

Εντοπισμός αντικειμένου σε εικόνα

Μάθημα: Υπολογιστική όραση

Κ. Δελήμπασης

Περιγραφείς σχημάτων

- Αναγνώριση αντικειμένων με χρήση ροπών
- Περιγραφείς Fourier
- Στατιστική διακύμανση σχημάτων
- Αναζήτηση σε εικόνα τυχαίου σχήματος
 - Ο μετασχηματισμός απόστασης

Αναγνώριση αντικειμένων με χρήση ροπών

Moment Image descriptors

- Εισήχθηκαν από τον Hu (1962), με ευρεία εφαρμογή
- Περιγράφουν σχήματα με μικρό πλήθος περιγραφέων
- Η υπολογιστική πολυπλοκότητα αυξάνει όσο αυξάνει η τάξη της ροπής.
- Μη ορθογώνιες ροπές
- Άλλες ορθογώνιες ροπές: Legendre moments, geometrical moments, Zernike moments, pseudo-Zernike moments and Orthogonal Fourier-Mellin moments are evaluated

Ορισμός

- Ροπές τάξης $p+q$ μίας 2D συνάρτησης $f(x,y)$ με πεδίο ορισμού D :

$$M_{pq} = \iint_D p_{pq}(x,y) f(x,y) dx dy$$

- Οι ροπές υπάρχουν αρκεί το D να είναι φραγμένο
- Αν $p_{pq}(x,y)=x^p y^q$, τότε προκύπτουν οι γεωμετρικές ροπές

$$M_{pq} = \iint_D x^p y^q f(x,y) dx dy$$

- Τα μονώνυμα αποτελούν τις συναρτήσεις βάσης. Αν η τάξη είναι $n=p+q : p+q \leq n$, πλήθους $(n+1)(n+2)/2$.

- Κεντρικές ροπές: αμετάβλητες κατά την μετατόπιση (translation invariant)

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y), \bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}}$$

- Ροπές αματάβλητες κατά την αλλαγή κλίμακας (scale invariant)

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \gamma = \frac{p+q}{2} + 1$$

- Ροπές Hu (1962) αμετάβλητες (invariant) σε affine μετασχηματισμούς

$$\varphi_1 = \eta_{20} + \eta_{02}$$

$$\varphi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2$$

$$\varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\varphi_4 = (\eta_{30} + \eta_{12}) + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned} \varphi_5 = & (\eta_{03} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{03} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] + \\ & (3\eta_{21} + \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \end{aligned}$$

$$\begin{aligned} \varphi_6 = & (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + \\ & 4\eta_{11} (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned}$$

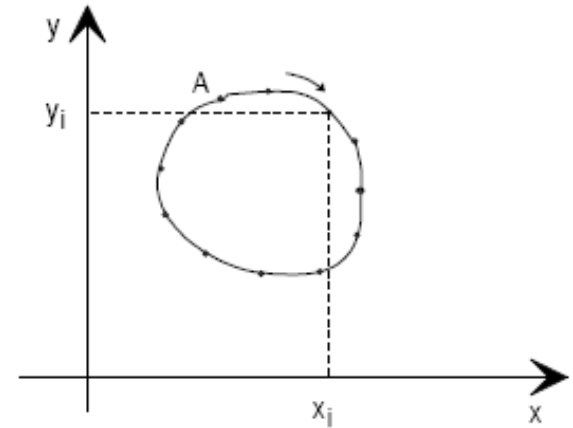
Πίνακας ροπών 2^{ης} τάξης

- Εστω σύνολο μη-μηδενικών pixel μίας εικόνας (πλήρες δυαδικό αντικείμενο, ή περίγραμμα)

$$\begin{pmatrix} \sum_{i=1}^N (x_i - \mu_x)^2 & \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \\ \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) & \sum_{i=1}^N (y_i - \mu_y)^2 \end{pmatrix}$$

Περιγραφείς Fourier (Fourier Descriptors -FD)

- Ισχυρή μέθοδος περιγραφής περιγραμμάτων και εξαγωγής χαρακτηριστικών των αντικειμένων
- Έστω ότι η καμπύλη του σχήματος αντιστοιχεί στο όριο ενός αντικειμένου. Ξεκινώντας από ένα αρχικό σημείο A , με ωρολογιακή φορά παίρνουμε μια ακολουθία σημείων (x_n, y_n) , $n=0,1,\dots,N-1$, με N το πλήθος των εικονοστοιχείων του ορίου.



Ορισμός FDs

- Εστω η μιγαδική ακολουθία:

$$z(n)=x(n)+jy(n), n=0,1,\dots,N-1$$

- Από τον ορισμό του διακριτού μετασχηματισμού Fourier (DFT) της $z(n)$:

$$a(k) = \sum_{n=0}^{N-1} z(n) \exp\left(\frac{-j2\pi kn}{N}\right)$$

- Με τον αντίστροφο DFT (IDFT) ανακατασκευάζεται η $z(n)$:

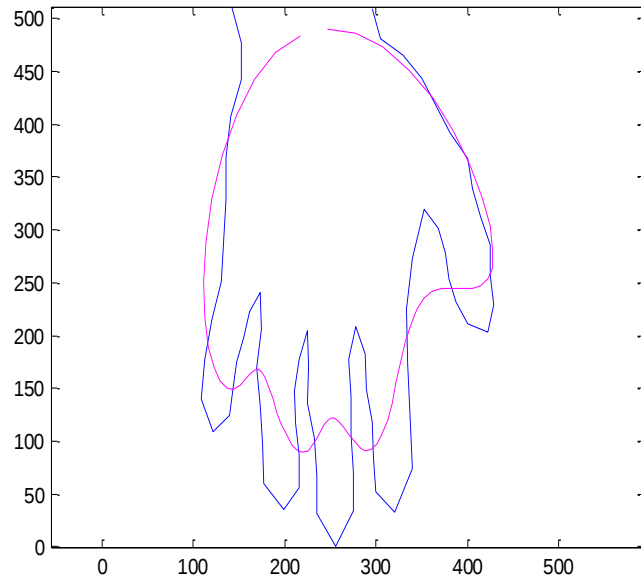
$$z(n) = \frac{1}{N} \sum_{k=0}^{N-1} a(k) \exp\left(\frac{j2\pi kn}{N}\right)$$

- Οι μιγαδικοί συντελεστές $\alpha(k)$ καλούνται περιγραφείς Fourier (Fourier Descriptors – FD).
- Για ομαλά σχήματα, η περισσότερη πληροφορία του σχήματος είναι αποθηκευμένη στις χαμηλότερες αρμονικές
- Όσους περισσότερους συντελεστές χρησιμοποιούμε, τόσο καλύτερη είναι η προσέγγιση και η αναπαράσταση / ανακατασκευή της καμπύλης.
- Η σύγκριση δύο σχημάτων για τα οποία έχουν προσδιοριστεί οι FDs μπορεί να γίνει με την άμεση σύγκριση των FDs, αφού ληφθεί υπόψη η αμεταβλητότητα / συμμεταβλητότητα ως προς γεωμετρικούς μετασχηματισμούς και αρίθμηση.

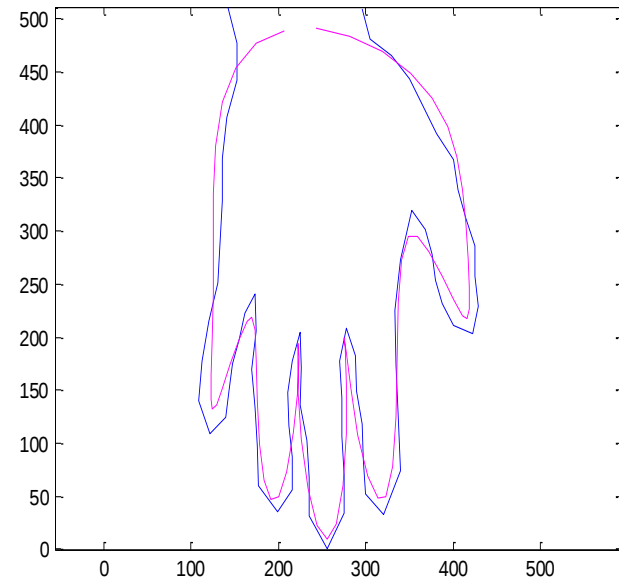
Αμεταβλητότητα / Συμμεταβλητότητα

Invariance / covariance

- Αν κάποια καμπύλη μετατεθεί κατά x_0, y_0 τότε οι FDs παραμένουν ίδιοι, εκτός της μηδενικής αρμονικής
- Αλλαγή κλίμακας του περιγράμματος συνεπάγεται ίδια αλλαγή κλίμακας των FDs.
 - Αν τα FDs κανονικοποιηθούν με τα μέτρα τους $a(k)/|a(k)|$ γίνονται ανεξάρτητα της αλλαγής κλίμακας
- Περιστροφή κατά γωνία θ του περιγράμματος: $a_k \leftarrow a_k e^{j\vartheta}$.
- Αλλαγή της αρίθμησης των σημείων:
$$a_k \leftarrow a_k \exp(2\pi k n_0 / N).$$



Αρχικό περίγραμμα,
Περίγραμμα με 5 συντελεστές



Αρχικό περίγραμμα,
Περίγραμμα με 8 συντελεστές

Περιγραφείς Fourier σε 3D

- Ως βάση 3D αντικειμένων χρησιμοποιούνται οι Σφαιρικές Αρμονικές (Spherical Harmonics)
- Είναι οι επιφάνειες που προκύπτουν από τη λύση της εξίσωσης: $\Delta S(u,v)=0$

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{2l+l(l-m)}{4\pi(l+m)!}} p_l^m(\cos \theta) e^{im\varphi}$$

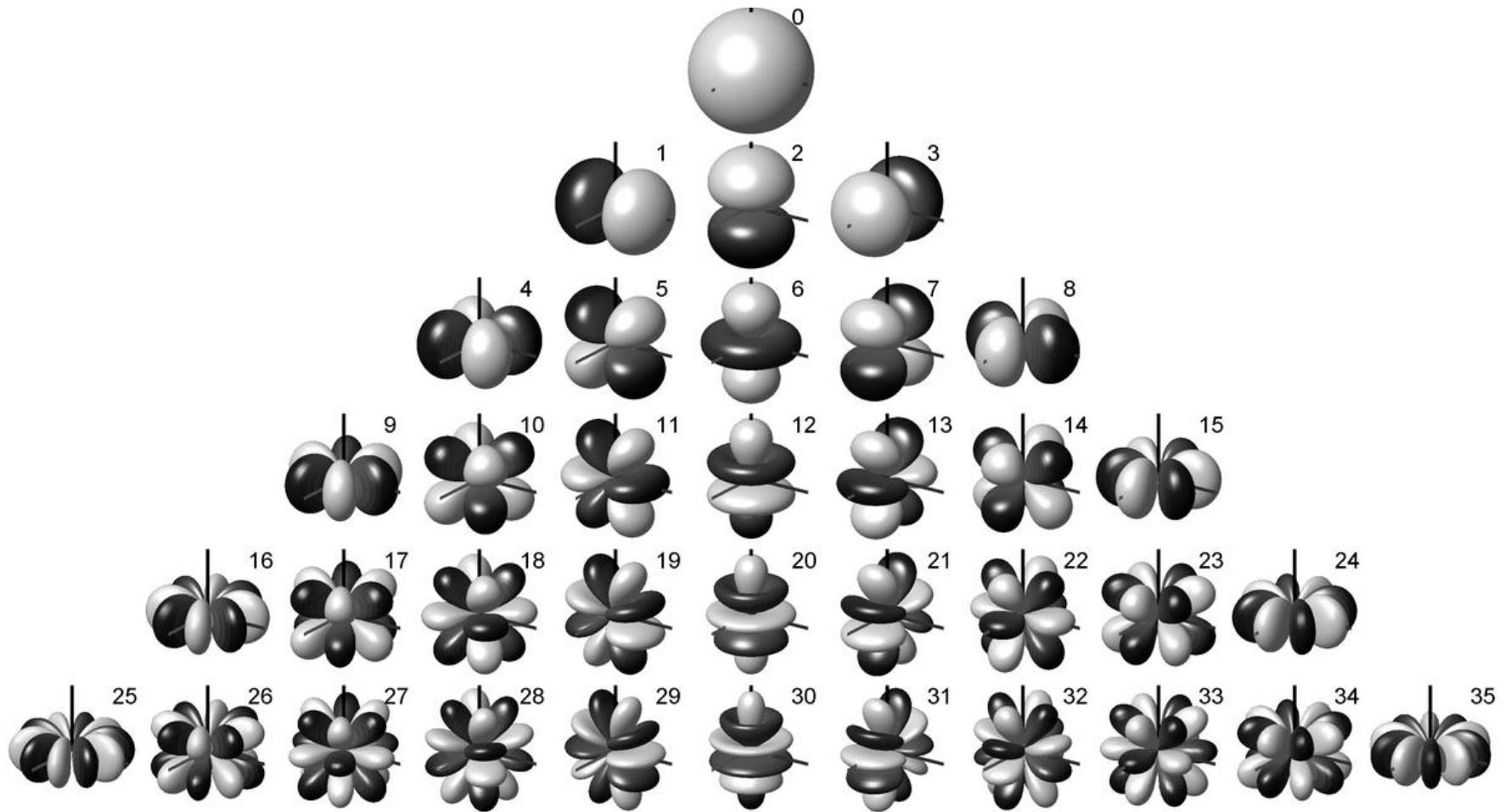
- l : βαθμός, $m=-l, -l+1, \dots, +l$: τάξη
- Legendre πολυώνυμα

$$p_l^m(x) = \frac{(-1)^m}{(2^l l!)} (1+x^2)^{\frac{m}{2}} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l$$

- Τα πολυώνυμα υπολογίζονται ταχύτερα με αναδρομικό τύπο

- SPHARM για $l=0,1,2,3$ και τις αντίστοιχες τιμές m .

l	$m = 0$	$m = 1$	$m = 2$	$m = 3$
0	1			
1	$\cos \theta$	$e^{i\phi} \sin \theta$		polar
2	$-1 + 3 \cos^2 \theta$	$e^{i\phi} \cos \theta \sin \theta$	$e^{2i\phi} \sin^2 \theta$	
3	$-3 \cos \theta + 5 \cos^3 \theta$	$e^{i\phi} (1 - 5 \cos^2 \theta) \sin \theta$	$e^{2i\phi} \cos \theta \sin^2 \theta$	$e^{3i\phi} \sin^3 \theta$
0	$1/\sqrt{4\pi}$			
1	$\sqrt{3/4\pi}$	$\sqrt{3/8\pi}$		
2	$\sqrt{5/16\pi}$	$\sqrt{15/8\pi}$	$\sqrt{15/32\pi}$	
3	$\sqrt{7/16\pi}$	$\sqrt{21/64\pi}$	$\sqrt{105/32\pi}$	$\sqrt{35/64\pi}$

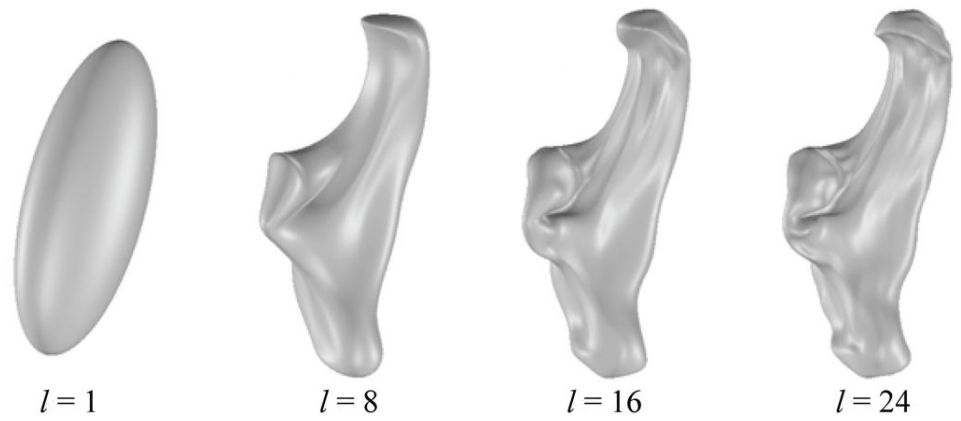
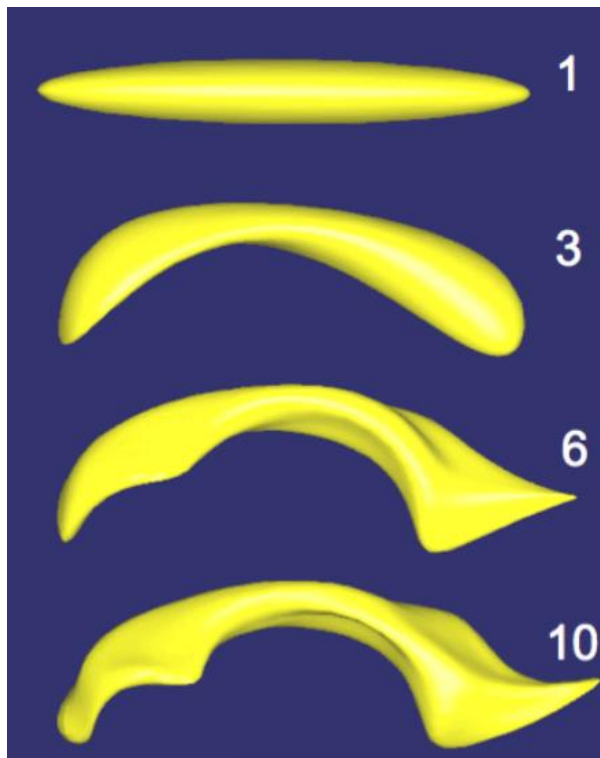


- The real part of the spherical harmonics, up to degree $l=5$.

- Any surface with topology equivalent to a sphere, can be decomposed into a weighted sum of SPHARM

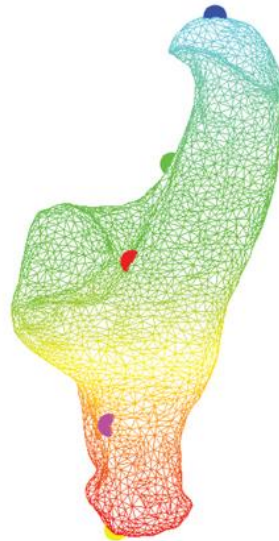
$$f(\theta, \varphi) = \sum_l \sum_m c_m^l Y_m^l(\theta, \varphi)$$

- The decomposition is hierarchical:
 - Larger degree SPHARM contain finer shape details
- If we truncate the summations, the set of $\{c_l^m\}$ is a 3D shape descriptor

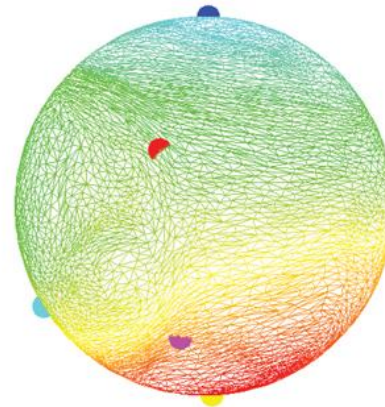


- Parameterization of closed surfaces for 3-D shape description
Ch.Brechbühl, 1996

Triangular Mesh



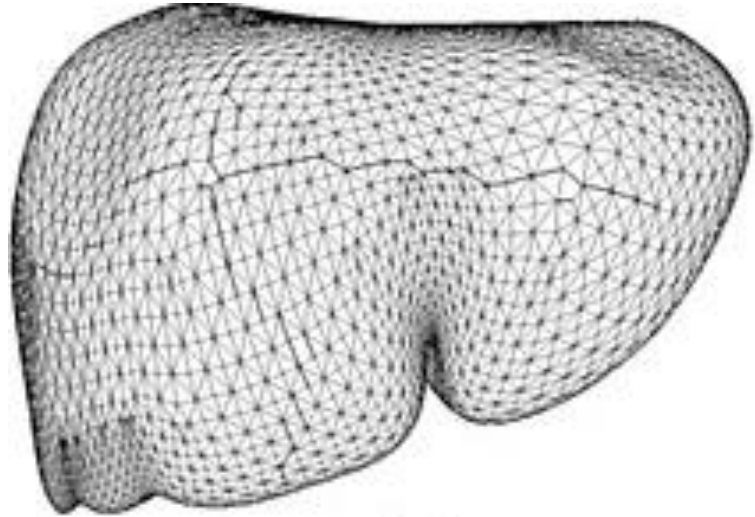
Unit Sphere
Representation



Types of 3D objects



Volumetric



Triangular meshes

- Triangular meshes

Στατιστική διακύμανση σχημάτων

- Συχνά τα αντικείμενα ενδιαφέροντος παρουσιάζουν μεταβολές στο σχήμα τους, πχ ίδια ανατομική δομή μεταξύ διαφορετικών ανθρώπων.
- Active Shape models –ASM: βασίζεται σε Principal Component Analysis -PCA, για να εκπαιδεύσει ένα σύστημα ώστε να παράγει στατιστικά πιθανά σχήματα, δεδομένου ενός συνόλου σχημάτων (αντικειμένων ίδιου τύπου)

Στατιστικές έννοιες

- Μέση τιμή $\mu_X = E(X)$:
$$\mu_x = E(X) = \frac{1}{n} \sum_i x_i$$
- Διασπορά
$$\sigma_x^2 = E\left(X - E[X]\right)^2 = E[X^2] - E[X]^2 = \frac{1}{n} \left(\sum_i x_i^2 - \left(\sum_i x_i \right)^2 \right)$$
- Συμμεταβλητότητα ((συνδιακύμανση ή συνδιασπορά) 2 μεταβλητών, X, Y):
$$\text{Cov}(X, Y) = E\left[(X - \mu_x)(Y - \mu_y)\right] = \frac{1}{N} \sum_{i=1}^{N-1} (x_i - \mu_x)(y_i - \mu_y)$$
- Σημ. Το $(1/N)$ μπορεί να αντικατασταθεί από $1/(N-1)$, αν θεωρούμε ότι τα διαθέσιμα είναι μέρος του συνόλου των σημείων του αντικείμενου

- Συσχέτιση δύο μεταβλητών
- Συνδιασπορά $>0 \rightarrow$ οι 2 μεταβλητές συσχετισμένες:
 - Όταν αυξάνει η μία, αυξάνει και η άλλη
- Συνδιασπορά $<0 \rightarrow$ οι 2 μεταβλητές συσχετισμένες:
 - Όταν αυξάνει (μειώνεται) η μία, μειώνεται (αυξάνει) και η άλλη
- Συντελεστής συσχέτισης: κανονικοποιούμε τη συνδιασπορά με τις αποκλίσεις των 2 μεταβλητών

$$\rho_{xy} = \frac{Cov(X, Y)}{\sigma_x \sigma_y} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sigma_x \sigma_y}$$

- Πίνακας συμμεταβλητότητας (συνδιακύμανση ή συνδιασπορά) δύο μεταβλητών X, Y με N τιμές έκαστη:
 - Αν μ_X, μ_Y οι μέσες τιμές, τότε ορίζεται ο συμμετρικός πίνακας 2x2:

$$C = \begin{pmatrix} x_1 - \mu_x & x_2 - \mu_x & \cdots & x_n - \mu_x \\ y_1 - \mu_y & y_2 - \mu_y & \cdots & y_n - \mu_y \end{pmatrix} \begin{pmatrix} x_1 - \mu_x & y_1 - \mu_y \\ x_2 - \mu_x & y_2 - \mu_y \\ \vdots & \vdots \\ x_n - \mu_x & y_n - \mu_y \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^N (x_i - \mu_x)^2 & \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \\ \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) & \sum_{i=1}^N (y_i - \mu_y)^2 \end{pmatrix}$$

- Αν ο πίνακας είναι διαγώνιος $\rightarrow$ οι X, Y ασυσχέτιστες

Συνδιακύμανση σχημάτων ίδιου αντικειμένου

- Εστω N αντικείμενα, κάθε ένα από τα οποία έχει D σημεία περιγράμματος

$$\left\{ \left(x_i^k, y_i^k \right) \right\}, \begin{matrix} i = 1, 2, \dots, D \\ k = 1, 2, \dots, N \end{matrix}$$

- Τα βήματα που ακολουθούν εφαρμόζονται είτε ξεχωριστά για τις X και Y συντεταγμένες, είτε ταυτόχρονα, θεωρώντας ένα διάνυσμα μήκους $2D$ (x πρώτα και y μετά)

- Εστω $\mathbf{p}_k = \begin{bmatrix} x_i^k, y_i^k \end{bmatrix}^T, 2N \times 1, \begin{matrix} k = 1, 2, \dots, D \\ i = 1, 2, \dots, N \end{matrix}$

- Μεταφέρουμε το σύστημα συντεταγμένων κάθε αντικειμένου στο κέντρο μάζας του $(\bar{x}^k, \bar{y}^k)$

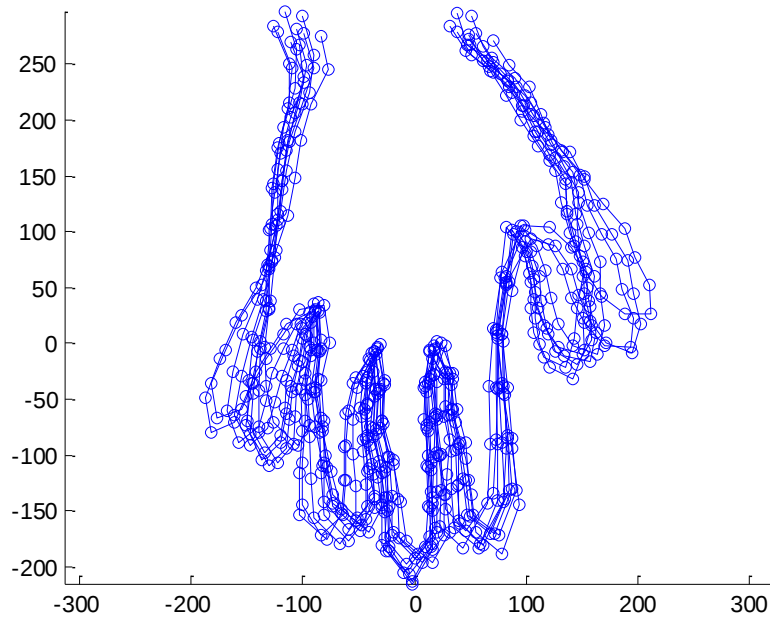
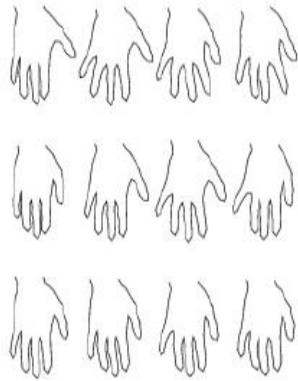
$$\mathbf{p}_k = \begin{bmatrix} x_i^k - \bar{x}^k, y_i^k - \bar{y}^k \end{bmatrix}^T, 2N \times 1$$

- Κατασκευάζουμε τον πίνακα

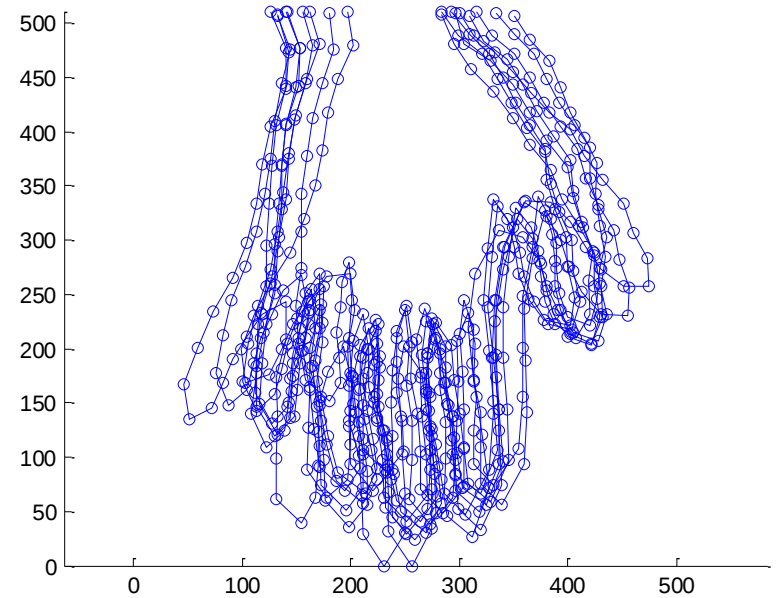
$$\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_D]^T, 2N \times D$$

Παράδειγμα

Σύνολο
περιγραμμάτων
ενός αντικειμένου
με μεταβολή στο
σχήμα



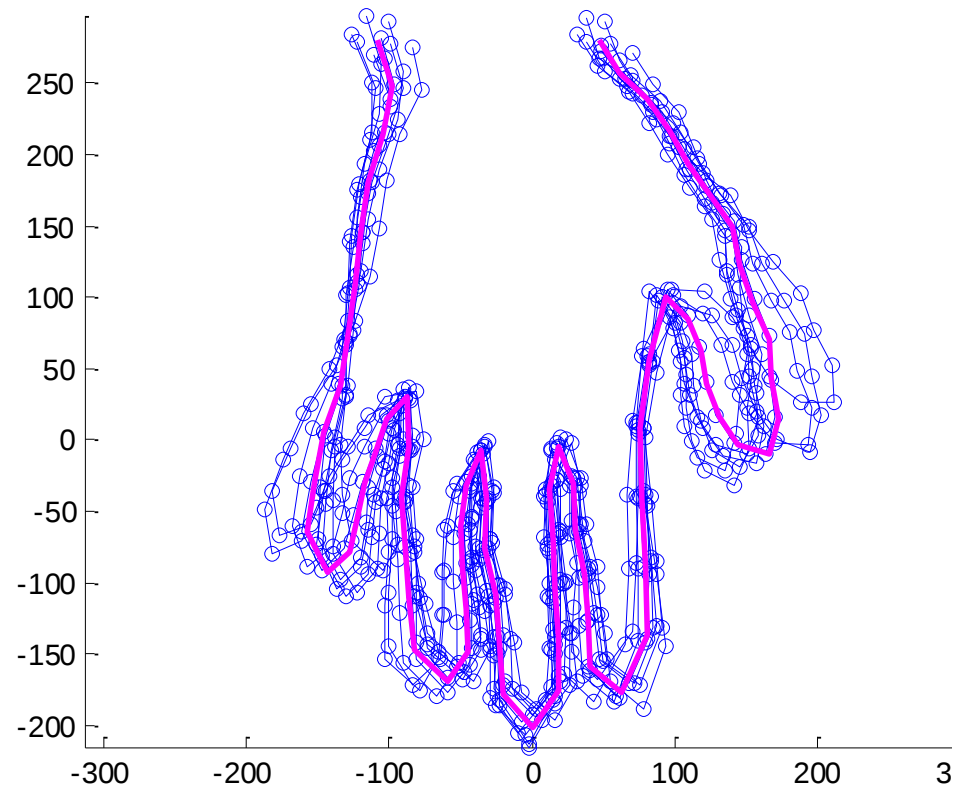
Τα περιγράμματα σε βαρυκεντρικές
συντεταγμένες



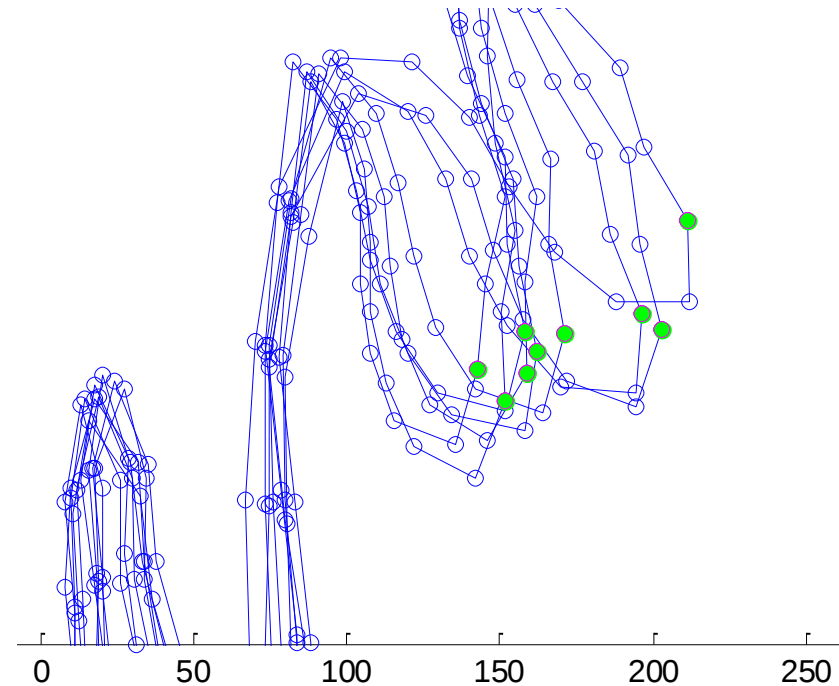
Παράδειγμα: $K=9$ περιγράμματα
χεριών, με $N=72$ σημεία το κάθε ένα

- Υπολογισμός μέσου σχήματος

$$\bar{\mathbf{p}} = \frac{1}{D} \sum_{k=1}^D \mathbf{p}_k, 2N \times 1$$



- Κατασκευάζεται πίνακας $\mathbf{P}$, διάστασης $2N \times D$.
- Ο πίνακας $\mathbf{M} = \mathbf{P} \cdot \mathbf{P}^T$ ($2N \times 2N$) δείχνει τη συμμεταβλητότητα των αντίστοιχων σημείων, μεταξύ των διαφορετικών χεριών
- Ο πίνακας $\mathbf{M} = \mathbf{P} \cdot \mathbf{P}^T$ ($2N \times 2N$) θα έχει $\text{rank}(\mathbf{M}) = 2N$, $\Leftrightarrow D > 2N$, Αλλιώς, $\text{rank}(\mathbf{M}) = D$



Διαφορετικές θέσεις του σημείου 12
για 10 περιγράμματα χεριών

Συμπύεση πληροφορίας σχήματος με χρήση του ASM

- Εστω $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_{2N})$ με μήκος $2N$ και
- $\mathbf{Q}$ ο πίνακας που οι στήλες του είναι τα ιδιοδιανύσματα του $\mathbf{M}$, σε φθίνουσα διάταξη των ιδιοτιμών του

$$\mathbf{Q} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{2N}]^T, 2N \times 2N,$$

$$\text{ιδιοτιμές του } \mathbf{M}: d_1 \geq d_2 \geq \dots \geq d_{2N}$$

- Για κάθε τυχαίο σχήμα $\mathbf{q}$, υπάρχουν $\Lambda = (\lambda_1, \lambda_2, \dots, \lambda_{2N})$, ώστε να μπορεί να γραφεί ως

$$\mathbf{q} = \bar{\mathbf{p}} + \mathbf{M} \cdot [\lambda_1, \lambda_2, \dots, \lambda_{2N}]^T, 2N \times 1 = \\ \bar{\mathbf{p}} + \mathbf{M} \cdot \Lambda$$

- Εστω ότι δεν χρησιμοποιούμε όλες τις ιδιοτιμές, αλλά τις T μεγαλύτερες, όπου

$$\sum_{i=1}^T d_i > 0.9 \sum_{i=1}^{2N} d_i$$

- Εστω ότι ο $\mathbf{Q}$ περιέχει τα ιδιοδιανύσματα με τις T ιδιοτιμές

$$\mathbf{Q} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_T]^T, 2N \times T,$$

$$\text{ιδιοτιμές του } \mathbf{M}: d_1 \geq d_2 \geq \dots \geq d_T \dots \geq d_{2N}$$

- Τώρα, ένα τυχαίο σχήμα $\mathbf{q}$, περιγράφεται από T αριθμούς $\Lambda=(\lambda_1, \lambda_2, ..., \lambda_T)$, ως

$$\mathbf{q} = \bar{\mathbf{p}} + \underset{2N \times T}{\mathbf{M}} \cdot \underset{2T \times 1}{\Lambda},$$

$$\Lambda = [\lambda_1, \lambda_2, ..., \lambda_T]^T$$

- Για να καλύπτουμε το 98% της μεταβολής του σχήματος, επιλέγουμε

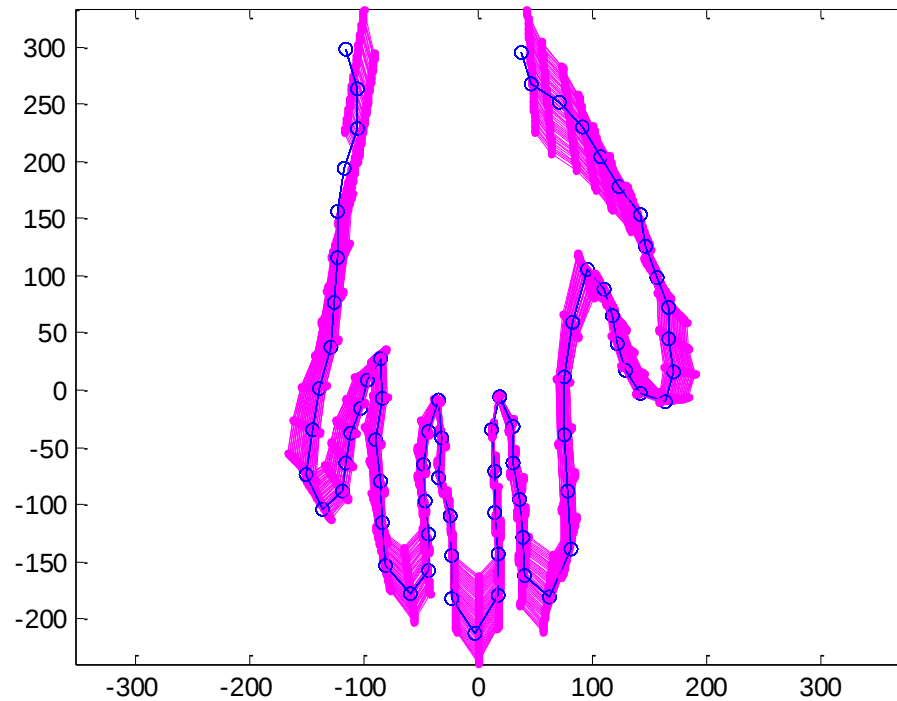
$$\lambda_i \in \left[d_i - 3\sqrt{d_i}, d_i + 3\sqrt{d_i} \right], i \leq T$$

- Αντίστροφα, δοθέντος του $\mathbf{q}$, τα Λ υπολογίζονται (διότι $\mathbf{M}^T \cdot \mathbf{M} = \mathbf{I}$)

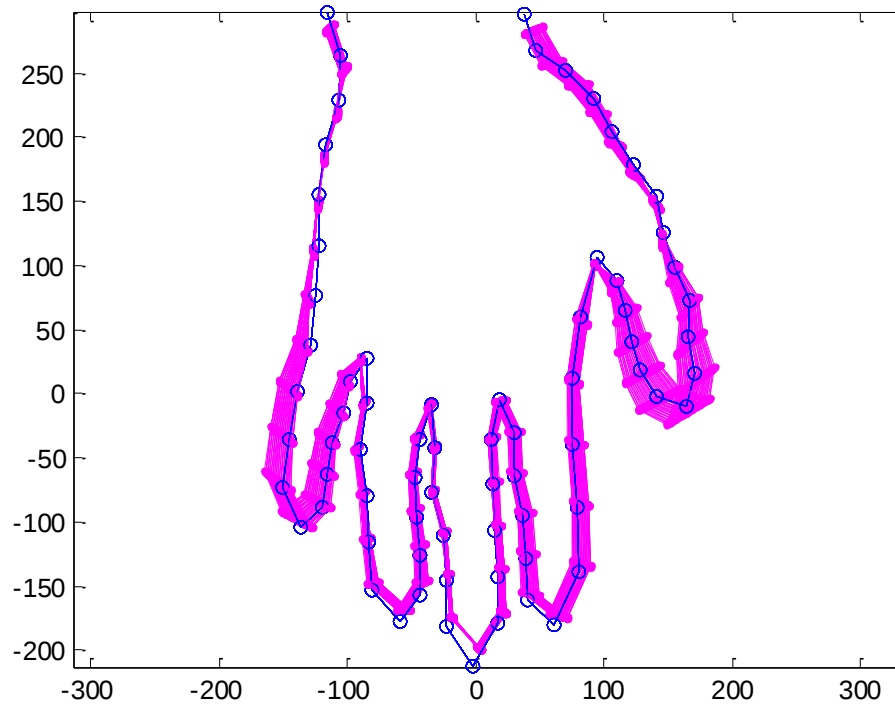
$$\Lambda = \mathbf{M}^T \cdot (\mathbf{q} - \bar{\mathbf{p}})$$

Συμπέρασμα

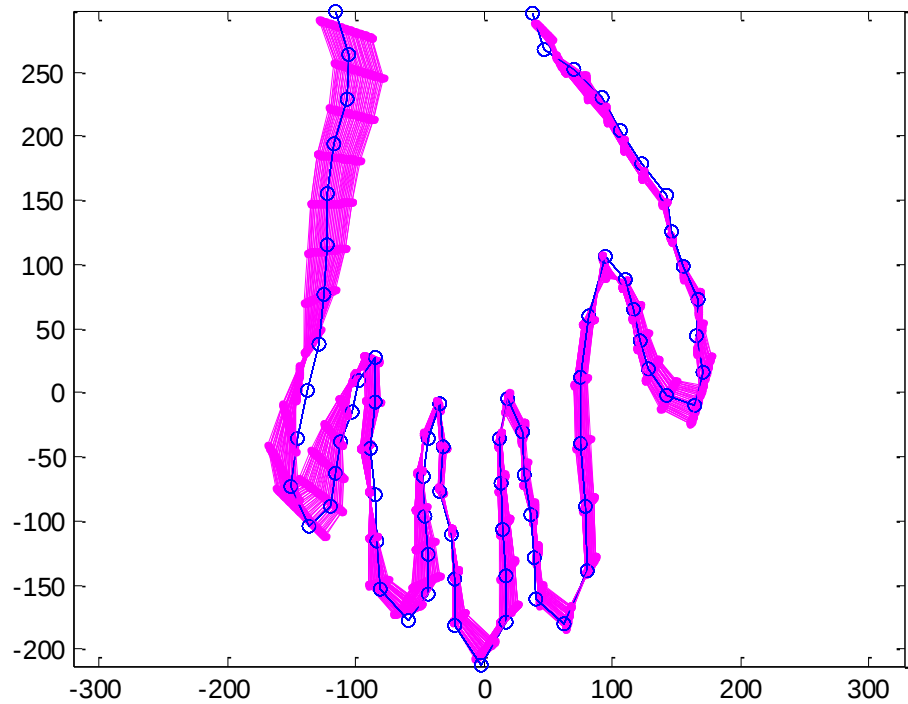
- Η περισσότερη πληροφορία του σχήματος υπάρχει συγκεντρωμένη στα ιδιοδυναύσματα του $\mathbf{M}$ που έχουν τις T μεγαλύτερες ιδιοτιμές
- Κάθε ιδιοδιάνυσμα «αποτυπώνει» ένα τρόπο μεταβολής (mode of variation)
- Για ομαλά σχήματα, $N \ll T$, $\rightarrow$ περιγράφουμε ένα σχήμα με πολύ λιγότερους αριθμούς από τις συντεταγμένες των σημείων του
- Μπορούμε να κατασκευάσουμε καινούργια σχήματα επιλέγοντας $(\lambda_1, \lambda_2, \dots, \lambda_T)$,



Μεταβολή σχήματος που κωδικοποιείται από το 1^ο
ιδιοδιάνυσμα



Μεταβολή σχήματος που κωδικοποιείται από το 2^ο
ιδιοδιάνυσμα



Μεταβολή σχήματος που κωδικοποιείται από το 3^ο
ιδιοδιάνυσμα

Αναζήτηση σε εικόνα τυχαίου σχήματος με μετατόπιση, αλλαγή κλίμακας και περιστροφή

- Κατασκευή του τυχαίου σχήματος με χρήση ASM (T παράμετροι)
- Περιστροφή, αλλαγή κλίμακας και μετατόπιση (περίπτωση 2D: $1 + 2 + 2 = 5$ παράμετροι)
- Υπολογισμός παραμέτρων με βελτιστοποίηση μίας συνάρτησης (ταιριάσματος, ή λάθους)
 - Μία πρακτική προσέγγιση: Μετασχηματισμός απόστασης / Chamfer matching

Chamfer Matching: Εντοπισμός αντικειμένου σε εικόνα

- Εστω ότι αναζητούμε ένα αντικείμενο (πχ παλάμη) σε μία εικόνα.
- Υπολογίζουμε τις ακμές της εικόνας
- Θεωρούμε ότι το αντικείμενο ανιχνεύεται όταν το περίγραμμα του βρίσκεται πολύ κοντά σε ακμές της εικόνας
- Για το λόγο αυτό χρησιμοποιείται ο Μετασχηματισμός Απόστασης (Distance Transform)

Συνάρτηση σφάλματος εντοπισμού αντικειμένου

- Εστω τα σημεία ενός αντικειμένου

$$(x_i, y_i), i = 1, 2, \dots, N$$

- Αν BW η δυαδική εικόνα ακμών της εικόνας I και $D=DT(I)$ ο μετασχ απόστασης
- Η συνάρτηση σφάλματος εντοπισμού του αντικειμένου μπορεί να οριστεί:

$$err = \frac{1}{N} \sum_{i=1}^N D(x_i, y_i)$$

Δυαδικές εικόνες: Μετασχηματισμός Απόστασης (Distance Transform)

- Έστω δυαδική εικόνα $I(i,j)$ η οποία περιέχει pixel ακμών με τιμή 0 (ενώ τιμή υποβάθρου θεωρείται η μεγιστη (big_value πχ 255). Σε περίπτωση που η εικόνα ακμών είναι δυαδική με τιμή 1 (ακμή) και 0 (υπόβαθρο), κάνουμε τον απαραίτητο μετασχηματισμό.
- Ο DT παράγει μία εικόνα της οποίας το κάθε pixel περιέχει την απόσταση του από το κοντινότερο pixel ακμών.
- Ο DT είναι χρήσιμος για την ταύτιση αντικειμένων με ψηφιακές εικόνες.
- “Αφελής (naïve) προσέγγιση”: Υπολογισμός απόστασης κάθε pixel μη-υποβάθρου από κάθε κάθε pixel υποβάθρου:
 - Brute force DT: $O(N^2)$

- Παρουσιάζονται 2 αλγόριθμοι:
 1. Μη βέλτιστος (χρειάζεται πολλαπλά περάσματα της εικόνας και δεν υπολογίζει ευκλείδεια απόσταση), παρουσιάζεται λόγω απλότητας.
 2. Βέλτιστος, 2 διπλά περάσματα, $O(N)$

Αλγόριθμος με πολλαπλά περάσματα της εικόνας

Αλγόριθμος Μετασχηματισμού απόστασης Manhattan

Εισοδος: I εικόνα ακμών (δυαδική)

Εξοδος: D εικόνα απόστασης (ακέραιες τιμές)

Αρχικοποίηση: για κάθε pixel $\mathbf{p}$, $D(\mathbf{p}) = \text{big_value}$

$k=0$

while (υπάρχουν pixel $\mathbf{p}$ με $D(\mathbf{p}) = \text{big_value}$)

$k=k+1$;

 Για κάθε pixel $\mathbf{p}$ της εικόνας

 Αν $D(\mathbf{p}) == \text{big_value}$ (pixel υποβάθρου)

 Για κάθε ένα από τους 4 γείτονες $\mathbf{n}$ του $\mathbf{p}$

 Αν $D(\mathbf{n}) == k-1$ then $D(\mathbf{p}) = I(\mathbf{n}) + 1$;

 end

 end

 end

end

Παράδειγμα εφαρμογής του αλγόριθμου DT

- Εστω η εικόνα I :

	1	2	3	4
1	1	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	1

	1	2	3	4
1	0	255	255	255
2	255	255	255	255
3	255	255	255	255
4	255	255	255	0

Αρχικοποίηση D
(big_value=255)

	1	2	3	4
1	0	1	255	255
2	1	255	255	255
3	255	255	255	1
4	255	255	1	0

Εικόνα D μετά το
1^ο βήμα ($k=1$).

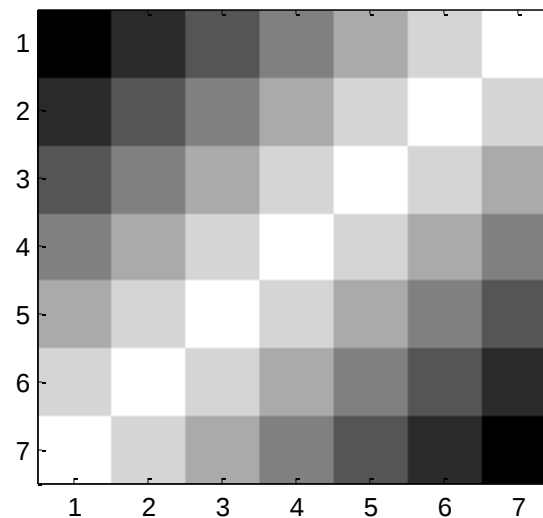
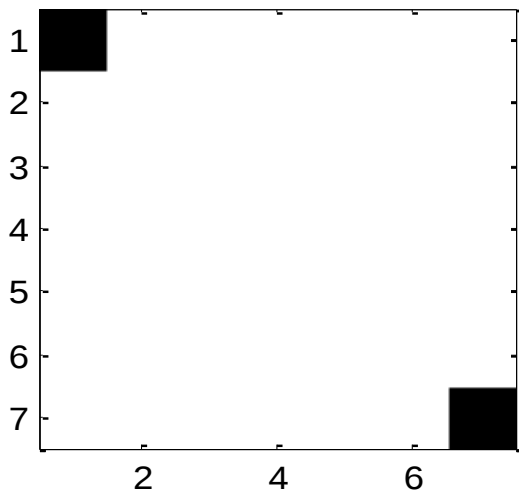
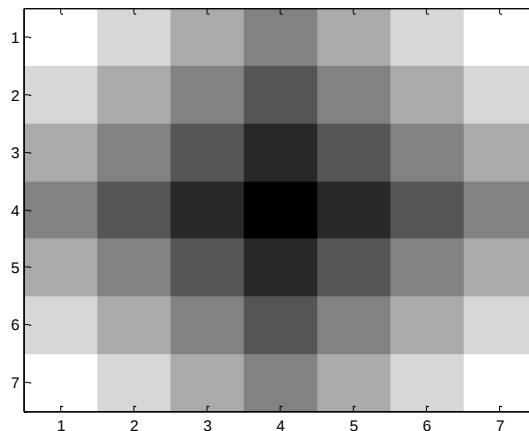
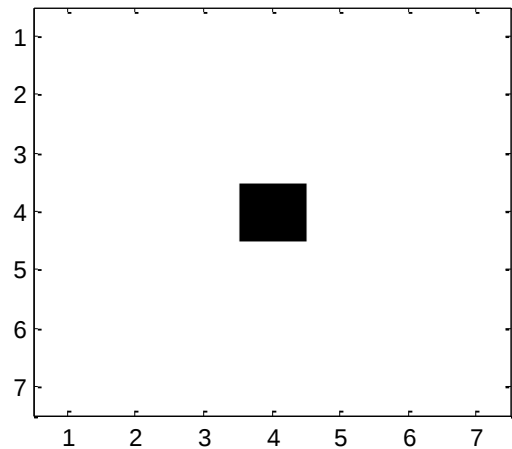
	1	2	3	4
1	0	1	2	255
2	1	2	255	2
3	2	255	2	1
4	255	2	1	0

Εικόνα D μετά το
2^ο βήμα ($k=2$).

	1	2	3	4
1	0	1	2	3
2	1	2	3	2
3	2	3	2	1
4	3	2	1	0

Εικόνα D μετά το
3^ο βήμα ($k=3$).

Εφαρμογή του αλγόριθμου DT



Distance Transform Αλγόριθμος: Sequential Euclidian 4-point

- 2 διπλά περάσματα, Ευκλείδεια απόσταση
- Επιπλέον, μπορεί να επιστρέψει το κοντινότερο pixel ακμής για κάθε pixel της εικόνας
- Αρχικοποίηση:
για κάθε pixel $\mathbf{p}$, $D(\mathbf{p}) = \text{big_value}$
- Επίσης θεωρούμε οτι ισχύει
 $\text{big_value} + 1 \rightarrow \text{big_value}$

Raw-wise, Forward

1. $\forall i=1:N$

1. $\forall j=2:M$ $\xrightarrow{\text{Forward}}$

$$D(i, j) = \min(D_{ij}, D_{i, j-1} + 1, D_{i-1, j} + 1)$$

2. $\forall j=M-1:1$ $\xleftarrow{\text{Column-wise, Backward}}$

$$D(i, j) = \min(D_{ij}, D_{i, j+1} + 1)$$

2. $\forall i=N-1:1$

1. $\forall j=2:M$ $\xrightarrow{\text{Forward}}$

$$D(i, j) = \min(D_{ij}, D_{i, j+1} + 1, D_{i+1, j} + 1)$$

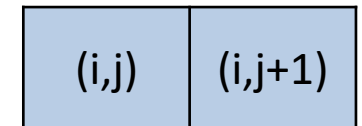
2. $\forall j=M-1:1$ $\xleftarrow{\text{Column-wise, Backward}}$

$$D(i, j) = \min(\min(D_{ij}, D_{i, j+1} + 1))$$

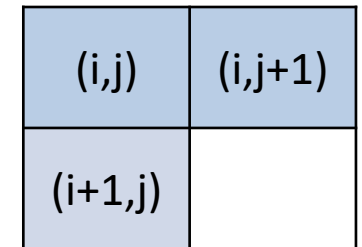
Raw-wise, Backward



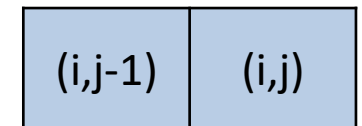
Mask 1



Mask 2



Mask 3



Mask 4

- Για την ευρεση του κοντινότερου Pixel ακμής για κάθε pixel της εικόνας:
 - Τηρούμε 2 πίνακες, διάστασης ίσης με την αρχική
 - P_i, P_j : parent γραμμή i , στήλη j , για κάθε (i, j)
 - P_i, P_j : Αρχικοποιούνται με NIL (πχ. -1)
 - Σε κάθε επανάληψη που ενημερωνεται η απόσταση του (i, j) , γίνεται αντίστοιχη ενημέρωση των.

- Image moments, Fourier Descriptors:
 - Gonzalez, Woods, Digital Image Processing, κεφ. 11.
- Distance Transform, Αρχικό paper:
 - Borgefors, Gunilla. "Hierarchical chamfer matching: A parametric edge matching algorithm." *IEEE Transactions on PAMI* 10.6 (1988): 849-865.
 - Tutorial with many algorithmic variations (eg. 4SSEDt)
 - <https://pdfs.semanticscholar.org/494f/6e6637b035308238b29323149739b0a7e502.pdf>, (p11)
- ASM Αρχικό paper
 - T. Cootes, et al (1995). "Active shape models - their training and application". *Computer Vision and Image Understanding* (61): 38–59
 - http://personalpages.manchester.ac.uk/staff/timothy.f.cootes/papers/cootes_cviu95.pdf
 - Tutorial:
 http://www2.compute.dtu.dk/courses/02511/docs/asm_overview.pdf