

# DIP\_06 – Συμπίεση εικόνας - JPEG

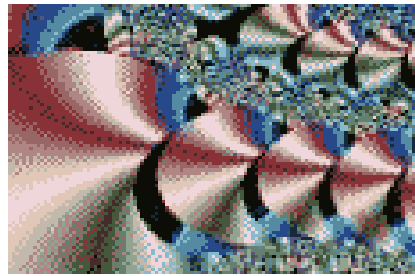
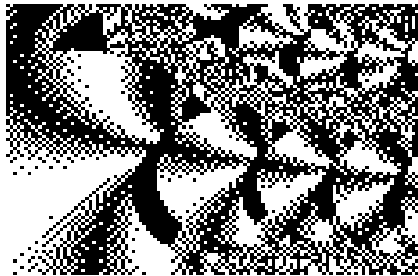
ΤΕΙ Κρήτης

# Συμπίεση εικόνας

- Το μέγεθος μιας εικόνας είναι πολύ μεγάλο
  - π.χ. Εικόνα μεγέθους A4 δημιουργημένη από ένα σαρωτή με 300 pixels ανά ίντσα και με χρήση του RGB μοντέλου (24 bits, 8 ανά χρώμα) καταλαμβάνει 20 Mbytes μνήμης
- Δύο κατηγορίες τεχνικών συμπίεσης
  - Χωρίς απώλειες (lossless): η εικόνα που παράγεται μετά από συμπίεση και αποσυμπίεση της αρχικής είναι της ίδιας ποιότητας
  - Με απώλειες (lossy): η εικόνα που παράγεται μετά από συμπίεση και αποσυμπίεση της αρχικής είναι κατώτερης ποιότητας
- Ο βαθμός αποτελεσματικότητας μιας τεχνικής συμπίεσης είναι συνήθως αντιστρόφως ανάλογος της διατήρησης ποιότητας της αρχικής εικόνας
- Ο τύπος (format) εικόνας σχετίζεται και με τον τρόπο συμπίεσής της
  - GIF (Graphics Interchange Format) – περιορίζεται στην αναπαράσταση εικόνων μέσω 256 χρωμάτων (8 bits of color depth)
  - JPEG (Joint Photographic Experts Group) – εκατομμύρια χρωμάτων (24 bits of color depth)

# Παραδείγματα GIF και JPEG

Εικόνα με 115 x 87 pixels



2 χρώματα  
1 bit  
GIF  
1,329 bytes

16 χρώματα  
4 bit  
GIF  
4,407 bytes

256 χρώματα  
8 bit  
GIF  
8,822 bytes

16,777,216 χρώματα  
24 bit  
JPEG  
4,321 bytes

Μέγιστο  
όριο GIF

Καλύτερη συμπίεση

Bit depth (βάθος ψηφίου) = αριθμός bit που χρησιμοποιείται για κάθε pixel  
Color depth (βάθος χρώματος) = αριθμός χρωμάτων που χρησιμοποιείται για  
κάθε pixel

# GIF

- Αναπτύχθηκε από την CompuServe με στόχο τη διευκόλυνση της ανταλλαγής εικόνων μέσω δικτύου.
- Υποστηρίζει συμπίεση εικόνας που βασίζεται
  - στο αλγόριθμο LZW (Lempel-Ziv-Welch). Ο αλγόριθμος αυτός ανήκει στην κατηγορία των διανυσματικών τεχνικών συμπίεσης και λειτουργεί χωρίς απώλειες. Ο λόγος συμπίεσης (compression ratio) που επιτυγχάνει είναι 4:1 αλλά περιορίζεται σε εικόνες βάθους χρώματος των 8bits (256 χρώματα).
- Προτερήματα:
  - λειτουργεί χωρίς απώλειες για εικόνες βάθους 8 bits
  - είναι ιδανικός για εικόνες με πολλές ακμές και γωνίες όπως γραμμικά σχέδια.
  - χρησιμοποιείται ευρέως και ελεύθερα
- Μειονεκτήματα:
  - δεν είναι κατάλληλο για εικόνες με πολλά χρώματα
  - οι λόγοι συμπίεσης είναι μικροί και δεν μπορούν να ανταλλαχθούν με ποιότητα εικόνας
  - δεν μπορεί να χρησιμοποιηθεί για κινούμενη εικόνα
  - δεν είναι ανεξάρτητος από την ανάλυση της εικόνας

# Αρχεία PNG

Έρθαν για να ανταγωνιστούν τα αρχεία GIF είναι μια patent-free απάντηση στην απαίτηση να πληρώνουν άδεια χρήσης όσοι κατασκευάζουν προγράμματα με έξοδο αρχεία gif. Πράγματι είναι καλύτερα κατά 25% σε συμπίεση από τα αρχεία gif υποστηρίζουν 48-bit truecolor και 16-bit grayscale φωτογραφίες. Η συμπίεση είναι η καλύτερη δυνατή χωρίς να έχουμε απώλεια δεδομένων

# Άλλοι τύποι αρχείων εικόνας

- **BMP**
  - Bitmap
  - τύπος που υποστηρίζουν τα MS Windows
  - ίσως ο πιο απλό τύπος - απλώς περιγράφει το τρόπο με τον οποίο θα αποθηκευτούν τα bits της εικόνας σε ένα αρχείο
  - δεν προσφέρει καμιά δυνατότητα συμπίεσης.
- **TIFF**
  - Tagged Image File Format
  - αναπτύχθηκε από τις εταιρείες Aldus και Microsoft
  - χρησιμοποιείται ευρέως στην ανταλλαγή εικόνων
  - αποτελείται από ένα σύνολο εικόνων, με μια κεφαλίδα (header) να καθορίζει τις παραμέτρους της κωδικοποίησης.
  - Δεν περιλαμβάνει αλγορίθμους συμπίεσης.

# Το πρότυπο JPEG

- Σχεδιάστηκε από την ομάδα Joint Photographic Expert Group σε συνεργασία με την Διεθνή Ένωση Τηλεπικοινωνιών (ITU-TS)
- Μπορεί να δώσει διαφορετικό αποτέλεσμα ανάλογα με τις απαιτήσεις που έχουμε για την ποιότητα της εικόνας και το λόγο συμπίεσης
  - 10:1 έως 20:1 – υψηλή ποιότητα
  - 30:1 έως 50:1 – μέτρια ποιότητα
  - 60:1 έως 100:1 – κακή ποιότητα
- Πετυχαίνει συμπίεση του αρχικού μεγέθους ακόμη και μικρότερο του ενός δέκατου χωρίς ορατές απώλειες στην ποιότητα της εικόνας.
- Βασίζεται στη
  - Μείωση της χρωματικής πληροφορίας
  - Συμπίεση των πληροφοριών φωτεινότητας και απόχρωσης

# JPEG

- Χρησιμοποιείται κυρίως για (ασπρόμαυρες ή έγχρωμες) εικόνες, αλλά και για βίντεο (Motion JPEG)
- Υλοποιείται μέσω είτε μόνο ειδικού λογισμικού είτε συνδυασμού λογισμικού και hardware (ειδικών καρτών)
  - Ο χρήστης μπορεί να ορίσει την ποιότητα της εικόνας που θα παραχθεί, το χρόνο επεξεργασίας της συμπίεσης και το μέγεθος της συμπιεσμένης εικόνας
- Εκτός από τις γενικές απαιτήσεις που πρέπει να ικανοποιεί κάθε τεχνική συμπίεσης, ισχύουν επίσης τα εξής:
  - Η τεχνική είναι ανεξάρτητη του μεγέθους ή του περιεχομένου της εικόνας και εφαρμόσιμη σε οποιοδήποτε είδος εικόνας
  - Η αναπαράσταση των χρωμάτων πρέπει να είναι ανεξάρτητη από τη συγκεκριμένη υλοποίηση της τεχνικής
  - Ο βαθμός συμπίεσης και η ποιότητα της εικόνας που επιτυγχάνεται πρέπει να είναι όσο το δυνατόν καλύτερα



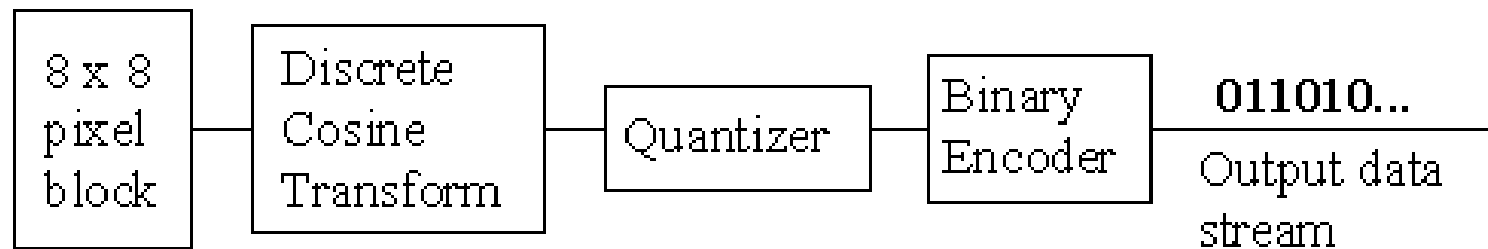
# Η Τεχνική JPEG

- Λειτουργεί καλλίτερα με φυσικές εικόνες
- Στη συνέχεια περιγράφεται η συμπίεση εικόνων Grayscale που
- Επεκτείνεται εύκολα σε έγχρωμες (αρκεί να θεωρηθούν σαν επίθεση τριών Grayscale - R, G & B)

# JPEG συμπίεση

1. Χωρίζεται η εικόνα σε blocks των 8X8 pixels
2. Υπολογίζεται για κάθε block ο Διακριτός Συνημιτονικός Μετασχηματισμός
3. Γίνεται κβαντοποίηση με χρήση της σχετικής μήτρας – Φάση απωλειών – Υψηλή συμπίεση
4. Γίνεται κωδικοποίηση των συντελεστών με χρήση κωδικοποίησης μεταβλητού μήκους και τα δεδομένα αποθηκεύονται σε αρχείο \*.jpg

# JPEG συμπίεση



# Ο Διακριτός Συνημιτονικός Μετασχηματισμός

- Βοηθά τον διαχωρισμό της εικόνας σε περιοχές (υπό-περιοχές φάσματος) με διαφορετική αξία (σε σχέση με την ποιότητα της εικόνας)
- Μοιάζει με τον μετασχηματισμό Fourier
- Μετασχηματίζει ένα σήμα ή μια εικόνα από την περιοχή του χώρου στην περιοχή της συχνότητας

# Ο Διακριτός Συνημιτονικός Μετασχηματισμός

$$B(k_1, k_2) = \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} 4 \cdot A(i, j) \cdot \cos\left[\frac{\pi \cdot k_1}{2 \cdot N_1} \cdot (2 \cdot i + 1)\right] \cdot \cos\left[\frac{\pi \cdot k_2}{2 \cdot N_2} \cdot (2 \cdot j + 1)\right]$$

# Ο Διακριτός Συνημιτονικός Μετασχηματισμός

- Έστω η εικόνα εισόδου με  $N_1$  γραμμές και  $N_2$  στήλες
- $A(i,j)$  η ένταση του pixel στη θέση  $(i,j)$
- $B(k_1,k_2)$  ο συντελεστής του ΔΣΜ στη θέση  $(k_1,k_2)$  της μήτρας ΔΣΜ

# Πως λειτουργεί;

- Η είσοδο στον μετασχηματιστή είναι μια μήτρα πραγματικών αριθμών 8X8
- Κάθε στοιχείο της είναι η τιμή Grayscale του συγκεκριμένου pixel
- Για 8 bit pixels έχουμε 256 τιμές
- Η έξοδος του μετασχηματιστή είναι μια μήτρα ακεραίων με τιμές στην περιοχή (-1024,+1023)

# Πως λειτουργεί;

- Για τις περισσότερες εικόνες, το μεγαλύτερο μέρος της ενέργειας του σήματος βρίσκεται στις χαμηλές συχνότητες που εμφανίζονται στην πάνω αριστερή πλευρά της μήτρας ΔΣΜ
- Η κάτω δεξιά πλευρά της μήτρας ΔΣΜ παρουσιάζει τις υψηλές συχνότητες των οποίων συνήθως οι τιμές είναι χαμηλές, ώστε να μπορούν να μηδενιστούν με μικρή ορατή παραμόρφωση στην εικόνα



# Η κβαντοποίηση

- Υπάρχει ένα tradeoff μεταξύ ποιότητας εικόνας και βαθμού κβαντοποίησης
- Ένα πολύ μεγάλο βήμα κβαντοποίησης μπορεί να οδηγήσει σε απαράδεκτη παραμόρφωση της εικόνας
- Ενώ, μικρότερο βήμα οδηγεί σε μικρότερη συμπίεση

# Η κβαντοποίηση

- Το πρόβλημα είναι να εντοπιστεί το ιδανικό βήμα για την μέγιστη συμπίεση με την ελάχιστη παραμόρφωση
- Γίνεται χρήση της ιδιομορφίας του ανθρώπινου ματιού το οποίο παρουσιάζει μικρότερη ευαισθησία στις υψηλές συχνότητες
- Έτσι, χρησιμοποιείται μικρό βήμα στις χαμηλές και μεγάλο βήμα στις υψηλές συχνότητες

# Η μήτρα κβαντοποίησης

- Η μήτρα αυτή διαστάσεων  $8 \times 8$  ορίζει τα βήματα κβαντοποίησης
- Τα βήματα είναι μικρότερα πάνω αριστερά (χαμηλές συχνότητες) και μεγαλύτερα κάτω-δεξιά (μεγάλες συχνότητες)
- Ο κβαντοποιητής διαιρεί τους συντελεστές της μήτρας  $\Delta \Sigma M$  με το αντίστοιχο βήμα και στρογγυλοποιεί την τιμή στον κοντινότερο ακέραιο
- Τα μεγάλα βήματα μηδενίζουν τους μικρούς συντελεστές

# DCT συντελεστές

- Το αποτέλεσμα είναι ότι πολλοί συντελεστές υψηλών συχνοτήτων μηδενίζονται και έτσι είναι πιο εύκολη η κωδικοποίησή τους
- Οι συντελεστές χαμηλών συχνοτήτων δέχονται πολύ μικρές αλλαγές

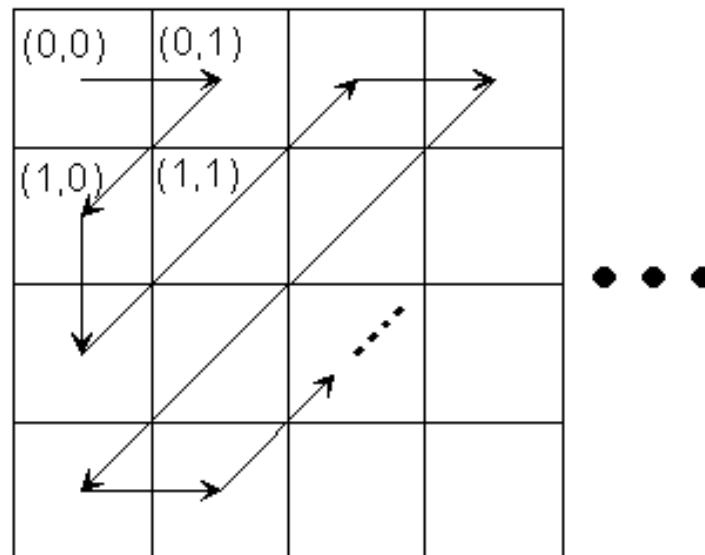
# DCT συντελεστές

- Συνήθως, μετά την κβαντοποίηση, πολλοί συντελεστές μηδενίζονται και γι' αυτό χρησιμοποιείται κωδικοποίηση μεταβλητού μήκους
- Για κάθε μη μηδενικό συντελεστή της μήτρας ΔΣΜ, κωδικοποιούνται τα μηδενικά που προηγούνται, ο αριθμός των bits που απαιτούνται για την παράσταση του συντελεστή και τέλος η τιμή του ίδιου του συντελεστή

# Zig-zag

Για καλλίτερη ομαδοποίηση των μηδενικών γίνεται καταγραφή των συντελεστών με μέθοδο zigzag

DCT  
MATRIX:



The sequence continues  
for the entire 8 by 8 block.

# Κωδικοποίηση

- Τα προηγούμενα του συντελεστή μηδενικά και ο αριθμός των bits που απαιτούνται για την παράσταση του συντελεστή δημιουργούν ένα ζευγάρι
- Κάθε ζευγάρι αναπαρίσταται με μια λέξη με χρήση κωδικοποίησης μεταβλητού μήκους (Huffman, Shannon-Fano ή Αριθμητικής Κωδικοποίησης)

# Κωδικοποίηση

- Στη συνέχεια κωδικοποιούνται τα ζευγάρια και οι τιμές των συντελεστών
- Στο τέλος κάθε block τοποθετείται μια ειδική σειρά χαρακτήρων end-of-block
- Στο τέλος όλων των block τοποθετείται ένας δείκτης end-of-file



# Επιδόσεις JPEG

- Στη πράξη χρησιμοποιούνται διάφοροι συντελεστές ποιότητας για την συμπίεση JPEG
- Συντελεστής ποιότητας 100% αντιστοιχεί στην ελάχιστη παραμόρφωση

*Table 1. Compression Ratios for JPEG Quality Factors*

Original size of color image (Kb)				313.076
Original size of B/W image (Kb)				104.437
Quality Factor	Color JPEG		B/W JPEG	
	File Size (Kb)	Comp. Ratio	File Size (Kb)	Comp. Ratio
75	23.039	13.59	21.02	4.97
20	8.457	37.02	7.599	13.74
5	4.009	78.09	3.257	32.07
3	3.268	95.80	2.522	41.41

# JPEG

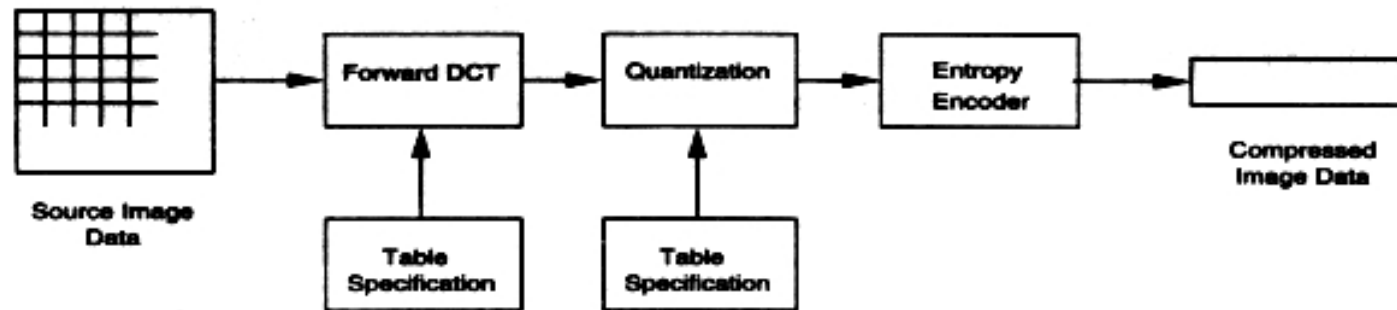
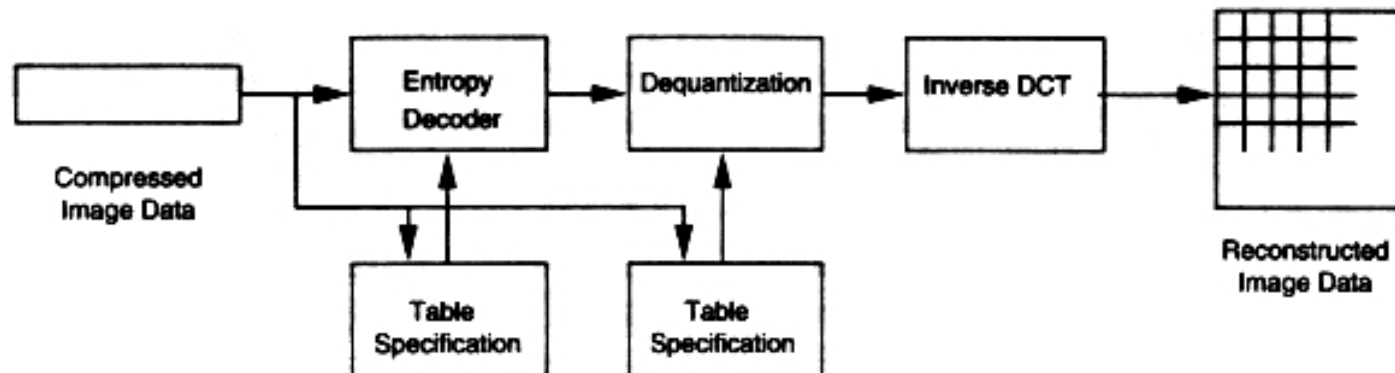


Figure 3-11 Processing Steps of Baseline Mode Encoder



# Δείκτες συμπίεσης

$$\text{Compression factor} = \frac{\text{Bits original image}}{\text{Bits compressed image}}$$

$$\text{Bits/pixel} = \frac{\text{Bits compressed image}}{\text{Number of pixels}}$$

# Δείκτες συμπίεσης



Original image  
256 x 256 x 8 bits



Compressed image  
40.000 bits

$$\text{Bits/pixel} = \frac{40.000}{256 \times 256} = 0.61 \text{ bpp}$$

$$\text{C. F.} = \frac{8 \text{ bpp}}{0.61 \text{ bpp}} = 13.1$$

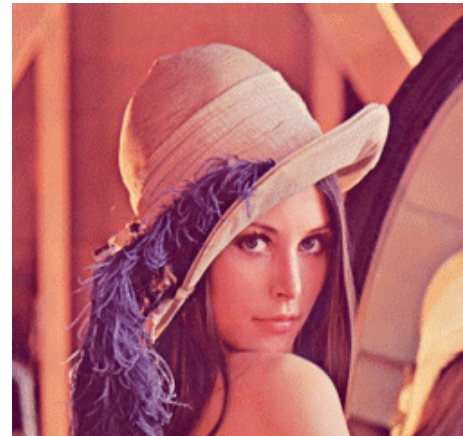
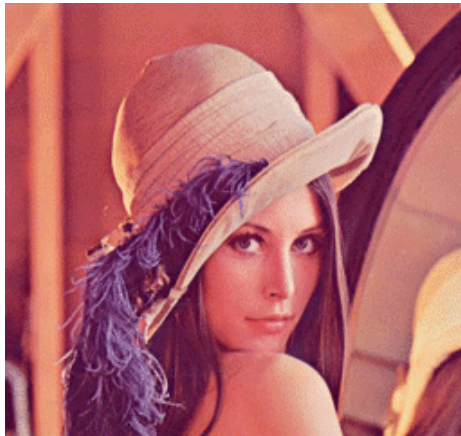
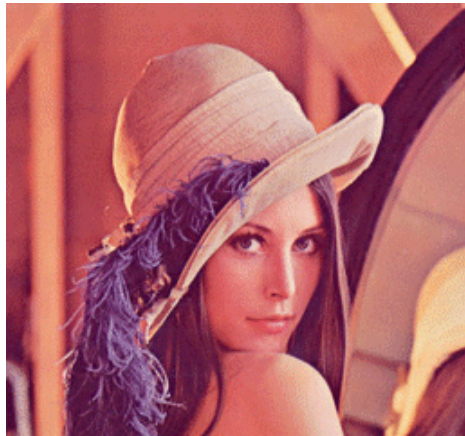
# Δείκτες συμπίεσης

JPEG (Joint Photographic Experts Group)

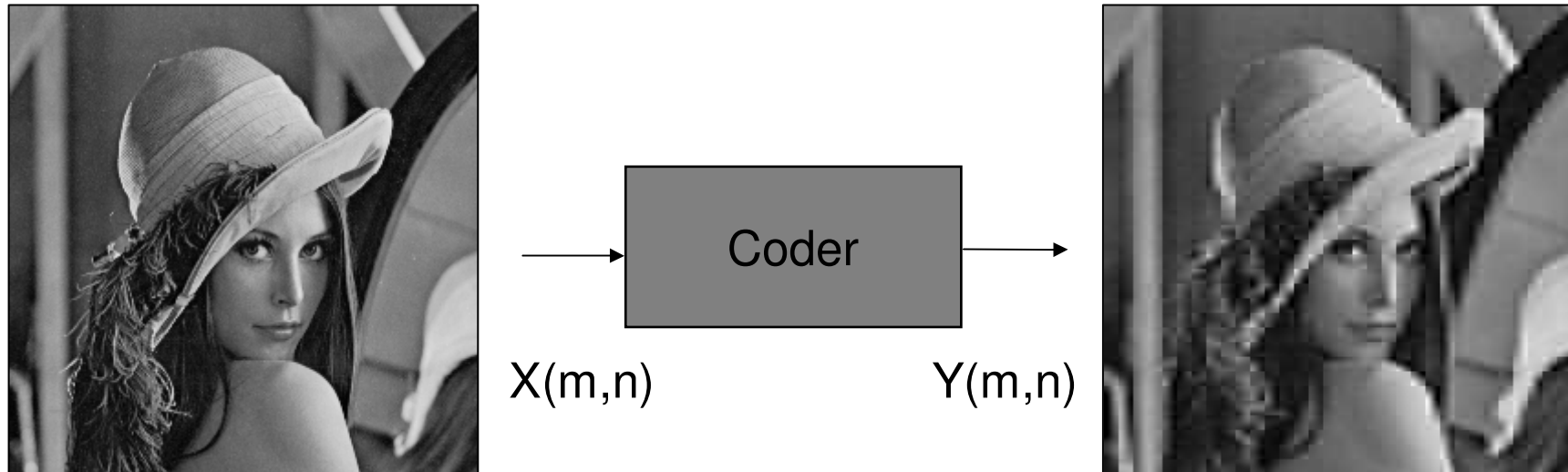
Original lena

5.75:1

22.35:1



# Δείκτες ποιότητας



Subjective measure  $\Rightarrow$

Opinions over a scale of 5

Objective measure  $\Rightarrow$

$$\text{PSNR(dB)} = 10 \log_{10} \frac{255^2}{\text{MSE}}$$

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (y_{ij} - x_{ij})^2$$