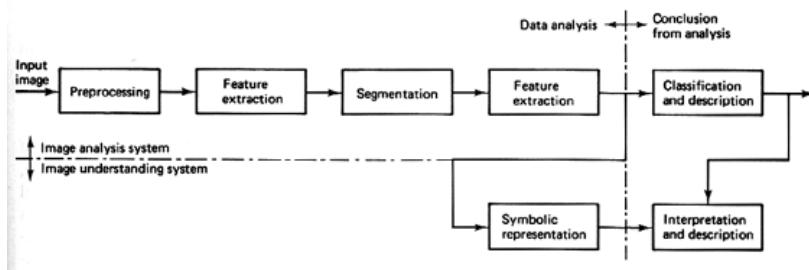


Επεξεργασία Πολυδιάστατων Δεδομένων – Υπολογιστική Όραση

Ηλίας Μαγκλογιάννης
Αναπληρωτής Καθηγητής
Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Ανάλυση ψηφιακών εικόνων

- Η ανάλυση ψηφιακών εικόνων σχετίζεται με την περιγραφή και την αναγνώριση του περιεχομένου μιας εικόνας.
- Οι τεχνικές ανάλυσης εικόνας αφορούν την πραγματοποίηση ποσοτικών ή ποιοτικών μετρήσεων σε μια εικόνα, για να παραχθεί μια περιγραφή της.
 - Μελέτη μεθόδων εξαγωγής χαρακτηριστικών (feature extraction), και ταξινόμησης (classification). Σκοπός είναι η εξαγωγή συγκεκριμένων χαρακτηριστικών που μπορούν να βοηθήσουν στην αναγνώριση αντικειμένων.
 - Οι ποσοτικές μετρήσεις των χαρακτηριστικών των αντικειμένων, επιτρέπουν την ταξινόμησή τους και την παραχωγή μιας περιγραφής και ερμηνείας της σκηνής από έναν υπολογιστή.

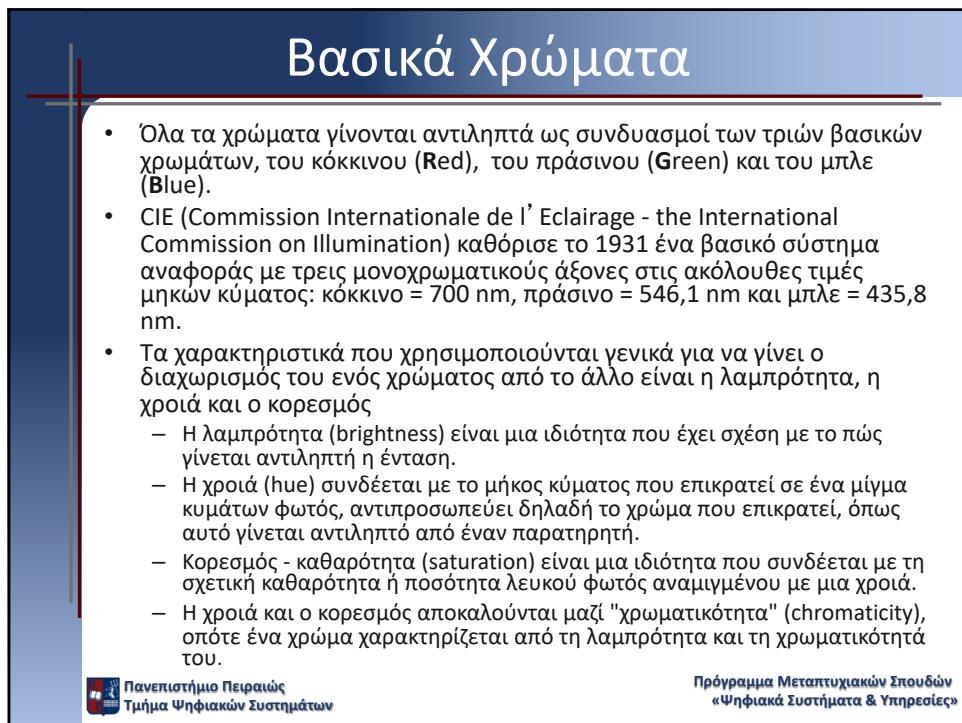
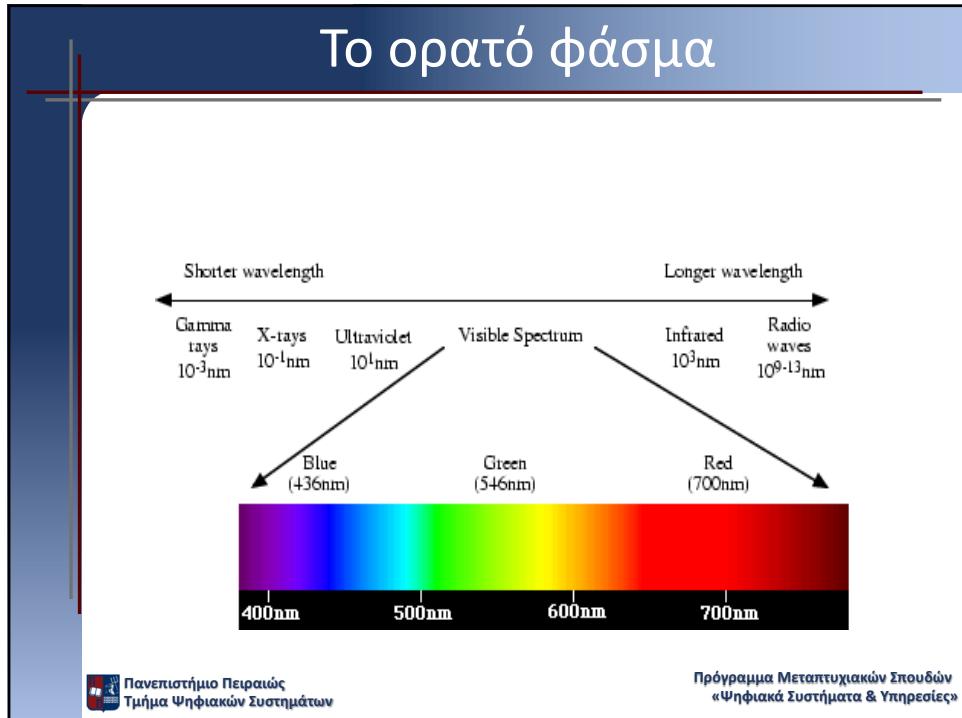


ΕΙΣΑΓΩΓΗ

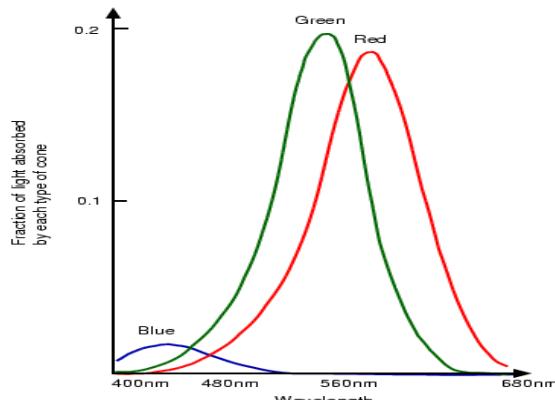
- Χρήση Χαρακτηριστικών:
 - Μείωση Πληροφορίας (Information Reduction)
 - Π.χ. Σύγκριση εικόνων βάση χαρακτηριστικών και όχι βάση εικονοστοιχείων
 - Απεικόνιση ή Αντιστοίχιση Πληροφορίας (Information Mapping)
 - Π.χ. Αντιστοίχιση χαρακτηριστικών -> κατηγοριοποίηση
 - Ονοματισμός Πληροφοριών (Information Labeling)
 - Π.χ. Γεωμετρικά χαρακτηριστικά
- Περιγραφή αντικειμένων με χαρακτηριστικά:
 - Απεικόνιση αντικειμένων στον χώρο των χαρακτηριστικών
 - -> Η αναγνώριση αντικειμένων = μέτρηση ομοιότητας μεταξύ χαρακτηριστικών
- Τα χαρακτηριστικά που επιλέγονται πρέπει:
 - Να διαχωρίζουν και να περιγράφουν μονοσήμαντα τα αντικείμενα

Χρώμα και ανθρώπινη αντίληψη

- Το χρώμα θεωρείται βασικό στοιχείο στην επεξεργασία εικόνας.
- Όταν το φως είναι άχρωμο, το μόνο χαρακτηριστικό του είναι η ένταση (intensity). Έτσι, ο όρος "επίπεδο του γκρι" (gray level) αναφέρεται σε ένα μέτρο της έντασης που κυμαίνεται από το μαύρο στο γκρι και τελικά στο άσπρο.
- Το έγχρωμο φως αντιστοιχεί σε μήκη κύματος 400-700 nm περίπου στο ηλεκτρομαγνητικό φάσμα.
- Τρεις είναι οι βασικές ποσότητες που χρησιμοποιούνται για να περιγράψουν τις ιδιότητες μιας έγχρωμης πηγής φωτός: η ακτινοβολία, η φωτεινότητα και η λαμπρότητα.
 - Η ακτινοβολία (radiance) είναι η συνολική ποσότητα ενέργειας που εκπέμπεται από την πηγή φωτός και συνήθως μετράται σε watt (W).
 - Η φωτεινότητα (luminance), που μετράται σε lumens (lm), δίνει ένα μέτρο της ποσότητας της ενέργειας που αντιλαμβάνεται ένας παρατηρητής από μια πηγή φωτός.
 - Η λαμπρότητα (brightness) είναι ένα υποκειμενικό χαρακτηριστικό που είναι πρακτικά αδύνατο να μετρηθεί και έχει σχέση με το πώς γίνεται αντιληπτή η ένταση.



Η απόκριση του ανθρώπινου ματιού



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Tristimulus τιμές

- Οι κανονικοποιημένες εντάσεις του κόκκινου, του πράσινου και του μπλε χρωματικού καναλιού που χρειάζονται για τη δημιουργία οποιουδήποτε χρώματος καλούνται tristimulus τιμές και ορίζονται αντίστοιχα ως εξής:

$$r = \frac{R}{R + G + B} \quad g = \frac{G}{R + G + B} \quad b = \frac{B}{R + G + B}$$

- Προφανώς ισχύει ότι:

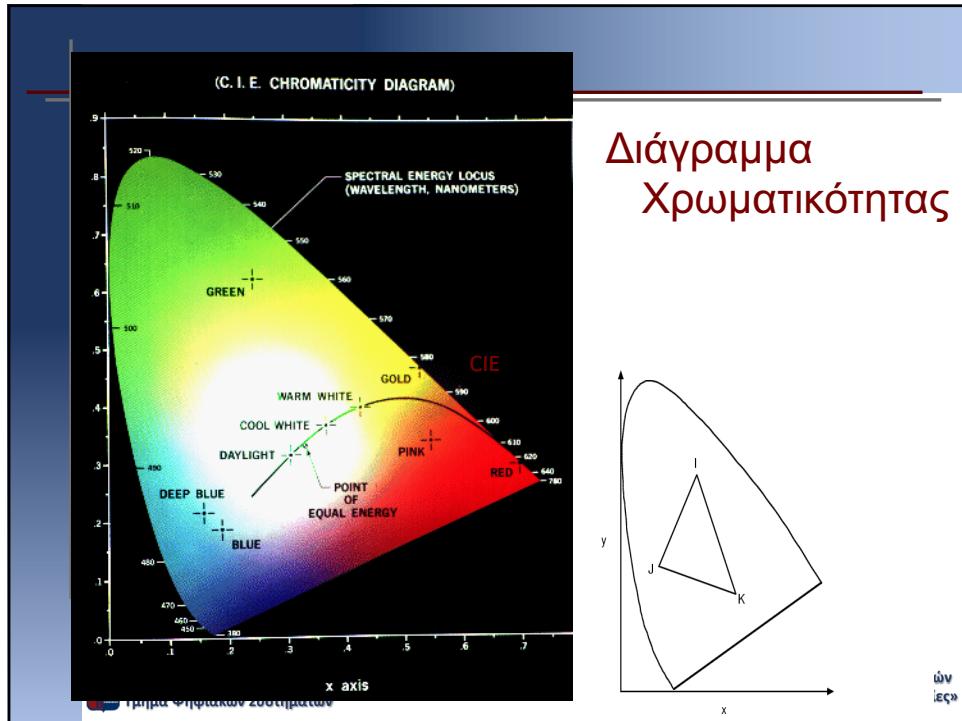
$$r + g + b = 1$$

- Οι τιμές των r , g , b που προκύπτουν από τον παραπάνω μη γραμμικό μετασχηματισμό καλούνται επίσης συνιστώσες χρωματικότητας (chromaticity coordinates). Για τη δημιουργία του χρώματος που αντιστοιχεί σε κάποιο μήκος κύματος φωτός στο ορατό φάσμα, οι tristimulus τιμές που απαιτούνται μπορούν να βρεθούν απ' ευθείας από εμπειρικούς πίνακες και καμπύλες.
- Διάγραμμα χρωματικότητας (chromaticity diagram)



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

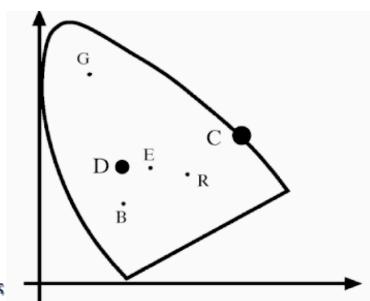


Ποιο είναι το κυρίαρχο χρώμα που αντιστοιχεί στο σημείο D.

Ποιο χρώμα αντιστοιχεί στο σημείο E.

Τι απεικονίζουν τα σημεία που είναι στο όριο του διαγράμματος (σχεδιασμένα με εντονη γραμή, πχ το σημείο C)

Τι απεικονίζουν τα σημεία που είναι έξω από αυτό.



Καταγραφή εικόνας

- Τι χρειαζόμαστε για να καταγράψουμε και να επεξεργαστούμε μια εικόνα;
 - Μια ψηφιακή κάμερα (CCD/CMOS), ένα σαρωτή (scanner), ή ένα μηχάνημα καταγραφής video
 - Ένα framegrabber για τη μετατροπή του αναλογικού ηλεκτρικού σήματος σε μια ψηφιακή μορφή
 - Έναν υπολογιστή και το κατάλληλο λογισμικό

Αισθητηρες εικονας (image sensors)

- Πιο συνηθισμένη μορφή: charge coupled devices – CCD
- Ορθογώνιο πλέγμα από φωτο-ευαίσθητα στοιχεία που απορροφούν φωτόνια και συγκεντρώνουν ηλεκτρικό ρεύμα
 - Η ένταση του ρεύματος είναι ανάλογη προς την ένταση της ακτινοβολίας

Αισθητήρες εικονας (image sensors)

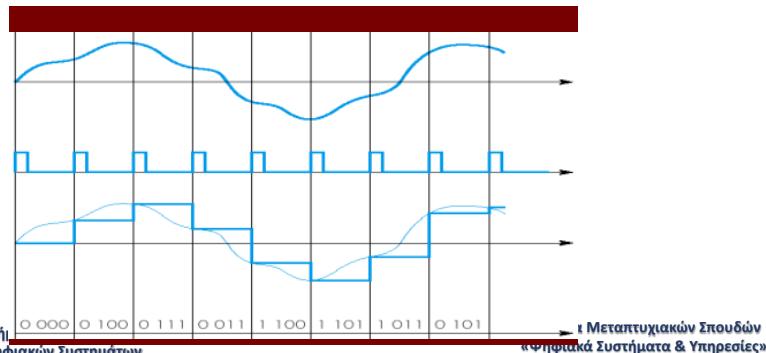
- Αισθητήρες τύπου CMOS
- Δεν έχουν την ίδια ποιότητα σε συνθήκες χαμηλού φωτισμού, αλλά...
- ...καταναλώνουν λιγότερη ενέργεια
- και παράγονται στη μορφή ολοκληρωμένων κυκλωμάτων



- Αριστερά: CMOS κάμερα με αναλογική και ψηφιακή έξοδο
- Δεξιά: Κάμερα CCD υπερυψηλής πιστότητας

Μετατροπέας Αναλογικού σε Ψηφιακού

- Διεξάγει **Δειγματοληψία** και **Κβαντοποίηση** για να μετατρέψει μια συνεχής κυματομορφή τάσης σε διακριτές τιμές
➔ σημαντική η **Συχνότητα Δειγματοληψίας** και το **Διάστημα Κβαντοποίησης**
- Οι κάρτες ψηφιοποίησης βίντεο (video digitizer board) συνήθως μπορούν να ενωθούν με την βιντεοκάμερα
- Οι νέες «εντελώς ψηφιακές» κάμερες περιλαμβάνουν ενσωματωμένο ADC

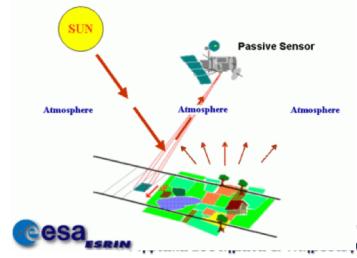
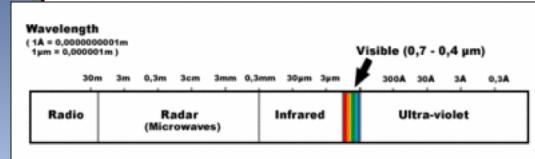


Οπτικοί Αισθητήρες

- Ο όρος αισθητήρας έχει προτιμηθεί έναντι του όρου κάμερα, γιατί αναφέρεται σε έναν ευρύτερο τρόπο λήψης δεδομένων
- Η κάμερα αναφέρεται συνήθως σε πληροφορίες ορατές, δηλαδή πληροφορίες που μπορούμε να τις δούμε μόνο με το ανθρώπινο μάτι
- Εφαρμογές Επεξεργασίας Εικόνας αφορούν πολλούς διαφορετικούς τύπους ακτινοβολίας του ηλεκτρομαγνητικού φάσματος.

Παθητικοί αισθητήρες

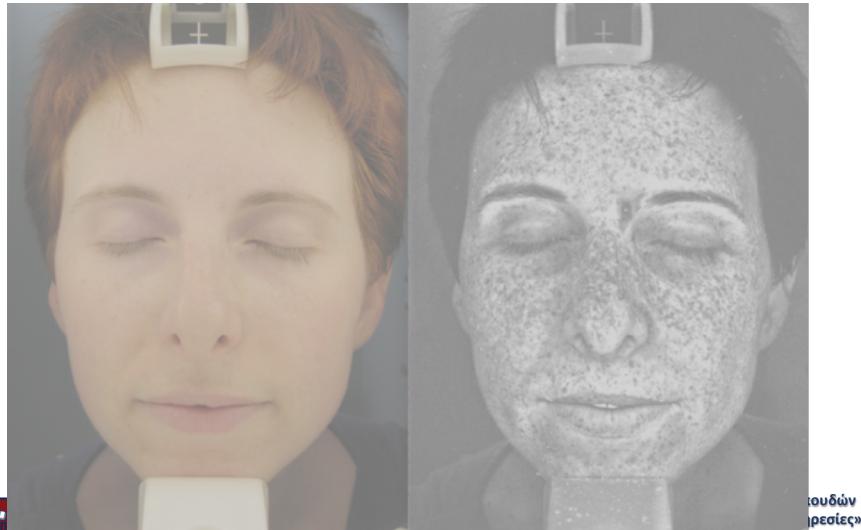
- Συλλέγουν ακτινοβολία στα διάφορα μήκη κύματος του ηλεκτρομαγνητικού φάσματος.
- Για παράδειγμα, οι υπέρυθρες διόπτρες είναι ειδικά φτιαγμένες για να «βλέπουν» την ακτινοβολία που εκπέμπεται (ακόμη και στο σκοτάδι) στο υπέρυθρο τμήμα του ηλεκτρομαγνητικού φάσματος.
- Γενικά, οι αισθητήρες που χρησιμοποιούν εξωτερικές πηγές ενέργειας για να «παρατηρήσουν» ένα αντικείμενο, όταν πχ χρησιμοποιούν το ηλιακό φως για να παρατηρήσουν τη γη, ονομάζονται παθητικοί αισθητήρες.



Ενεργητικοί αισθητήρες

- Οι «ενεργητικοί αισθητήρες» επειδή βασίζονται σε δικές τους πηγές ακτινοβολίας για να «φωτίσουν» τα αντικείμενα και να μπορέσουν στη συνέχεια να μετρήσουν την ενέργεια που ανακλάται και επιστρέφει στον αισθητήρα.
- Οι πιο κοινοί ενεργητικοί αισθητήρες είναι οι αισθητήρες απόστασης ή 'ραντάρ'.

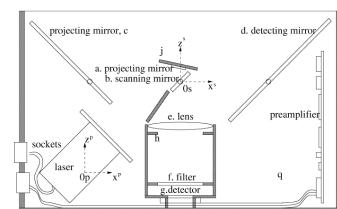
Εικόνα Δερμάτος στο Υπεριώδες



Αισθητήρες (απόστασης laser)

- Βασίζεται σε τεχνολογία laser, όπου μια ακτίνα προβάλλεται μέσω κατόπτρων στο χώρο και κάνει γραμμική σάρωση. Μετριέται ο χρόνος αντανάκλασης της ακτίνας και επομένως η απόσταση

- + αποφεύγεται επαφή
- + δεδομένα σε 3Δ
- + ανεξαρτησία από φωτισμό



- εποπτεία του χώρου μόνο κατά μια γραμμή
- χρειάζεται μετακίνηση για σάρωση χώρου
- κόστος



Πρόγραμμα «Ψηφί»
Σπουδών Ιππείσες

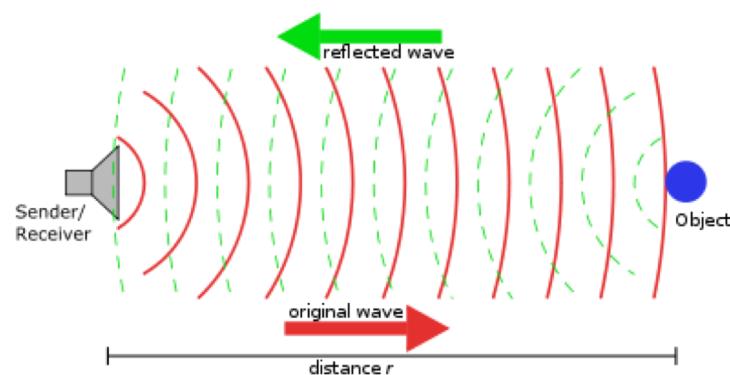
Αισθητήρες υπερήχων

- Βασίζονται στο φαινόμενο της αντανάκλασης των ηχητικών κυμάτων
 - + καλή και φτηνή λύση για ανίχνευση προσέγγισης μεγάλων όγκων
 - αυξημένος θόρυβος



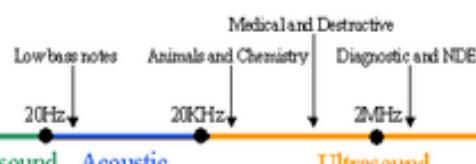
Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Ι
«Ψηφιακά Συστήματα & Υπηρεσίες»



Παν
Τμήμα Ψηφιακών Συστημάτων

υχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»



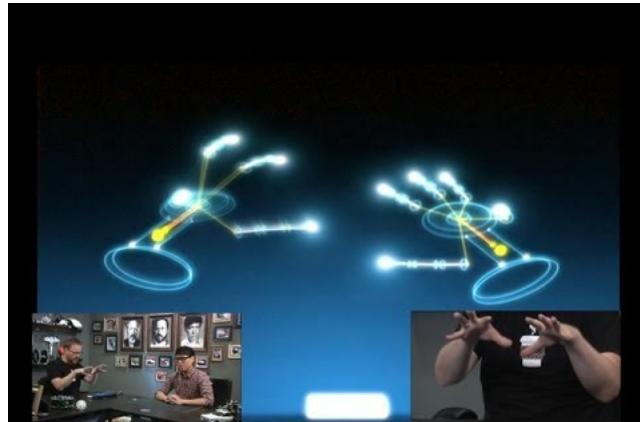
Εικόνες Βάθους – Αισθητήρας Kinect

- Η συγκεκριμένη συσκευή καταγραφής, εκτός από "κλασσική" RGB εικόνα (και βίντεο), μπορεί επίσης να καταγράψει "εικόνες βάθους", στις οποίες οι τιμή του κάθε pixel μεταφράζεται σε μία προσέγγιση της απόστασης του αντικειμένου από την κάμερα.
- Ενεργητικός αισθητήρας με χρήση υπέρυθρης ακτινοβολίας

Εικόνες Βάθους



Αισθητήρας Leap



Πανεπιστήμιο Πειραιώς

Τμήμα Ψηφιακών Συστημάτων

Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Χωρική Αναπαράσταση Εικόνων

- Οι εικόνες αποτελούνται από ένα σύνολο στοιχειωδών πληροφοριών:
 - Αντιστοιχούν στην τιμή ακτινοβολίας ή μίας μετρούμενης παραμέτρου
 - Οργανώνονται σε ορθογωνικές περιοχές
 - Ανάλογα με τη διάσταση των εικόνων, η στοιχειώδης πληροφορία ορίζεται:
 - 2D Εικόνες: Εικονοστοιχείο ή Pixel (Picture Element)
 - 3D Εικόνες: Χωροστοιχείο ή Voxel (Volume Element)
 - 4D Εικόνες: Υπερ-χωροστοιχείο ή Hypervoxel (Hyper Volume Element)
- Ως τέταρτη διάσταση θεωρείται ο χρόνος (t)

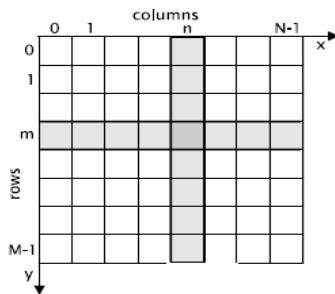
Πανεπιστήμιο Πειραιώς

Τμήμα Ψηφιακών Συστημάτων

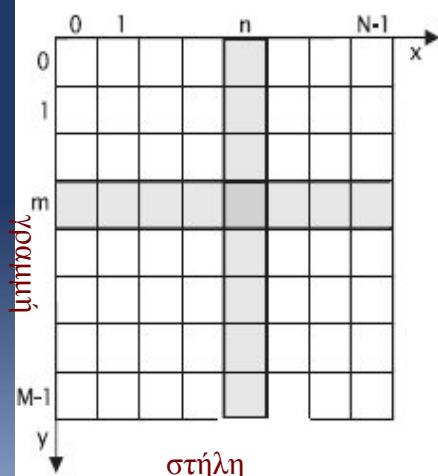
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Χωρική αναπαράσταση

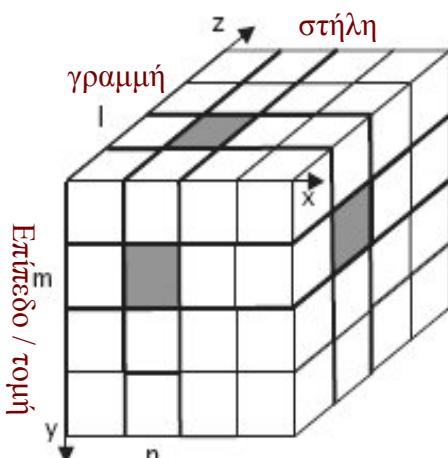
- Ένα pixel αντιστοιχεί στην τιμή της ακτινοβολίας σε μια στοιχειώδη επιφάνεια και όχι σε ένα μεμονωμένο σημείο
- Πόσα pixels είναι αρκετά



Χωρική 3Δ Αναπαράσταση Εικόνων



2D Εικόνα: Pixels
 Πανεπιστήμιο Πειραιώς
 Τμήμα Ψηφιακών Συστημάτων



3D Εικόνα: Voxels
 Πρόγραμμα Μεταπυχιακών Σπουδών
 «Ψηφιακά Συστήματα & Υπηρεσίες»

Ανάλυση εικόνας

c

d

e

f

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

«Ψηφιακά Συστήματα & Υπηρεσίες»

Ανάλυση εικόνας

3x4

48x64

12x16

192x256

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

«Ψηφιακά Συστήματα & Υπηρεσίες»

Κβαντισμός (Quantization) – βάθος

- Αντιπροσωπεύει το πόσο καλά μπορούν να διακρινονται διαφορές έντασης σε μία εικόνα (διακριτική ικανότητας αντίθεσης)
 - Συνδέεται με τον αριθμό των διακριτικών τιμών – στάθμες – στις οποίες αποθηκεύεται κάθε pixel της εικόνας
 - Χαμηλός αριθμός από στάθμες → μείωση ακμών

$$256 \Sigma \text{πάθμες} \Rightarrow \log_2(256) = 8 \text{bit}$$

Κβαντισμός -> Βάθος χρώματος



16 στάθμες

8 στάθμες

Κβαντισμός -> Βάθος χρώματος



4 στάθμες

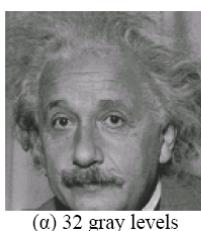


2 στάθμες

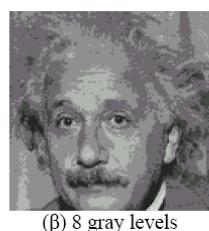
 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

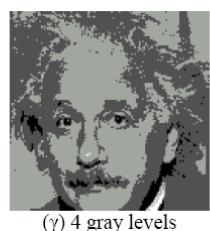
Βάθος χρώματος



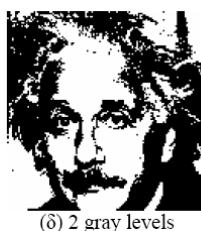
(α) 32 gray levels



(β) 8 gray levels



(γ) 4 gray levels



(δ) 2 gray levels

 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Ανάλυση εικόνας

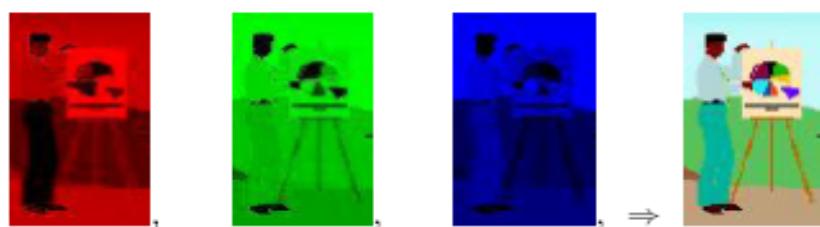
- Γενικά, η ανάλυση περιορίζεται από τους διαθέσιμους αισθητήρες και όχι από τις ανάγκες των εφαρμογών
- Οι διαθέσιμες αναλύσεις έχουν μάλλον μικρότερη διακριτιτητα από αυτές που χρησιμοποιούμε σε άλλα μεγέθη



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

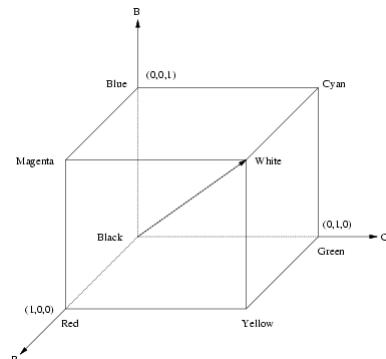
Δειγματοληψία



- Έγχρωμες εικόνες: επανάληψη της διαδικασίας τρεις φορές
 - Βασικά χρώματα: κόκκινο (R), πράσινο (G), μπλε (B)

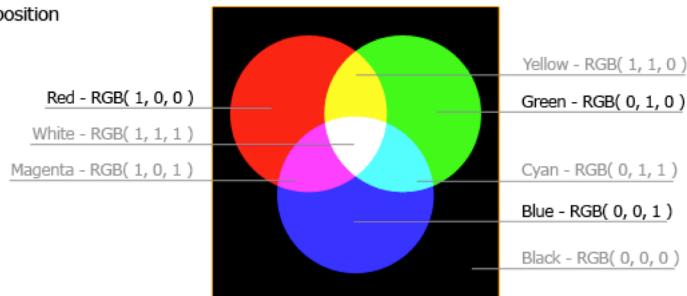
Χρωματικά Μοντέλα - RGB

- Βασίζεται σε ένα καρτεσιανό σύστημα συντεταγμένων
- Κάθε χρώμα αναλύεται στα βασικά φασματικά συστατικά
- Δημιουργήθηκε από την ψηφιοποίηση έγχρωμων εικόνων με τη χρήση κόκκινου, πράσινου και μπλε φίλτρου και μιας μονόχρωμης κάμερας.
- Η ορθογωνική γεωμετρία που ακολουθείται από το μοντέλο επιτρέπει την αναπαράσταση κάθε pixel από το αντίστοιχο διάνυσμα των συνιστωσών RGB ή από οποιονδήποτε γραμμικό ή μη γραμμικό μετασχηματισμό του



Additive Color Composition

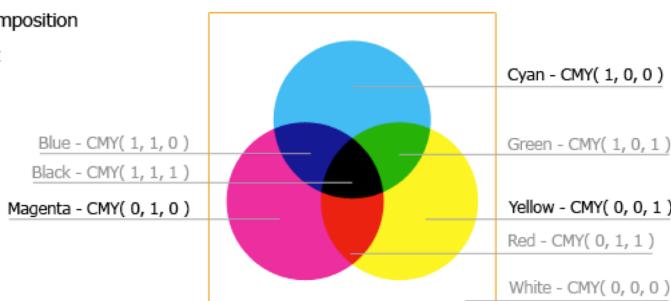
RGB Color Format



www.equasys.de

Subtractive Color Composition

CMY(K) Color Format



www.equasys.de

Χρωματικά Μοντέλα – YIQ

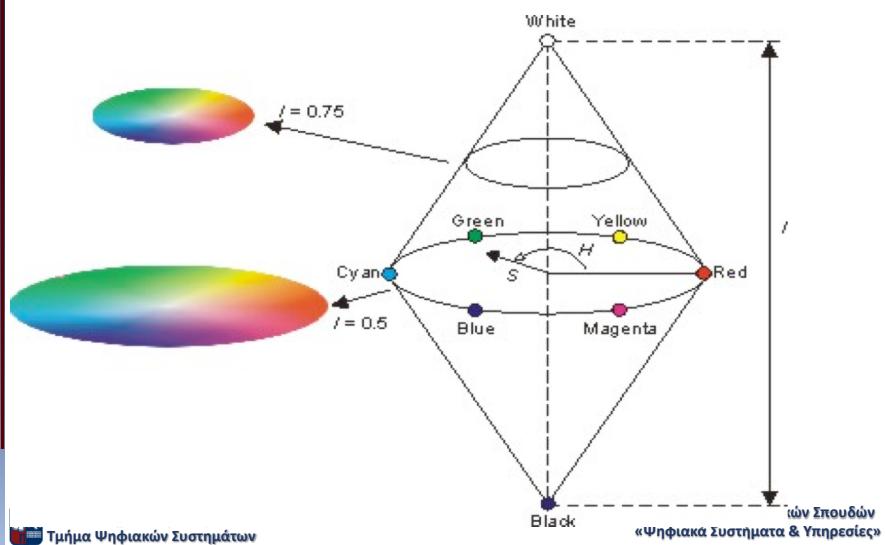
- Σχεδιάστηκε με σκοπό να εκμεταλλευτεί το γεγονός ότι η όραση του ανθρώπου παρουσιάζει μεγαλύτερη ευαισθησία στις μεταβολές της φωτεινότητας από ότι στις μεταβολές της χροιάς ή του κορεσμού.
- Ο μετασχηματισμός από το μοντέλο RGB στο YIQ είναι ο ακόλουθος

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Χρωματικά Μοντέλα - HIS



Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Μετατροπή

RGB to HSI

First, we convert RGB color space image to HSI space beginning with normalizing RGB values:

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B}.$$

Each normalized H, S and I components are then obtained by,

$$h = \cos^{-1} \left\{ \frac{0.5 \cdot [(r-g)+(r-b)]}{\sqrt{[(r-g)^2 + (r-b)(g-b)]}} \right\} \quad h \in [0, \pi] \text{ for } b \leq g$$

$$h = 2\pi - \cos^{-1} \left\{ \frac{0.5 \cdot [(r-g)+(r-b)]}{\sqrt{[(r-g)^2 + (r-b)(g-b)]}} \right\} \quad h \in [\pi, 2\pi] \text{ for } b > g$$

$$s = 1 - 3 \cdot \min(r, g, b) \quad s \in [0, 1]$$

$$i = (R + G + B) / (3 \cdot 255) \quad i \in [0, 1].$$

For convenience, h, s and i values are converted in the ranges of [0,360], [0,100], [0, 255], respectively by: $H = h \times 180/\pi$; $S = s \times 100$ and $I = i \times 255$.

HSI to RGB

$h = H \cdot \pi / 180 ; s = S / 100 ; i = I / 255$
 $x = i \cdot (1 - s)$
 $y = i \cdot \left[1 + \frac{s \cdot \cos(h)}{\cos(\pi / 3 - h)} \right]$
 $z = 3i - (x + y);$
when $h < 2\pi / 3$, $b = x$; $r = y$ and $g = z$.
when $2\pi / 3 \leq h < 4\pi / 3$, $h = h - 2\pi / 3$, and $r = x$; $g = y$ and $b = z$.
when $4\pi / 3 \leq h < 2\pi$, $h = h - 4\pi / 3$, and $g = x$; $b = y$ and $r = z$.
The result r, g and b are normalized values which are in the ranges of [0,1], therefore, they should be multiplied by 255 for displaying.

«Ψηφιακά Συστήματα & Υπηρεσίες»

Αναπαράσταση με πίνακες

```

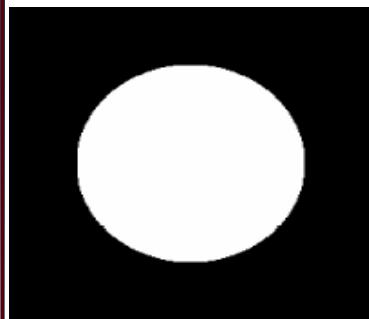
>> for i = 1 : 256
    for j = 1 : 256
        A(i,j) = j - 1;
    end
end
>> image(A);
>> colormap(gray(256));
>> axis('image');

```

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Αναπαράσταση με πίνακες



```
>> for i = 1 : 256
    for j = 1 : 256
        dist = ((i - 128)^2 + (j - 128)^2)^(.5);
        if (dist < 80)
            B(i, j) = 255;
        else
            B(i, j) = 0;
        end
    end
>> image(B);
>> colormap(gray(256));
>> axis('image');
```

Αναστροφή του πίνακα



```
>> for i = 1 : 512
    for j = 1 : 512
        B(j, i) = A(i, j);
    end
end
```

Αναστροφή της εικόνας



```
>> clear B;
>> for i = 1 : 512
    for j = 1 : 512
        B(i, 512 + 1 - j) = A(i, j);
    end
end
```

 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Η έννοια της γειτονιάς του pixel

2D Εικόνες:

- Γειτονικά (neighbors) Pixels θεωρούνται όσα έχουν κοινή ακμή ή τουλάχιστον μία κοινή γωνία
- Pixels μπορούν να έχουν 4 ή 8 γείτονες (4-connected και 8-connected-)

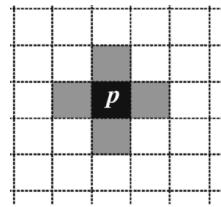
• 3D Εικόνες:

- Voxels μπορούν να έχουν 6 ή 18 ή 26 γείτονες σε κυβικό πλέγμα

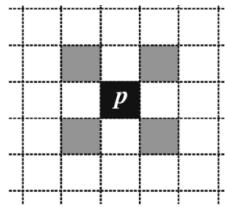
 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

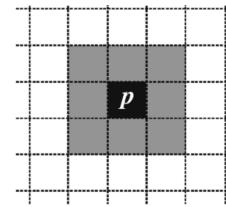
Γειτονιες εικονοστοιχείων



(a)



(b)



(c)

4-neighbors (a),

d-neighbors(b),

8-neighbors(c)



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Γειτονιες εικονοστοιχείων

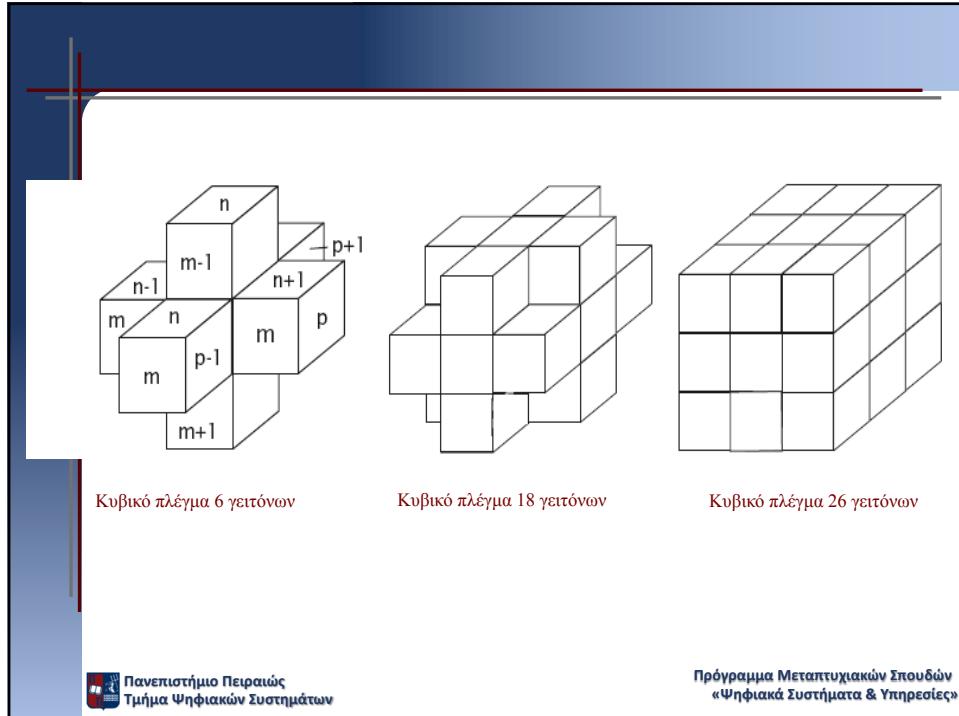
	$y - 1$	y	$y + 1$
$x - 1$	$(x-1, y-1)$	$(x-1, y)$	$(x-1, y+1)$
x	$(x, y-1)$	(x, y)	$(x, y+1)$
$x + 1$	$(x+1, y-1)$	$(x+1, y)$	$(x+1, y+1)$

$$N_8(p) = \{(x-1,y-1), (x-1,y), (x-1,y+1), (x,y-1), (x,y+1), (x+1,y-1), (x+1,y), (x+1,y+1)\}$$



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»



Μετρική της απόστασης σε εικόνες

- Έστω τα pixels p_1 και p_2 της εικόνας I $p_1(x_1, y_1), p_2(x_2, y_2) \in I$
- Ευκλείδεια απόσταση $D(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$
- D_4 μετρική απόστασης (city block distance):

$$D_4(p_1, p_2) = |x_1 - x_2| + |y_1 - y_2|$$

D_4

$Eυκλείδεια$

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Γειτονιες εικονοστοιχείων

$$N_D(p) = \{(x-1,y-1), (x-1,y+1), (x+1,y-1), (x+1,y+1)\}$$

$$N_8(p) = N_4(p) \cup N_D(p).$$

Οι Τεχνικές Φιλτραρίσματος στο Πεδίο του Χώρου εφαρμόζονται σε περιοχές εικόνας

Οι τελεστές αυτοί χρησιμοποιούν τον γενικό αλγόριθμο της **κυλιόμενης (ή κινητής) μάσκας –moving mask**.

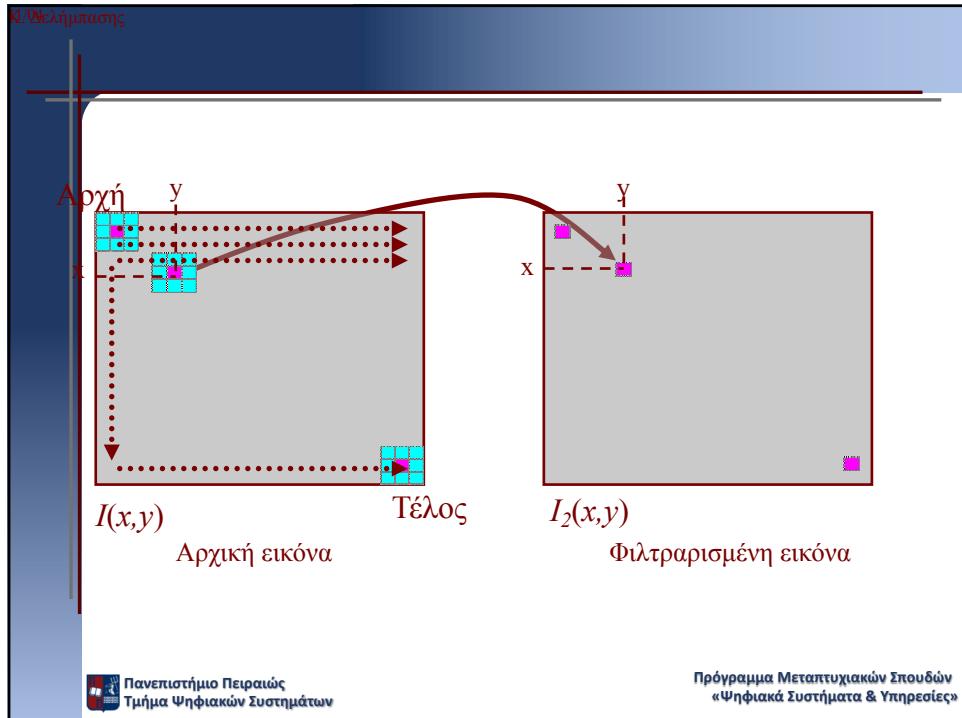
Εφαρμογή για εικόνα $I \in N \times M$ (γραμμές x στήλες) και τετραγωνική περιοχή (μάσκα) $2n+1 \times 2m+1$ pixels και αποθήκευση του αποτελέσματος στην εικόνα I_2

Για κάθε γραμμή x από $n+1$ ως $N_x - (n+1)$

Για κάθε στήλη y από $m+1$ ως $N_y - (m+1)$

Υπολόγισε την τιμή a του τελεστή στην περιοχή γύρω από το (x, y) ;

$$I_2(x, y) = a;$$



Συνέλιξη εικόνων

- Έστω εικόνα $I(x,y)$, $x=0..N-1$, $y=0..M-1$ και μάσκα M με περιττό αριθμό γραμμών και στηλών $(2n+1)x(2m+1)$: $M(k,l)$, $k=-n..n$, $l=-m..m$.
- Η συνέλιξη $I * M$, στη θέση (x,y) ορίζεται (κατ' αναλογία με την μονοδιάστατη περίπτωση των σημάτων) ως εξής:

$$I * M = c(x,y) = \sum_{k=-n}^n \sum_{l=-m}^m I(x-k, y-l) M(k, l)$$

- Στον ορισμό αυτό οι δείκτες m και n διατρέχουν τη μάσκα M .
- Στην υπολογιστική υλοποίηση της συνέλιξης πρέπει να ληφθεί μέριμνα ώστε οι δείκτες να μην έχουν αρνητικές τιμές

• Έστω εικόνα $f(x,y)$ και μάσκα $m(i,j)$ 3x3. Ο υπολογισμός της συνέλιξης τους στο σημείο (x_0, y_0) γίνεται ως εξής:

- Το κεντρικό pixel της m τοποθετείται «πάνω» από το pixel (x_0, y_0) της f .
- Υπολογίζεται το άθροισμα των γινομένων «χιαστί» στις 2 διαστάσεις, με τρόπο ανάλογο της μιας διάστασης (σήματα).
- Το αποτέλεσμα (άθροισμα των γινομένων) αποθηκεύεται στο pixel (x_0, y_0) μίας νέας εικόνας f_2 .
- Η μάσκα m κινείται ώστε το κεντρικό pixel της m να βρεθεί πάνω από το επόμενο pixel της f και επαναλαμβάνεται η διαδικασία.

Πανεπιστήμιο Αιγαίου
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΦΙΛΤΡΟ ΜΕΣΗΣ ΤΙΜΗΣ (mean FILTER)

Η λειτουργία του φίλτρου μέσης τιμής συνίσταται με την αντικατάσταση της φωτεινότητας σε κάθε εικονοστοιχείο με τη μέση φωτεινότητα σε μια γειτονιά του.

Αν N είναι η γειτονιά του εικονοστοιχείου (i,j) μιας εικόνας I , τότε η τιμή του εικονοστοιχείου (i,j) αντικαθίσταται με τη βοήθεια της σχέσης:

$$I'(i,j) = \frac{1}{M} \sum_{(x,y) \in N} I(x,y)$$

όπου M το πλήθος των εικονοστοιχείων της γειτονιάς N .

ΦΙΛΤΡΟ ΜΕΣΗΣ ΤΙΜΗΣ (mean FILTER)

Η γειτονιά Ν είναι συνήθως καθορισμένη για κάθε επεξεργασία και συνήθως αντιστοιχεί σε τετράγωνες μάσκες. Έτσι για ακτίνα ίση με ένα έχουμε ουσιαστικά μια γειτονιά διαστάσεων 3×3 . Ένα 3×3 φίλτρο μέσης τιμής μπορεί πρακτικά να υλοποιηθεί με μια μάσκα της μορφής:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

57

ΦΙΛΤΡΟ ΜΕΣΗΣ ΤΙΜΗΣ (mean FILTER)

Το φίλτρο μέσης τιμής μπορεί να θεωρηθεί ως ένα κατωδιαβατό φίλτρο.

Αν θέλουμε να τονίσουμε περισσότερο τη συνεισφορά των εικονοστοιχείων ανάλογα με την απόστασή τους, τότε μπορούμε να χρησιμοποιήσουμε μάσκες εξομάλυνσης όπως η παρακάτω

$$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

58

39

Φίλτρα ανίχνευσης ακμών

	Roberts	Prewitt	Sobel	Robinson	Frei-Chen
$G_R(n,m)$	$\begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix}$
$G_c(n,m)$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & \sqrt{2} & -1 \end{bmatrix}$

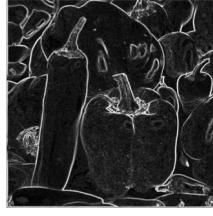
Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

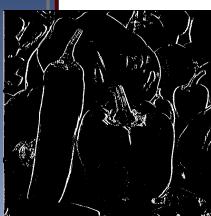
50

Εφαρμογή των μασκών του Sobel



(α) αρχική εικόνα
(β) Sobel G_R
(γ) Sobel G_C
(δ) Sobel $|G(n,m)|$






(ε) Sobel G_R με $T=96$
(στ) Sobel G_C με $T=98$
(ζ) Sobel $|G(n,m)|$ με $T=112$
(η) Sobel $\max(|G_R|, |G_C|)$, με $T=112$

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

51

Εφαρμογή άλλων μασκών



(a) Roberts |G(n,m)|
με T=25

(b) Prewitt |G(n,m)|
με T=99

(γ) Robinson |G(n,m)|
με T=102

(δ) Frei-Chen |G(n,m)|
με T=105

 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Καθηγητής
Νικόλαος Παπαμάρκος

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Παράδειγμα Prewitt

109	112	113	108	106	115	108	105
106	111	116	110	111	107	112	113
118	113	112	110	113	115	112	109
110	111	114	117	110	115	107	110
190	191	179	180	191	185	183	181
178	180	185	187	190	188	185	183
190	191	179	182	191	190	187	186
183	179	184	189	194	189	188	183

⊗

-1	-1	-1
0	0	0
1	1	1

=

9	2	8	9	11	8
2	5	4	14	2	0
217	215	215	218	219	213
208	210	221	223	231	224
0	2	2	7	9	14
3	0	5	7	8	4

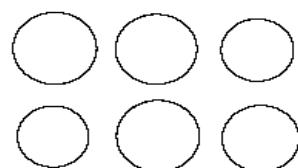
 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Ανθρώπινη και μηχανική όραση

- Ποιες διαφορές στην ένταση μπορούμε να διακρίνουμε;
- Ποια είναι η χωρική ανάλυση του ματιού;
- Με πόση ακρίβεια μπορούμε να μετρήσουμε αποστάσεις και επιφάνειες;
- Πώς αντιλαμβανόμαστε τα χρώματα;
- Με βάση ποιες ιδιότητες διακρίνουμε και αναγνωρίζουμε τα αντικείμενα;

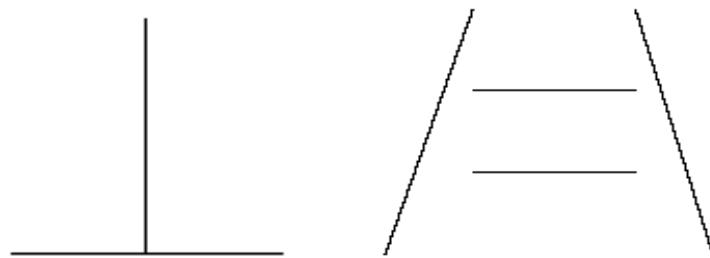
Ανθρώπινη και μηχανική όραση



Ανθρώπινη και μηχανική όραση

- Οι μικρές διαφορές στο μήκος των παράλληλων γραμμών είναι σχετικά ευδιάκριτες
- Δε συμβαίνει το ίδιο με τις επιφάνειες των κύκλων!

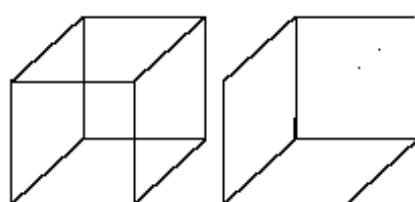
Ανθρώπινη και μηχανική όραση



Ανθρώπινη και μηχανική όραση

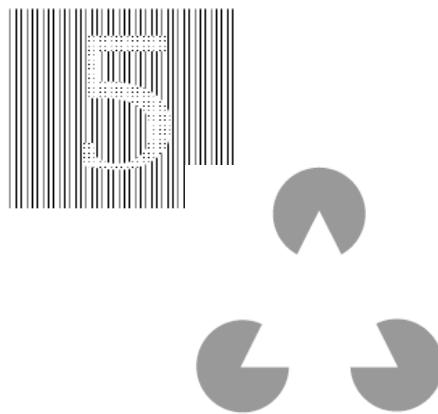
- Οι γραμμές έχουν το ίδιο μήκος!
- Αλλά ο προσανατολισμός και η θέση τους μας δίνει άλλη εντύπωση
- Όπως και οι πλάγιες γραμμές...
- ...καταργώντας την αίσθηση της παραλληλίας

Ανθρώπινη και μηχανική όραση



- Δεν υπάρχει η πληροφορία ορατών ή όχι πλευρών
- Μπορούμε όμως να εκτιμήσουμε την τρίτη διάσταση των αντικειμένων

Ανθρώπινη και μηχανική όραση



- Απουσία συνεχούς χρωματισμένης περιοχής
- Μπορούμε εύκολα να κατανοήσουμε το περιεχόμενο...
- ...και να φανταστούμε το ισοπλευρο τρίγωνο, ακόμα και χωρίς τις πλευρές του

Χαρακτηριστικά Εικόνας

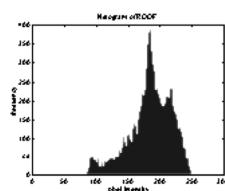
- Τύποι χαρακτηριστικών
 - Μετρήσιμα: Συνεχή και Ασυνεχή
 - Συμβολικά: Π.χ. Αλφαριθμητικές μεταβλητές (ΝΑΙ-ΟΧΙ, 0-1, χρώμα εικόνων, κτλ.)
- Κατηγορίες χαρακτηριστικών
 - Χαρακτηριστικά χώρου (spatial features)
 - Γεωμετρικά χαρακτηριστικά (geometrical features)
 - Στατιστικά χαρακτηριστικά όπως είναι οι ροπές (moments)
 - Χαρακτηριστικά υφής (texture features)
 - Χαρακτηριστικά από μετασχηματισμό (transform features)
 - Π.χ. Η εικόνα μετασχηματίζεται και τα χαρακτηριστικά ορίζονται και εξάγονται στο πεδίο της συχνότητας

ΕΙΣΑΓΩΓΗ

- Ιδιότητες χαρακτηριστικών:
 - Διακριτικότητα: Το σύνολο των χαρακτηριστικών θα πρέπει να είναι ικανό να διαχωρίζει τα αντικείμενα μεταξύ τους
 - Αξιοπιστία: Για παρόμοια αντικείμενα να προκύπτουν κοντινές τιμές για τα χαρακτηριστικά
 - Μικρό υπολογιστικό κόστος (πολυπλοκότητα)
 - Καλή προσαρμογή κατά την ταξινόμηση (classification): Να ταιριάζουν τα χαρακτηριστικά με τη μέθοδο ταξινόμησης
 - Ευαισθησία στο θόρυβο

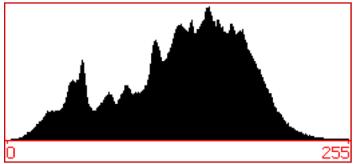
ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Μορφή Ιστογράμματος:
 - Σημαντικές πληροφορίες για το χαρακτήρα εικόνας: π.χ. Χαμηλή αντίθεση -> στενός λοβός ιστογράμματος



ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Μορφή Ιστογράμματος:
 - Σημαντικές πληροφορίες για το χαρακτήρα εικόνας: π.χ. Υψηλή αντίθεση (διαφορετικές αποχρώσεις, αντικειμένα, διαφορετικό φόντο) -> διαφορετικοί λοβοί, «απλωμένο ιστόγραμμα»


 \rightarrow




Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Συνήθη χαρακτηριστικά που εξάγονται από ιστόγραμμα:
 - Ροπές
 - Απόλυτες Ροπές
 - Κεντρικές Ροπές
 - Απόλυτα Κεντρικές Ροπές
 - Εντροπία



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

$$m_i = \sum_{b=0}^{L-1} b^i P(b)$$

$$\hat{m}_i = \sum_{b=0}^{L-1} |b|^i P(b)$$

$$\mu_i = \sum_{b=0}^{L-1} (b - m_1)^i P(b)$$

$$\hat{\mu}_i = \sum_{b=0}^{L-1} |b - m_1|^i P(b)$$

$$H = - \sum_{b=0}^{L-1} P(b) \log_2 [P(b)]$$

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΟΡΟΥ

- Συνήθη χαρακτηριστικά που εξάγονται από ιστόγραμμα:
 - Μέση τιμή

$$m_i = \sum_{b=0}^{L-1} b P(b)$$
 - Τυπική απόκλιση

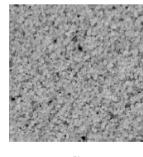
$$s = \left[\sum_{b=0}^{L-1} (b - m_i)^2 P(b) \right]^{1/2}$$
 - Μεταβλητότητα

$$\mu_2 = \sum_{b=0}^{L-1} (b - m_i)^2 P(b)$$

 Πανεπιστήμιο Πειραιώς
 Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
 «Ψηφιακά Συστήματα & Υπηρεσίες»

Παραδείγματα Ροπών



α

$m = 154.6$
 $\mu_2 = 0.8 \cdot 10^3$
 $\mu_3 = -2.5 \cdot 10^4$
 $\mu_4 = 3.0 \cdot 10^6$



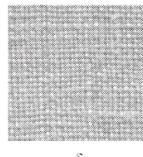
β

$m = 91.4$
 $\mu_2 = 5.7 \cdot 10^3$
 $\mu_3 = 3.3 \cdot 10^5$
 $\mu_4 = 7.4 \cdot 10^7$



γ

$m = 156.3$
 $\mu_2 = 1.5 \cdot 10^3$
 $\mu_3 = -4.8 \cdot 10^4$
 $\mu_4 = 7.8 \cdot 10^6$



δ

$m = 178.7$
 $\mu_2 = 1.0 \cdot 10^3$
 $\mu_3 = -1.0 \cdot 10^3$
 $\mu_4 = 2.5 \cdot 10^6$



ϵ

$m = 107.0$
 $\mu_2 = 3.3 \cdot 10^3$
 $\mu_3 = 1.5 \cdot 10^5$
 $\mu_4 = 3.0 \cdot 10^7$

 Πανεπιστήμιο Πειραιώς
 Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
 «Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Συνήθη χαρακτηριστικά που εξάγονται από ιστόγραμμα:

- Λοξότης

$$\mu_3 = \sum_{b=0}^{L-1} (b - m_l)^3 P(b)$$

- Κύρτωση

$$\mu_4 - 3 = \sum_{b=0}^{L-1} (b - m_l)^4 P(b) - 3$$

- Ενέργεια

$$E = \left[\sum_{b=0}^{L-1} P(b) \right]^2$$

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Σε περίπτωση κινούμενου παραθύρου (π.χ. Εξετάζουμε γειτονιές της εικόνας για εύρεση χαρακτηριστικών) είναι δυνατό να υπολογιστούν χαρακτηριστικά ιστογράμματος χωρίς να υπολογιστεί το συνολικό ιστόγραμμα:

$$m_i(k, \ell) = \frac{1}{N} \sum_{(m,n) \in W} \sum [u(m-k, n-\ell)]^i$$

$$\mu_i(k, \ell) = \frac{1}{N} \sum_{(m,n) \in W} \sum [u(m-k, n-\ell) - m_l(k, \ell)]^i$$

- $i = 1, 2, \dots, N$ πλήθος εικονοστοιχείων στο παράθυρο

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

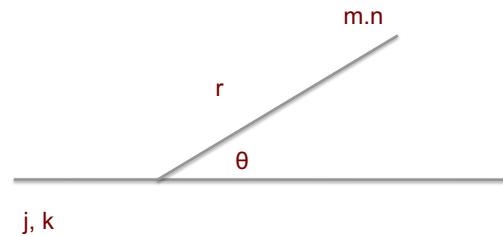
- Συνδεδεμένες πιθανότητες
- 2^ο βαθμού (second order joint probabilities)

$$P(b_1, b_2) = \frac{N(b_1, b_2)}{M(b_1, b_2)}$$

- όπου:
 - $N(b_1, b_2)$: Το σύνολο ζευγών των εικονοστοιχείων που ικανοποιούν κάποια σχέση ή ιδιοτητα
 - $M(b_1, b_2)$: Το σύνολο των διασυνδεδεμένων σημείων στην περιοχή

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΧΩΡΟΥ

- Συνδεδεμένες πιθανότητες
- Η σχέση που συνδέει τα εικονοστοιχεία μπορεί να είναι οποιαδήποτε χωρική σχέση:



ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΠΟ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟ

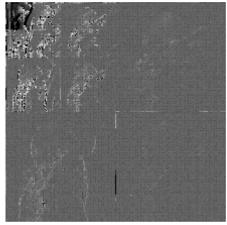
- DCT, DFT, Wavelets, ...
- Πληροφορίες εικόνας στο πεδίο συχνότητας
- DCT + Wavelets: Προσεγγίζουν καλύτερα το φάσμα συχνότητας με λιγότερους συντελεστές



Original



DC Decomposition



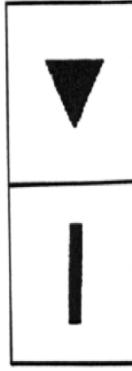
3 Level DW Decomposition



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΠΟ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟ












Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

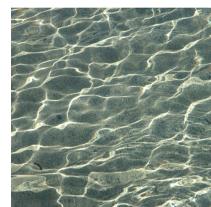
Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- **Ύπαρξη Υφής (texture):**
 - Κάποια τοπική «δομή» επαναλαμβάνεται σε μια περιοχή εικόνας που είναι μεγάλη σε σχέση με το μέγεθος της δομής
 - Η δομή συνίσταται σε μη τυχαία ρύθμιση των στοιχειωδών μερών της
 - Τα μέρη είναι κατά προσέγγιση ομοιόμορφες οντότητες που έχουν περίπου τις ίδιες διαστάσεις παντού μέσα στις περιοχές που υπάρχει υφή
- Η υφή παίζει σημαντικό ρόλο σε θέματα τεχνητής όρασης:
 - Επιθεώρηση επιφάνειας, ανάκτηση & ταξινόμηση εικόνων, προσδιορισμός σχημάτων, κτλ.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Γενικά: Ένα συνδεδεμένο σύνολο εικονοστοιχείων το οποίο ικανοποιεί μια συγκεκριμένη ιδιότητα που λαμβάνει χώρα επανειλημμένα σε μια περιοχή της εικόνας δημιουργεί μια **περιοχή υφής**.
- Παραδείγματα εικόνων υφής



ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Σημαντικότερα χαρακτηριστικά υφής:
 - Πίνακας συνεμφάνισης (co-occurrence matrix)
 - Αυτοσυσχέτιση (Autocorrelation)
 - Χαρακτηριστικά υφής του Laws

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Πίνακας συνεμφάνισης $P[i,j]$
- Διάνυσμα διασύνδεσης $d = (dx, dy)$
- Μετράμε όλα τα ζευγάρια εικονοστοιχείων τα οποία διασυνδέονται μεταξύ τους με το d και έχουν γκρι αποχρώσεις i και j .
- Δηλαδή:
 - Αν έχουμε 3 επίπεδα γκρι (0, 1, 2) ο πίνακας $P[i,j]$ θα είναι 3x3
 - Αν έχουμε $d(1,1)$: το εικονοστοιχείο $r(i,j)$ διασυνδέεται με το εικονοστοιχείο $r(i+1, j-1)$, δηλαδή μετακινούμαστε διαγώνια κάτω δεξιά

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Παράδειγμα (για $d = (1,1)$):

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

2	3	1
1	0	4
2	2	1

- Στο συγκεκριμένο παράδειγμα, υπάρχουν τρία ζευγάρια εικονοστοιχείων τα οποία έχουν τιμές $[0,1]$ και τα οποία διασυνδέονται από τη συγκεκριμένο διάνυσμα, οπότε το στοιχείο $P[0,1]$ λαμβάνει την τιμή 3
- Μπορούμε να κανονικοποιήσουμε τον πίνακα $P[i,j]$ και να τον χρησιμοποιήσουμε ως συνάρτηση πιθανότητας (εφόσον και το άθροισμα των στοιχείων του = 1)

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Ένα χαρακτηριστικό το οποίο μετρά την τυχαιότητα των φωτεινοτήτων στην εικόνα είναι η εντροπία

$$\text{Entropy} = - \sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} P[i,i] \log_2 P[i,j]$$

- Όπου $N_c \times N_c$ το μέγεθος του πίνακα συνεμφάνισης

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Επιπλέον χαρακτηριστικά που εξάγονται από τον πίνακα συνεμφάνισης:

Μέγιστη τιμή:

$$\max_{i,j} (p_{i,j})$$

Ομοιογένεια (καλείται και Inverse Difference Moment):

$$Homogeneity = \sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} \frac{p_{i,j}}{1 + (i-j)^2}$$

Ενέργεια:

$$Energy = \sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} p_{i,j}^2$$

Ανομοιογένεια:

$$Dissimilarity = \sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} |i - j| p_{i,j}$$

Αντίθεση:

$$Contrast = \sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} (i - j)^2 p_{i,j}$$

Συσχέτιση:

$$Correlation = \frac{\sum_{i=0}^{N_c-1} \sum_{j=0}^{N_c-1} (i - \mu_x)(j - \mu_y) p_{i,j}}{\sigma_x \sigma_y}$$



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Αυτοσυσχέτιση
- $p(k,l)$ για εικόνα $N \times M$ ορίζεται ως:

$$p(k, \ell) = \frac{MN}{(N-k)(M-\ell)} \frac{\sum_{i=1}^{N-k} \sum_{j=1}^{M-\ell} f(i, j)f(i+k, j+\ell)}{\sum_{i=1}^N \sum_{j=1}^M f^2(i, j)},$$

- όπου

$$0 \leq k \leq N-1, \quad 0 \leq \ell \leq M-1$$



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Για εικόνες με επαναλαμβανόμενα σχήματα υφής:
 - Η συνάρτηση παρουσιάζει περιοδική συμπεριφορά
 - Περίοδος = διάστημα μεταξύ γειτονικών στοιχείων της υφής
- Χρησιμοποιείται ως μέτρο περιοδικότητας υφής και ως μέτρο κλίμακας των στοιχείων της υφής

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Χαρακτηριστικά υφής Laws
- Γνωστά ως φίλτρα μέτρησης της ενέργειας υφής
- Η διαμόρφωση των χαρακτηριστικών προκύπτει από 3 διανύσματα:
 - $L_3=(1,2,1)$
 - $E_3=(-1,0,1)$
 - $S_3=(-1,2,-1)$
- Το διάνυσμα L_3 σχετίζεται με τον τονισμό της μέσης τιμή, το διάνυσμα E_3 με σημεία ακμής, ενώ το διάνυσμα S_3 τονίζει τα στίγματα (spots) μια γραμμή.

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Χαρακτηριστικά υφής Laws
- Από τη συνέλιξη των διανυσμάτων προκύπτουν τα ακόλουθα 5 διανύσματα του Laws:
 - $L_5 = L_3 * L_3 = (1,4,6,4,1)$
 - $E_5 = L_3 * E_3 = (-1,-2,0,2,1)$
 - $S_5 = L_3 * S_3 = (-1,0,2,0,-1)$
 - $W_5 = -E_3 * S_3 = (-1,2,0,-2,1)$
 - $R_5 = S_3 * S_3 = (1,-4,6,-4,1)$

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΦΗΣ

- Από τα μονοδιάστατα διανύσματα παράγονται 25 δισδιάστατες μάσκες ως εξής:

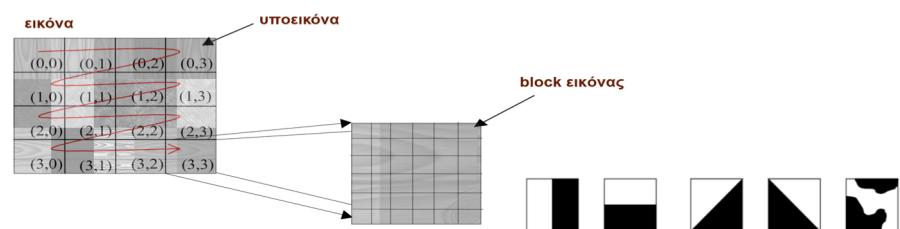
$$LE = I_3^T E_3 = \begin{bmatrix} 1 \\ 4 \\ 6 \\ 4 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

- Λίστα ονομάτων 25 μασκών:
- Με τη χρήση των μασκών παράγονται χαρακτηριστικά υφής που από την εφαρμογή (συνέλιξη) των μασκών αυτών με την εικόνα και τη μέτρηση στη συνέχεια της ενέργειας που περιέχουν οι εικόνες που προκύπτουν

LL	EL	SL	WL	RL
LE	EE	SE	WE	RE
LS	ES	SS	WS	RS
LW	EW	SW	WW	RW
LR	ER	SR	WR	RR

Περιγραφέας ιστογράμματος ακμών

- Η εικόνα διαιρείται αρχικά σε 4×4 υποεικόνες και σε κάθε υποεικόνα υπολογίζεται, με τη μορφή ιστογράμματος, η τοπική κατανομή των ακμών.
- Κάθε υποεικόνα διαιρείται σε blocks και στη συνέχεια προσδιορίζεται ο τύπος των ακμών σε κάθε block.
- Οι ακμές κατηγοριοποιούνται στους ακόλουθους πέντε τύπους: κάθετες, οριζόντιες, διαγώνιες 45°, διαγώνιες 135° και σε μη κατευθυντικές ακμές
- Στην πραγματικότητα εντοπίζονται όλες οι ακμές και για καθε block υπολογίζεται η επικρατέστερη ακμή,

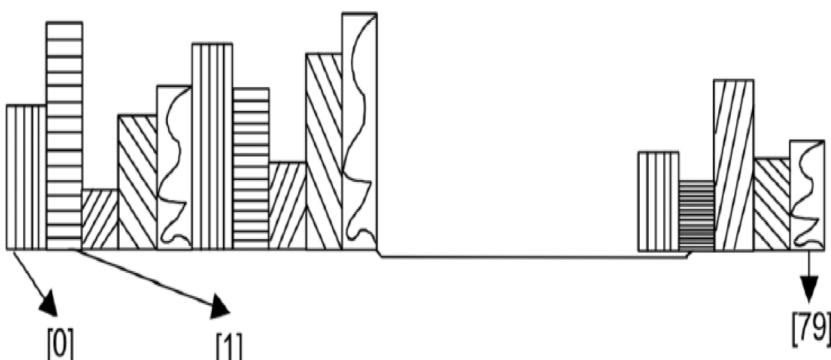


Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Παράδειγμα ιστογράμματος ακμών

Ιστόγραμμα ακμών



Τμήμα Ψηφιακών Συστημάτων

«Ψηφιακά Συστήματα & Υπηρεσίες»

HOG feature vector for one block

Angle				Magnitude			
0	15	25	25	5	20	20	10
10	15	25	30	5	10	10	5
45	95	101	110	20	30	30	40
47	97	101	120	50	70	70	80

Binary voting

Orientation Range	Count
0-19	4
20-39	4
40-59	2
60-79	0
80-99	2
100-119	3
120-139	1
140-159	0
160-179	0

Magnitude voting

Orientation Range	Count
0-19	40
20-39	45
40-59	70
60-79	0
80-99	95
100-119	140
120-139	75
140-159	0
160-179	0

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

(a)

(b)

(c)

(d)

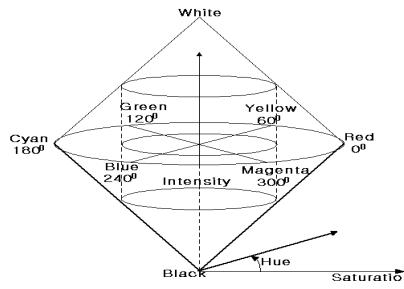
In each triplet: (1) the input image, (2) the corresponding R-HOG feature vector (only the dominant orientation of each cell is shown), (3) the dominant orientations selected by the SVM (obtained by multiplying the feature vector by the corresponding weights from the linear SVM).

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Χαρακτηριστικά Χρώματος

- Ιστόγραμμα χρώματος
- Μετατροπές σε χρωματικά κανάλια (πχ HIS, CMY, YCbCr, Lab κλπ



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Ιδιαίτερα χρήσιμο σύνολο στατιστικών χωρικών χαρακτηριστικών αντικειμένων
- Εφαρμόζονται σε:
 - Ανάλυση εικόνων & σχημάτων
 - Ανάκτηση εικόνων
 - Οπτική αναγνώριση χαρακτήρων
 - ...

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Οι διδιάστατες συνεχείς γεωμετρικές ροπές τάξης $p+q$ μιας εικόνας $f(x, y)$ δίνονται από την σχέση:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

- Τα $p+q$ παίρνουν μη αρνητικές ακέραιες τιμές (περιορισμένο σύνολο ροπών)
- Περισσότερες ροπές -> καλύτερη περιγραφή εικόνας
- Μια εικόνα αντιστοιχίζεται σε συγκεκριμένο σύνολο ροπών και αντιστρόφως



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Σε διακριτές εικόνες: $m_{pq} = \sum_x \sum_y x^p y^q f(x, y)$
- Κεντρικές Ροπές
 - Κέντρο βάρους αντικειμένου $\bar{x} = \frac{m_{10}}{m_{00}}$ $\bar{y} = \frac{m_{01}}{m_{00}}$
 - Ροπές ανεξάρτητες από τη θέση (μετατόπιση) αντικειμένων:

$$\mu_{qp} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Κανονικοποιημένες Κεντρικές Ροπές

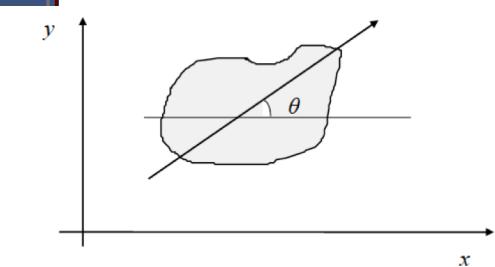
$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}$$

όπου $\gamma = \frac{p+q}{2} + 1$

για $p+q = 2, 3, \dots$

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Κύριος Άξονας
- Υπολογίζεται με τη βοήθεια ροπών
- Προσδιορίζει προσανατολισμό αντικειμένου



$$\tan(2\theta) = \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}$$

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Ανεξάρτητες Ροπές (Ροπές Hu)
- 7 ροπές με σχετική ανεξαρτησία σε
 - Κλιμάκωση
 - Περιστροφή
 - Μετατόπιση
- φ_1 - φ_6 και το μέτρο της φ_7 παρουσιάζουν αμεταβλητότητα και σε ανάκλαση

$$\begin{aligned}\varphi_1 &= \eta_{20} + \eta_{02} \\ \varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \varphi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \varphi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ &\quad - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \varphi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \varphi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Ανεξάρτητες Ροπές (Ροπές Hu)

	φ_1	φ_2	φ_3	φ_4	φ_5	φ_6	φ_7
$\Delta \Pi \Theta$	0.862	0.475	$7.55 \cdot 10^{-3}$	$5.055 \cdot 10^{-4}$	$-8.974 \cdot 10^{-7}$	$-3.483 \cdot 10^{-4}$	$-7.811 \cdot 10^{-8}$
$\Delta \Pi \varnothing$	0.859	0.461	0.011	$1.068 \cdot 10^{-3}$	$-2.164 \cdot 10^{-6}$	$-7.33 \cdot 10^{-4}$	$-4.866 \cdot 10^{-7}$
$\Delta \Pi \psi$	0.862	0.475	$7.55 \cdot 10^{-3}$	$5.055 \cdot 10^{-4}$	$-8.974 \cdot 10^{-7}$	$-3.483 \cdot 10^{-4}$	$-7.811 \cdot 10^{-8}$
$\Theta \Pi \nabla$	0.862	0.475	$7.55 \cdot 10^{-3}$	$5.055 \cdot 10^{-4}$	$-8.974 \cdot 10^{-7}$	$-3.483 \cdot 10^{-4}$	$-1.835 \cdot 10^{-7}$
$\Delta \Pi \sigma$	1.016	0.692	0.012	$2.695 \cdot 10^{-3}$	$1.047 \cdot 10^{-5}$	$7.672 \cdot 10^{-4}$	$-7.529 \cdot 10^{-6}$
$\Delta \Pi \rho$	0.848	0.468	0.013	$1.828 \cdot 10^{-3}$	$2.299 \cdot 10^{-6}$	$-3.477 \cdot 10^{-4}$	$-1.024 \cdot 10^{-5}$
$\Delta \Pi \theta$	0.858	0.463	$7.171 \cdot 10^{-3}$	$7.955 \cdot 10^{-4}$	$-7.62 \cdot 10^{-9}$	$-2.978 \cdot 10^{-4}$	$-2.102 \cdot 10^{-6}$

ΔΝ
ΕΣΥ

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Ροπές Zernike
- Χαμηλής τάξης ροπές -> ελάχιστες πληροφορίες για λεπτομέρειες εικόνας
- Υψηλής τάξης ροπές -> ευαισθησία στο θόρυβο
- Ορθογώνιες ροπές:
 - Μπορούν να περιγράψουν εικόνα με περιορισμένο πλήθος ροπών
 - Μπορούν να εκμεταλλευτούν τη χρήση ροπών υψηλής τάξης

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Ροπές Zernike
- Πολυώνυμα Zernike
 - Οικογένεια συναρτήσεων ορθογωνίων στο μοναδιαίο δίσκο $x^2+y^2 < 1$
- Συνάρτηση Zernike $W_{pq}(r,\vartheta)$ τάξης (q, p) ορίζεται στο πολικό σύστημα συντεταγμένων (r,ϑ)

$$W_{pq}(x, y) = W_{pq}(r, \theta) = R_{pq} e^{iq\theta}, \quad r^2 = x^2 + y^2 \leq 1, \quad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

$$R_{p,\pm q}(r) = \sum_{k=q, p-|k|=\text{ορτοσ}}^p B_{pqk} r^k \quad B_{pqk} = \frac{(-1)^{(p-k)/2} ((p+k)/2)!}{((p-k)/2)! ((q+k)/2)! ((k-q)/2)!}$$

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- **Ροπές Zernike**

$$W_{pq}(r, \theta) = \left[\sum_{k=q-p, |k|=ap+ta}^p \frac{(-1)^{(p-k)/2} ((p+k)/2)!}{((p-k)/2)! ((q+k)/2)! ((k-q)/2)!} r^k \right] e^{iq\theta}$$

- Υψηλό υπολογιστικό κόστος -> χρησιμοποιείται απευθείας η ακριβής ανάπτυξή τους

(p, q)	$R_{p,q}(r)$	(p, q)	$R_{p,q}(r)$
(0, 0)	1	(1, 1)	$r^2 - 2r$
(2, 0)	$2r^2 - 1$	(3, 1)	r^3
(2, 2)	$2r^4 - 4r^2 + 1$	(3, 3)	$16r^5 - 12r^3 + 3r$
(4, 0)	$4r^4 - 8r^2 + 1$	(5, 1)	$5r^5 - 4r^3$
(4, 2)	$4r^6 - 12r^4 + 7r^2 + 1$	(5, 3)	$35r^7 - 60r^5 + 30r^3 - 4r$
(4, 4)	r^8	(7, 1)	$21r^9 - 30r^7 + 10r^5$
(6, 0)	$70r^8 - 140r^6 + 90r^4 - 20r^2 + 1$	(7, 3)	$7r^9 - 6r^7$
(6, 2)	$15r^10 - 20r^8 + 6r^6$	(7, 5)	$126r^{10} - 280r^8 + 210r^6 - 60r^4 + 5r^2$
(6, 4)	$6r^10 - 5r^8 + 6r^6 - 10r^2$	(9, 3)	$84r^{12} - 168r^{10} + 105r^8 - 20r^6$
(6, 6)	r^{12}	(9, 5)	$36r^{14} - 56r^{12} + 21r^10$
(8, 0)	$70r^{12} - 140r^{10} + 90r^8 - 20r^6 + 1$	(9, 7)	$9r^{14} - 5r^{12}$
(8, 2)	$28r^{14} - 42r^{12} + 15r^{10}$	(9, 9)	r^{16}
(8, 4)	$8r^8 - 7r^6$	(11, 1)	$465r^{18} - 1500r^{16} + 1260r^{14} - 560r^{12} + 165r^{10} - 6r^8$
(8, 6)	$8r^8 - 7r^6$	(11, 3)	$210r^{18} - 563r^{16} + 423r^{14} - 140r^{12} + 15r^{10}$
(10, 0)	$252r^{10} - 630r^8 + 560r^6 - 210r^4 + 30r^2 - 1$	(11, 5)	$129r^{18} - 252r^{16} + 168r^{14} - 35r^{12}$
(10, 2)	$210r^{12} - 563r^{10} + 423r^8 - 140r^6 + 15r^4$	(11, 7)	$165r^{14} - 366r^{12} + 252r^{10} - 56r^8$
(10, 4)	$129r^{14} - 252r^{12} + 168r^{10} - 35r^8$	(11, 9)	$55r^{14} - 90r^{12} + 36r^{10}$
(10, 6)	$45r^{16} - 72r^{14} + 28r^{12}$	(11, 11)	$11r^{14} - 10r^{12}$
(10, 8)	$45r^{16} - 9r^{14}$	(12, 1)	r^{14}
(12, 0)	$924r^{12} - 2772r^{10} + 3150r^8 - 1680r^6 + 420r^4 - 42r^2 + 1$		
(12, 2)	$792r^{14} - 2310r^{12} + 2520r^{10} - 1260r^8 + 280r^6 - 21r^2$		
(12, 4)	$495r^{14} - 1320r^{12} + 1260r^{10} - 504r^8 + 70r^4$		
(12, 6)	$220r^{16} - 495r^{14} + 360r^{12} - 84r^6$		
(12, 8)	$66r^{12} - 110r^{10} + 45r^8$		
(12, 10)	$12r^{14} - 11r^{12}$		
		(12, 12)	r^{12}

**Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων**

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Υπολογισμός Ροπών Zernike
- Σε μια εικόνα $f(x,y)$: Οι προβολές της εικόνας στην ορθογωνική βάση των συναρτήσεων
- Για βαθμό p και επανάληψη q :

$$Z_{pq} = \frac{p+1}{\pi} \iint_{x^2+y^2 \leq 1} f(x, y) W_{pq}(r, \theta) dx dy$$

- Σε ψηφιακή (διακριτή) εικόνα:

$$Z_{pq} = \frac{p+1}{\pi} \sum_x \sum_y f(x, y) W_{pq}(r, \theta), \quad x^2 + y^2 \leq 1$$

$x = r\cos\theta, y = r\sin\theta$

**Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων**

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

ΓΕΩΜΕΤΡΙΚΕΣ ΡΟΠΕΣ

- Υπολογισμός Ροπών Zernike
- Κέντρο εικόνας: Σημείο (0,0)
- Συντεταγμένες εικονοστοιχείων: περιοχή μοναδιαίου κύκλου
- Σφάλματα κβαντισμού
- Υψηλό υπολογιστικό κόστος
- Λύση: Γρήγοροι αλγόριθμοι υπολογισμού ροπών, μελέτες για την ακρίβεια υπολογισμών

Ψηφιακή Αναπαρασταση vs Ανθρώπινη Όραση

- Στόχος: Αυτοματοποίηση της Αναγνώρισης Εικονικών Προτύπων
- Εξαγωγή χαρακτηριστικών από την εικόνα στην οποία βρίσκονται (Πρότυπα χαρακτηριστικών).
- Ταξινόμηση των εικονικών χαρακτηριστικών βάση των πρότυπων χαρακτηριστικών

Εξαγωγή Χαρακτηριστικού

- Αυτόματη Εξαγωγή Χαρακτηριστικού
 - Μέθοδοι κατάτμησης ενός αντικειμένου από το background
 - Αναγνώριση μίας περιοχής της εικόνας που αντιστοιχεί στο αντικείμενο
 - Αναγνώριση του ορίου του αντικειμένου και μετά εξαγωγή της περιβαλλόμενης περιοχής



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Παράδειγμα Αναγνώριση Μελανώματος

Issues to be addressed

- How can we acquire good images or correct already captured ?
- How can we segment medical images? i.e. separate ROIs in images (image segmentation)
- How can we register medical images for follow ups?
- How are the image features defined, i.e., what are we looking for?
- How are these features detected in the image? (trivial for humans but non-trivial for machines).
- Which are the proper features to use and how many are they? (feature selection).
- How do we use the features to design the classifier for the specific task,
- How can we assess the performance of a classifier?



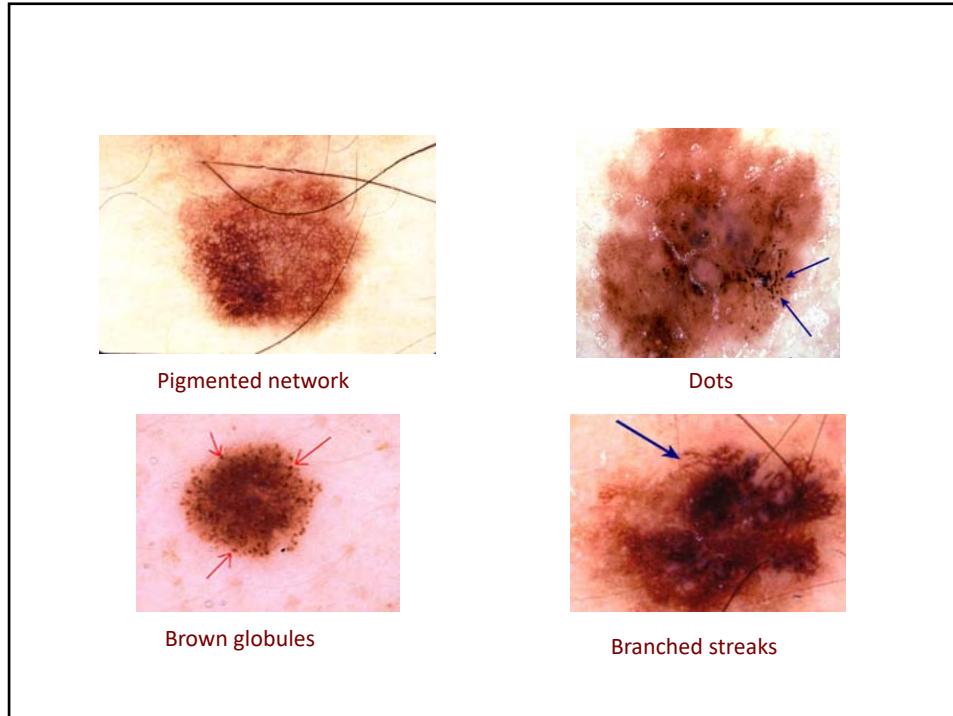
Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Image Acquisition

- Main techniques used for this purpose are
 - The epiluminescence microscopy (ELM or dermoscopy) and
 - The image acquisition using still or video cameras.
 - Multispectral images
- Problem
 - Lack of reproducibility and accuracy due to dependency on equipment and environmental constraints, such as image resolution, image noise, illumination, skin reflectivity and pose uncertainty
- Requirement
 - Image acquisition Standardization and Camera Calibration





Asymmetry Features

- Examines symmetry with respect to all mentioned features
- Computed by overlapping the two halves of the tumor along the principal axes of inertia and dividing the non-overlapping area differences of the two halves by the total area of the tumor

- Generic region-wise feature: $R_i = \frac{Q_i}{\sum Q_i}$

- Q_i : individual feature of lesion slice, defined by symmetry axes

Border Features

Examples of border features used:

- Irregularity

$$\text{Irregularity} = \frac{\text{Perimeter}}{\text{Area}}$$

- Thinness ratio

$$\text{ThinnessRatio} = 4\pi \frac{\text{Area}}{(\text{Perimeter})^2}$$

- Variance of the distance of the border lesion points from the centroid

Noise sensitivity => need for holistic border descriptors, not indices

Proposed methods: fractals, fourier series, wavelets

Other feature class: transition from lesion to the skin: minimum, maximum, average and variance responses of the gradient operator, applied on the intensity image along the lesion border

Color Features

- Color spaces: RGB, YUV, HSI etc
- Minimum, maximum, average and standard deviations of the selected channel values, chromatic differences inside the lesion
- Colors that are typical for skin lesions may be detected after a color quantization and the percentage of each color in the lesion can be exploited as a feature.
- Color entropy

$$C_v = - \sum_{c1} \sum_{c2} \sum_{c3} p_{clc2c3} \log(p_{clc2c3})$$

- Color homogeneity

$$C_h = \sqrt{\frac{1}{n} \sum_i (\mu_{c_v} - c_v^i)^2}$$

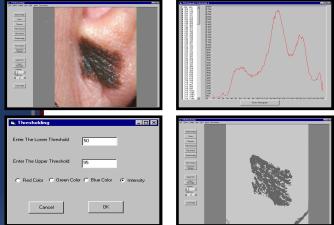
Differential Structure Features

- Very rare in literature despite their significance in conventional diagnosis
- Potential use of pigment network feature
 - Pigmentary structures extracted in the X-plane of the CIE-xyz space.
 - Calculate a skeleton of the pigment network, and the orientation of the peripheral elements of this skeleton is tested to detect radial streaming and pseudopods.
 - Skeleton is also used for assessment of network hole sizes.
 - Globules extracted in similar fashion

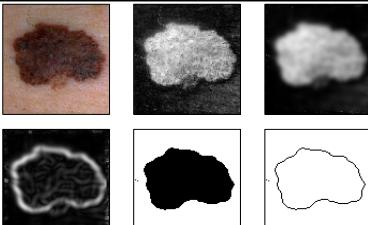
Feature extraction using image processing

- Image segmentation: involves the separation of the skin lesion from the healthy skin.
- Based on one of two basic properties: discontinuity and similarity.
 - Detection of discontinuities between the skin lesion and the surrounding healthy skin.
 - Pixels, which belong in skin lesion, have different color attributes from pixel corresponding to healthy skin
 - Thresholding is implemented by choosing an upper and a lower value and then isolating the pixels which have values in this range.
 - Region Growing is a procedure that groups pixels or subregions into larger regions.
 - Clustering initially divides the image into rectangular regions small enough to be considered as having only a single color. This is followed by conservative merging, where adjacent regions whose colors are similar are connected.

Examples



A) Thresholding



D) PCT Transform



B) Use of weighted functions



C) Region Growing

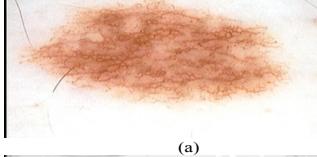


E) Edge Mapping

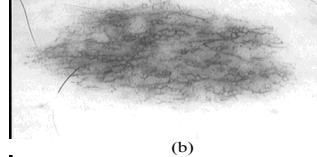
 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

 Πρόγραμμα Μεταπτυχιακών Σπουδών
 «Ψηφιακά Συστήματα & Υπηρεσίες»

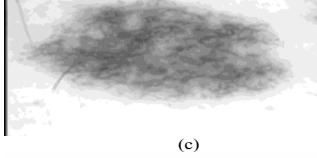
Adaptive Thresholding Segmentation (a) original image (b) grey scale image (c) smoothed image (d) adaptive thresholding (e) object extraction (f) segmented image



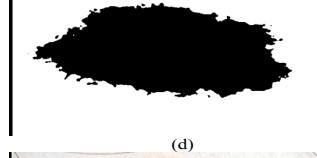
(a)



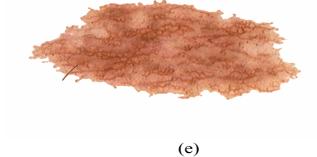
(b)



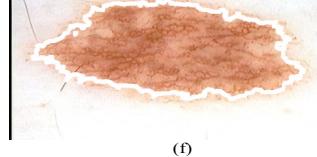
(c)



(d)



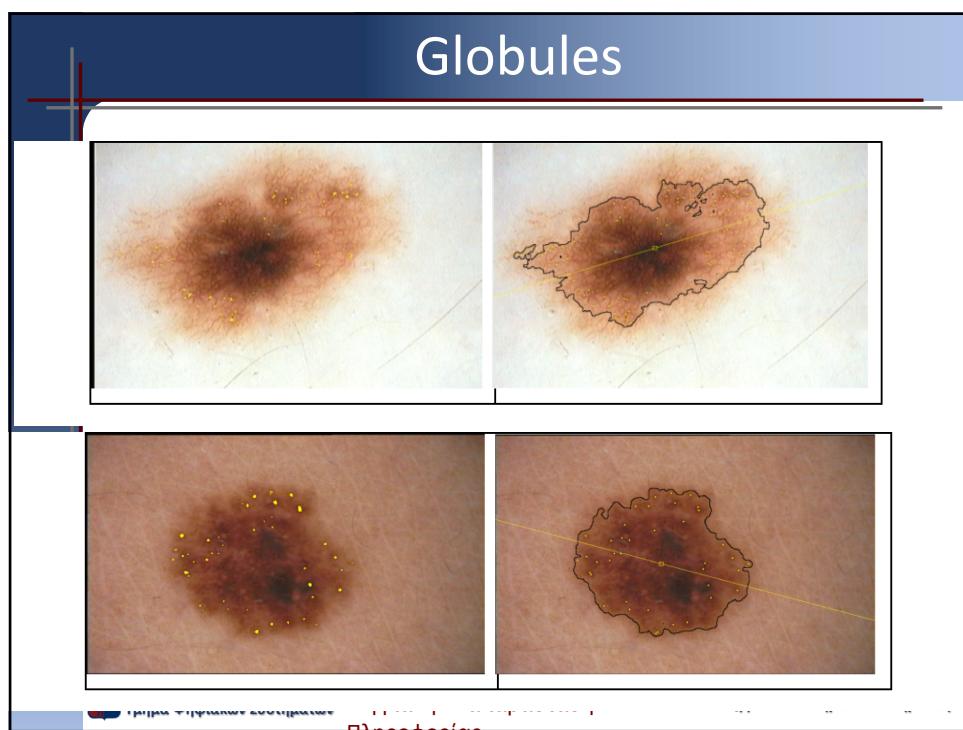
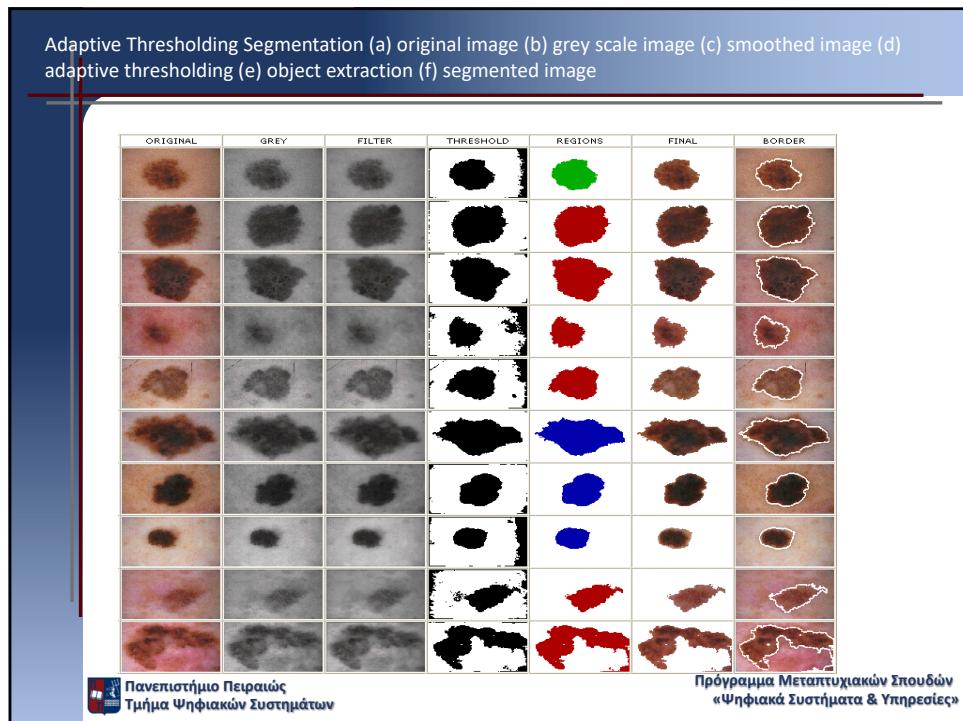
(e)



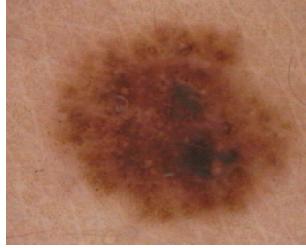
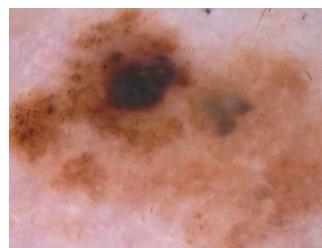
(f)

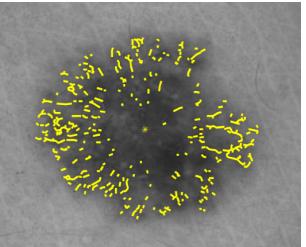
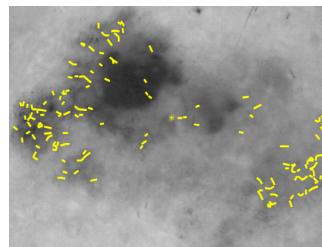
 Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

 Πρόγραμμα Μεταπτυχιακών Σπουδών
 «Ψηφιακά Συστήματα & Υπηρεσίες»



Streaks

Exemplar results



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Streak-based features

- Number of
 - Individual detected streaks,
 - Pixels in streaks
- Asymmetry of azimuthial angle and radial distribution of streaks (A_θ, A_r)
 - The centroid of each streak \mathbf{c}_i is located, defined as its median pixel.
 - The azimuthial angle of the centroid ϑ_i and its distance from the boundary $r_i = DT(\mathbf{c}_i)$ is calculated for each streak $i=1,2,\dots,K$.
 - The histogram of ϑ_i and r_i is calculated (H_θ and H_r , respectively) using 10 bins

$$A_\theta = \frac{\max(H_\theta) - \min(H_\theta)}{\max(H_\theta)}, A_r = \frac{\max(H_r) - \min(H_r)}{\max(H_r)}$$



Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Using images captured by the Molemax II

SELECTED FEATURES					
Feature	MM (std)	MD (std)	Feature	MM (std)	MD (std)
mean-R	116.65 (33.65)	157.29(28.09)	Complexity	10.89 (16.87)	8.07 (12.37)
I-mean	75.72 (22.04)	101.66 (22.32)	ASM	8949.96 (7505.5)	7247.92 (6716.8)
L-mean	141.86 (40.50)	190.46 (38.45)	Dissimilarity	3430017(2571071)	2781110(2571071)
mean-G	62.46 (19.60)	83.79(21.74)	Perimeter	2640.49 (1874.6)	2252.06(1592.4)
mean-B	48.03 (16.68)	63.90 (20.31)	Area	68924.59 (25955)	64009.45(23396)
GMSM-mean	140.25 (36.11)	134.51 (32.18)	Eccentricity	1.68 (0.42)	1.77(0.48)
S-mean	93.48 (22.91)	100.56 (18.74)	Asymmetry	30.53 (18.63)	29.68 (16.99)
H-mean	27.66 (22.06)	25.96 (28.56)	Grad-mean	1.26 (0.52)	1.23 (0.45)
B-mean	40.41 (5.48)	38.88 (4.44)	A-mean	98.8 (5.35)	100.52 (4.31)

Selected features for the construction of the training and test set, where MM and MD the mean values for the melanoma and the nevus cases respectively.

Performance of the SVM algorithm using the exponential radial basis function with various values of sigma

SIGMA	ERRORS	TP	TN	FP	FN	ACCURACY	SPECIFICITY	SENSITIVITY
7	85	63	893	79	6	91.84%	91.87%	91.30%
8	87	62	892	80	7	91.64%	91.77%	89.86%
6	88	62	891	81	7	91.55%	91.67%	89.86%
9	90	62	889	83	7	91.35%	91.46%	89.86%
10	91	62	888	84	7	91.26%	91.36%	89.86%
12	97	62	882	90	7	90.68%	90.74%	89.86%
5	99	61	881	91	8	90.49%	90.64%	88.41%

TP: True Positive (melanoma instances actually classified as melanoma by the SVM algorithm)

TN: True Negative (dysplastic nevus instances actually classified as nevus by the SVM algorithm)

FP: False Positive (melanoma instances classified as nevus by the SVM algorithm)

FN: False Negative (dysplastic nevus instances classified as melanoma by the SVM algorithm)

Deep Learning

Attempt to solve the feature extraction problem

Faces Cars Elephants Chairs

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

Assume that we feed them with face images

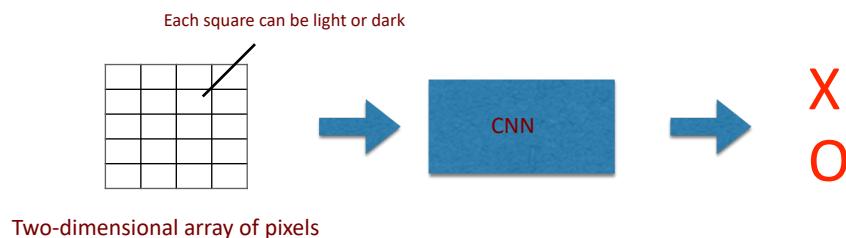
- They are multilayer neural networks
- First layer: they learn some basic things
- e.g. edges, spots, dark spots, bright spots
- Second layer: they learn more abstract concepts
- e.g. eyes, noses, mouths
- Third layer: they keep going with the abstractions
- e.g. faces

Πανεπιστήμιο Πειραιώς
Τμήμα Ψηφιακών Συστημάτων

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Ψηφιακά Συστήματα & Υπηρεσίες»

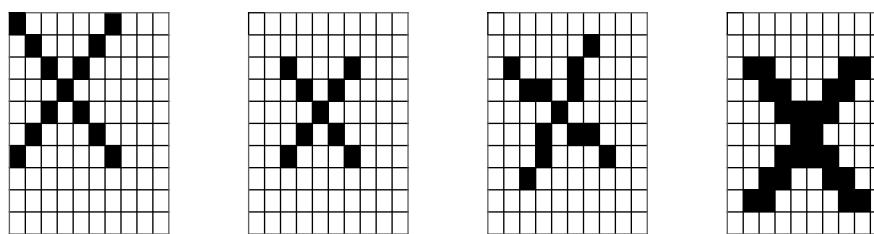
Is the picture X or O?

CNNs seems magic but actually they apply a chain of well known operations to provide decisions



Trickier cases

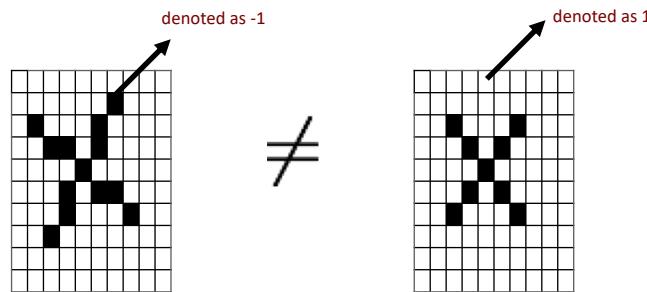
translation, scaling, rotation, weight



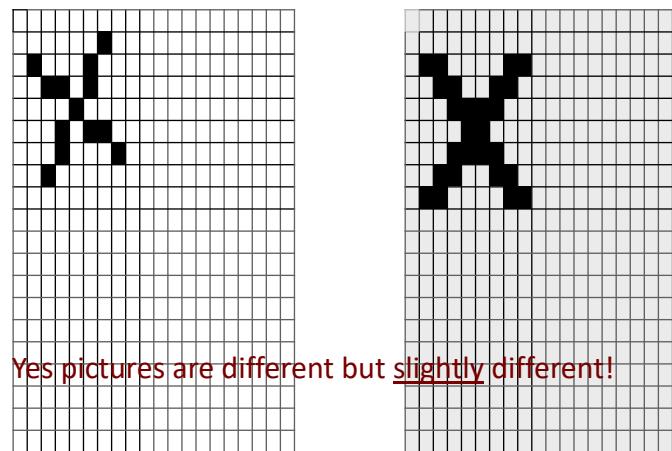
All of them are Xs!

Why is tricky?

Because computer sees in an absolute and literal way!

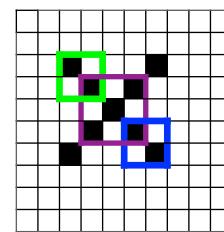
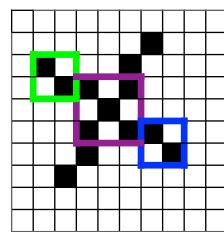


What computer sees?



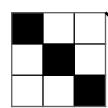
Yes pictures are different but slightly different!

CNNs matches pieces of the image

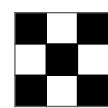


Features match pieces of the image

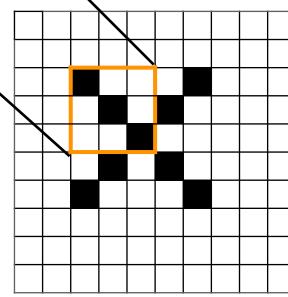
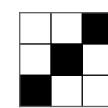
Left diagonal line



An X

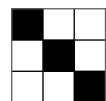


Right diagonal line

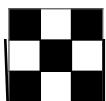


Features match pieces of the image

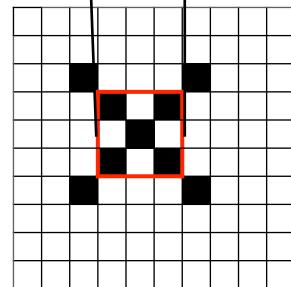
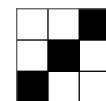
Left diagonal line



An X

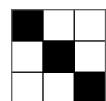


Right diagonal line

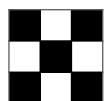


Features match pieces of the image

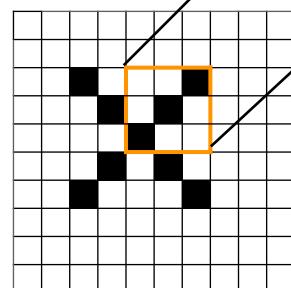
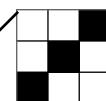
Left diagonal line



An X



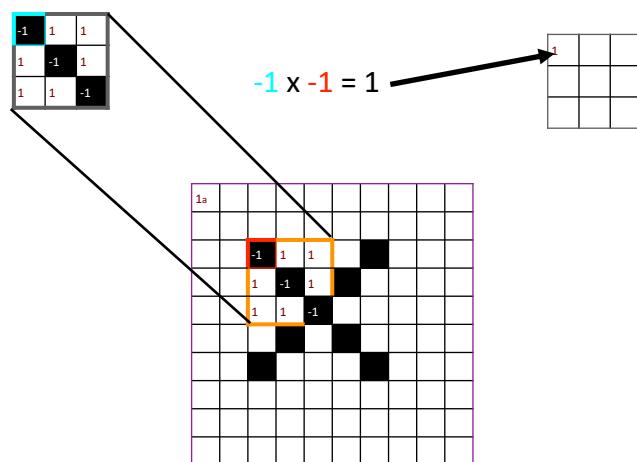
Right diagonal line

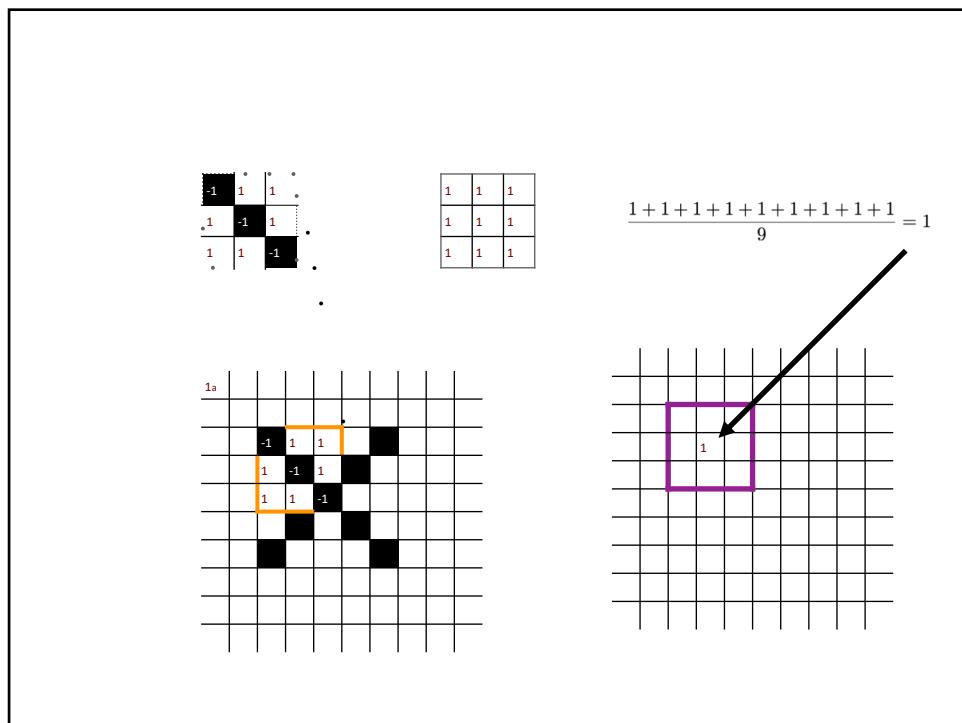
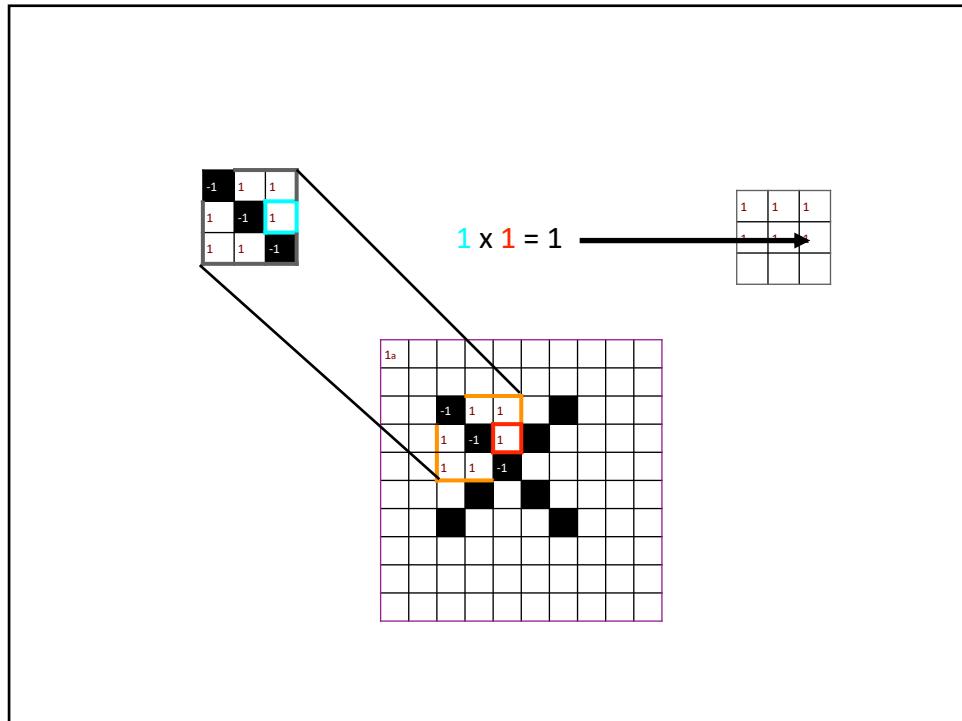


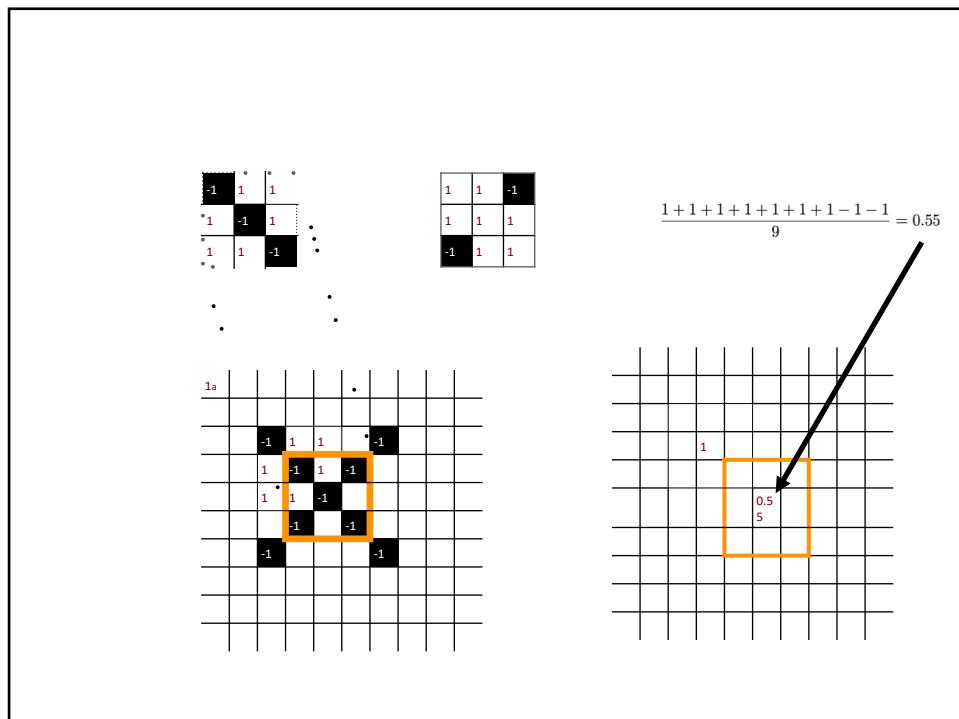
The math behind that: “Filtering”

1. Line up the feature with some image patch
2. Multiply each image pixel by the corresponding feature path
3. Add them up
4. Divide by the total number of pixels in the feature

Filtering







Convolution

By moving a filter to any possible position we get the feature map of the image and that's convolution!

-1	1	1	.
.	1	-1	1
1	1	-1	.



0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	1.00

Repeat with any other filter

The diagram illustrates three separate convolution operations on a 5x5 input image (represented by a grid of 1s) using 3x3 filters. Each operation is shown with a multiplication symbol (*) and an equals sign (=).

- Filter 1:** A 3x3 matrix with values: $\begin{matrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.77 & -0.11 & 0.11 \\ -0.11 & 1.00 & -0.11 \\ 0.11 & -0.11 & 1.00 \end{matrix}$.
- Filter 2:** A 3x3 matrix with values: $\begin{matrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.33 & 0.33 & -0.33 \\ 0.33 & 0.55 & -0.33 \\ -0.33 & -0.33 & 0.33 \end{matrix}$.
- Filter 3:** A 3x3 matrix with values: $\begin{matrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.55 & -0.11 & 0.11 \\ -0.11 & 0.11 & -0.11 \\ 0.11 & -0.11 & 0.11 \end{matrix}$.

Ellipses (...) indicate that there are more filters and operations.

Convolution layer

One image becomes a stack of filtered images

The diagram shows a single 5x5 input image being processed by three different 3x3 filters. The output is a stack of three 3x3 feature maps.

- Filter 1:** A 3x3 matrix with values: $\begin{matrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.77 & -0.11 & 0.11 \\ -0.11 & 1.00 & -0.11 \\ 0.11 & -0.11 & 1.00 \end{matrix}$.
- Filter 2:** A 3x3 matrix with values: $\begin{matrix} -1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.33 & 0.33 & -0.33 \\ 0.33 & 0.55 & -0.33 \\ -0.33 & -0.33 & 0.33 \end{matrix}$.
- Filter 3:** A 3x3 matrix with values: $\begin{matrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{matrix}$. The result is a 3x3 matrix of values: $\begin{matrix} 0.55 & -0.11 & 0.11 \\ -0.11 & 0.11 & -0.11 \\ 0.11 & -0.11 & 0.11 \end{matrix}$.

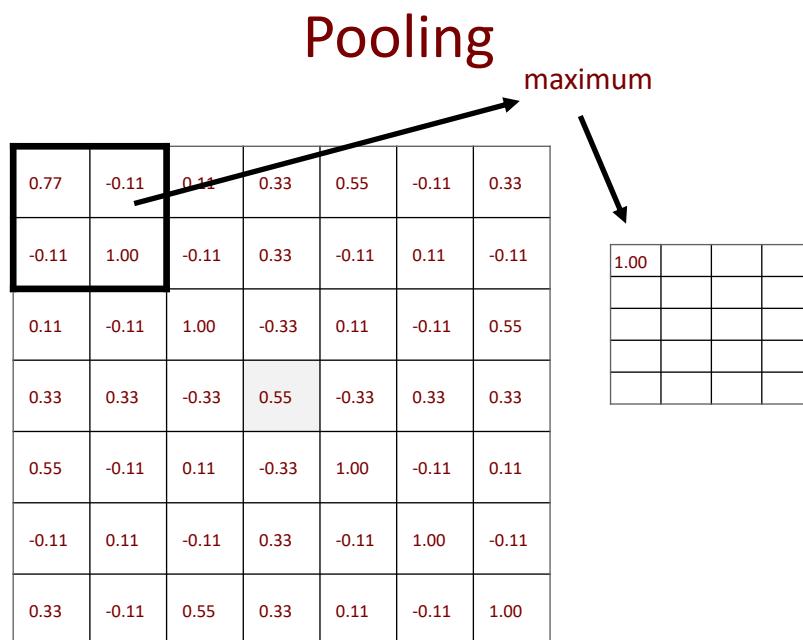
Ellipses (...) indicate that there are more filters and operations.

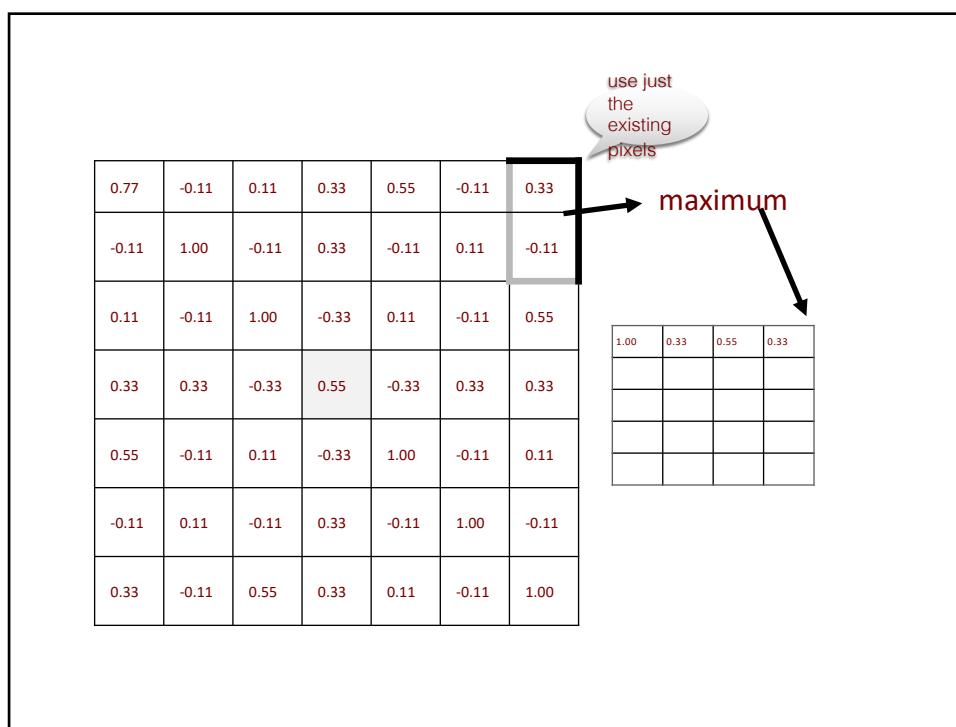
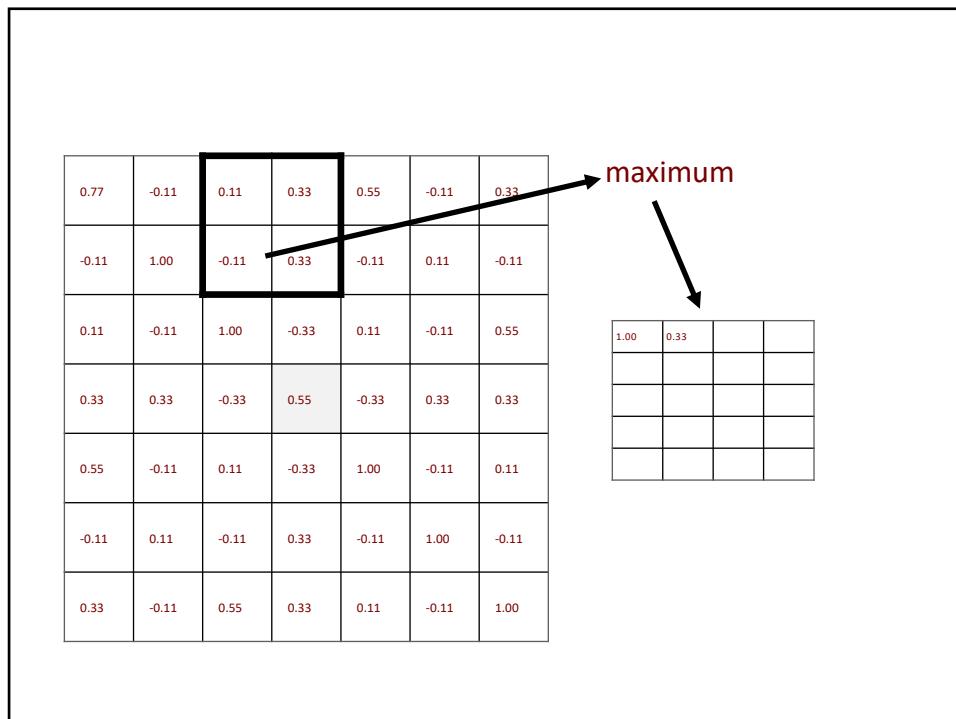
`#stack_images = #filters = #features`

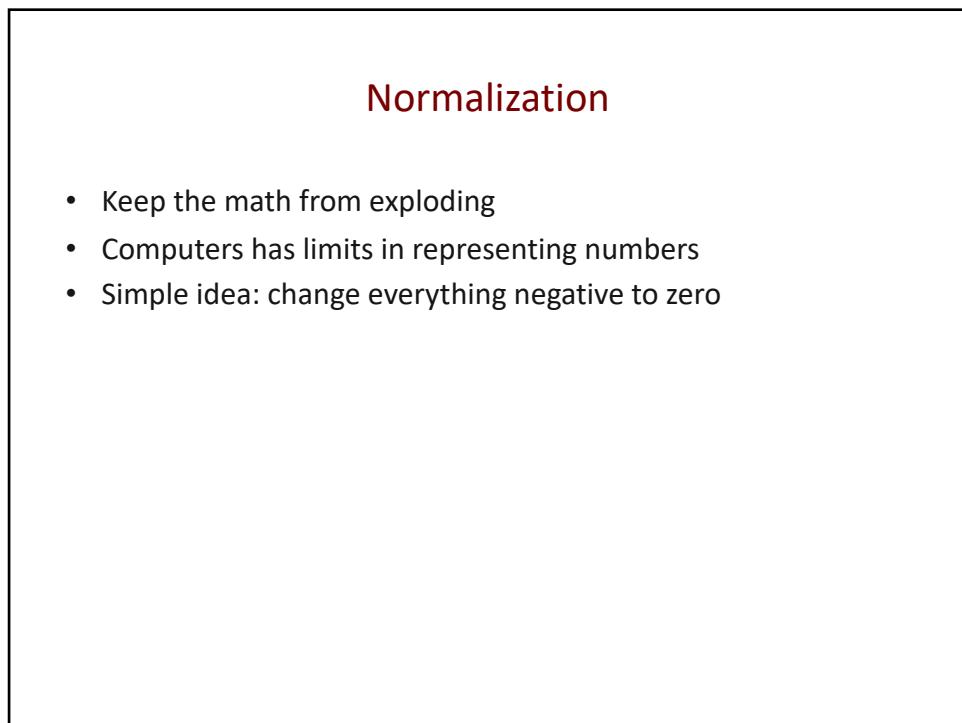
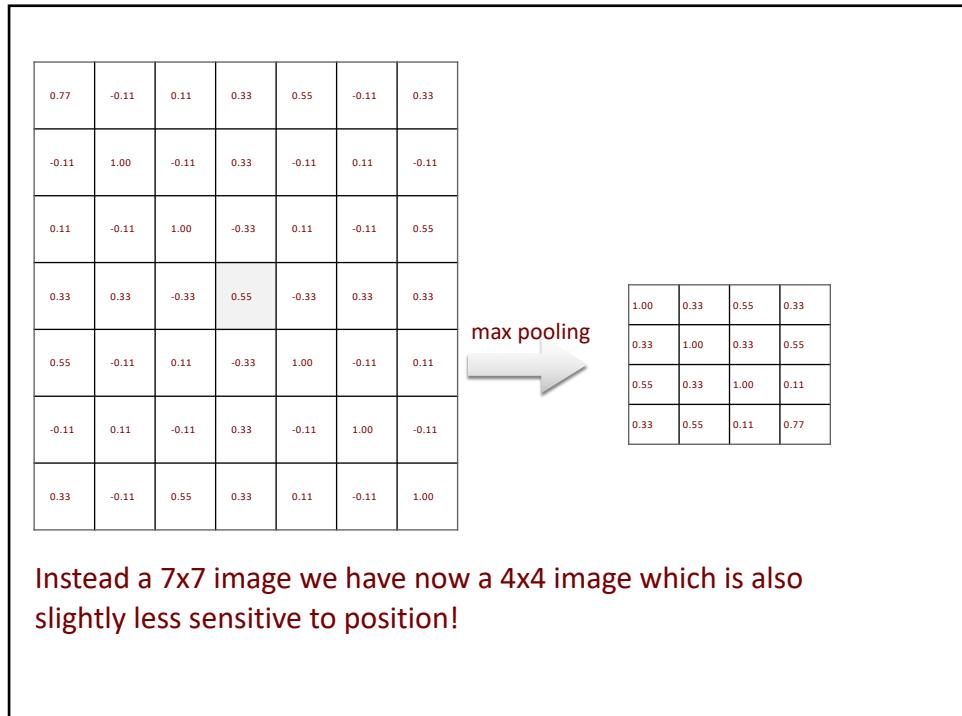
Pooling: Shrinking the image stack horizontally

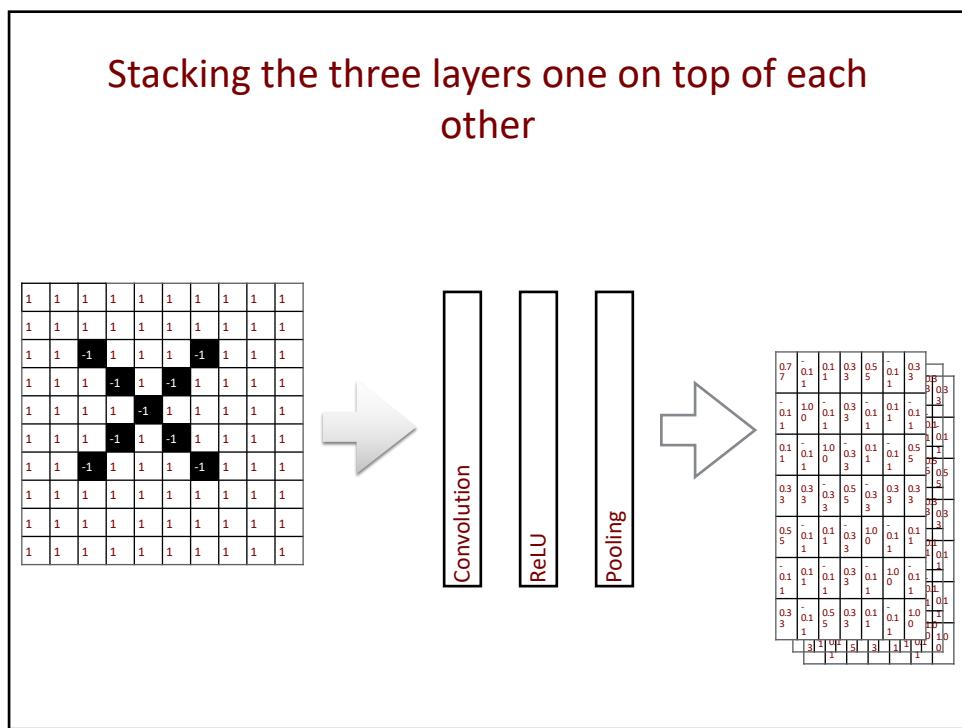
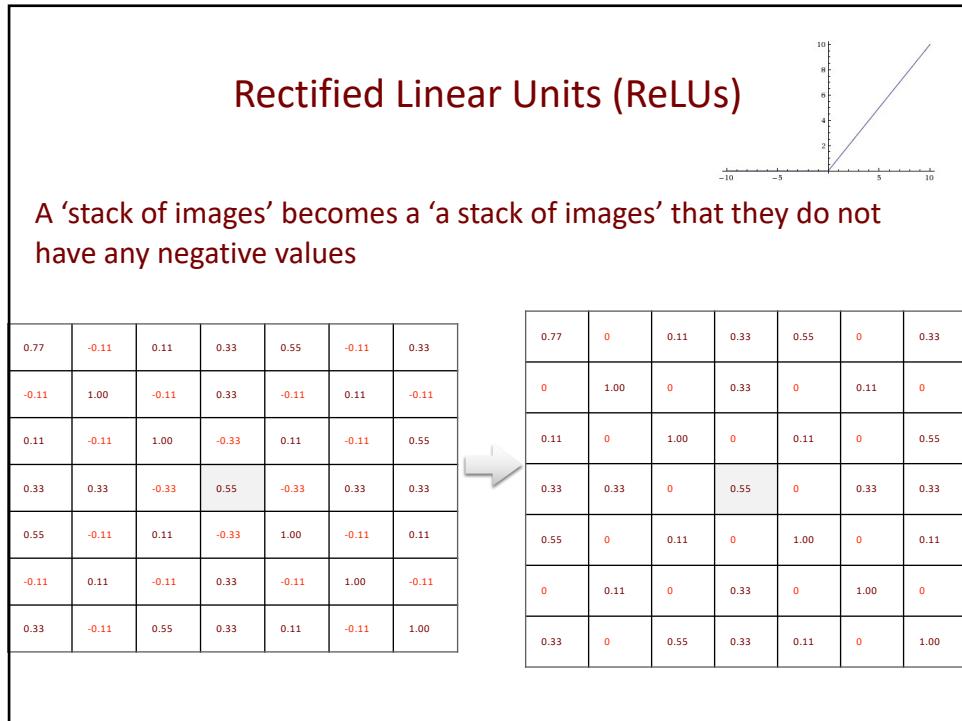
Shrink each image of the stack

1. Pick a window size (usually 2 or 3)
2. Pick a stride (usually 2)
3. Walk your window across your filtered images
4. From each window, take the maximum value

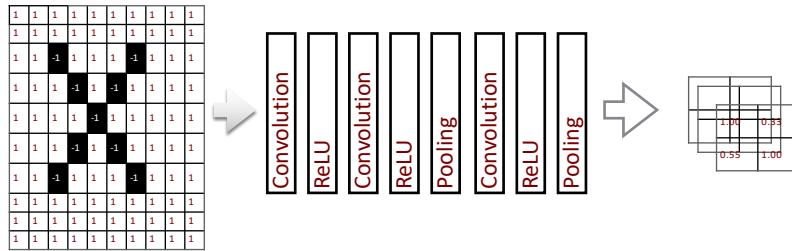








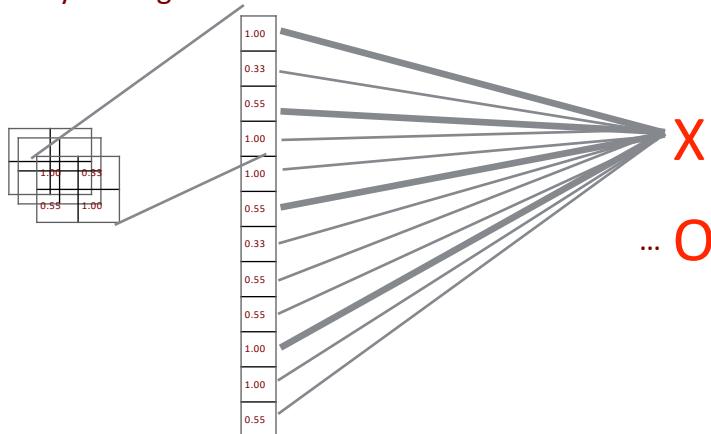
We can go even deeper



Each time the image get more filtered as passing though the convolution layers and smaller as passing from the pooling layers

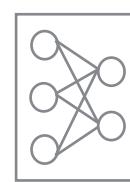
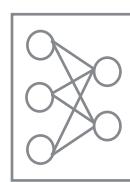
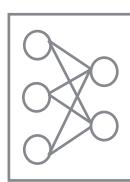
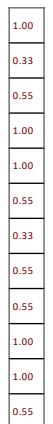
Finally we have a fully connected layer

Every value gets a vote



Vote depends on how strongly a value depends X or O

These layers can also be stacked ...



X
O

Learning

- What about all the “magic numbers” that are used in the whole chain? (features in convolutional layers, weights in fully connected layers)
 - Backpropagation training

CNN parameters

- These include all that architectural features that cannot be learned!
 - Convolution
 - Number of features
 - Size of features
 - Pooling
 - Window size
 - Window stride
 - Fully Connected: Number of neurons
- Architecture
- Type
 - Order
 - Number of layers

Applications

Any 2D and 3D data!

Keep in mind that things closer together are more closely related than things far away!

Limitations

CNNs only capture local spatial patterns in data.

Data has to be made to look like an image.

