

Ψηφιακή Επεξεργασία Εικόνας

Οι θεμελιώδεις αρχές των ψηφιακών εικόνων

Γιώργος Σφήκας
sfikas@cs.uoi.gr

Images taken and/or material adapted from, unless otherwise denoted:

R. Gonzalez and R. Woods. Digital Image Processing, Prentice Hall, 2008;
Digital Image Processing course by Brian Mac Namee, Dublin Institute of Technology;
Publications in <http://www.cs.uoi.gr/~sfikas>

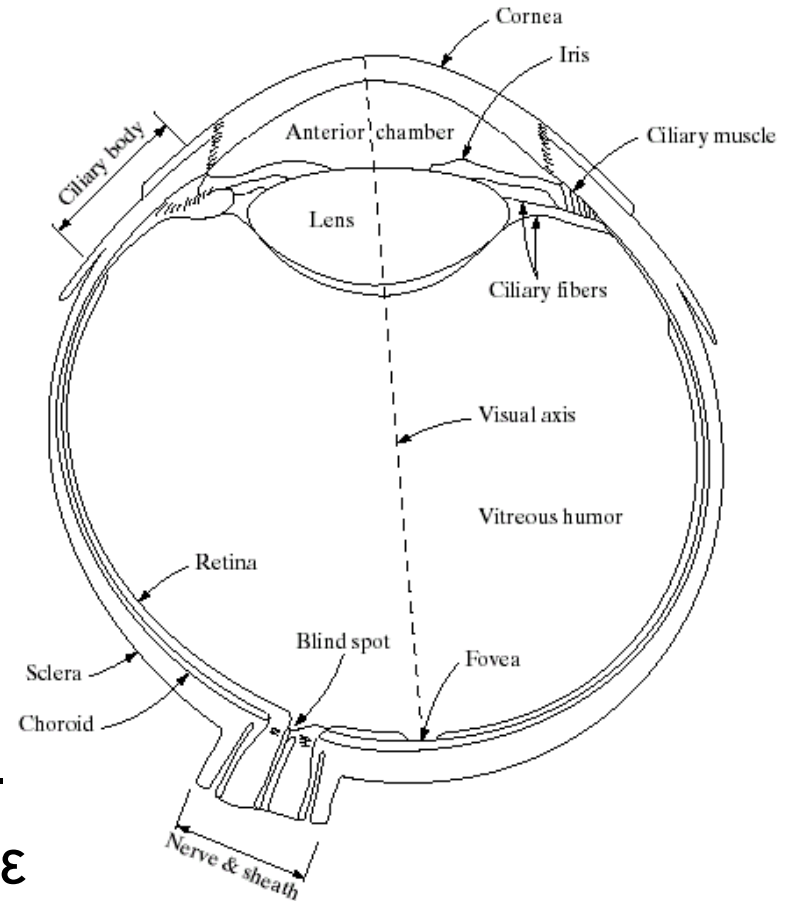
Αυτή η διάλεξη πραγματεύεται:

- Το ανθρώπινο οπτικό σύστημα
- Φως και ηλεκτρομαγνητικό φάσμα
- Αναπαράσταση εικόνας
- Λήψη εικόνας
- Δειγματοληψία, κβαντισμός και ανάλυση

- Η γνώση της διαδικασίας σχηματισμού των εικόνων στον οφθαλμό μπορεί να μας βοηθήσει με την επεξεργασία εικόνων
- Θα δούμε εν συντομία τα βασικά του μηχανισμού της όρασης στον άνθρωπο

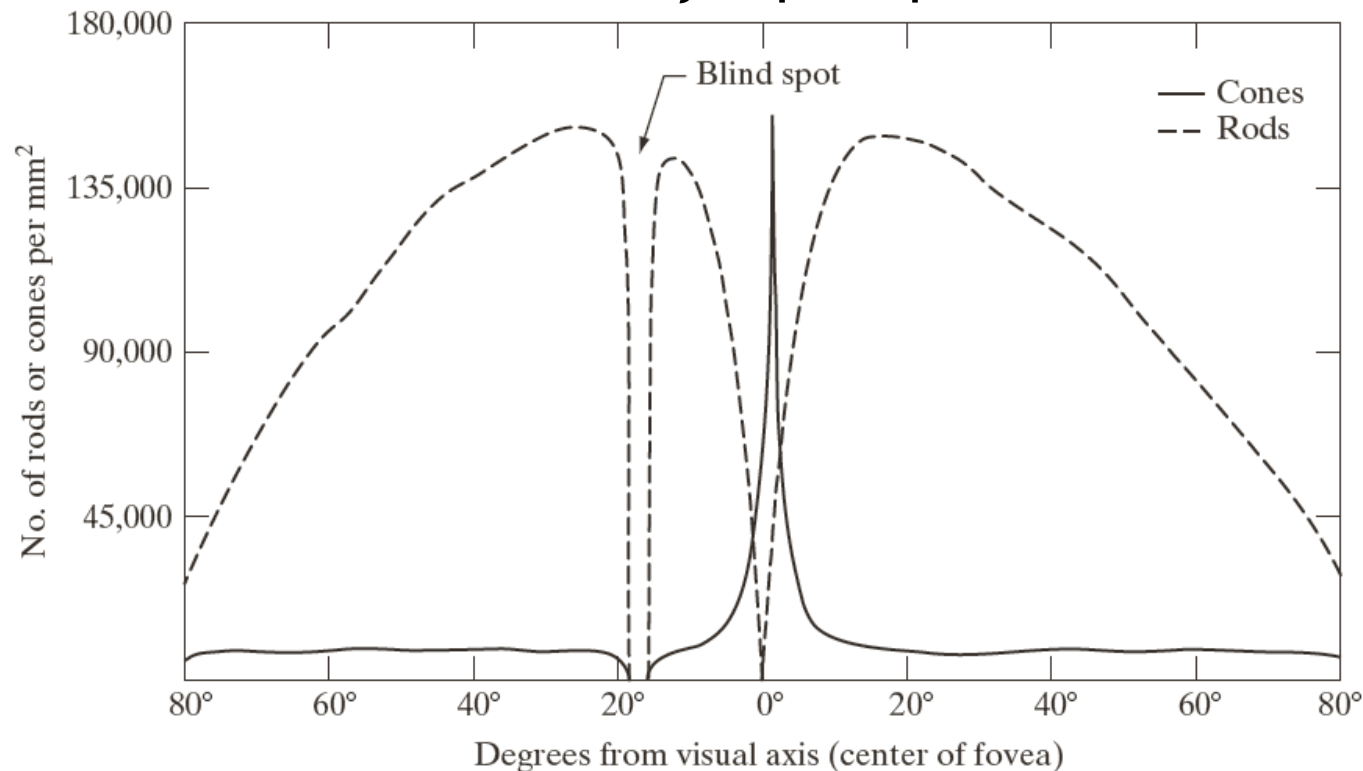
Η δομή του ανθρώπινου οφθαλμού

- Ο φακός εστιάζει φως από τα αντικείμενα στην επιφάνεια του αμφιβληστροειδούς
- Ο αμφιβληστροειδής (retina) καλύπτεται από φωτο-αισθητήρες που ονομάζονται κωνία (6-7 εκ.) και ραβδία (75-150 εκ.) (cones, rods)
- Τα κωνία συγκεντρώνονται γύρω από το βοθρίο (fovea) και είναι ευαίσθητα στο χρώμα
- Τα ραβδία είναι πιο αραιά κατανεμημένα και είναι ευαίσθητα σε χαμηλά επιπέδα φωτεινότητας



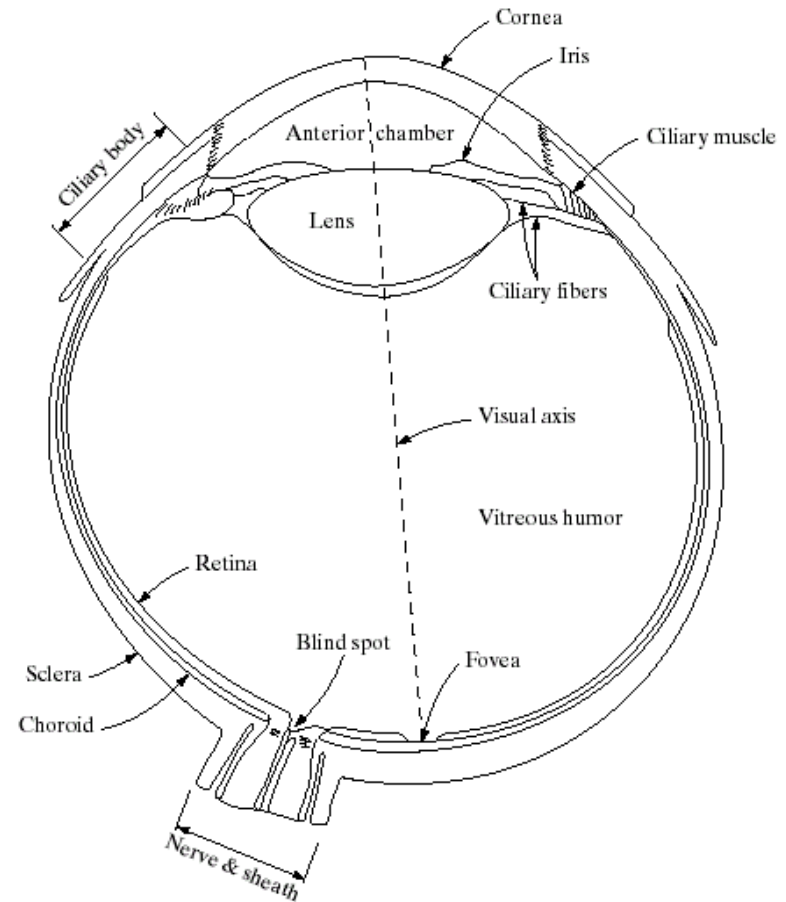
Η δομή του ανθρώπινου οφθαλμού

Κατανομή των κωνίων και των ραβδίων κατά μήκος τομής στον δεξί οφθαλμό



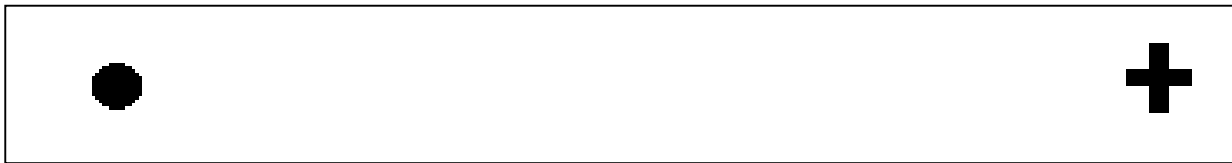
Η δομή του ανθρώπινου οφθαλμού

- Κάθε κωνίο συνδέεται στη δική του νευρική απόληξη
 - Μπορούν να αναλύσουν λεπτομέρειες της εικόνας
 - Ευαίσθητα στο χρώμα (φωτοοπική όραση)
- Πολλά ραβδία αντιστοιχούν σε ένα νεύρο
 - Χαμηλότερη ανάλυση
 - Όχι ευαίσθητα στο χρώμα
 - Ευαίσθητα σε χαμηλά επίπεδα φωτεινότητας (σκοτοπική όραση)



Το «τυφλό σημείο»

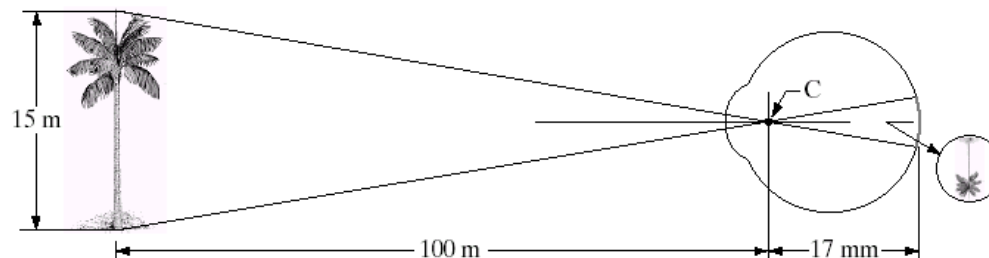
- Σχεδιάστε το παρακάτω σκίτσο σε ένα κομμάτι χαρτί (η κουκκίδα και ο σταυρός να απέχουν γύρω στα 15 εκατοστά)



- Με κλειστό το δεξί μάτι, εστιάστε στον σταυρό με το αριστερό μάτι
- Κρατήστε το κομμάτι χαρτί περίπου μισό μέτρο μακριά και σιγά-σιγά πλησιάστε το κοντά σας
- Η κουκκίδα θα εξαφανιστεί !

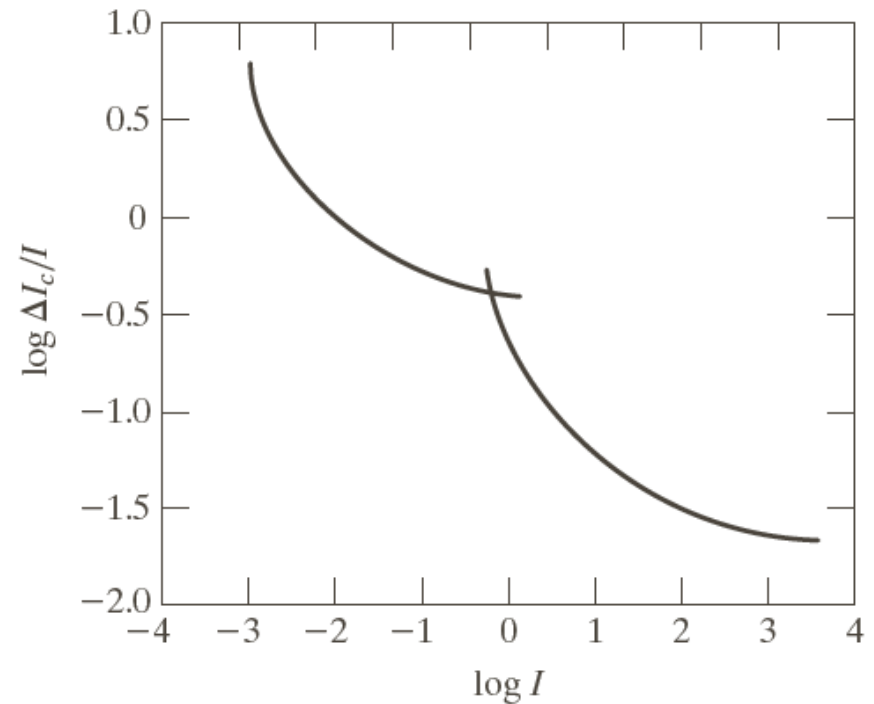
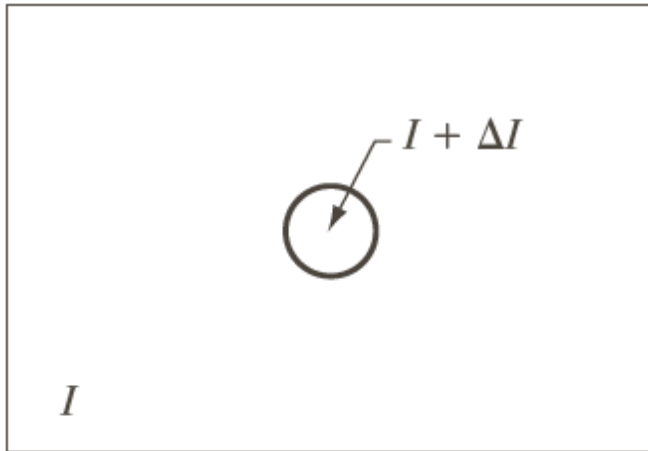
Σχηματισμός εικόνας στον οφθαλμό

- Οι οφθαλμικοί μύες μπορούν να αλλάξουν το σχήμα του φακού, επιτρέποντας στον άνθρωπο να εστιάσει σε μακρινά ή κοντινά αντικείμενα κατά βούληση
- Οι φωτογραφικές μηχανές αντίθετα ρυθμίζουν το σημείο εστίασης μετακινώντας τον φακό
- Μια εικόνα προβάλλεται στον αμφιβληστροειδή ερεθίζοντας τα ραβδία και τα κωνία, τα οποία με τη σειρά τους στέλνουν σήματα στον εγκέφαλο

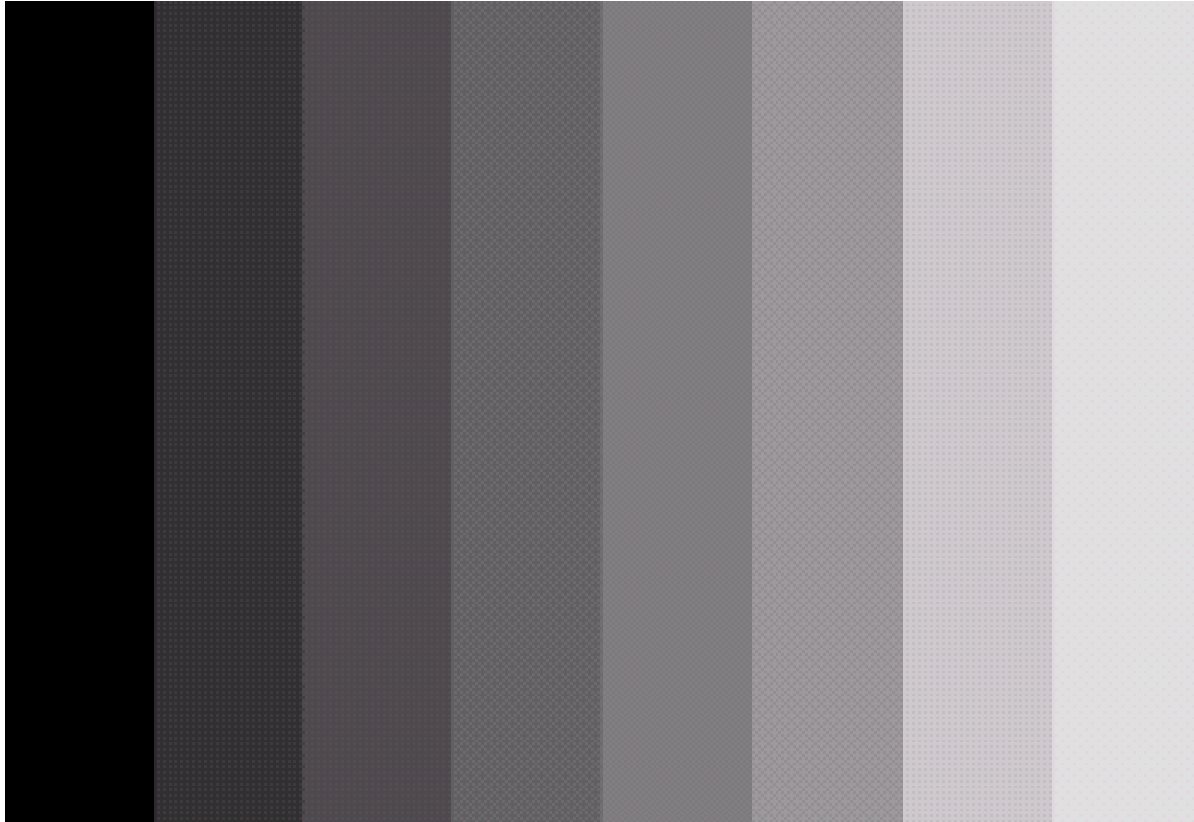


- Το ανθρώπινο οπτικό σύστημα μπορεί να αντιληφθεί σχεδόν 10^{10} διαφορετικά επίπεδα φωτεινότητας.
- Σε μια οποιαδήποτε χρονική στιγμή, μπορούμε να διακρίνουμε μεταξύ ενός πολύ μικρότερου αριθμού επιπέδων / έντασης φωτεινότητας.
- Η αισθητή ένταση μιας περιοχής σχετίζεται με τις φωτεινότητες των γύρω περιοχών.

Λόγος Weber

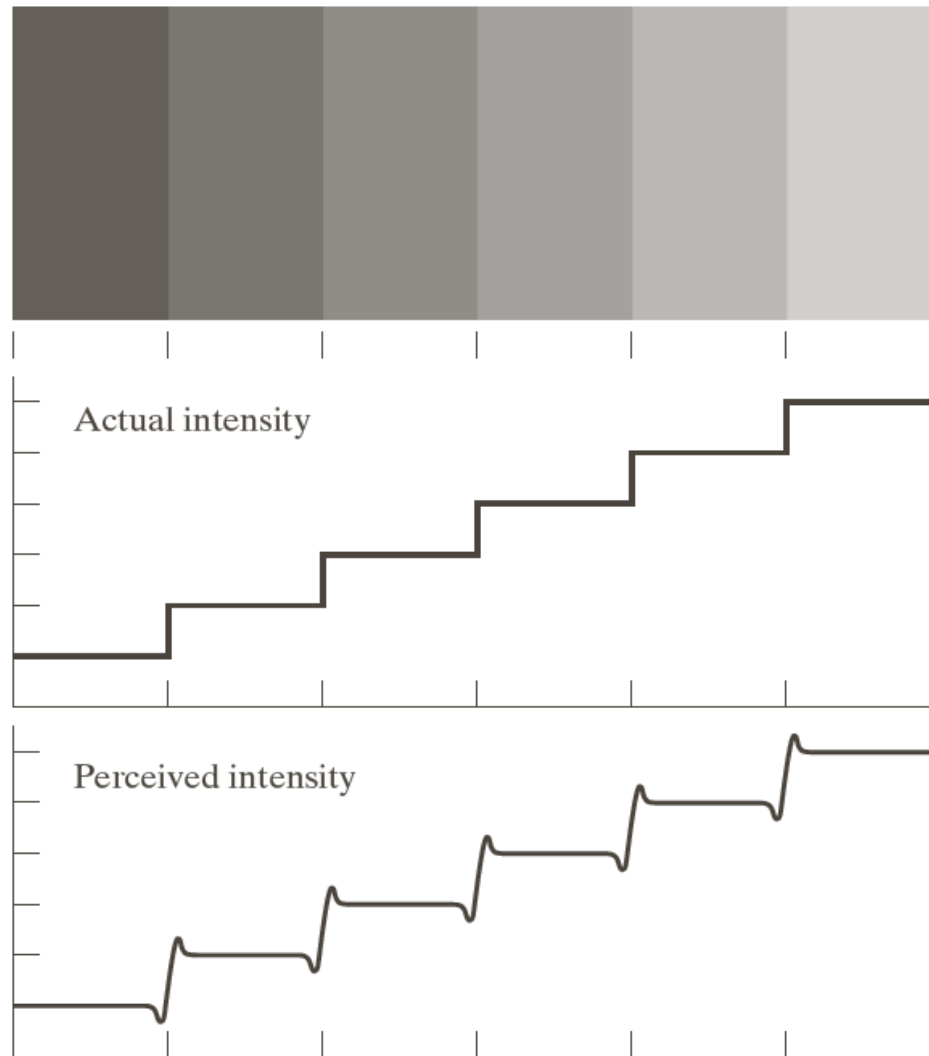


Προσαρμογή στη φωτεινότητα



Ζώνες Mach

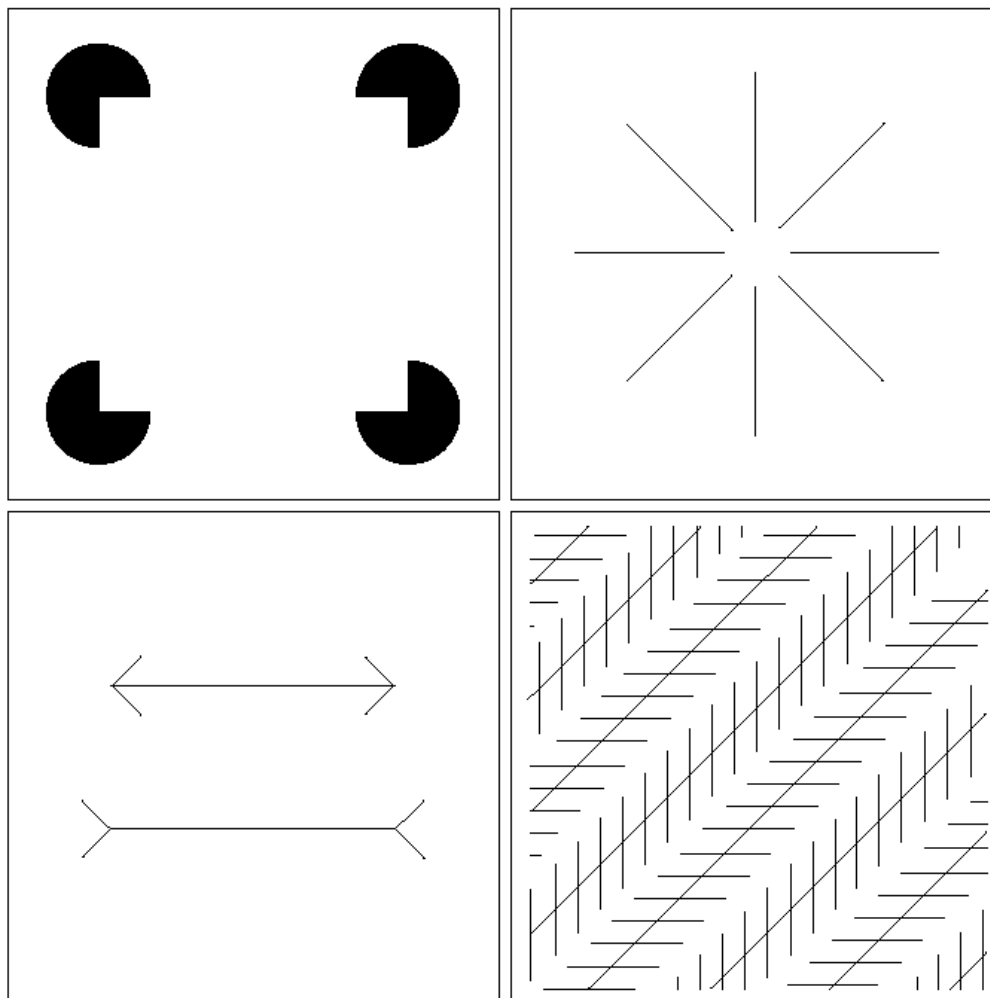
Brightness Adaptation & Discrimination (cont...)

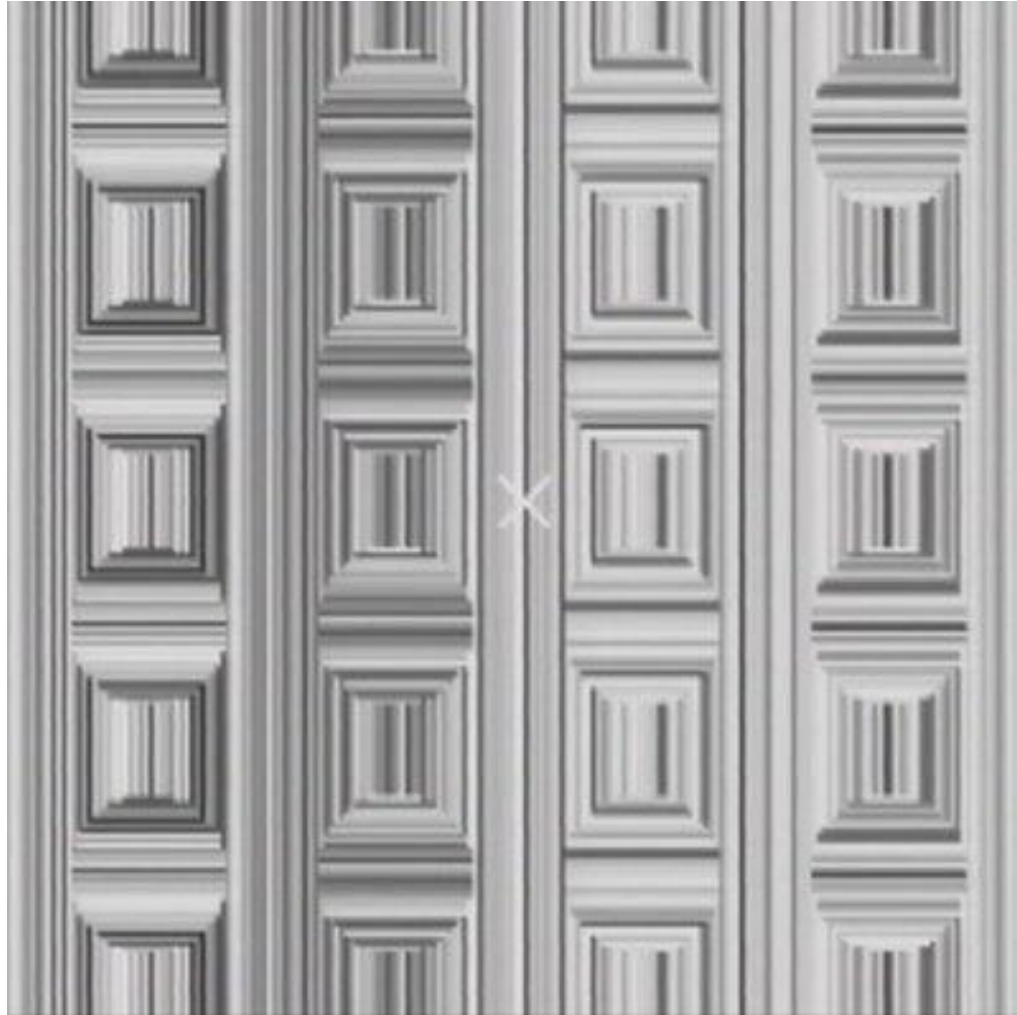


Προσαρμογή στη φωτεινότητα

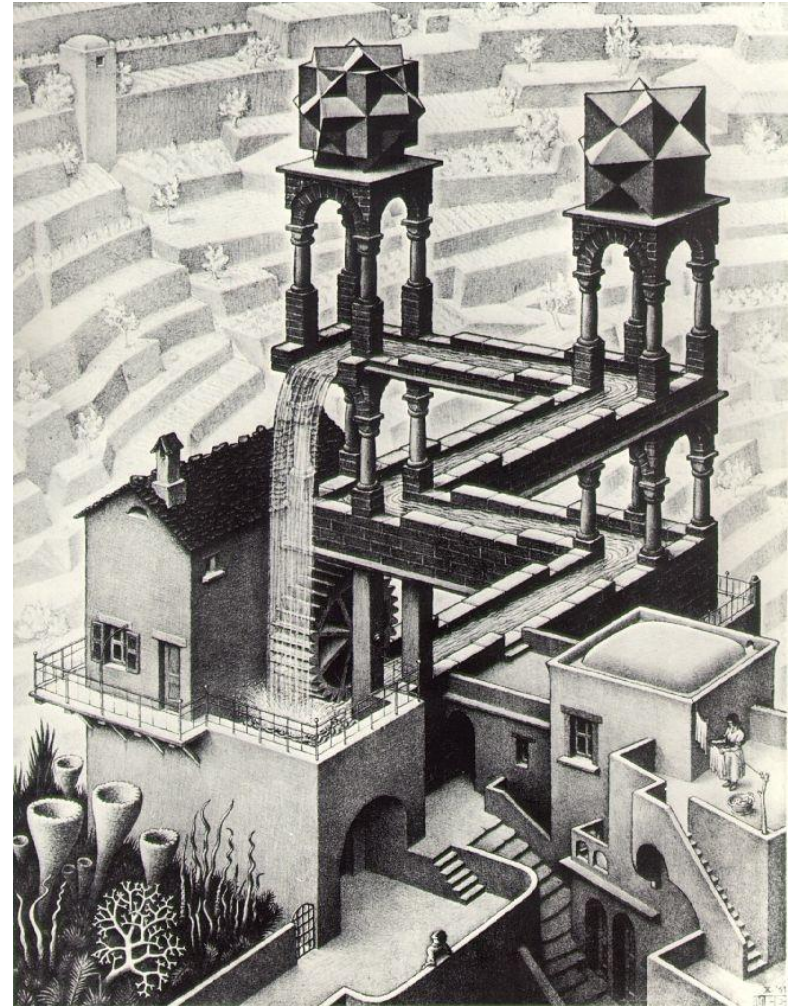
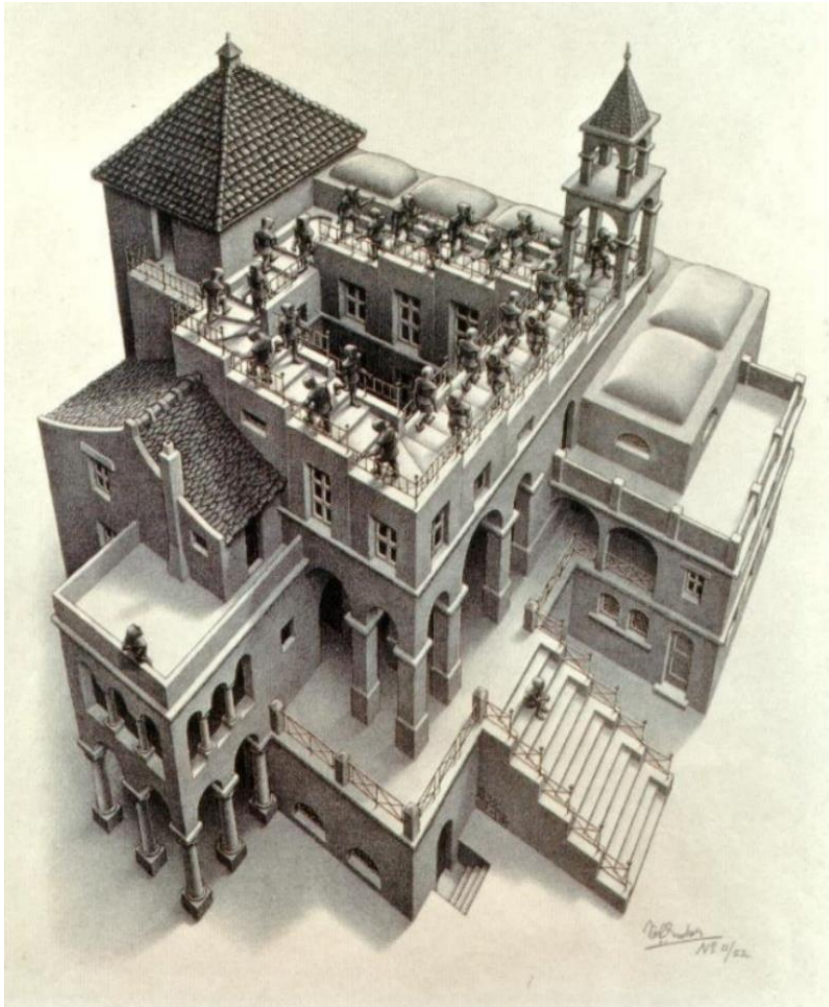


Η ένταση του κεντρικού τετραγώνου είναι στην πραγματικότητα σταθερή



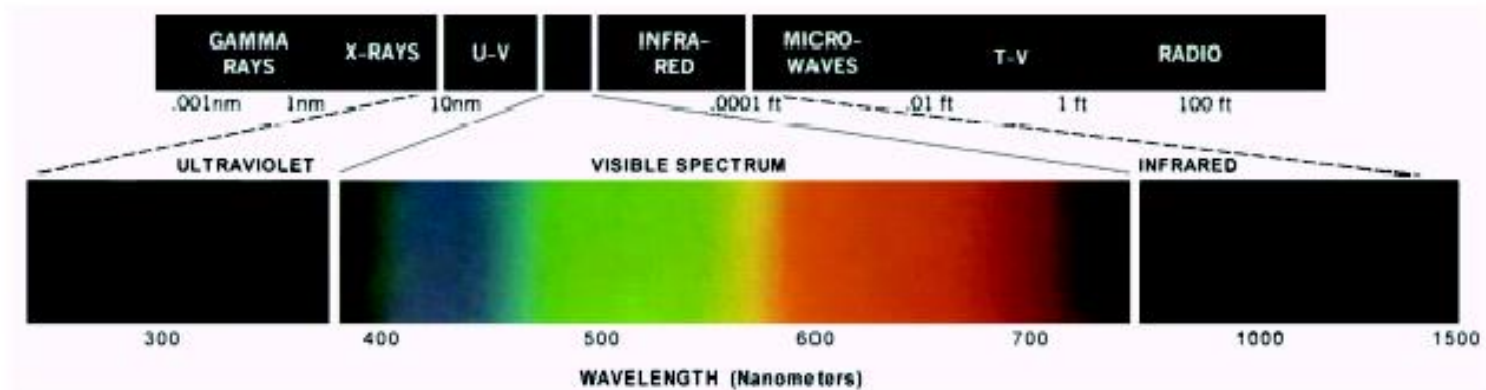


Εστιάστε στον
σταυρό που
βρίσκεται στο
κέντρο της εικόνας
και σκεφτείτε
κύκλους

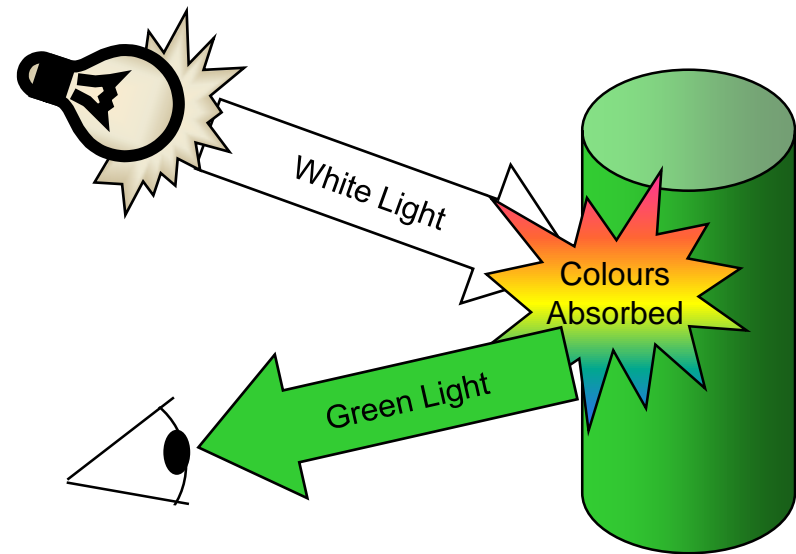


Φως και το ηλεκτρομαγνητικό φάσμα

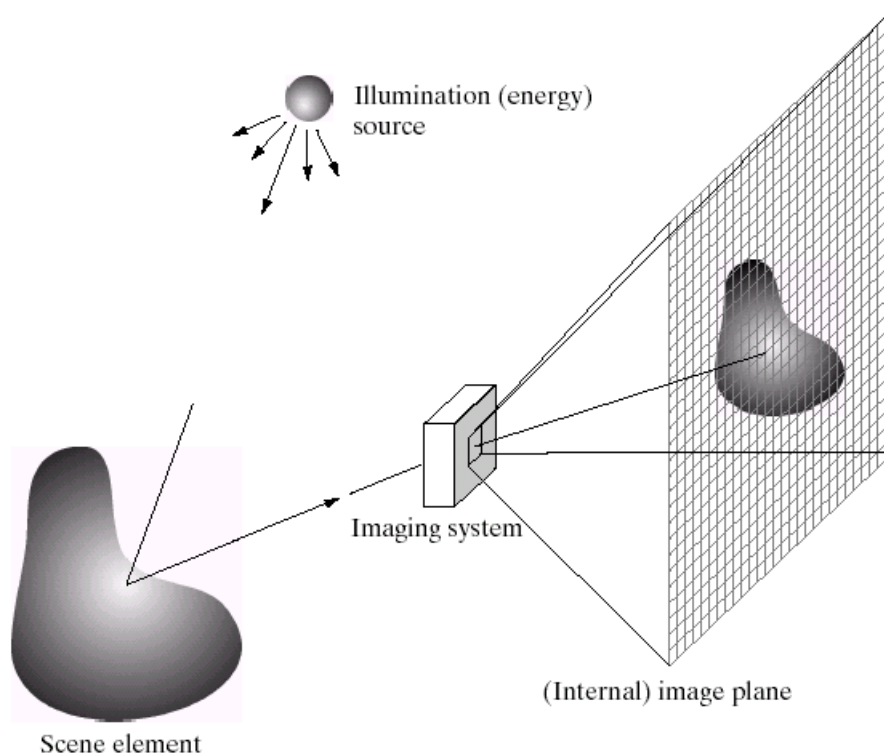
- Το φως αντιστοιχεί σε μια ζώνη ακτινοβολίας του ηλεκτρομαγνητικού φάσματος (ΗΜΦ) που είναι οπτικά αντιληπτή
- Το ΗΜΦ διαιρείται σε ζώνες ανάλογα με το μήκος κύματος και την φέρουσα ενέργεια που αντιστοιχεί κατά περίπτωση



- Το φως που αντιλαμβανόμαστε εξαρτάται από τις ιδιότητες του φωτός που ανακλάται από το εκάστοτε αντικείμενο
- Για παράδειγμα, αν λευκό φως προσπέσει σε ένα πράσινο αντικείμενο, τα περισσότερα μήκη κύματος απορροφώνται ενώ το πράσινο φως ανακλάται



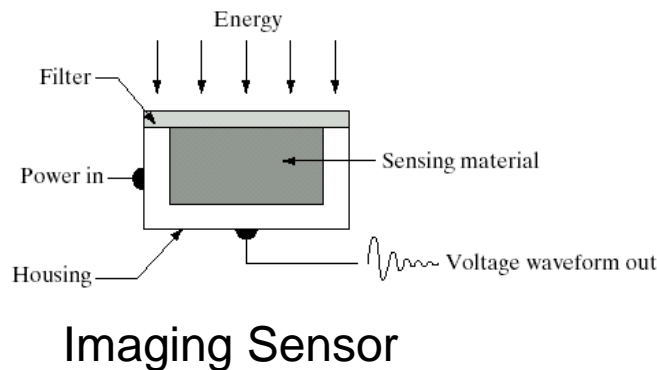
Τυπικά μια εικόνα παράγεται φωτίζοντας μια σκηνή και καταγράφοντας την ανακλώμενη ακτινοβολία



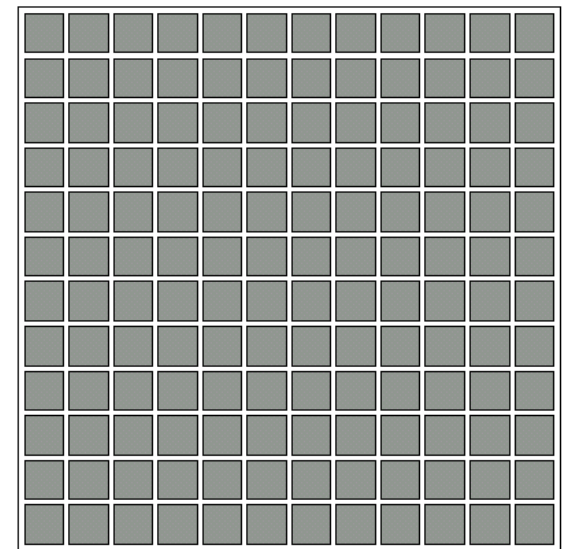
– Το παραπάνω υπόδειγμα λήψης δεν είναι όμως το μοναδικό:

- Εικόνες ακτίνων X
- Υπέρηχοι
- Ηλεκτρομικροσκοπικές εικόνες μορίων

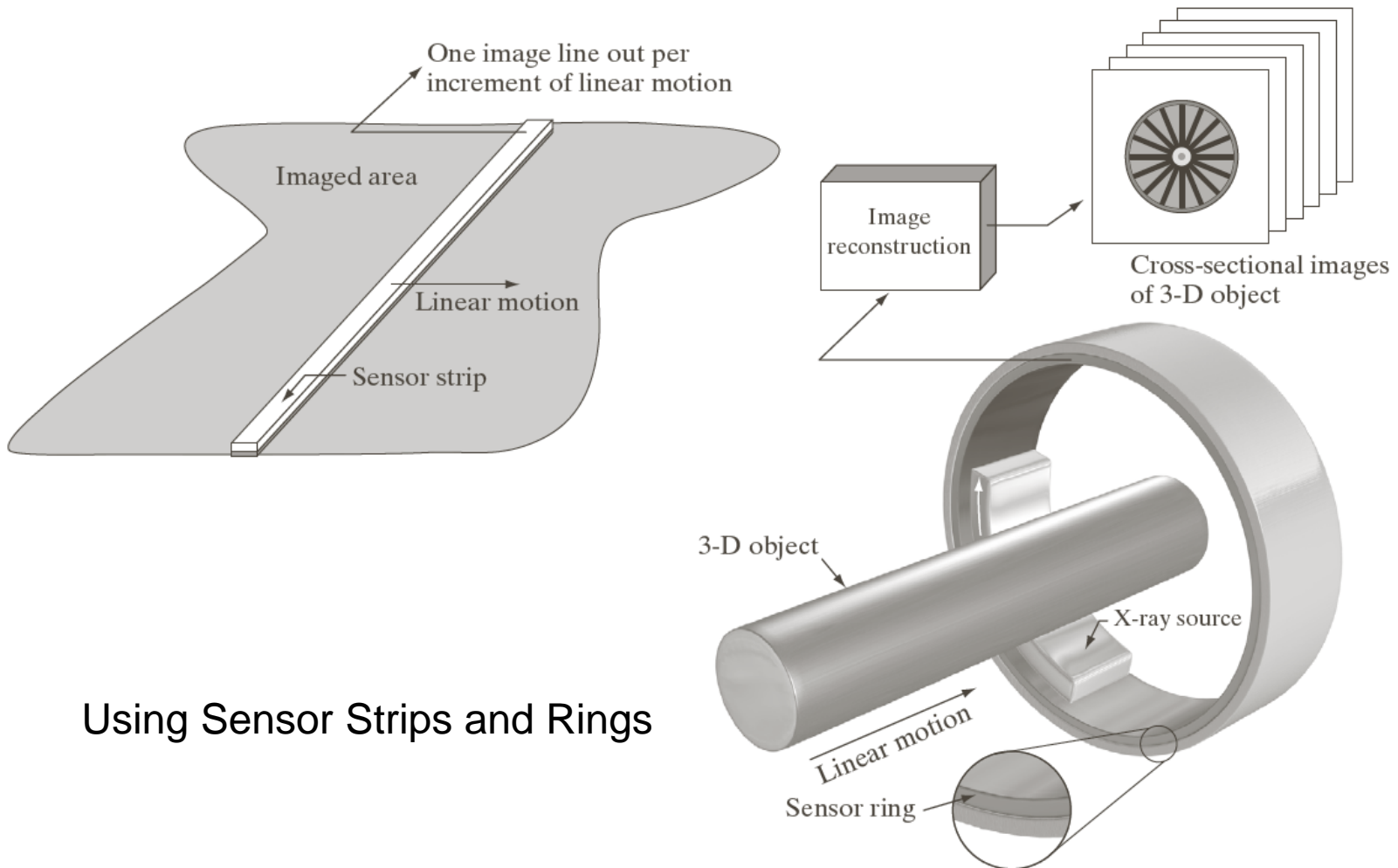
- Αισθητήρες μετατρέπουν την εισερχόμενη ακτινοβολία σε ηλεκτρική τάση, η οποία στη συνέχεια κβαντίζεται / ψηφιοποιείται



Line of Image Sensors



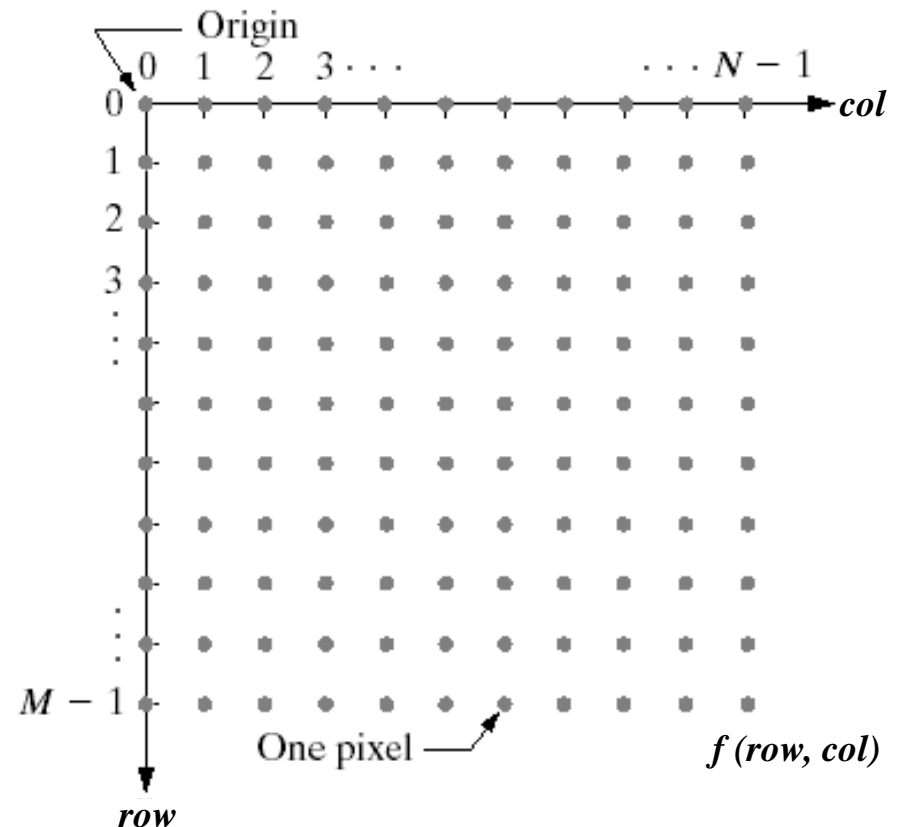
Array of Image Sensors

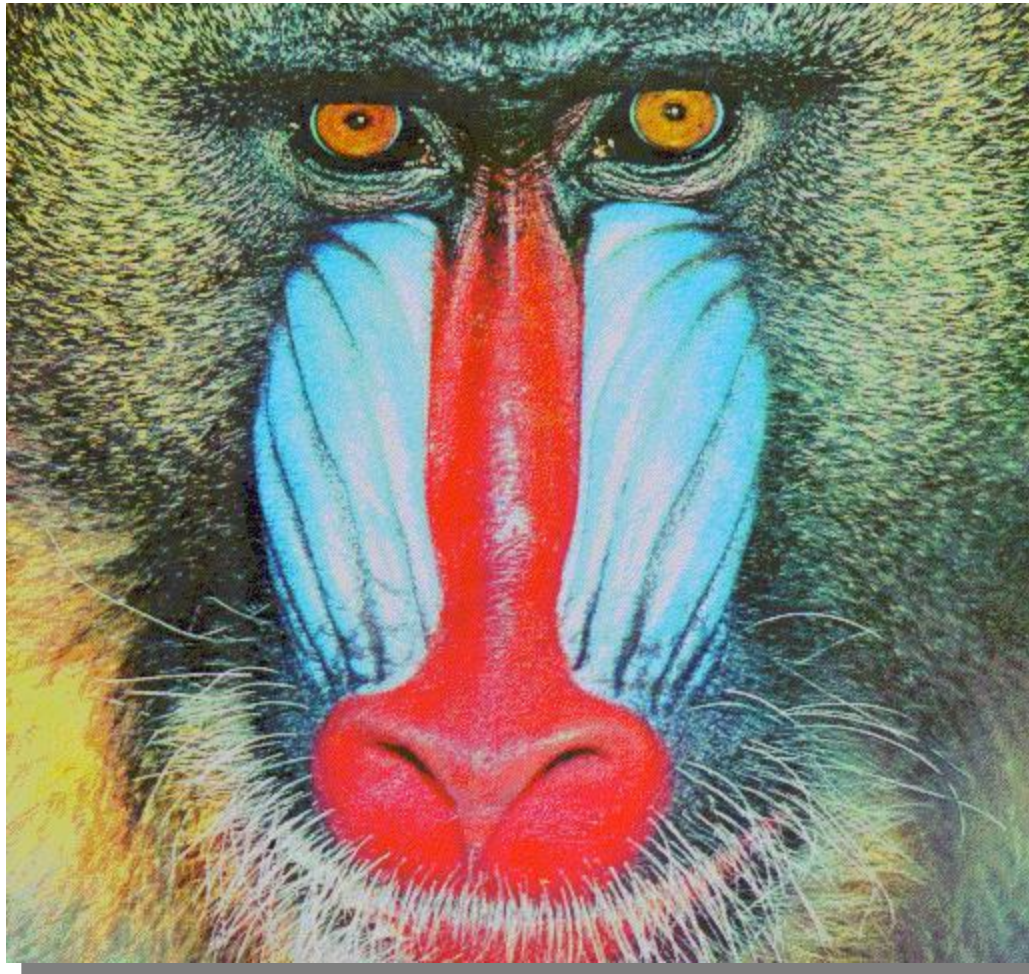


Using Sensor Strips and Rings

Αναπαράσταση εικόνας

- Μια ψηφιακή εικόνα αποτελείται από M γραμμές και N στήλες εικονοστοιχείων
- Οι τιμές των εικονοστοιχείων έχουν εύρος 0..255 (μαύρο-λευκό)
- Οι εικόνες αναπαριστώνται σαν πίνακες

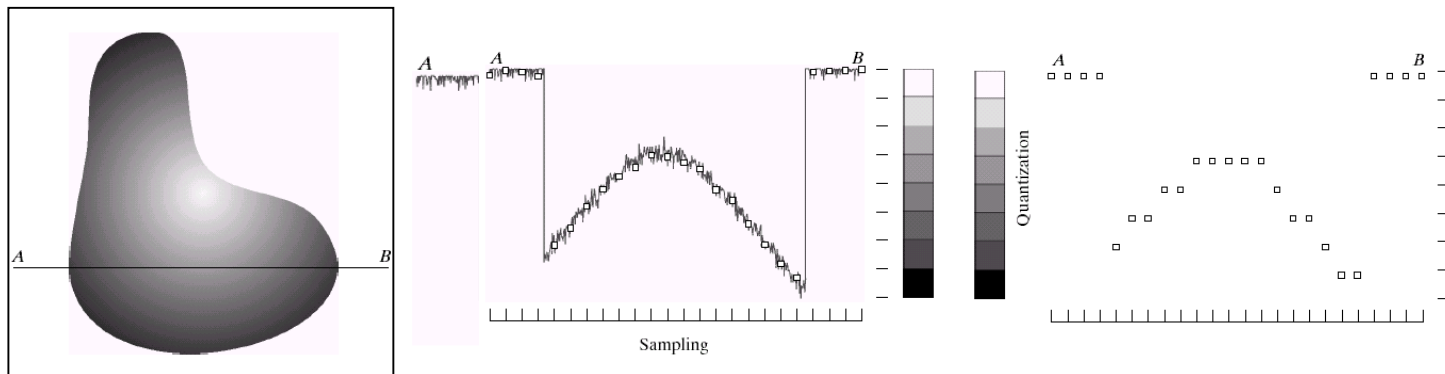






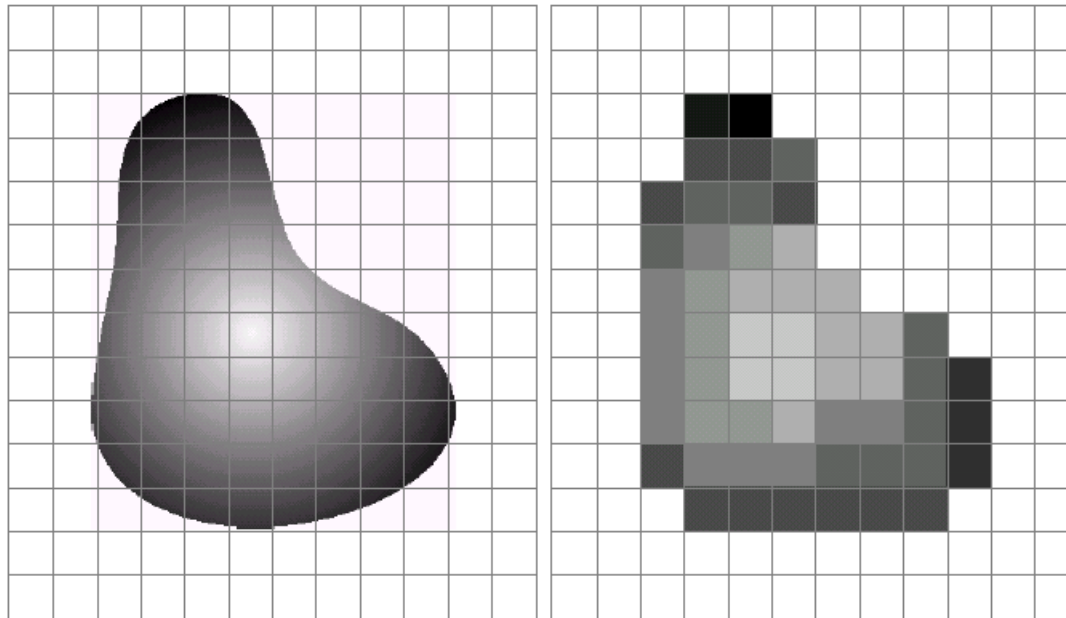
Δειγματοληψία και κβαντισμός εικόνας

- Ένας ψηφιακός αισθητήρας μπορεί να μετρήσει μόνο έναν περιορισμένο αριθμό δειγμάτων σε ένα διακριτό σύνολο επιπέδων ενέργειας
- Κβαντισμός είναι η διαδικασία μετατροπής του αναλογικού σήματος σε ψηφιακό

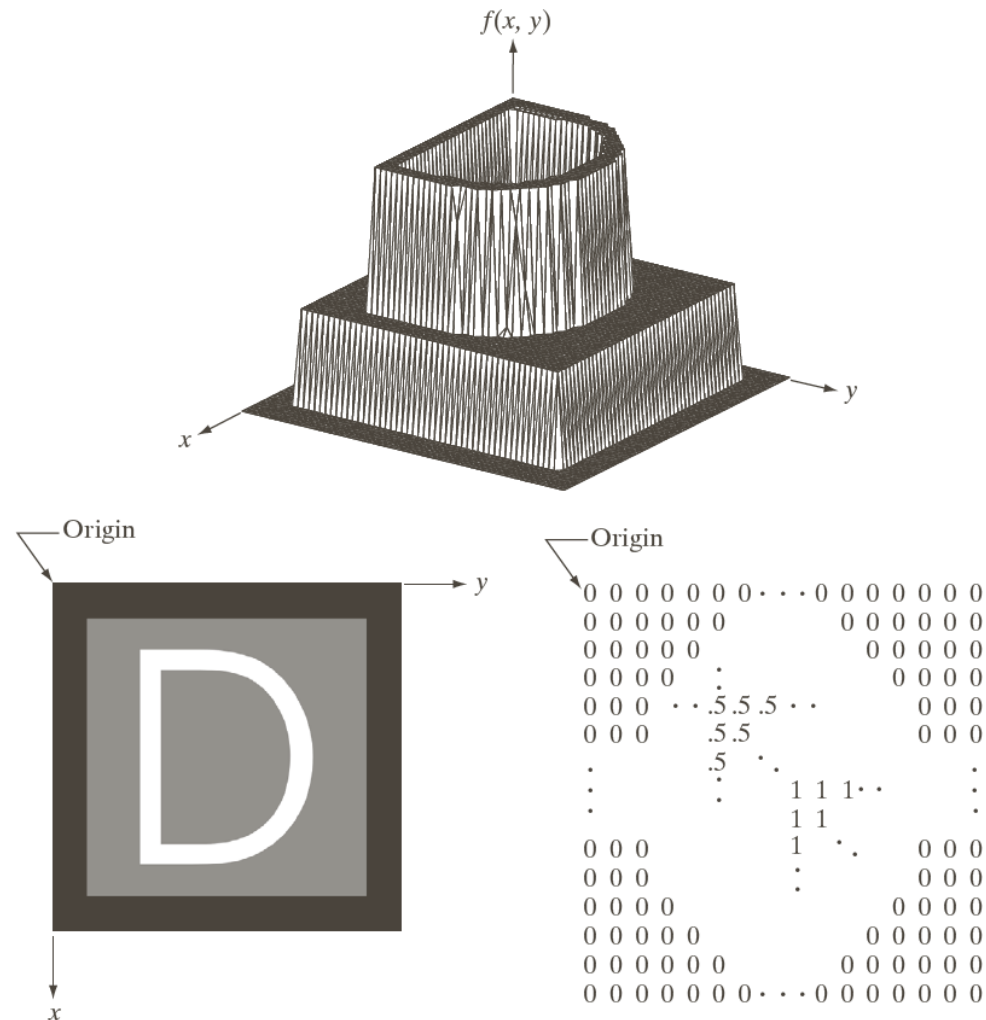


Δειγματοληψία και κβαντισμός εικόνας

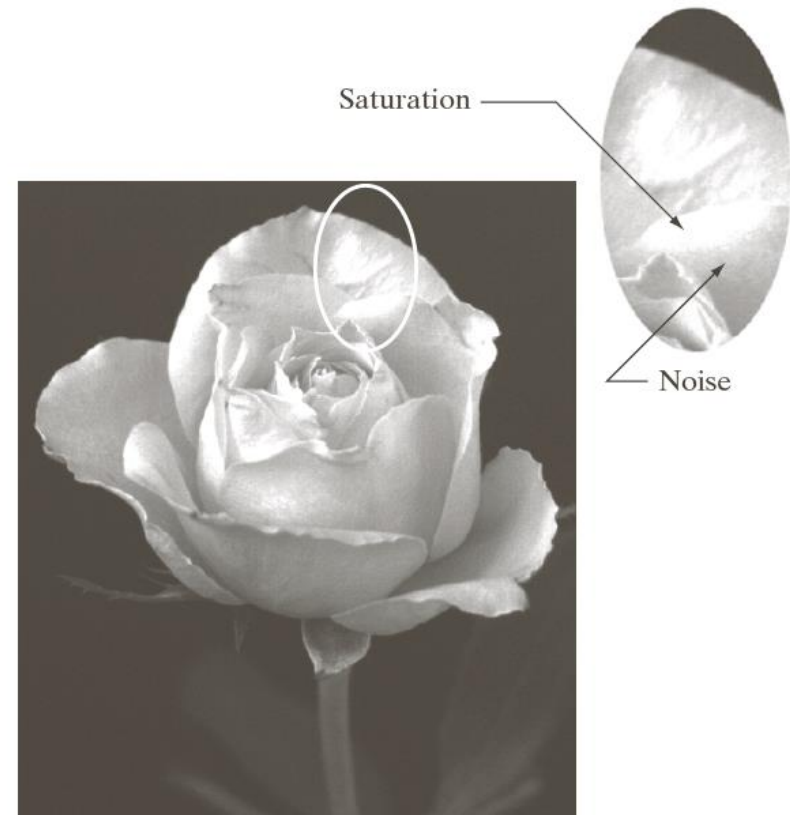
- Μια ψηφιακή εικόνα είναι πάντα μια κατά προσέγγιση καταγραφή της πραγματικότητας



Αναπαράσταση εικόνας



- **Δυναμικό εύρος:** Ο λόγος της μέγιστης (κορεσμός) προς την ελάχιστη (θόρυβος) αντιληπτή ένταση φωτεινότητας.
- Ο θόρυβος γενικά εμφανίζεται σαν μια κοκκώδης υφή στις σκουρότερες περιοχές τις εικόνας, καλύπτοντας το πραγματικό ελάχιστο που μπορεί να καταγραφεί



- Η χωρική ανάλυση μιας εικόνας εξαρτάται από τις παράμετρους της δειγματοληψίας
- Η χωρική ανάλυση αναφέρεται στο μικρότερο δυνατό επίπεδο λεπτομέρειας που μπορεί να διακριθεί
 - Μέγεθος εικονοστοιχείων
 - Dots per inch (DPI)





32

1024

512

256

128

64

1024 * 1024



512 * 512



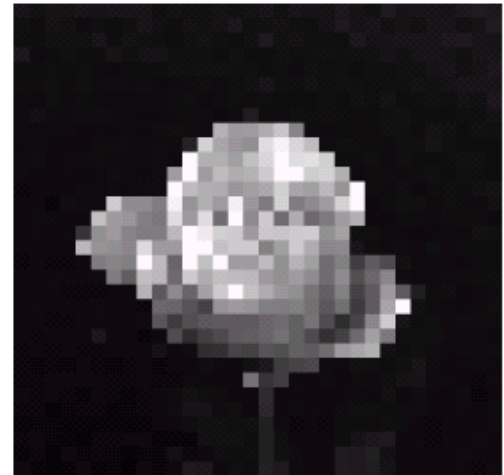
256 * 256



128 * 128



64 * 64



32 * 32



Ανάλυση επιπέδων έντασης

- Η ανάλυση επιπέδων έντασης αναφέρεται στον αριθμό των δυνατών επιπέδων έντασης (intensity levels) στην εικόνα
 - Περισσότερα επίπεδα έντασης συνεπάγονται καλύτερο επίπεδο λεπτομέρειας που μπορεί να διακριθεί
 - Η ανάλυση επιπέδων έντασης εκφράζεται συνήθως σαν αριθμός bits ανά εικονοστοιχείο

Αριθμός bits	Αριθμός intensity levels	Παραδείγματα
1	2	0, 1
2	4	00, 01, 10, 11
4	16	0000, 0101, 1111
8	256	00110011, 01010101
16	65,536	1010101010101010

Ανάλυση επιπέδων έντασης

256 grey levels (8 bits per pixel)



128 grey levels (7 bpp)



64 grey levels (6 bpp)



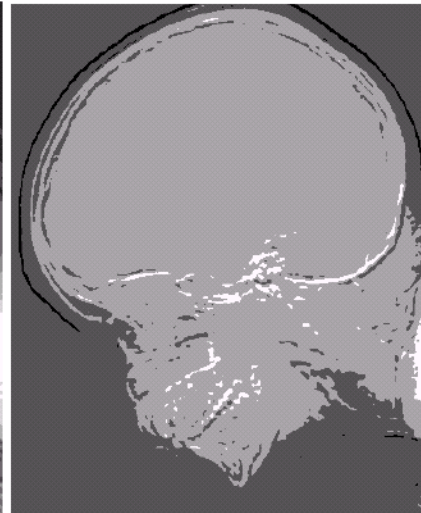
32 grey levels (5 bpp)



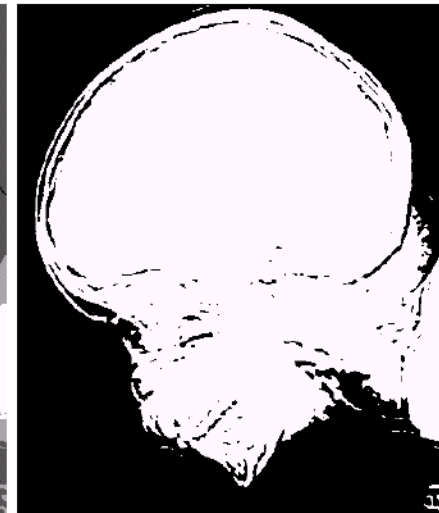
16 grey levels (4 bpp)



8 grey levels (3 bpp)



4 grey levels (2 bpp)



2 grey levels (1 bpp)

Ανάλυση επιπέδων έντασης



Low Detail

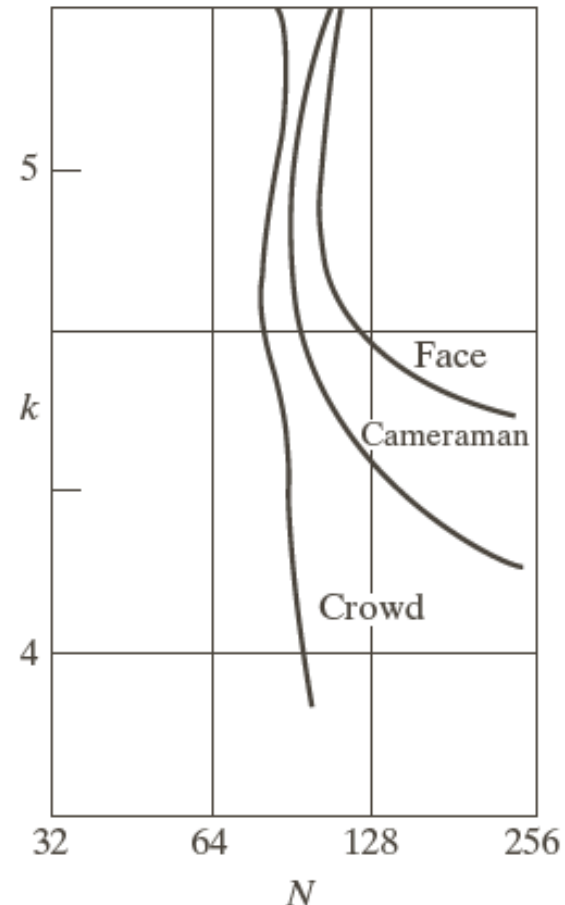


Medium Detail



High Detail

- Οι καμπύλες *ίσης προτίμησης* (*Isopreference curves*) αναπαριστούν την σχέση μεταξύ έντασης και χωρικής ανάλυσης
 - Τα σημεία μιας καμπύλης αναπαριστούν εικόνες «ίσης» υποκειμενικής ποιότητας
 - Οι καμπύλες γίνονται πιο κατακόρυφες όσο ο βαθμός λεπτομέρειας αυξάνεται (περισσότερη λεπτομέρεια απαιτεί λιγότερα επίπεδα έντασης).



Το μεγάλο ερώτημα σχετικά με την ανάλυση είναι πόση ανάλυση είναι αρκετή;

- Αυτό εξαρτάται από το τι απεικονίζεται αλλά και το τι θέλουμε να κάνουμε με την εικόνα
- Συνήθως μας ενδιαφέρει:
 - Η εικόνα φαίνεται αισθητικά καλή;
 - Μπορούμε να διακρίνουμε ό,τι κατά περίπτωση χρειάζεται να διακρίνουμε στην εικόνα;

Ανάλυση: Πόση χρειαζόμαστε;



Η εικόνα στα δεξιά αρκεί για να μετρήσουμε τον αριθμό των οχημάτων στην εικόνα, αλλά όχι και για να μπορέσουμε να διαβάσουμε τον αριθμό της πινακίδας

Μαθηματικές πράξεις στην ψηφιακή επεξεργασία εικόνας

- Αριθμητικές πράξεις (π.χ. Αφαίρεση εντάσεων ανά εικονοστοιχείο)
- Πράξεις πινάκων και διανυσμάτων
- Γραμμικές (sum) και μη-γραμμικές πράξεις (min, max)
- Λογικές πράξεις και πράξεις σε σύνολα
- Χωρικές πράξεις (π.χ. Τοπικό μέσο)
- Γεωμετρικοί χωρικοί μετασχηματισμοί (π.χ. περιστροφή)

- Για pixels $p(x,y)$, $q(s,t)$ and $z(v,w)$, η D είναι μια συνάρτηση απόστασης ή μετρική εάν:

$$a) D(p, q) \geq 0 \text{ (} D(p, q) = 0 \text{ iff } p = q),$$

$$b) D(p, q) = D(q, p),$$

$$c) D(p, z) \leq D(p, q) + D(q, z).$$

- Η Ευκλείδεια απόσταση μεταξύ p και q δίνεται ως:

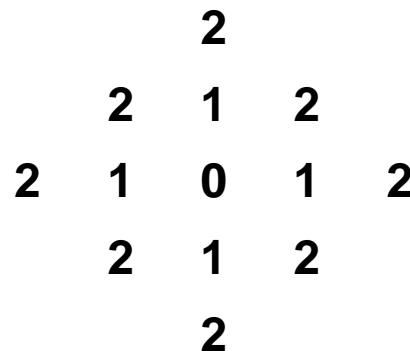
$$D_e(p, q) = \left[(x-s)^2 + (y-t)^2 \right]^{\frac{1}{2}}$$

Αποστάσεις μεταξύ σημείων / εικονοστοιχείων

- Η οικοδομική (city-block) ή D_4 απόσταση μεταξύ p και q δίνεται ως:

$$D_4(p, q) = |x - s| + |y - t|$$

- Pixels που έχουν μια city-block απόσταση από ένα pixel (x, y) μικρότερη ή ίση κάποιας τιμής T , σχηματίζουν ένα «διαμάντι» με κέντρο το (x, y) .

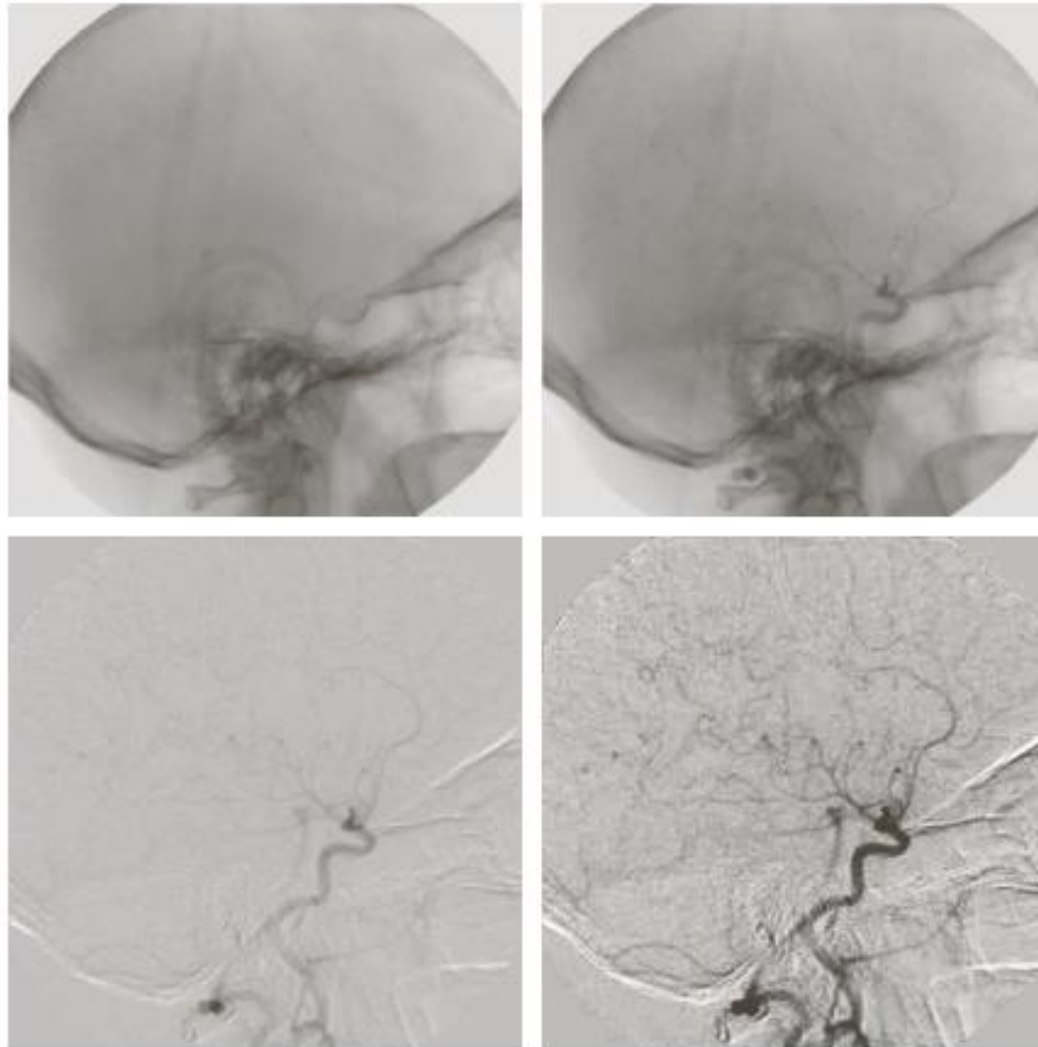


- Η chessboard ή απόσταση D_8 μεταξύ p και q ορίζεται ως:

$$D_8(p, q) = \max(|x - s|, |y - t|)$$

- Pixels που έχουν μια chessboard απόσταση από ένα pixel (x, y) μικρότερη ή ίση κάποιας τιμής T , σχηματίζουν ένα «τετράγωνο» με κέντρο το (x, y)

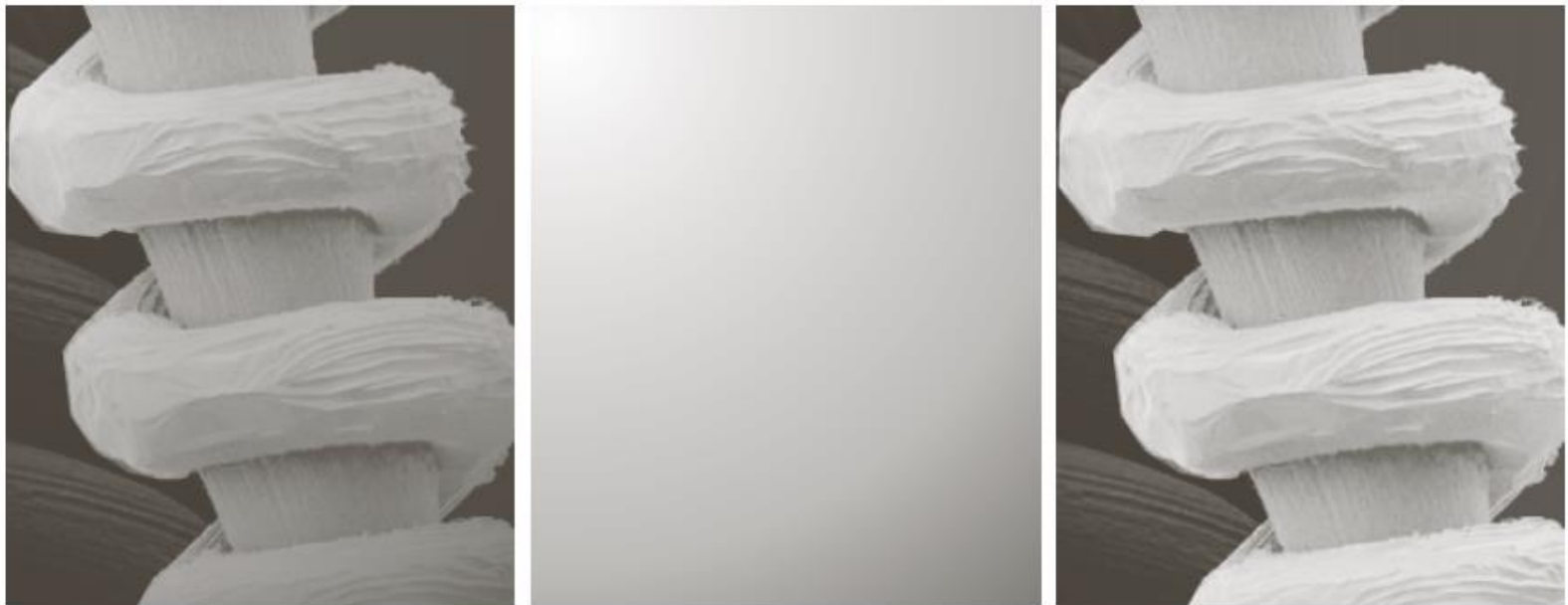
2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2



a b
c d

FIGURE 2.28

Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)



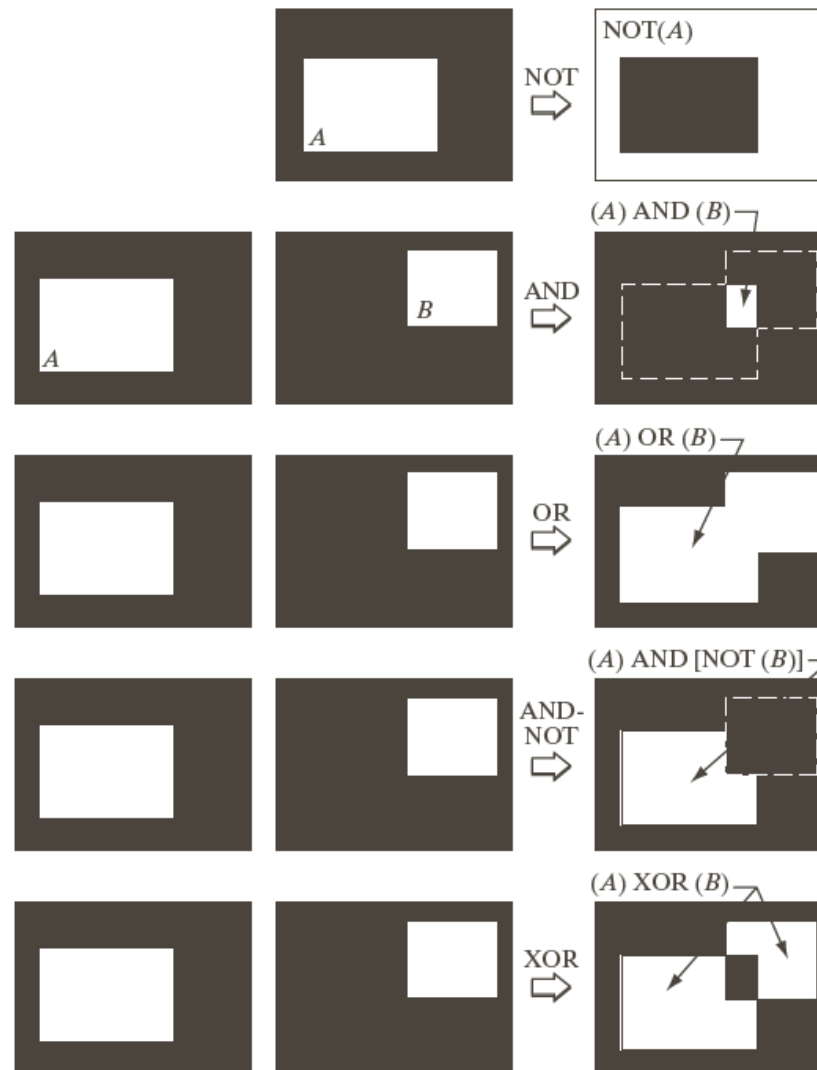
a b c

FIGURE 2.29 Shading correction. (a) Shaded SEM image of a tungsten filament and support, magnified approximately 130 times. (b) The shading pattern. (c) Product of (a) by the reciprocal of (b). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

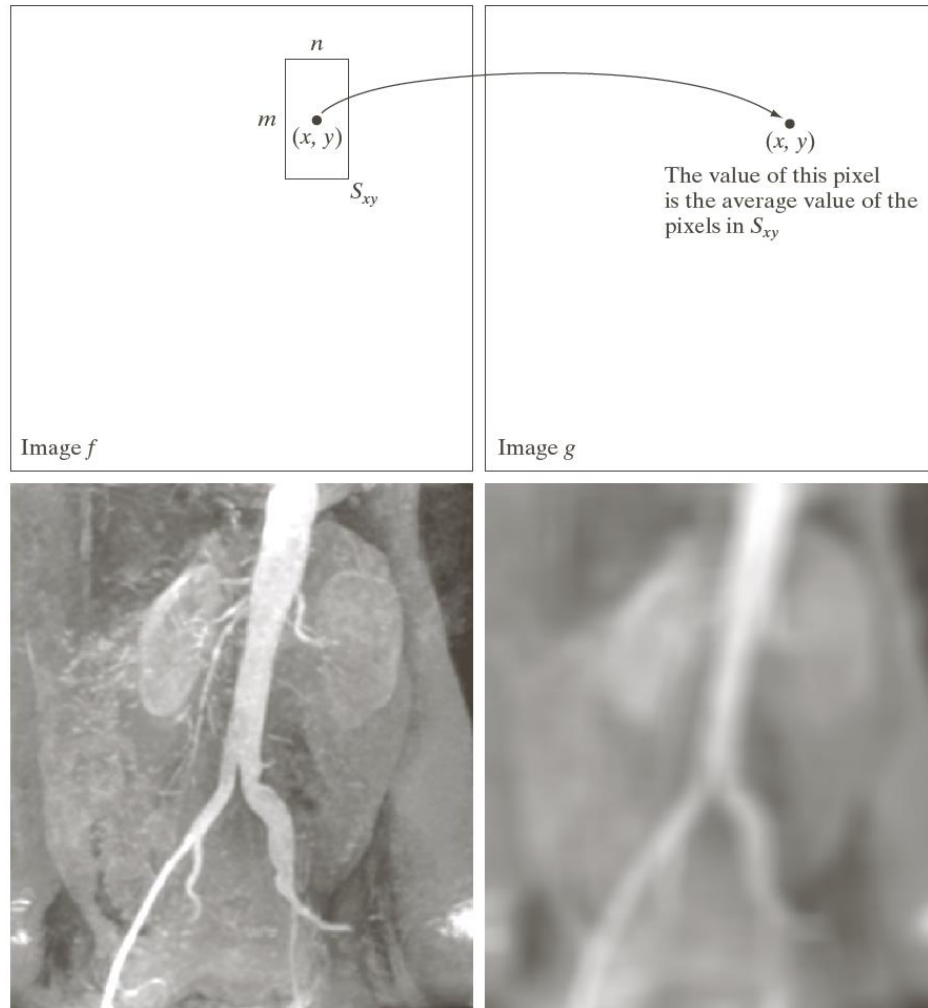


a b c

FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).



Πράξεις επί γειτονιάς σημείων



Σχετικά με τις αριθμητικές πράξεις

- Οι περισσότερες εικόνες που θα δούμε στο μάθημα είναι σε 8 bits (0-255).
- Όταν μια ψηφιακή εικόνα αποθηκεύεται στον δίσκο σε format όπως tiff ή jpeg, η μετατροπή σε αυτό το εύρος είναι αυτόματη.
- Ωστόσο, ο τρόπος μετατροπής γενικά εξαρτάται από το πακέτο λογισμικού που κάθε φορά χρησιμοποιείται.
 - Η διαφορά δυο εικόνων είναι στο εύρος [-255, 255] ενώ το άθροισμα τους στο εύρος [0, 510].
 - Κάποια πακέτα λογισμικού θέτουν τις αρνητικές τιμές ίσες με 0 και όλες τις τιμές άνω του 255 σε 255 – αυτό θέλουμε να το αποφύγουμε !

Σχετικά με τις αριθμητικές πράξεις

- Μια διαδικασία που μπορούμε να ακολουθήσουμε για να κρατήσουμε το πλήρες εύρος σε ένα σταθερό αριθμό bits είναι :
- Πρώτα κάνουμε την ελάχιστη τιμή ίση με μηδέν:

$$f_m = f - \min(f)$$

- Στη συνέχεια κλιμακώνουμε στο εύρος $[0, K]$

$$f_s = \frac{f_m}{\max(f_m)} K$$