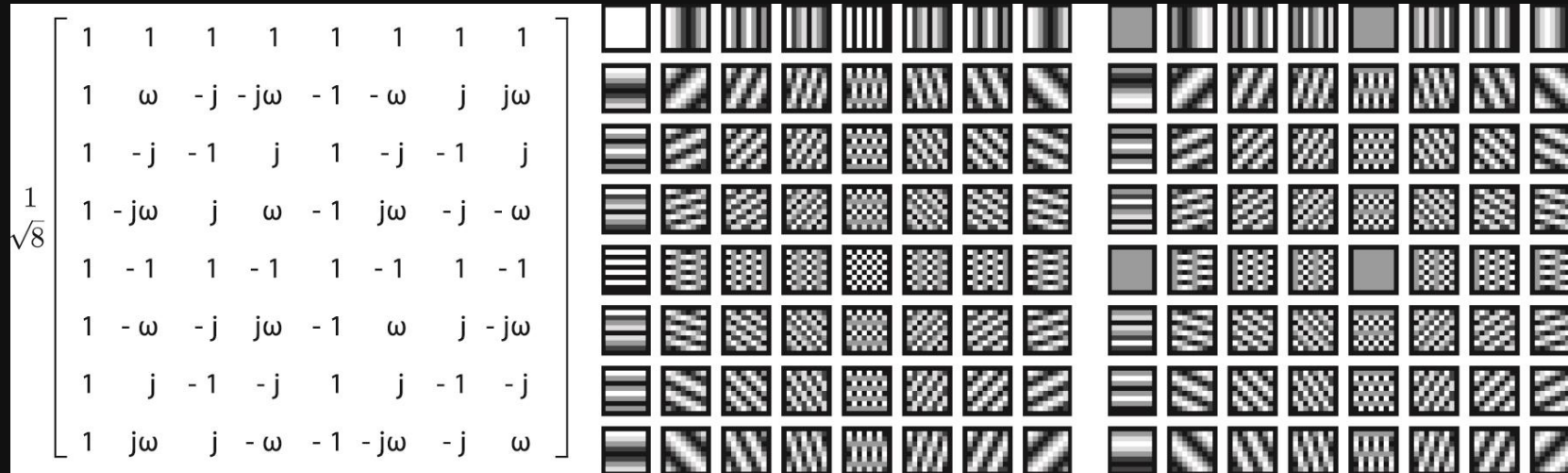
# Κωδικοποίηση ομάδας με DFT, WHT, DCT



Προσέγγιση της εικόνας χρησιμοποιώντας (α) μετασχηματισμό Fourier (β) μετασχηματισμό Walsh-Hadamard (γ) μετασχηματισμό συνημιτόνου με τις αντίστοιχες κλιμακούμενες εικόνες σφάλματος

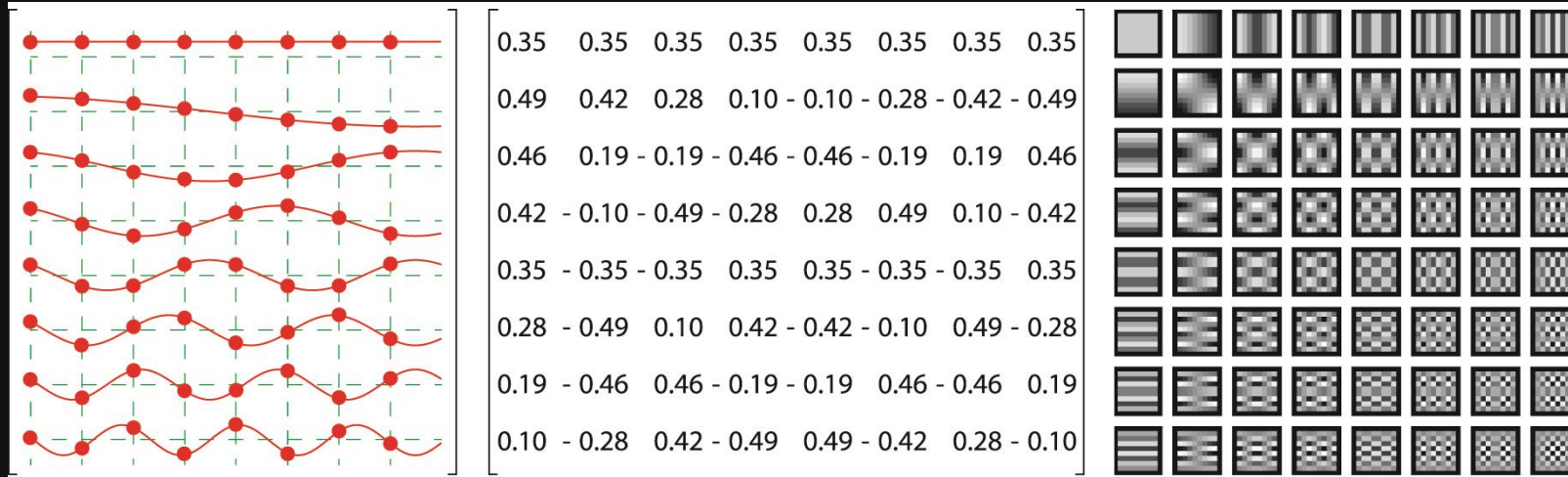


# Εικόνες βάσης μετασχηματισμών - DFT



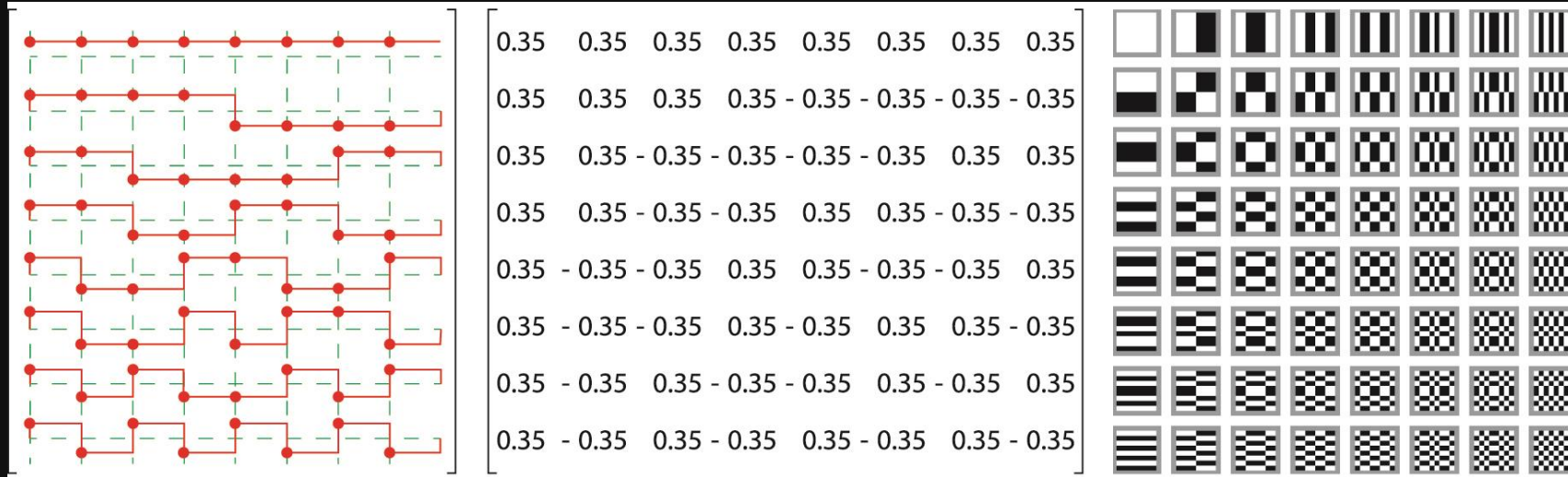
(α) Πίνακας μετασχηματισμού του διάκριτου μετασχηματισμού Fourier για  $N = 8$  (β)-(γ) το πραγματικό και φανταστικό μέρος των εικόνων βάσης του DFT

# Εικόνες βάσης μετασχηματισμών - DCT



(α) γραφική αναπαράσταση του ορθογώνιου πίνακα μετασχηματισμού του διάκριτου συνημιτόνου για  $N = 8$  (β) ο πίνακας μετασχηματισμού (γ) οι εικόνες βάσης του DCT

# Εικόνες βάσης μετασχηματισμών - WHT



(α) γραφική αναπαράσταση του ορθογώνιου πίνακα μετασχηματισμού Walsh-Hadamard για  $N = 8$  (β)  
ο πίνακας μετασχηματισμού (γ) οι εικόνες βάσης του WHT

# Παράδειγμα κωδικοποίησης ομάδας

- Από το προηγούμενο παράδειγμα, η απόδοση του μετασχηματισμού DCT για συμπίεση πληροφορίας είναι ανώτερη από τους DFT, WHT
- Αυτό ισχύει συνήθως, αλλά ο βέλτιστος μετασχηματισμός είναι ο Karhunen-Loeve
- Το πρόβλημα με αυτόν είναι ότι εξαρτάται από τα χρησιμοποιούμενα σε κάθε περίπτωση δεδομένα
- Ο υπολογισμός των εικόνων βάσης του μετασχηματισμού αυτού για κάθε υποεικόνα παρουσιάζει μεγάλο υπολογιστικό κόστος.
- Για τον λόγο αυτό δεν χρησιμοποιείται αυτός, αλλά οι υπόλοιποι τρεις που έχουν σταθερές εικόνες βάσης
- Πιο εύκολοι μετασχηματισμοί είναι ο WHT, πιο κοντά στον βελτιστο KLT, είναι οι ημιτονοειδείς μετασχηματισμοί (DFT, DCT)

# Κωδικοποίηση ομάδας

- Στην πράξη χρησιμοποιούμε τον DCT με καλό συμβιβασμό μεταξύ πολυπλοκότητας και συμπίεσης πληροφορίας.
- Έχει καταστεί διεθνές πρότυπο για την κωδικοποίηση αυτού του τύπου.
- Πλεονεκτήματα:
  - Υλοποιείται με τη βοήθεια ενός απλού ολοκληρωμένου κυκλώματος
  - Συσκευάζει τη μεγαλύτερη δυνατή ποσότητα πληροφορίας τους λιγότερους δυνατούς συντελεστές (για τις περισσότερες εικόνες)
  - Ελαχιστοποιεί την εμφάνιση παραμορφώσεων στα όρια των blocks (blocking artifact) και προκύπτει όταν τα όρια ανάμεσα στις υπο-εικόνες καθίστανται ορατά.

# Συναρτήσεις βάσης μετασχηματισμού Walsh-Hadamard

$$r(x, y, u, v) = e^{-j2\pi(ux+vy)/n}$$

$$s(x, y, u, v) = \frac{1}{n^2} e^{-j2\pi(ux+vy)/n}$$

# Συναρτήσεις βάσης μετασχηματισμού DCT

$$r(x, y, u, v) = s(x, y, u, v) = a(u)a(v) \cos \left[ \frac{(2x+1)u\pi}{2n} \right] \cos \left[ \frac{(2y+1)v\pi}{2n} \right]$$

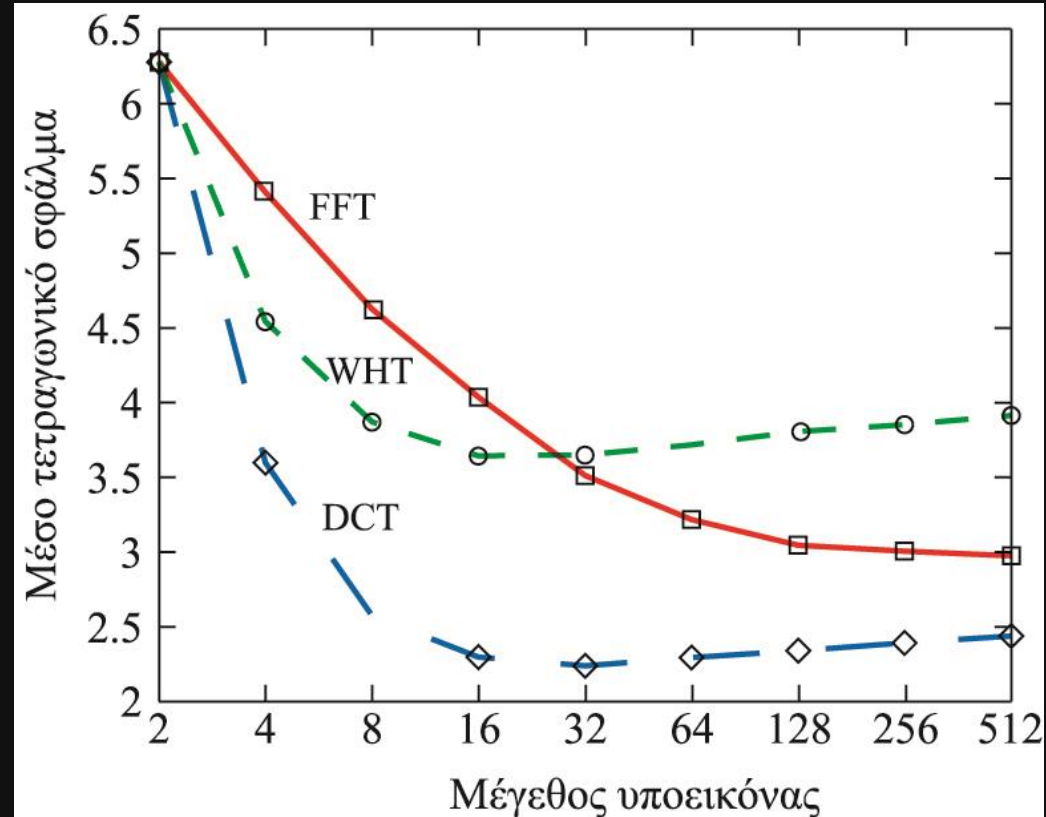
$$a(u) = \begin{cases} \sqrt{\frac{1}{n}}, u = 0 \\ \sqrt{\frac{2}{n}}, u = 1, 2, \dots, n-1 \end{cases}$$

# Επιλογή μεγέθους υπο-εικόνας

- Οι εικόνες υποδιαιρούνται σε μικρότερες εικόνες ώστε η συσχέτιση (πλεονασμός) ανάμεσα σε διαδοχικές υποεικόνες να ελαττώνεται σε κάποιο αποδεκτό επίπεδο (με η δύναμη του 2).
- Το επίπεδο της συμπίεσης και η υπολογιστική πολυπλοκότητα αυξάνονται με την αύξηση του μεγέθους της υποεικόνας.
- Τα πιο δημοφιλή μεγέθη είναι τα 8X8, 16X16

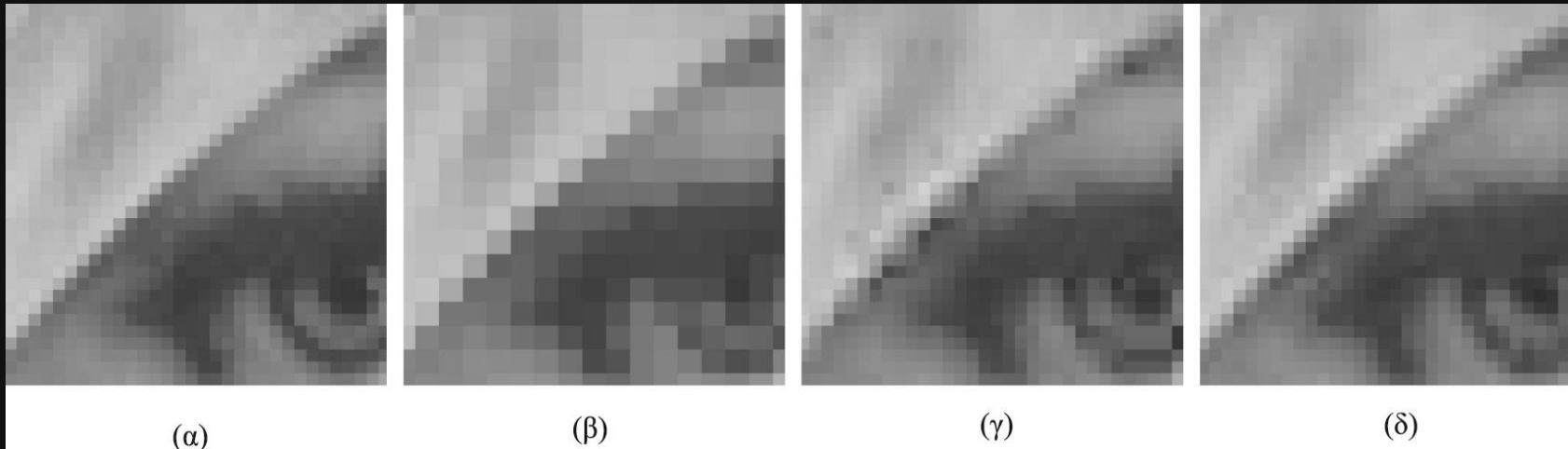


# Επίδραση του μεγέθους της υπο-εικόνας στην κωδικοποίηση μετασχηματισμού



Το σφάλμα ανακατασκευής σε συνάρτηση με το μέγεθος της υπο-εικόνας. Οι καμπύλες WHT, DCT οριζοντιώνονται για μέγεθος μεγαλύτερο από 8x8, ο DFT συνεχίζει και μειώνεται

# Επίδραση του μεγέθους της υπο-εικόνας στην κωδικοποίηση μετασχηματισμού



Προσεγγίσεις της εικόνας χρησιμοποιώντας το 25% των συντελεστών DCT και υποεικόνες διαστάσεων (β) 2X2, (γ) 4X4 και (δ) 8X8. Η αρχική εικόνα είναι ένα μεγενθυμένο τμήμα της αρχικής εικόνας.

# Κατανομή των bits

- Το σφάλμα ανακατασκευής είναι συνάρτηση του πλήθους και της σχετικής σημασίας των συντελεστών του μετασχηματισμού που απορρίπτονται και της ακρίβειας που χρησιμοποιείται για την αναπαράσταση των διατηρούμενων συντελεστών
- Οι συντελεστές που παραμένουν επιλέγονται με κριτήρια όπως:
  - ▣ η μέγιστη διακύμανση (κωδικοποίηση ζώνης)
  - ▣ μέγιστο μέτρο (κωδικοποίηση κατωφλίου).
- Η διαδικασία που περιλαμβάνει την απόρριψη συντελεστών, τον κβαντισμό και την κωδικοποίηση είναι γνωστή ως κατανομή bits.

# Παράδειγμα DCT – (12.5% συντελεστών)

Εικόνες από τις οποίες έχει απομακρυνθεί το 87.5% των συντελεστών μετασχηματισμού DCT της κάθε υπο-εικόνας διαστάσεων 8x8

(α)-(β) αποτελέσματα που έχουν προκύψει από την κωδικοποίηση κατωφλίου (8 μεγαλύτεροι συντελεστές)

(γ)-(δ) αποτελέσματα που έχουν προκύψει από την κωδικοποίηση ζώνης.



# Υλοποίηση της κωδικοποίησης ζώνης

- Οι συντελεστές του μετασχηματισμού που χαρακτηρίζονται από τη μέγιστη διακύμανση, μεταφέρουν το μεγαλύτερο μέρος της πληροφορίας
- Αυτοί θα πρέπει να διατηρηθούν κατά τη διαδικασία της κωδικοποίησης
- Οι συντελεστές που διατηρούνται πρέπει να περάσουν από το στάδιο του κβαντισμού και της κωδικοποίησης
- Κωδικοποιούνται με το ίδιο πλήθος bits
  - ▣ Ακολουθεί κανονικοποίηση με την τυπική τους απόκλιση και ομοιόμορφη διαδικασία κβαντισμού
- ή με διαφορετικό πλήθος bits
  - ▣ Για κάθε συντελεστή σχεδιάζεται και χρησιμοποιείται ένα σύστημα κβαντισμού

# Υλοποίηση κωδικοποίησης κατωφλίου

- Η κωδικοποίηση ζώνης υλοποιείται με χρήση απλής μάσκας που είναι ίδια για όλες τις υποεικόνες
- Η κωδικοποίηση κατωφλίου είναι από την φύση της προσαρμοσμένη αφού η θέση των συντελεστών που διατηρούνται μεταβάλλονται από την μία υποεικόνα στην άλλη.
- Για την κάθε υποεικόνα, οι συντελεστές που χαρακτηρίζονται από το μεγαλύτερο μέτρο έχουν την μεγαλύτερη συνεισφορά στην ανακατασκευασμένη εικόνα.
- Τα στοιχεία που προκύπτουν από τον μετασχηματισμό αναδιατάσσονται με κάποιο πρότυπο ώστε να μπορούν να κωδικοποιηθούν αποδοτικά με τον RLE.

# Κωδικοποίηση ζώνης / κατωφλίου

- (α) Μια τυπική μάσκα ζώνης
- (β) κατανομή ζώνης των bit
- (γ) μάσκα κατωφλίου
- (δ) ακολουθία ταξινόμησης των συντελεστών κατωφλίωσης.

Τα σκιασμένα τετράγωνα δείχνουν τους συντελεστές που διατηρήθηκαν για περαιτέρω επεξεργασία.

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

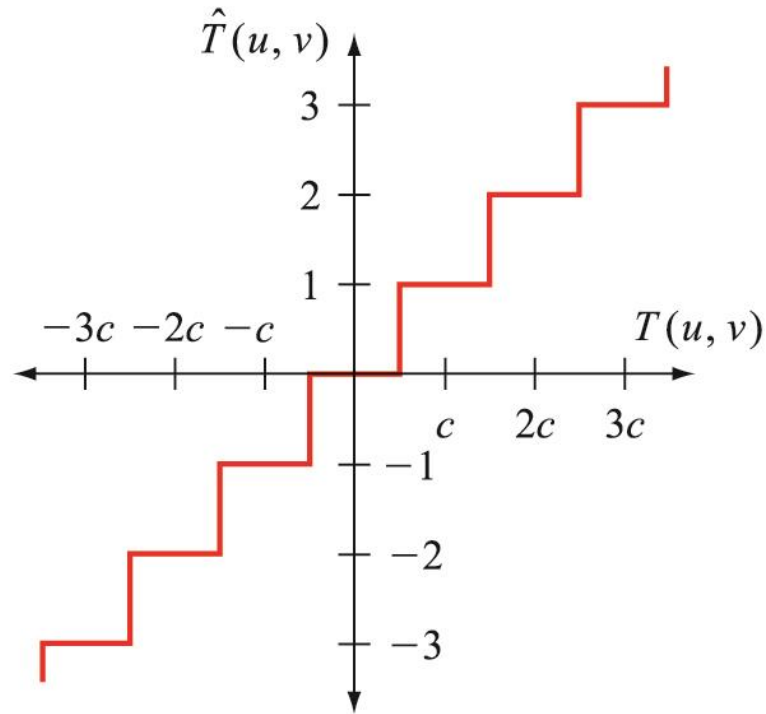
# Υλοποίηση κωδικοποίησης κατωφλίου

- Υπάρχουν τρεις τρόποι εφαρμογής του αλγόριθμου σε μια υπο-εικόνα:
  - Εφαρμογή καθολικού κατωφλίου σε όλες τις υπο-εικόνες (μεταβλητός ρυθμός κωδικοποίησης)
  - Εφαρμογή διαφορετικού κατωφλίου σε κάθε υπο-εικόνα (σταθερός ρυθμός κωδικοποίησης, γνωστός εκ των προτέρων)
  - Χρήση κατωφλίου το οποίο μεταβάλλεται ως συνάρτηση της θέσης του συντελεστή μέσα στην υπο-εικόνα (μεταβλητός ρυθμός κωδικοποίησης).
  - Σε αυτή την περίπτωση κανονικοποιούμε τους συντελεστές με έναν πίνακα κανονικοποίησης  $Z(u, v)$  και προκύπτει:

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



# Υλοποίηση κωδικοποίησης κατωφλίου



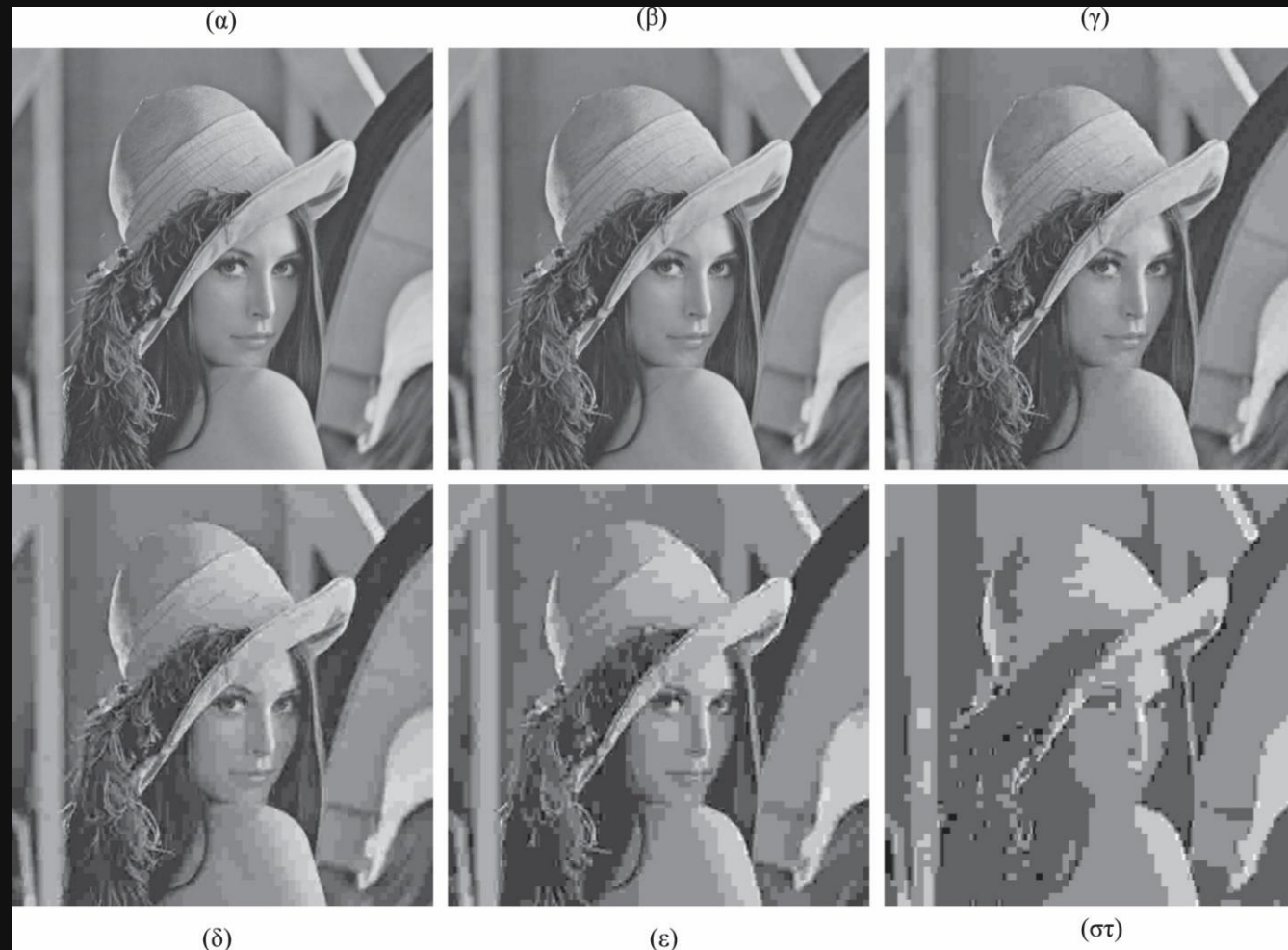
(α)

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

(β)

(α) Μια καμπύλη κβαντισμού που σχετίζεται με τη διαδικασία κωδικοποίησης κατωφλίου (β) ένας τυπικός πίνακας κανονικοποίησης (πρότυπο JPEG)

# Επίδειξη της κωδικοποίησης κατωφλίου



Προσεγγίσεις της δοκιμαστικής εικόνας με χρήση του DCT και του προηγούμενου πίνακα κανονικοποίησης (α)  $Z$ , (β)  $2Z$  (γ)  $4Z$  (δ)  $8Z$  (ε)  $16Z$  (στ)  $32Z$

# Το πρότυπο JPEG

- Περιλαμβάνει τρία διαφορετικά συστήματα κωδικοποίησης:
  - βασικό σύστημα με απώλειες που στηρίζεται στον DCT για τις περισσότερες εφαρμογές
  - εκτεταμένο σύστημα για εφαρμογές με υψηλότερο βαθμό συμπίεσης, μεγαλύτερη ακρίβεια και προοδευτική ανακατασκευή εικόνων
  - ανεξάρτητο σύστημα κωδικοποίησης χωρίς απώλειες για αντιστρέψιμη συμπίεση

# Συμβατότητα με το πρότυπο JPEG

- Για να είναι ένα σύστημα ή προϊόν συμβατό με το πρότυπο JPEG θα πρέπει να προσφέρει υποστήριξη τουλάχιστον για το βασικό σύστημα.
- Δεν καθορίζει κάποιον συγκεκριμένο τύπο διαμόρφωσης για αρχεία, κάποια τιμή χωρικής ανάλυσης και κάποιο μοντέλο για τον χρωματικό χώρο.
- Το βασικό σύστημα είναι και γνωστό ως ακολουθιακό βασικό σύστημα.

# Βασικό σύστημα με απώλειες

- Η ακρίβεια των δεδομένων στην είσοδο και την έξοδο περιορίζεται στα **8 bits**
- Οι κβαντισμένες τιμές του μετασχηματισμού DCT περιορίζονται στα **11 bits**.
- Η συμπίεση γίνεται σε 3 βήματα:
  - υπολογισμός του DCT
  - κβαντισμός
  - εκχώρηση κωδικών μεταβλητού μήκους

# Υπολογισμός JPEG

- Υποδιαίρεση εικόνας σε περιοχές 8X8 με επεξεργασία από αριστερά προς δεξιά και από πάνω προς τα κάτω
- Μετατόπιση των 64 εικονοστοιχείων προς τα αριστερά με αφαίρεση της ποσότητας  $2^{k-1}$
- Υπολογισμός μετασχηματισμού DCT
- Κβαντισμός συντελεστών με τον τύπο

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

- Αναδιάταξη συντελεστών με πρότυπο ταξινόμησης zigzag για την δημιουργία μιας μονοδιάστατης ακολουθίας κβαντισμένων συντελεστών.

# Υπολογισμός JPEG

- Οι μη μηδενικοί AC συντελεστές κωδικοποιούνται με κώδικα μεταβλητού μήκους που ορίζει τις τιμές των συντελεστών και το πλήθος των **μη μηδενικών** που προηγούνται.
- Η διαφορά της DC με την DC του προηγούμενου block (υπο-εικόνας) κωδικοποιείται.
- Οι κωδικοποιήσεις γίνονται βάσει πινάκων
- Ο πίνακας κβαντισμού της φωτεινότητας που προτείνεται από το JPEG (βλ. Προηγούμενη διαφάνεια) μπορεί να κλιμακωθεί για να επιτύχουμε διαφορετικά επίπεδα συμπίεσης.

# Παράδειγμα κωδικοποίησης JPEG

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	63	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94



# Βήμα 1: Μετατόπιση κατά $2^7$ αριστερά

-76	-73	-67	-62	-58	-67	-64	-55
-65	-69	-62	-38	-19	-43	-59	-56
-66	-69	-60	-15	16	-24	-62	-55
-65	-70	-57	-6	0	-22	-58	-59
-61	-67	-60	-24	-2	-40	-60	-58
-49	-63	-68	-58	-51	-65	-70	-53
-43	-57	-64	-69	-73	-67	-63	-45
-41	-49	-59	-60	-63	-52	-50	-34

## Βήμα 2: Εφαρμογή DCT

-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

## Βήμα 3: Κανονικοποίηση σύμφωνα με το πρότυπο JPEG

-26	-3	-6	2	2	0	0	0
1	-2	-4	0	0	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	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

## Βήμα 4: Αναδιάταξη με βάση το πρότυπο zigzag

[-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0 0 -1 2 0 0 0 0 0 -1 -1 EOB]

↓  
συντελεστής DC, θα κωδικοποιηθεί η **διαφορά** του από τον DC του προηγούμενου 8x8 block

# Κωδικοποίηση των συντελεστών

- Οι DC συντελεστές κωδικοποιούνται με βάση τον DPCM
  - Τα υπόλοιπα bits κωδικοποίησης αντιστοιχούν στα  $K$  λιγότερο σημαντικά bit της διαφοράς των DC συντελεστών ή της αρνητικής διαφοράς των συντελεστών
- Οι μη μηδενικοί AC συντελεστές κωδικοποιούνται με βάση πίνακες. Εξαρτώνται από:
  - το πλήθος των μη μηδενικών που προηγούνται
  - κατηγορία του μεγέθους του μη μηδενικού συντελεστή

## Βήμα 5: Κωδικοποίηση από πίνακες

```
1010110 0100 001 0100 0101 100001 0110 100011 001 100011 001  
001 100101 11100110 110110 0110 11110100 000 1010
```

# Αποκωδικοποίηση της εικόνας

-26	-3	-6	2	2	0	0	0
1	-2	-4	0	0	0	0	0
-3	1	5	-1	-1	0	0	0
-4	1	2	-1	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

# Άρση της κανονικοποίησης

-416	-33	-60	32	48	0	0	0
12	-24	-56	0	0	0	0	0
-42	13	80	-24	-40	0	0	0
-56	17	44	-29	0	0	0	0
18	0	03	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



# Αντίστροφος μετασχηματισμός DCT μήκους 8

-70	-64	-61	-64	-69	-66	-58	-50
-72	-73	-61	-39	-30	-40	-54	-59
-68	-78	-58	-9	31	-12	-48	-64
-59	-77	-57	0	22	-13	-51	-60
-54	-75	-64	-23	-13	-44	-63	-56
-52	-71	-72	-54	-54	-71	-71	-54
-45	-59	-70	-68	-67	-67	-61	-50
-35	-47	-61	-66	-60	-48	-44	-44

# Μετατόπιση εικονοστοιχείων κατά $2^7 (+128)$

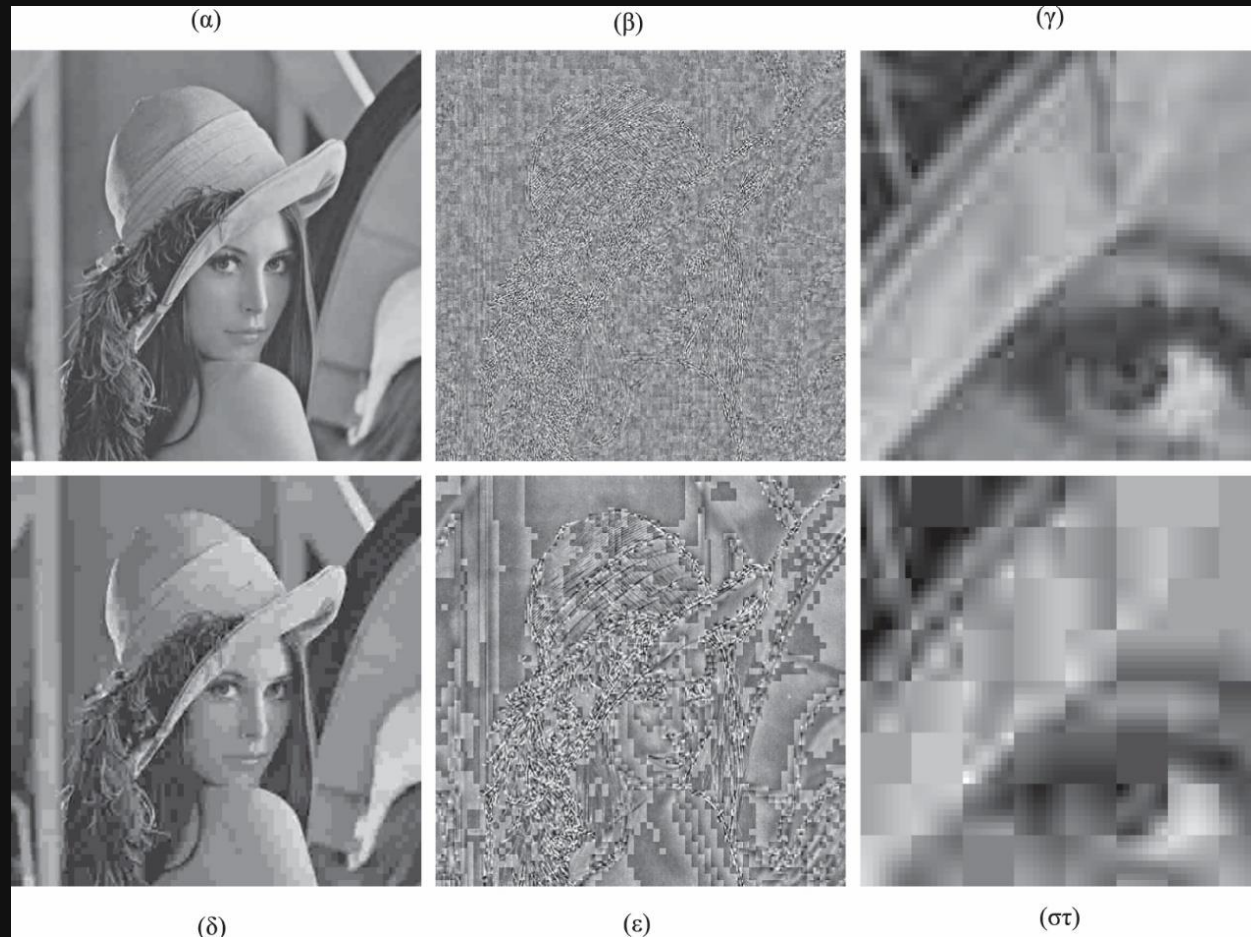
58	64	67	64	59	62	70	78
56	55	67	89	98	88	74	69
60	50	70	119	141	116	80	64
69	51	71	128	149	115	77	68
74	53	64	105	115	84	65	72
76	57	56	74	75	57	57	74
83	69	59	60	61	61	67	78
93	81	67	62	69	80	84	84

## Διαφορές αρχικής – ανακατασκευασμένης εικόνας

-6	-9	-6	2	11	-1	-6	-5
7	4	-1	1	11	-3	-5	3
2	9	-2	-6	-3	-12	-14	9
-6	7	0	-4	-5	-9	-7	1
-7	8	4	-1	6	4	3	-2
3	8	4	-4	2	6	1	1
2	2	5	-1	-6	0	-2	5
-6	-2	2	6	-4	-4	-6	10

Μέσο τετραγωνικό σφάλμα: 5.8 επίπεδα έντασης

# Επίδειξη κωδικοποίησης JPEG



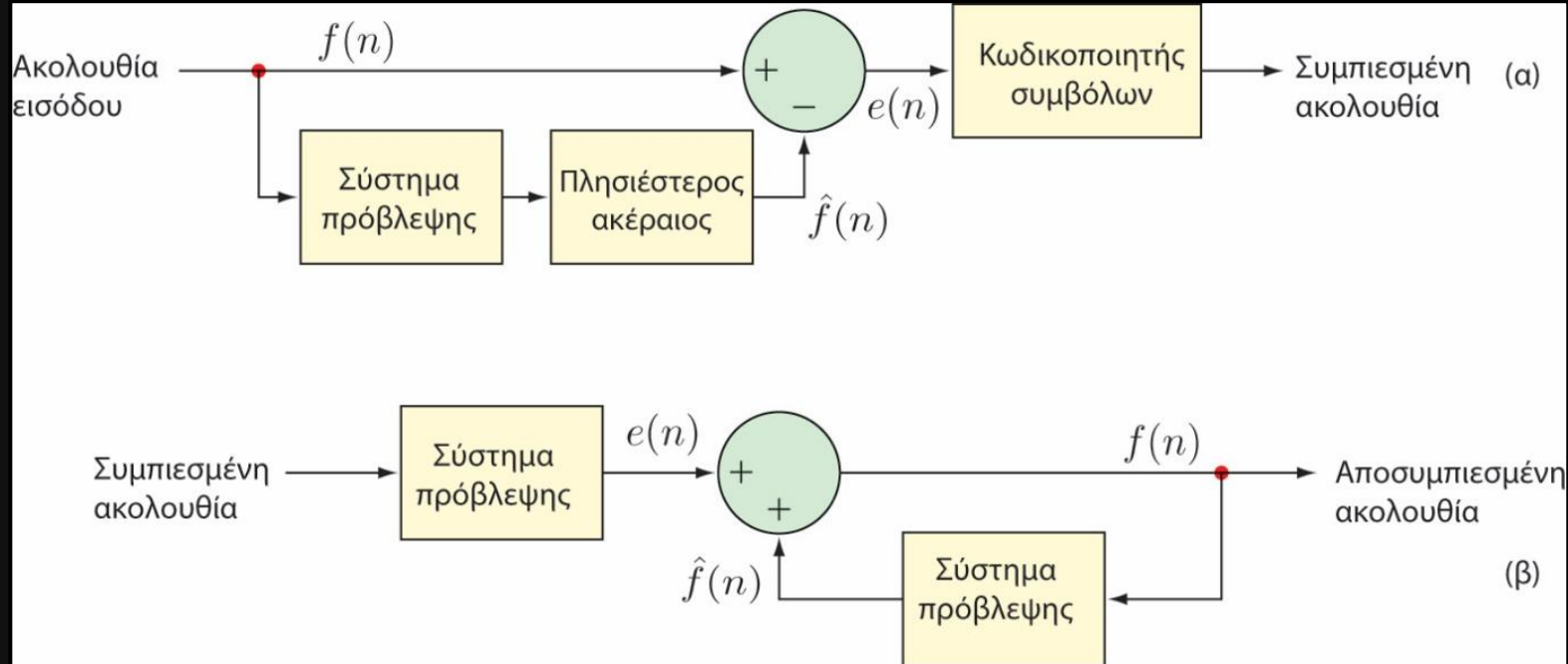
Δύο προσεγγίσεις (25:1 και 52:1) κατά JPEG της δοκιμαστικής εικόνας. Η κάθε γραμμή περιέχει το αποτέλεσμα εφαρμογής των διαδικασιών συμπίεσης και ανακατασκευής, τη μεγεθυμένη διαφορά ανάμεσα στο αποτέλεσμα και την αρχική εικόνα και ένα μεγεθυμένο τμήμα της ανακατασκευασμένης εικόνας.

Προβλεπτική κωδικοποίηση

# Προβλεπτική κωδικοποίηση

- Η προβλεπτική κωδικοποίηση στηρίζεται στην απαλοιφή του πλεονασμού που σχετίζεται με γειτονικά εικονοστοιχεία
- Εξάγει και κωδικοποιεί μόνο τη νέα πληροφορία που περιλαμβάνεται στο κάθε εικονοστοιχείο.
- Η νέα πληροφορία του εικονοστοιχείου ορίζεται ως η διαφορά ανάμεσα στην πραγματική και την προβλεφθείσα τιμή του εικονοστοιχείου
- Μπορεί να είναι με απώλειες ή χωρίς απώλειες.

# Προβλεπτική κωδικοποίηση χωρίς απώλειες



Μοντέλο προβλεπτικής κωδικοποίησης χωρίς σφάλμα. (α) κωδικοποιητής (β) αποκωδικοποιητής

# Προβλεπτική κωδικοποίηση και χωρικός πλεονασμός

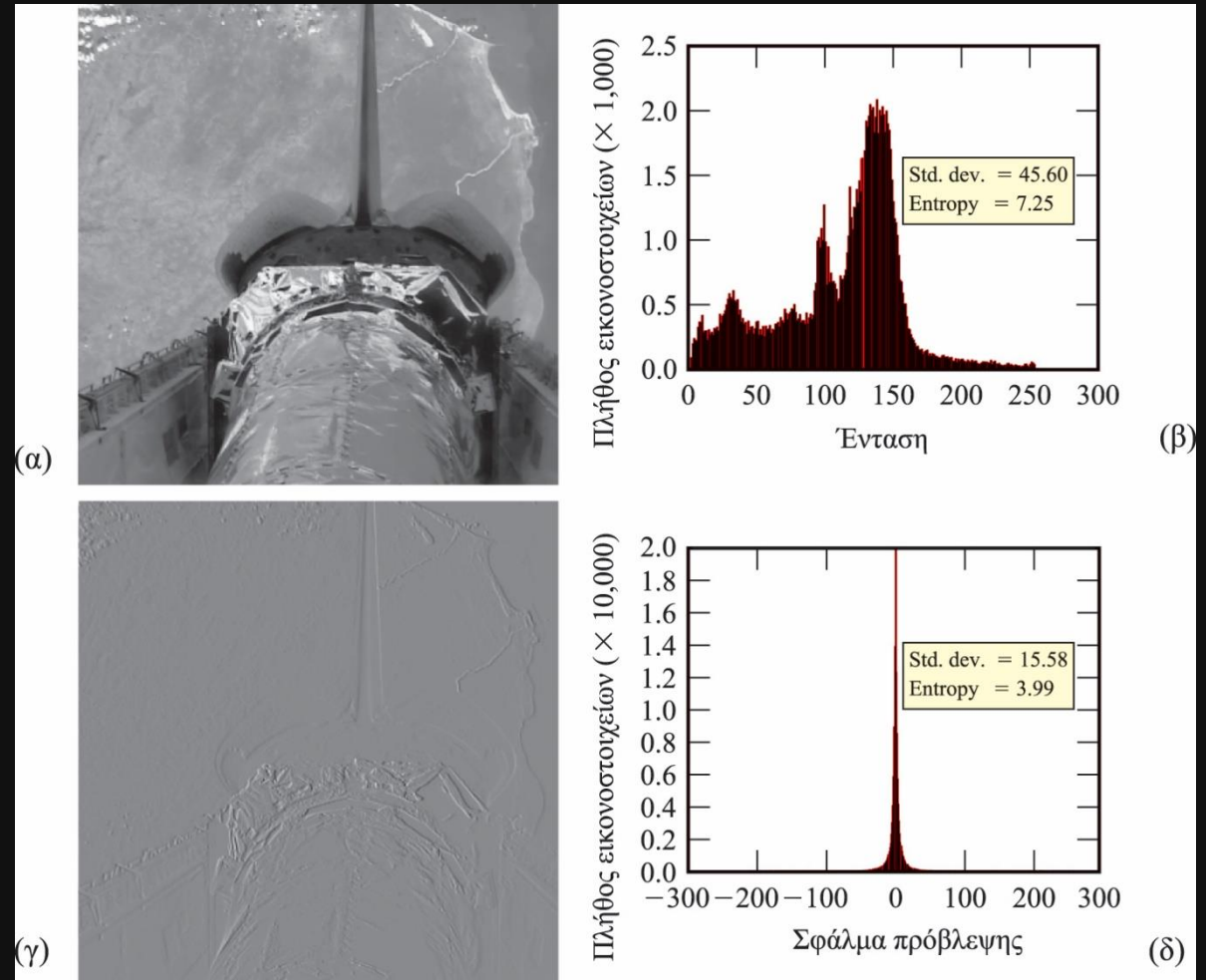
(α) άποψη της Γης από σκάφος σε τροχιά

(β) το ιστόγραμμα έντασης της εικόνας

(γ) εικόνα που αντιστοιχεί στο σφάλμα πρόβλεψης και προκύπτει από την πρόβλεψη 1<sup>ης</sup> τάξης

$$\hat{f}(x, y) = \text{round}[af(x, y - 1)]$$

(δ) ιστόγραμμα του σφάλματος πρόβλεψης

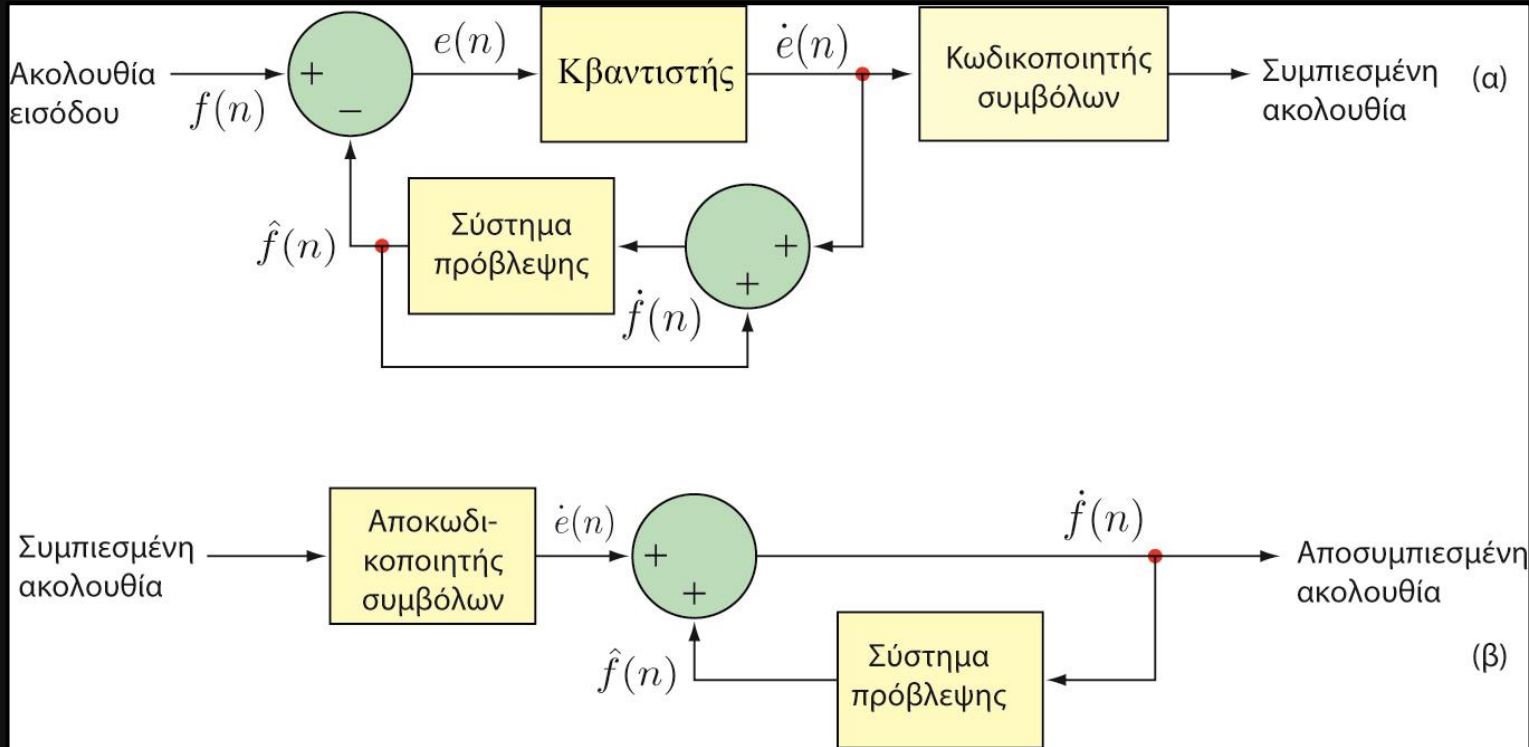




# Προβλεπτική κωδικοποίηση με απώλειες

- Δημιουργείται προσθέτοντας έναν κβαντιστή στο προηγούμενο μοντέλο χωρίς απώλειες
- Ο κβαντιστής αντικαθιστά τη συνάρτηση υπολογισμού του πλησιέστερου ακέραιου
- Χρειάζεται για να απεικονίσει το σφάλμα πρόβλεψης σε μια πεπερασμένη περιοχή εξόδων και στον καθορισμό της ποσότητας της συμπίεσης και της παραμόρφωσης που επιτυγχάνεται

# Προβλεπτική κωδικοποίηση με απώλειες



Μοντέλο προβλεπτικής κωδικοποίησης με απώλειες. (α) κωδικοποιητής (β) αποκωδικοποιητής



<http://www.sippre-group.com>