

DIP_01 – Εισαγωγή στην ψηφιακή εικόνα

ΤΕΙ Κρήτης

Πληροφορίες Μαθήματος

- Διαλέξεις
 - Πέμπτη 12:15 – 15:00
 - Αιθουσα Γ7
- Διδάσκων: Δ. Κοσμόπουλος
- Γραφείο: Κ23-0-15 (ισόγειο κλειστού γυμναστηρίου)
- Ωρες γραφείου Τε 16:00 – 18:00, Πε 10:00-12:00, και κατόπιν συνεννοήσεως
- Επικοινωνία dkosmo@ie.teicrete.gr

Πληροφορίες Μαθήματος

- Προαπαιτούμενες γνώσεις
 - Προγραμματισμός
 - Γραμμική άλγεβρα
 - Αλγόριθμοι
- Σχετικά μαθήματα
 - Επεξεργασία εικόνας
 - Αναγνώριση προτύπων
 - Τεχνητή νοημοσύνη

Αξιολόγηση (ενδεικτικά)

- Ασκήσεις 15%
- Πρόοδος 35%
- Τελική εξέταση 50%

Η εκπόνηση όλων των εργασιών όπως και η συμμετοχή σε όλες τις εξετάσεις είναι υποχρεωτική, καθώς κρίνεται απολύτως αναγκαία για την επιτυχή ολοκλήρωση του μαθήματος.

Ψηφιακή εικόνα

- Ψηφιακή εικόνα = αναλογική εικόνα μετά από δειγματοληψία στο χώρο (x και y διευθύνσεις)
- Αναπαριστάνεται από έναν ή περισσότερους 2D πίνακες
- Μπορεί να είναι δυαδική (binary), επιπέδων γκρι, ή έγχρωμη
- Ο όγκος οπτικών πληροφοριών είναι τεράστιος
 - Αποδοτική αποθήκευση ή/και μετάδοση
 - Αποδοτικοί αλγόριθμοι επεξεργασίας και ανάλυσης
- Η εξαγωγή συμπερασμάτων σε εφαρμογές απαιτεί:
 - Ταχύτητα
 - Αξιοπιστία

Ψηφιακή επεξεργασία εικόνων

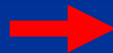
- Ένα σύστημα επεξεργασίας εικόνων ...
 - Καταγράφει εικόνες (μία προς μία ή/και ακολουθία)
 - Επεξεργάζεται εικόνες με ψηφιακό τρόπο
 - Εξάγει νέες εικόνες
- Παραδείγματα επεξεργασιών εικόνων
 - Βελτίωση ποιότητας
 - Φιλτράρισμα θορύβου
 - Αποκατάσταση υποβαθμισμένων εικόνων
 - Συμπίεση

Ψηφιακή ανάλυση εικόνων

- Ένα σύστημα ανάλυσης εικόνων
 - Λαμβάνει εικόνες
 - Περιγράφει και αναγνωρίζει το περιεχόμενο τους
 - Εξάγει περιγραφές
- Προσπαθεί να μιμηθεί την ανθρώπινη όραση
- Χρησιμοποιείται κυρίως σε εξειδικευμένες εφαρμογές (βιομηχανία, ιατρική, γεωλογία, κίνηση οχήματος κ.λπ.) αλλά σύντομα θα δούμε και εφαρμογές της καθημερινής ζωής

Επεξεργασία & Ανάλυση

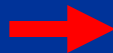
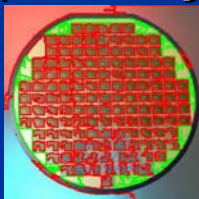
Επεξεργασία εικόνας :



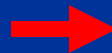
Αλγόριθμοι
επεξεργασίας
εικόνας



Ανάλυση εικόνας :



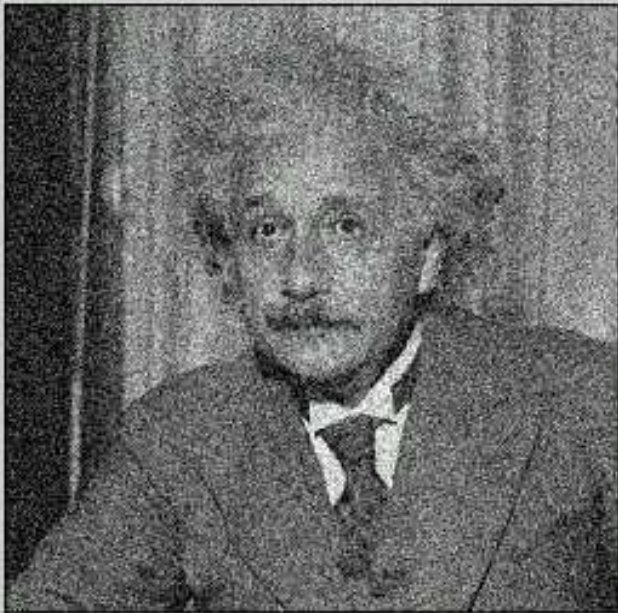
Αλγόριθμοι
ανάλυσης
εικόνας



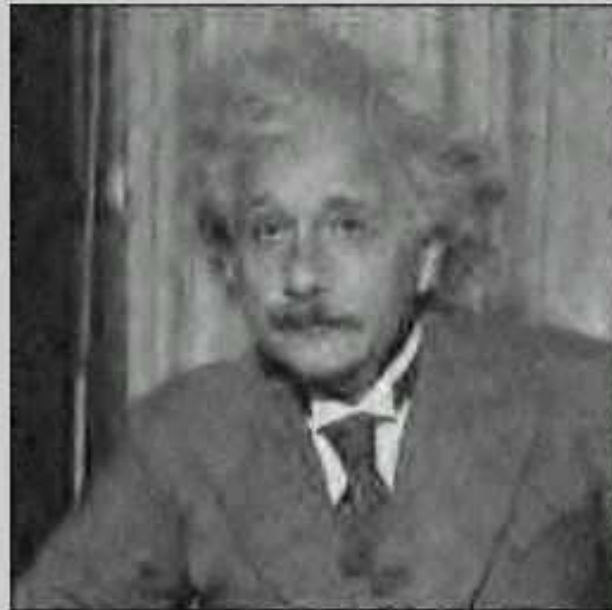
254 πυρήνες

Εφαρμογές – Αποθορυβοποίηση (denoising)

Noisy Image



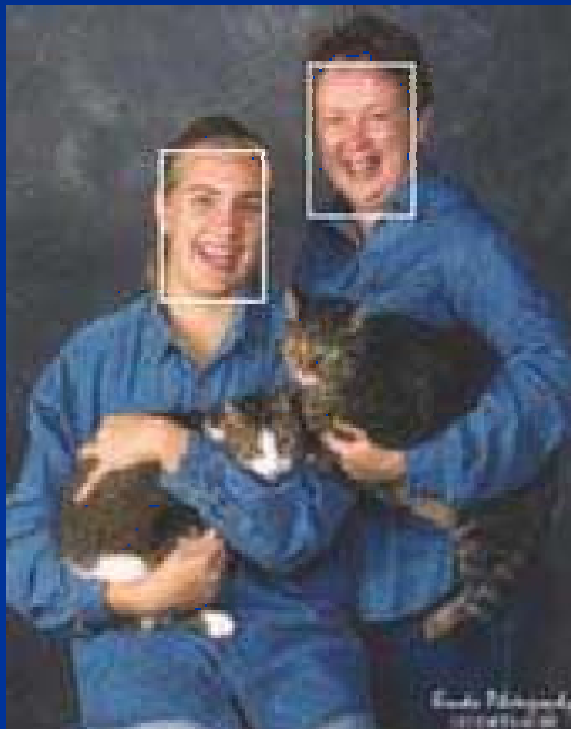
BayesJoint Estimator - QMF



Εφαρμογές – Τμηματοποίηση (segmentation)



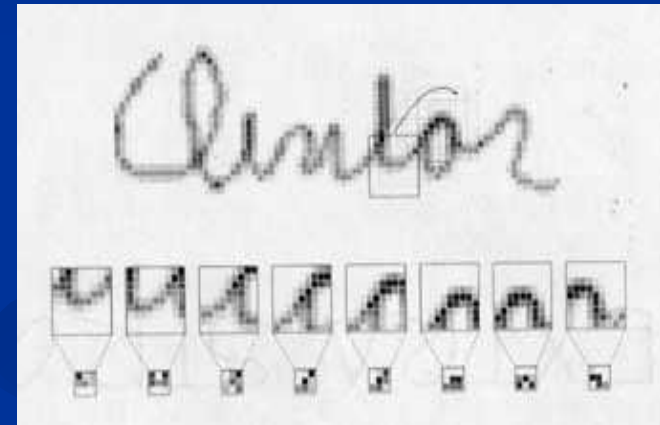
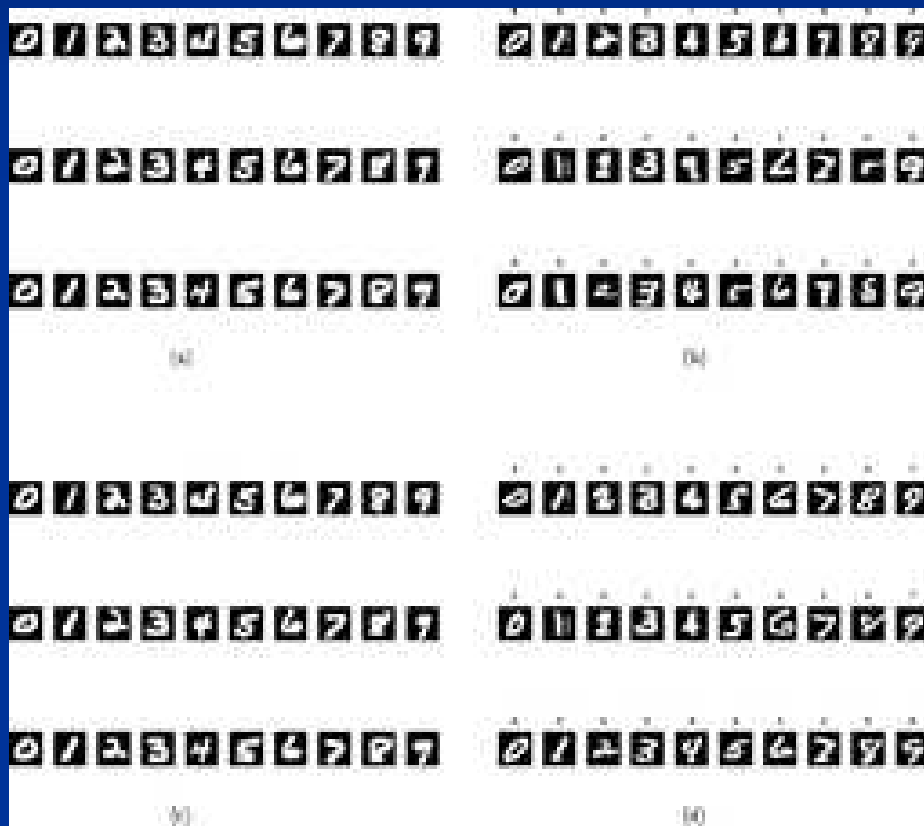
Εφαρμογές – Ανίχνευση προσώπου (face detection)



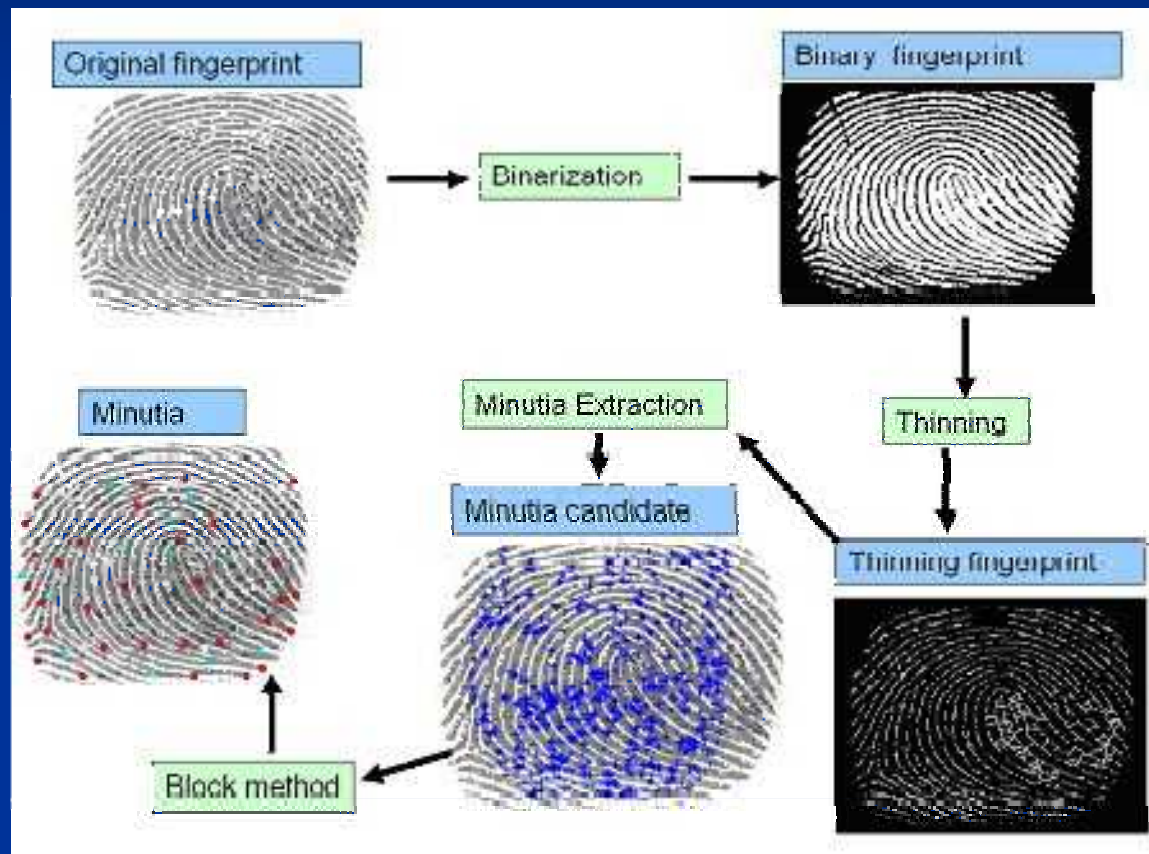
Εφαρμογές – Παρακολούθηση προσώπου (face tracking)



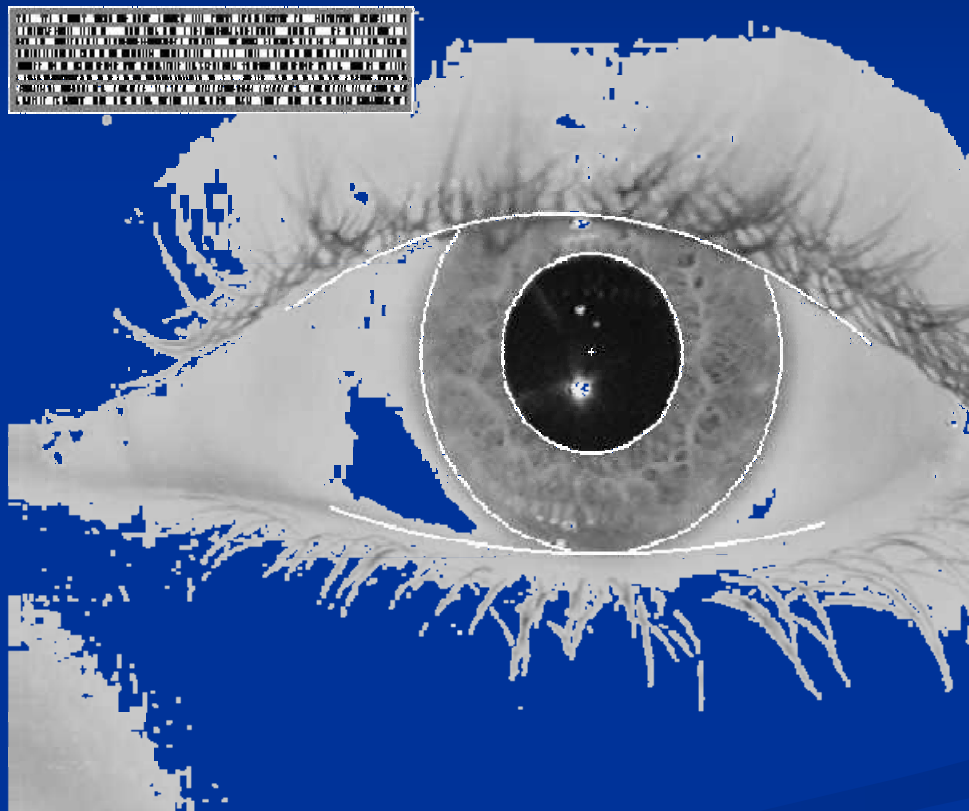
Εφαρμογές - Αναγνώριση χαρακτήρων (character recognition)



Εφαρμογές – Ταυτοποίηση δακτυλικών αποτυπωμάτων



Εφαρμογές – Ταυτοποίηση ίριδας



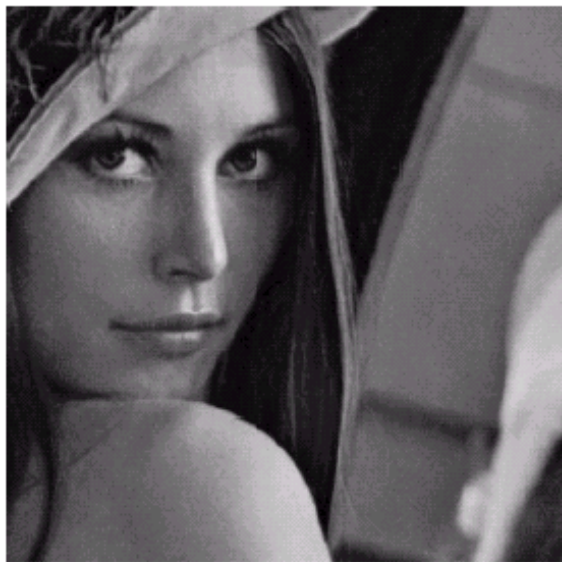
Σχετικές Επιστημονικές περιοχές

- Ψηφιακή επεξεργασία σήματος
 - Η εικόνα είναι 2Δ σήμα
 - Θεωρία και τεχνικές επεξεργασίας
- Αναγνώριση προτύπων
 - Κατηγοριοποίηση αντικειμένων που απεικονίζονται σε εικόνες
- Γραφική
 - Ψηφιακή σύνθεση εικόνας από συμβολική περιγραφή

Σχετικές Επιστημονικές περιοχές

- Τεχνητή νοημοσύνη
 - Κατανόηση εικόνας από μηχανή (ανάλυση εικόνας)
- Επικοινωνίες, Διαδίκτυο
 - Αποδοτική χρήση τηλεπικοινωνιακών πόρων (κωδικοποίηση)
 - Προστασία δικαιωμάτων (watermarking)
- Βάσεις δεδομένων
 - Αποθήκευση (επεξεργασία εικόνας)
 - Ανάκτηση (ανάλυση εικόνας)

Εικόνες δοκιμής



Lena Sjooblom



Το 1972, οι φοιτητές του ερευνητικού τμήματος στο Πανεπιστήμιο της Νότιας Καλιφόρνιας, προσπαθούσαν να αναπτύξουν αλγόριθμους συμπίεσης προκειμένου να επιτύχουν ταχύτερη μετάδοση εικόνων στο τότε Arpanet, το μετέπειτα Internet.

Η ώρα περνούσε, κανείς δεν αποφάσιζε να αναλάβει την πρωτοβουλία, ώσπου ένας από τους παρευρισκόμενους, ο οποίος παραμένει ανώνυμος μέχρι σήμερα, έβγαλε το γνωστό περιοδικό "Playboy" και "σάρωσε" τη Miss November 1972 Lena Sjooblom.



Με τη πάροδο του χρόνου, η εικόνα έγινε πρότυπο σε όλα τα ερευνητικά κέντρα επεξεργασίας εικόνας. Τον Μάϊο του 1997, 25 χρόνια μετά, κατάφεραν να εντοπίσουν τη Lena Sjooblom στη Σουηδία, παντρεμένη με 2 παιδιά, και να της εξηγήσουν πόσο πολύ έχει συμβάλλει στην έρευνα, εν αγνοία της.

Full Lena image: http://www.lenna.org/full/len_full.html

Αναλογική vs Ψηφιακή εικόνα

- **Αναλογική εικόνα:** 2D συνάρτηση της φωτεινότητας $f(x,y)$ σε κάθε σημείο (x,y)
- **Ψηφιακή εικόνα:** 2D διακριτή (discrete) αναπαράσταση της αναλογικής εικόνας, μετά από δειγματοληψία των τιμών $f(x,y)$ σε 2D πίνακα $I(x,y)$

$$x = 1, \dots, N,$$

$$y = 1, \dots, M$$

για $N \times M$ εικόνα. Συνήθως τα N, M είναι δυνάμεις του 2

Ψηφιακή εικόνα

- Εικόνα διαστάσεων $N \times M$ = διδιάστατος πίνακας ακεραίων $I(i,j)$
 - $i = 1, \dots, N$, $j = 1, \dots, M$, $0 \leq I(i,j) \leq G-1$
 - G : πλήθος αποχρώσεων: $G=2^m$
 - m : αριθμός bit/pixel (πχ 8 bit/pixel σημαίνουν 256 διαφορετικά επίπεδα: $G=0..255$)
 - πίνακας $I(i,j)$ = διακεκριμένη συνάρτηση έκφρασης την ένταση φωτεινότητας κάθε εικονοστοιχείου (pixel)

Pixel (ΕΙΚΟΝΟΣΤΟΙΧΕΙΟ)

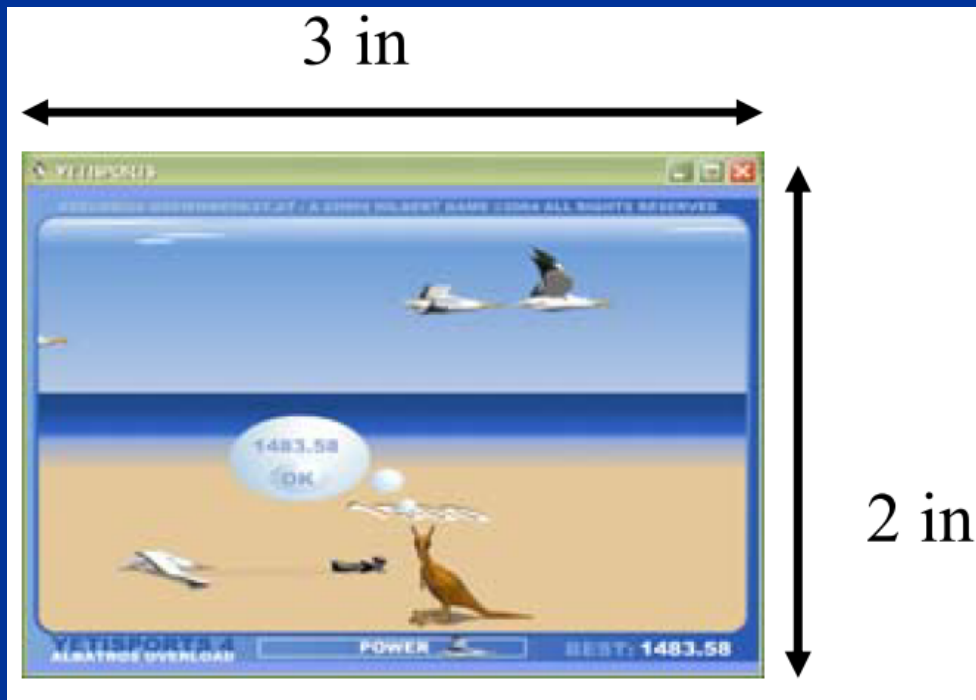
- Κάθε στοιχείο του πίνακα των τιμών μετά την δειγματοληψία λέγεται **pixel** (picture element). Αναπαριστάται με 2D πίνακα $I(x,y)$
- Οι εικόνες μπορούν να είναι
 - Δυαδικές, ασπρόμαυρες - Binary
 - Γκρι επιπέδων - Grayscale
 - Έγχρωμες - Colored

Χαρακτηριστικά εικονοστοιχείου

1. Θέση: Συντεταγμένες πάνω στο ορθογώνιο πλέγμα
2. Τιμή (χρώμα)
3. Μέγεθος (προκύπτει από την ανάλυση p_{pi})
4. Βάθος: Αριθμός των πιθανών χρωμάτων που μπορεί να έχει.

(1) Μέγεθος εικονοστοιχείων ppi (pixels per inch)

- Διάσταση εικόνας: Είναι ο αριθμός των εικονοστοιχείων στο μήκος και το πλάτος μιας εικόνας.



Αριθμός pixels στην
εικόνα:

$$(2 \times 72) \times (3 \times 72) = 31104$$

Pixels: τετραγωνικά

72 ppi



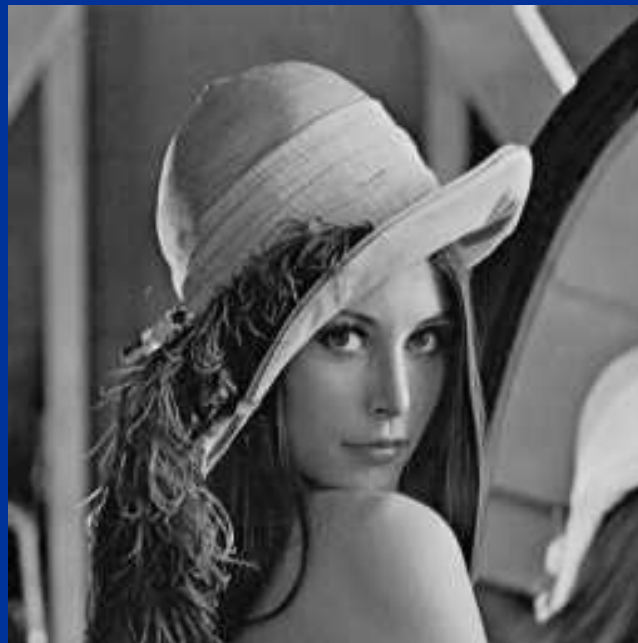
Σε 1 in -> 72 pixels

(2) Βάθος εικόνας

- Σημαίνει πόσα bits χρησιμοποιούνται για την αποθήκευση της πληροφορίας
- (Θυμηθείτε το βήμα κβάντισης που είπαμε στα βήματα της δειγματοληψίας)
- Τιμές που λαμβάνει:
 - 1-bit
 - 8-bit
 - 16-bit
 - 24-bit (Υψηλής ανάλυσης)

ΕΙΚΟΝΕΣ ΕΠΙΤΠΕΔΩΝ γκρι

- Για γκρι εικόνα, με m επίπεδα του γκρι, οι τιμές $I(x,y)$ είναι από 0 ως 2^m .
- Συνήθως $m=8$, οι $I(x,y)$ είναι από 0 ως 255



Βάθος εικόνας 1-bit

- Ομαδοποίηση χρωμάτων:
- Ανοιχτά χρώματα στο λευκό ενώ τα σκούρα στο μαύρο.
- Είναι εικόνες με χαμηλή πληροφορία, μικρό μέγεθος.



Βάθος εικόνας 4-bit

Τα εικονοστοιχεία
είναι μαύρα (0),
λευκά (15) ή 14
ενδιάμεσες τιμές
του γκρι.

0-> 0000

1-> 0001

.....

15-> 1111



Βάθος εικόνας 8-bit

Τα εικονοστοιχεία
είναι μαύρα (0),
λευκά (255) ή 254
ενδιάμεσες τιμές
του γκρι.

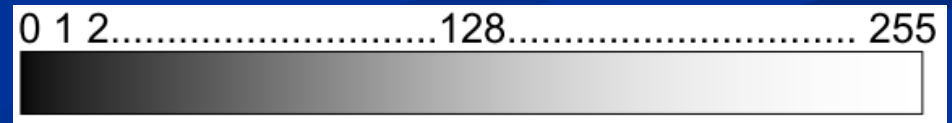
0-> 00000000

1-> 00000001

.....

254-> 11111110

255-> 11111111



Διάφορα βάθη χρώματος



4 bits



2 bits



3 bits



1 bit

Δυαδικές εικόνες

- Δυαδικές εικόνες αναπαριστούν θέση, μορφή (shape) αντικειμένων Αναγνώριση χαρακτήρων, υπογραφών, δακτυλικού αποτυπώματος



Έγχρωμες εικόνες

- Έγχρωμες: 3 πίνακες για κάθε βασικό χρώμα, $I_c(x,y)$, $c = 1,2,3$, οπότε το χρώμα σε κάθε σημείο δίνεται από 3 χρωματικές συνιστώσες:

$$\text{Color}(x,y) = [I_1(x,y), I_2(x,y), I_3(x,y)]$$

Στο RGB, $I_1=\text{Red}$, $I_2=\text{Green}$, $I_3=\text{Blue}$



Βάθος έγχρωμης εικόνας 8-bit

Τα 256 χρώματα δεν
είναι ικανοποιητικά
για αποτύπωση
έγχρωμων εικόνων.
Απαιτείται μεγαλύτερο
βάθος



Βάθος έγχρωμης εικόνας 24-bit

- Κάθε pixel περιγράφεται από 3 σετ 8-bit αριθμών.
- Κάθε σετ αντιπροσωπεύει 3 κανάλια (Red-Green-Blue)



Όγκος μιας εικόνας

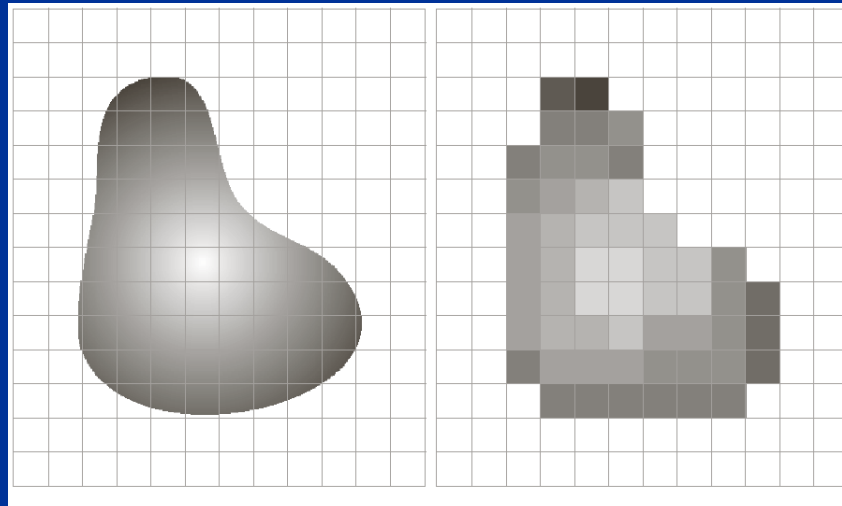
- Εικόνα $N \times M$, πλήθους αποχρώσεων $G=2^m$

$$b = N \times M \times m$$

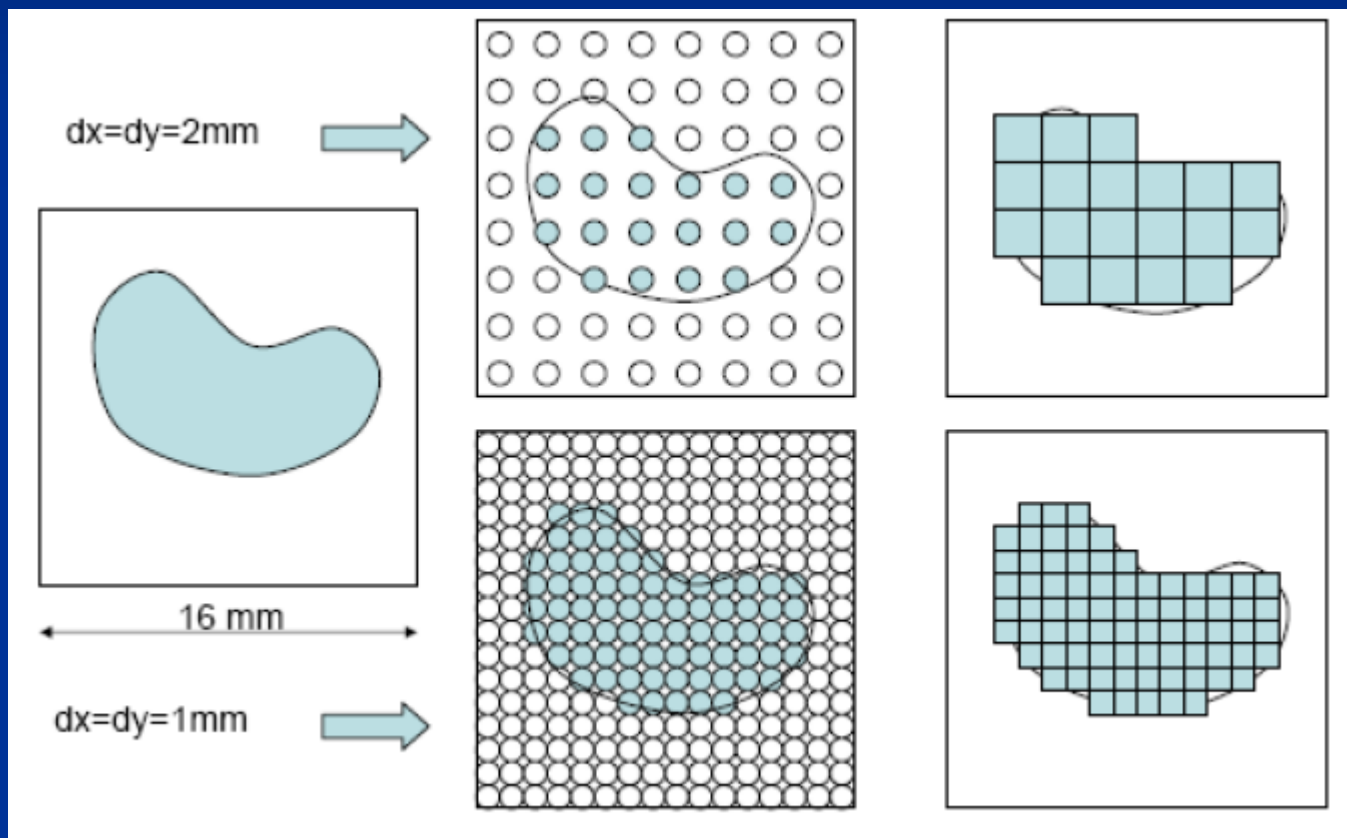
- m : βάθος χρώματος
 - εύρος χρωματικής πληροφορίας εικονοστοιχείου

Τύπος εικόνας	N	M	m	bits	bytes
δυναδική	100	100	1	10.000	1.250
αποχρώσεων γκρι	100	100	8	80.000	10.000
έγχρωμη	100	100	24	240.000	30.000

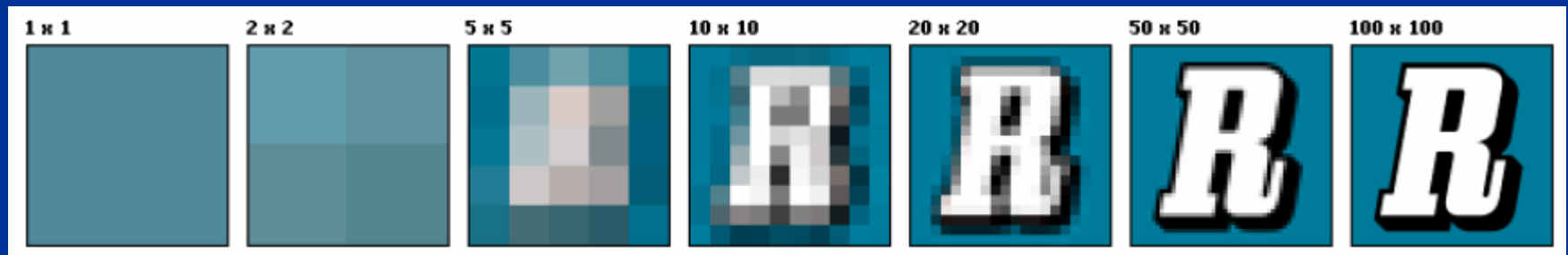
Αναλογική εικόνα – Ψηφιακή εικόνα (δειγματοληψία & κβαντισμός)



Δειγματοληψία



Δειγματοληψία

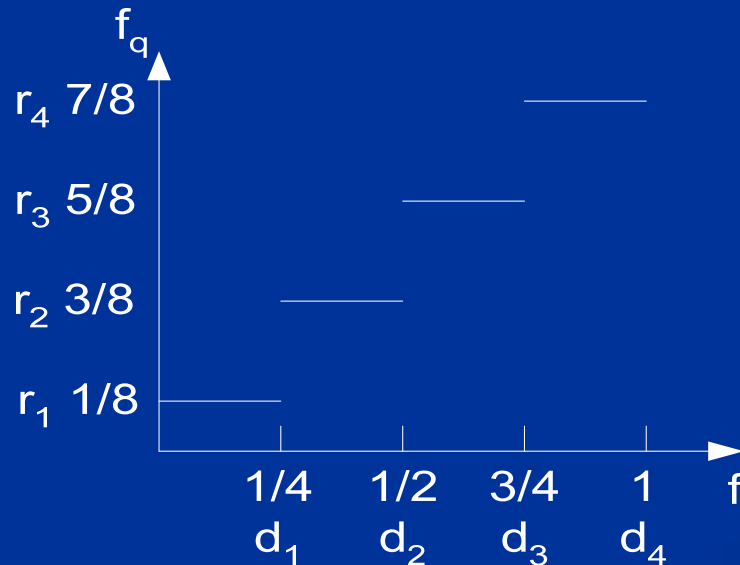


Κβαντισμός

- Ο κβαντισμός αναφέρεται στη διακριτοποίηση της τιμής f του κάθε εικονοστοιχείου (δηλαδή του πεδίου τιμών)
- Διαδικασία κβαντισμού
 - Το πεδίο τιμών της f χωρίζεται σε L υποδιαστήματα
 - Οι ακραίες τιμές κάθε υποδιαστήματος είναι τα όρια απόφασης d_{i-1} , d_i και μια τιμή μεταξύ αυτών ονομάζεται επίπεδο κβάντισης r_i
 - Αν $d_{i-1} < f \leq d_i$ τότε το $f_q = Q(f) = r_i$

Κβαντισμός

- Θόρυβος κβαντισμού: $f = f_q + e_q$
 - Το e_q εξαρτάται από το σήμα και το L
- Ομοιόμορφος κβαντισμός
 - $d_i - d_{i-1} = \Delta$, $r_i = (d_i + d_{i-1})/2$, $1 \leq i \leq L$, $\Delta = (d_L - d_0)/2^{(\# \text{ ψηφίων})}$



Παράδειγμα κβαντισμού

- Επανακβαντισμός εύρους τιμών $[0:255]$ στα πέντε επίπεδα και αντιστοίχιση στην κεντρική τιμή
- Παρατηρήστε ότι στην νέα εικόνα εμφανίζονται μόνο πέντε επίπεδα του γκρι

