



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΑΝΟΙΚΤΑ ακαδημαϊκά
μαθήματα ΠΠ

Τεχνητή Νοημοσύνη II

Ενότητα 3: Αντίληψη

Μουστάκας Κωνσταντίνος

Πολυτεχνική Σχολή

Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών

Σκοποί ενότητας

- Αντίληψη



Περιεχόμενα ενότητας

➤ Αντίληψη



Αντίληψη

Τεχνητή Νοημοσύνη II

Αντίληψη (2)

Ανίχνευση ακμών

- **Στόχος:** Ανίχνευση απότομων αλλαγών (ασυνέχειες) σε μία εικόνα
 - Γενικά, η σημασιολογική πληροφορία αλλά και το σχήμα των αντικειμένων κωδικοποιείται στις ακμές
 - Πιο συμπαγής αναπαράσταση από ότι τα εικονοστοιχεία
- **Ιδανικό:** Σκαρίφημα σχεδιαστή (και αυτός βασίζεται σε λογική αντικειμένου)

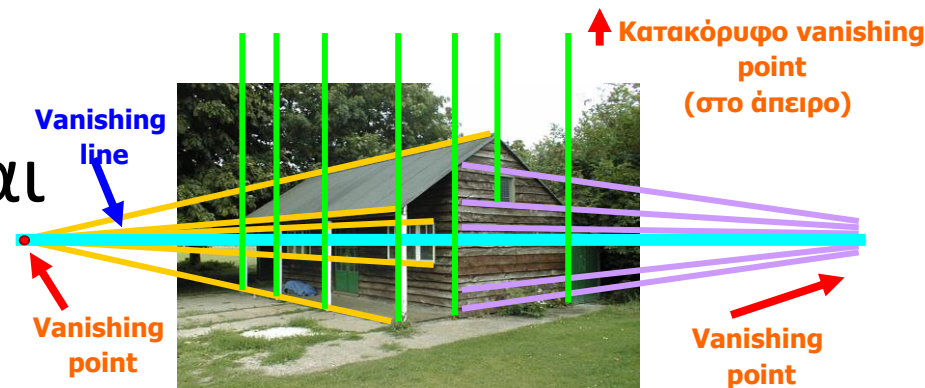


Τι μας χρειάζονται οι ακμές;

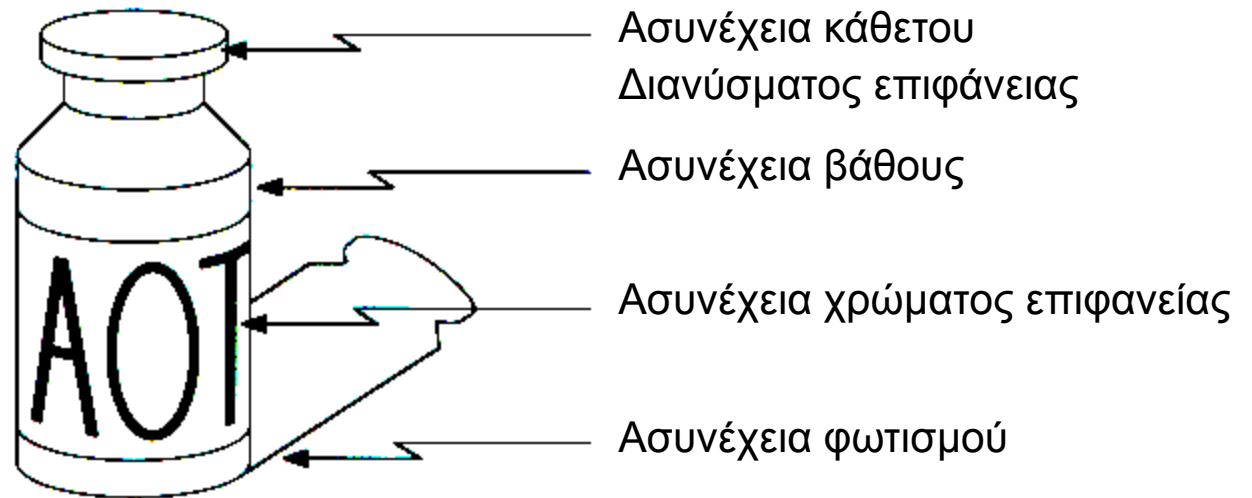
- Εξαγωγή πληροφορίας, αναγνώριση αντικειμένων



- Εξαγωγή γεωμετρίας και σημείου παρατήρησης



Από που προέρχονται οι ακμές;



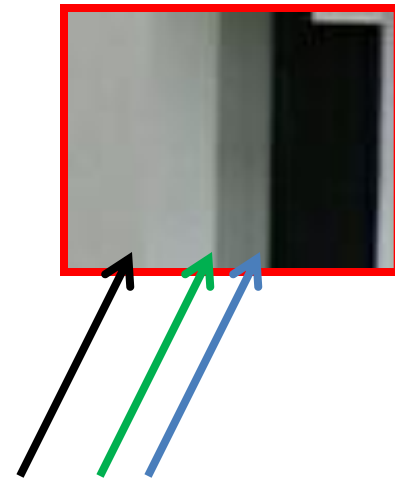
- Οι ακμές οφείλονται σε πολλούς παράγοντες



Ακμές σε μεγέθυνση



Ακμές σε μεγέθυνση



Ακμές σε μεγέθυνση

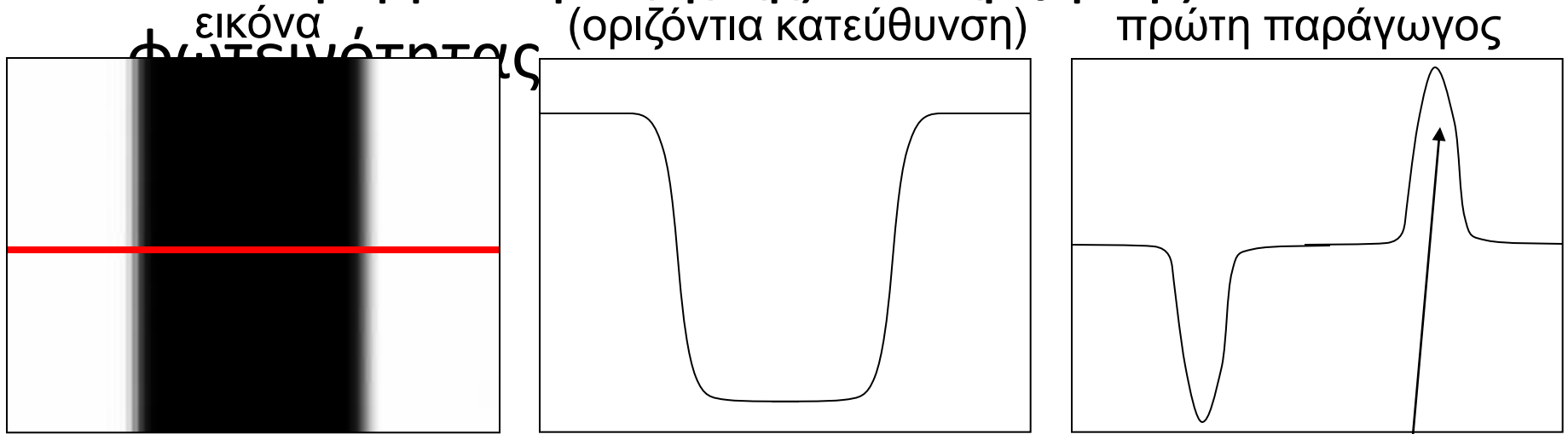


Ακμές σε μεγέθυνση



Χαρακτηρισμός ακμών

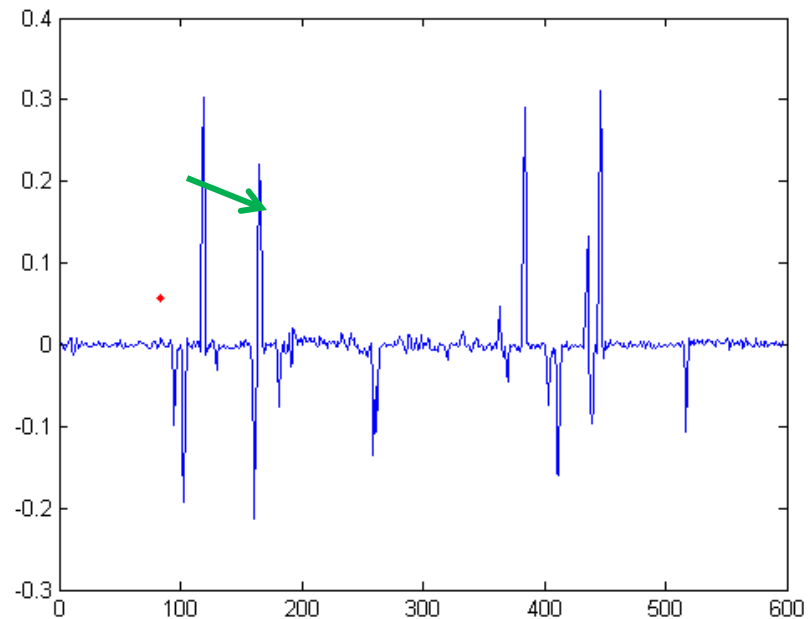
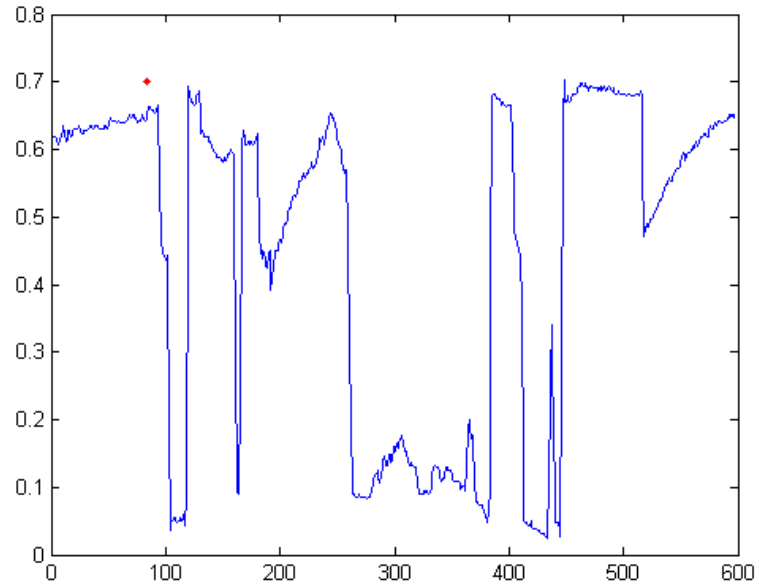
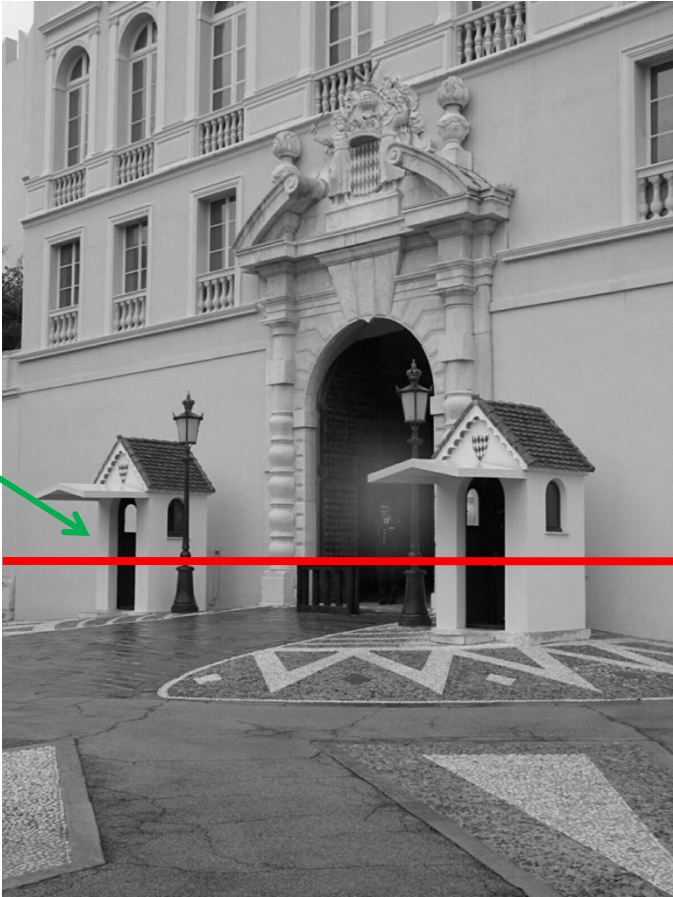
- Μία ακμή είναι μία περιοχή της εικόνας με απότομη μεταβολή της συνάρτησης



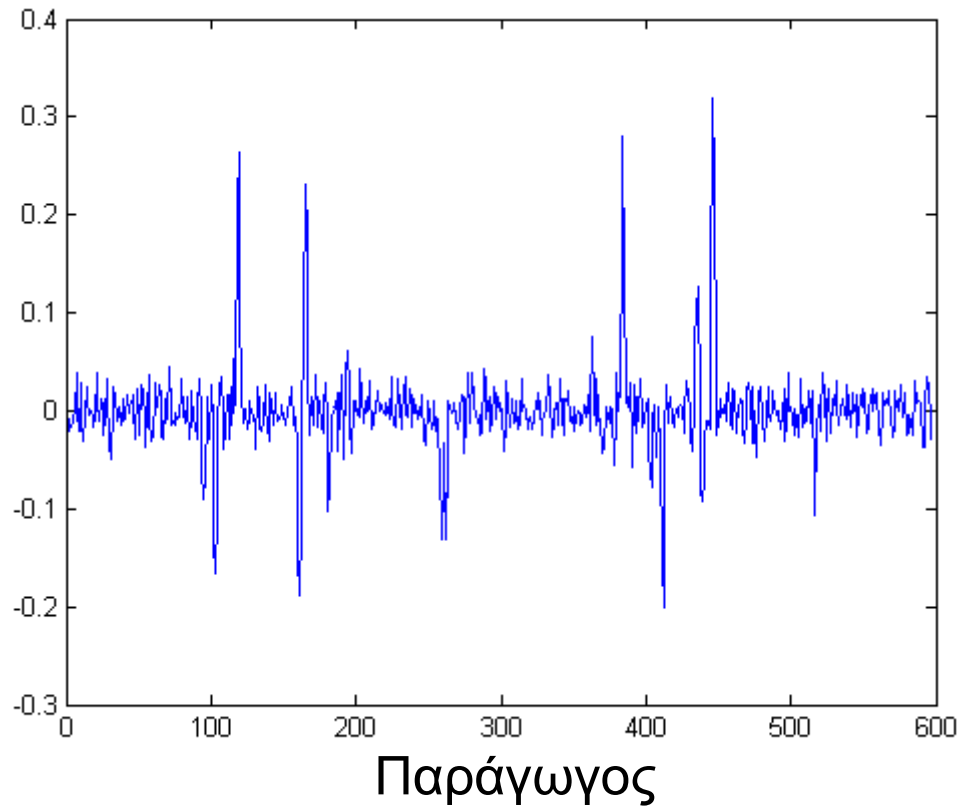
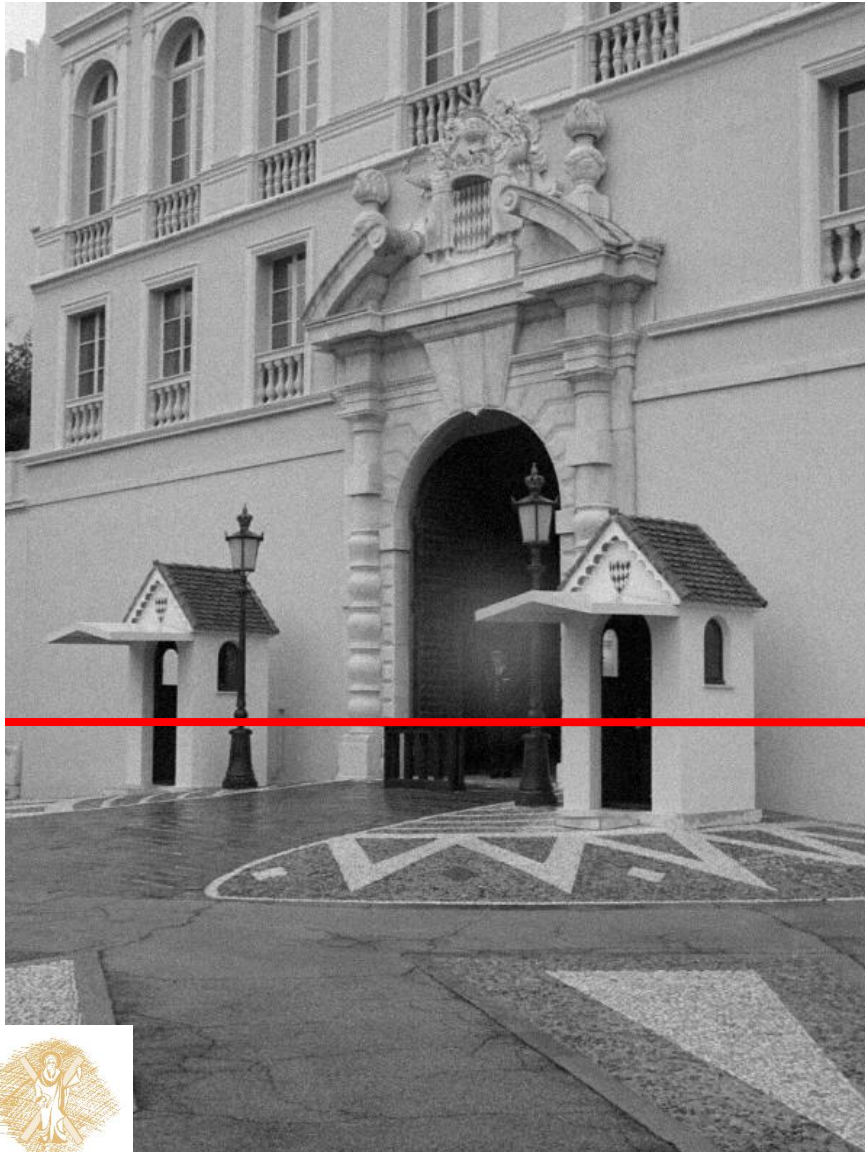
οι ακμές αντιστοιχούν
σε ακρότατα της παραγώγου



Προφίλ φωτεινότητας

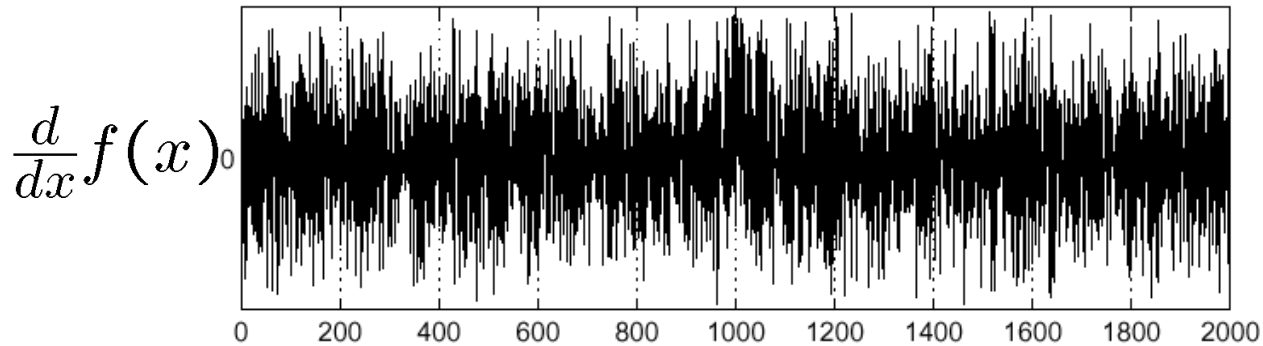
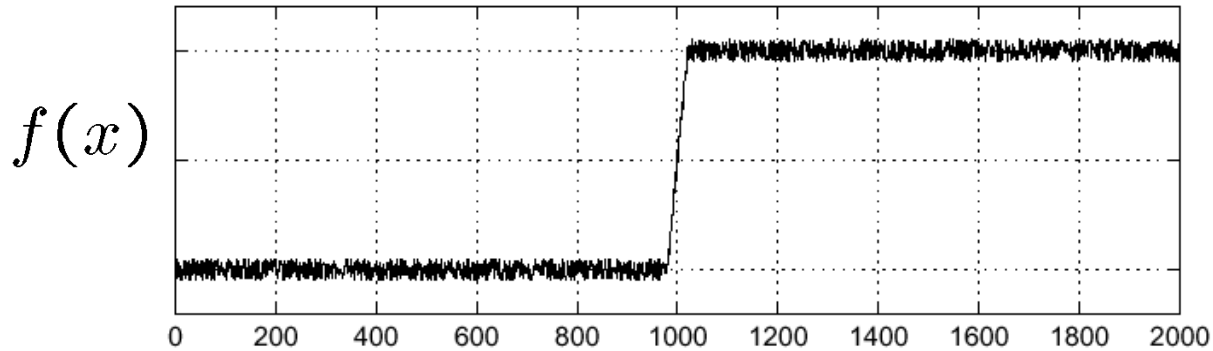


Με λίγο Gaussian θόρυβο



Αποτέλεσμα του θορύβου

- Έστω μία ακμή ή στήλη της εικόνας
 - Η απεικόνιση της φωτεινότητας ως συνάρτηση της θέσης δίνει ένα σήμα



Που είναι η ακμή;

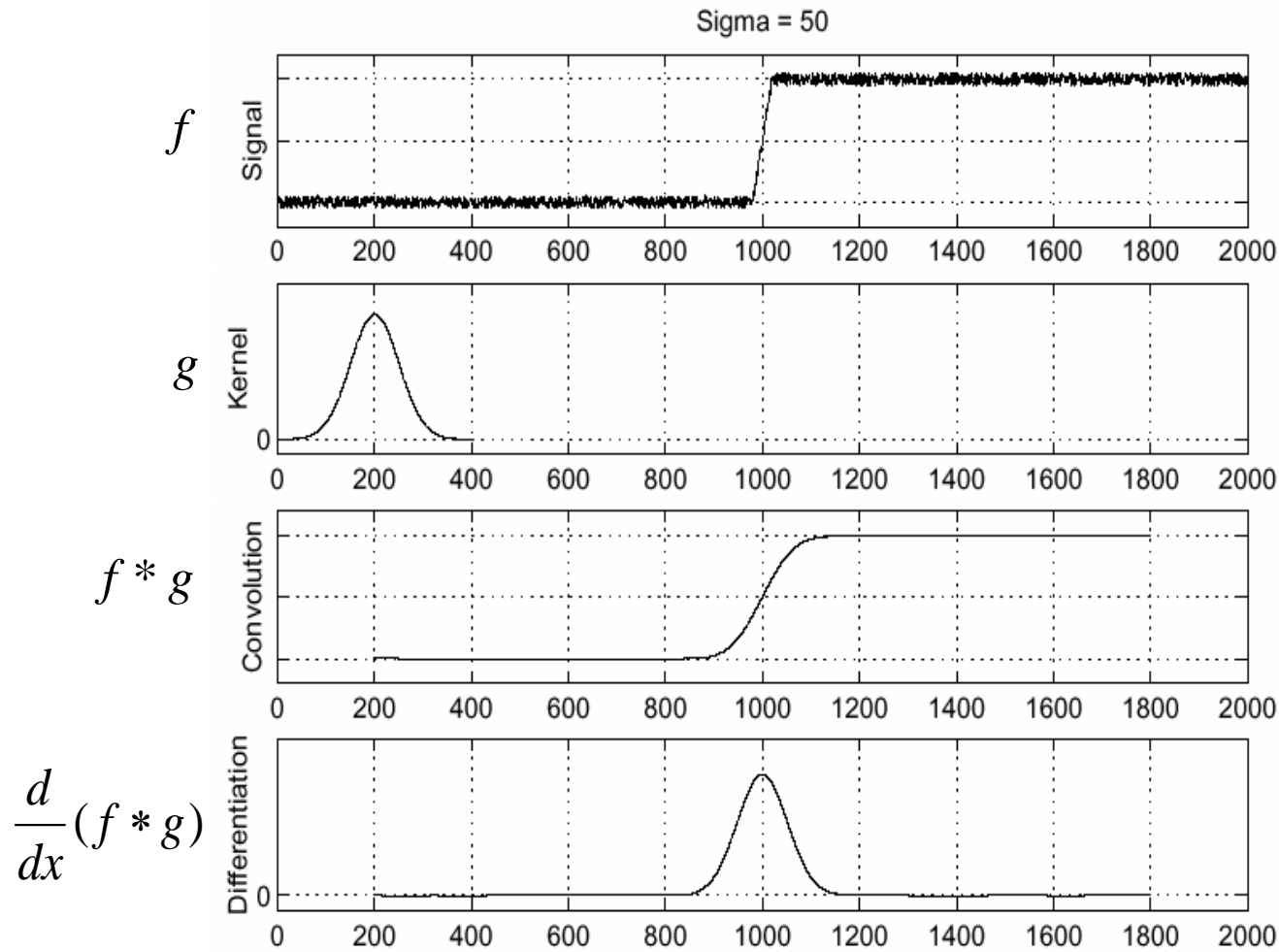


Αποτέλεσμα του θορύβου

- Τα φίλτρα διαφορών επηρεάζονται πολύ από το θόρυβο
 - Ο θόρυβος στις εικόνες έχει ως αποτέλεσμα εικονοστοιχεία τα οποία φαίνονται πολύ διαφορετικά από τους γείτονές τους
 - Γενικά, όσο μεγαλύτερος ο θόρυβος τόσο μεγαλύτερη η απόκρισή του στο φίλτρο διαφορών
- Οπότε?



Λύση: εξομάλυνση

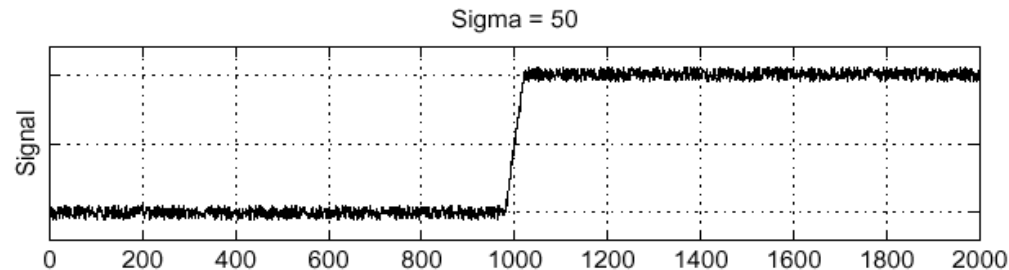


- Ψάχνουμε για ακμές, στις κορυφές του $\frac{d}{dx}(f * g)$

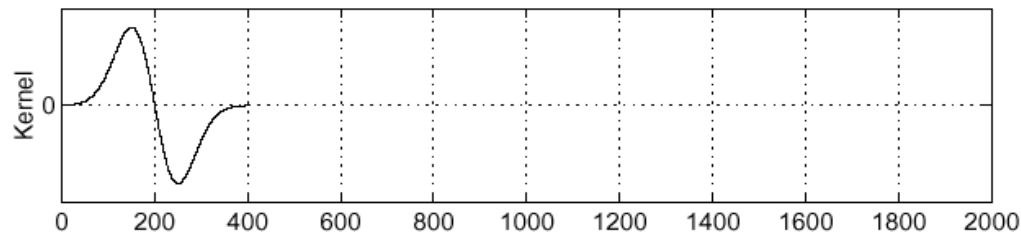
Θεώρημα παραγώγισης συνέλιξης

- Παραγώγιση συνέλιξης, η συνέλιξη είναι προσεταιριστική: $\frac{d}{dx}(f * g) = f * \frac{d}{dx}g$

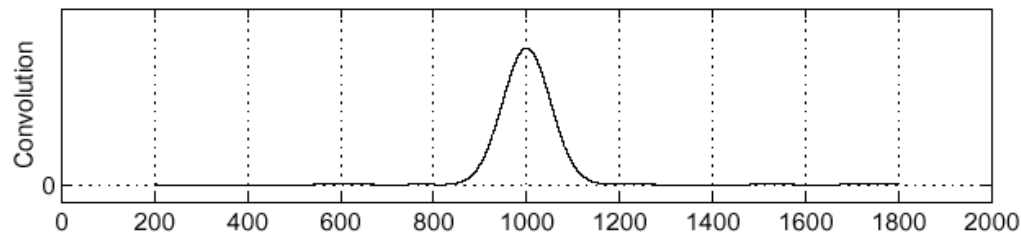
- Οπότε:
 f



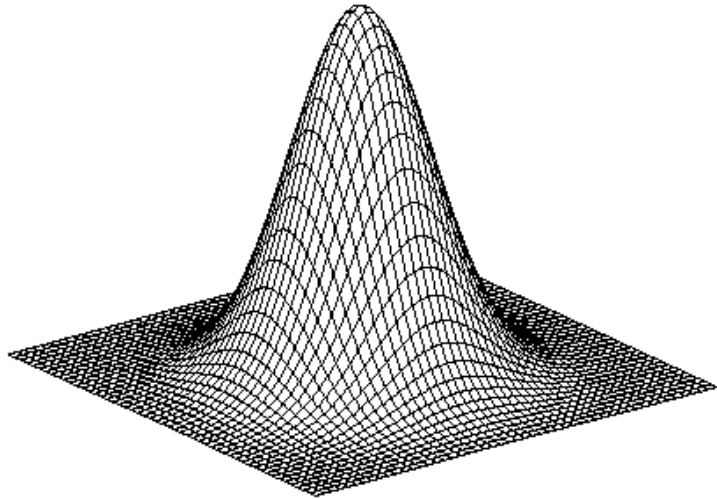
$$\frac{d}{dx}g$$



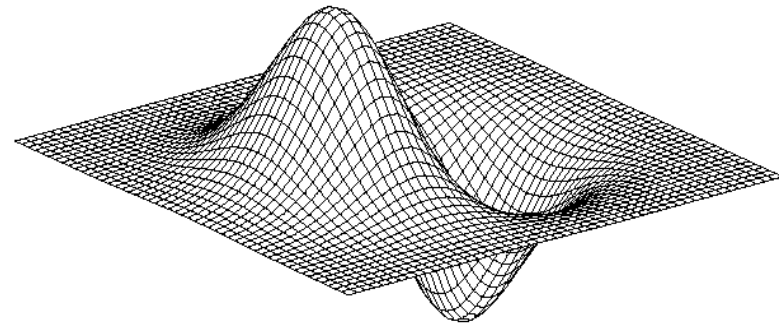
$$f * \frac{d}{dx}g$$



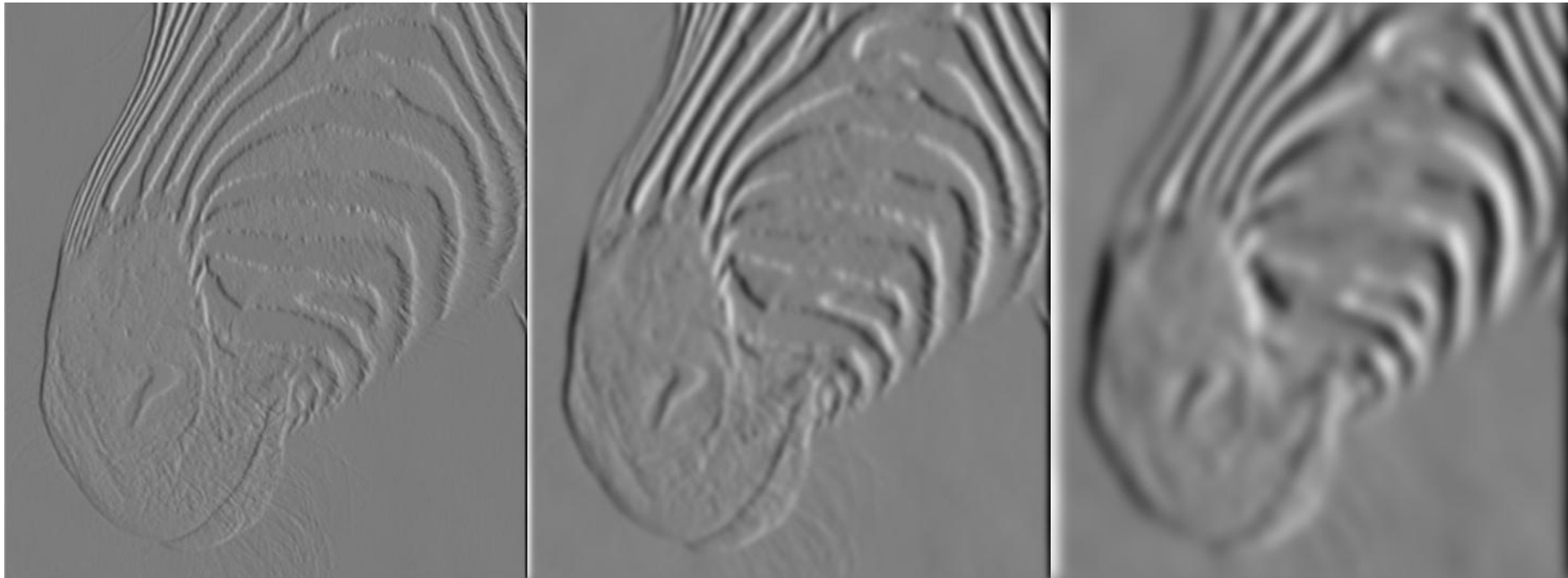
Παράγωγος Gaussian φίλτρου



$$* [1 \ -1] =$$



Εξομάλυνση και contrast



1 pixel

3 pixels

7 pixels

- Η εξομαλυμένη παράγωγος αφαιρεί το θόρυβο αλλά θολώνει τις ακμές. Μπορούν ακόμα να βρεθούν ακμές σε διάφορες κλίμακες. Πώς;



Σχεδιασμός ενός ανιχνευτή ακμών

- Κριτήρια ενός καλού ανιχνευτή:
 - **Σωστή ανίχνευση:** ο ιδανικός ανιχνευτής πρέπει να βρίσκει όλες τις πραγματικές ακμές αγνοώντας το θόρυβο
 - **Σωστή χωροθέτηση**
 - Οι ακμές που ανιχνεύονται πρέπει να είναι όσο πιο κοντά γίνεται στις πραγματικές ακμές
 - Ο ανιχνευτής πρέπει να επιστρέφει μόνο ένα σημείο για κάθε σημείο της ακμής
- Κλειδιά ανίχνευσης ακμών
 - Διαφορές στο χρώμα, στη φωτεινότητα, ή στην υφή κατά μήκος του ορίου
 - Συνέχεια και κλειστές καμπύλες
 - Σημασιολογική γνώση



Ανιχνευτής ακμών Canny

- Ίσως ο πιο ευρέως χρησιμοποιούμενος ανιχνευτής στην υπολογιστική όραση
- Θεωρητικό μοντέλο: ανίχνευση ακμών υπό προσθετικό Gaussian θόρυβο
- Ο Canny έδειξε ότι η πρώτη παράγωγος του Gaussian φίλτρου προσεγγίζει τον τελεστή που βελτιστοποιεί το γινόμενο *signal-to-noise ratio* και χωροθέτησης

J. Canny, [**A Computational Approach To Edge Detection**](#), IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986.



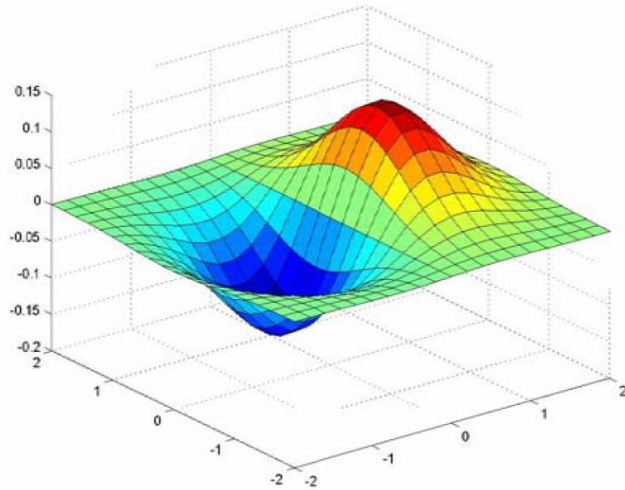
Παράδειγμα



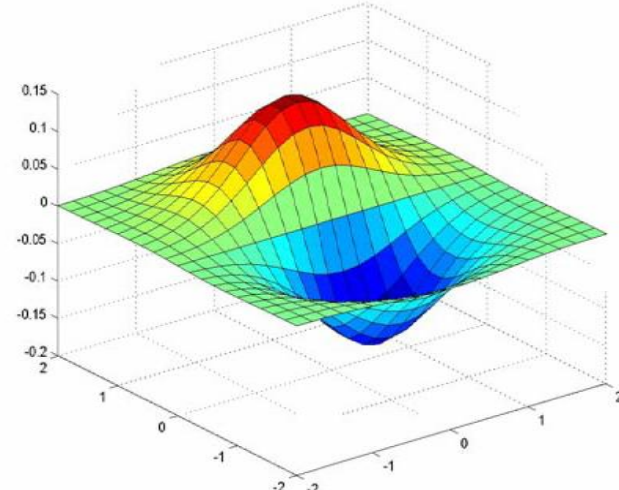
αρχική εικόνα (Lena)



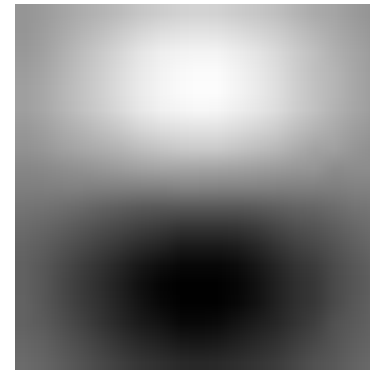
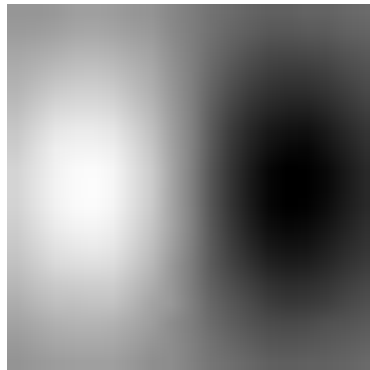
Παράγωγος Gaussian φίλτρου



χ-κατεύθυνση



γ-κατεύθυνση



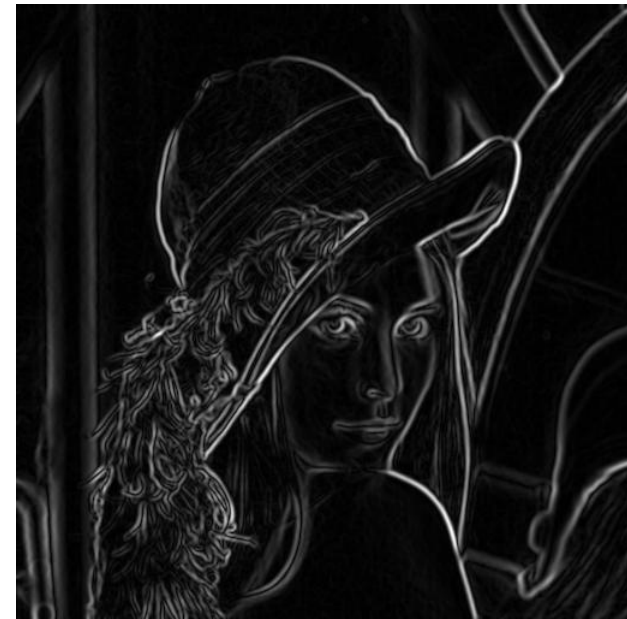
Υπολογισμός παραγώγων (DoG)



X-Derivative of Gaussian



Y-Derivative of Gaussian



Πλάτος Gradient



Επιρροή του σ



αρχική



Canny με $\sigma = 1$



Canny με $\sigma = 2$

Η επιλογή του σ εξαρτάται από την επιθυμητή συμπεριφορά

- μεγάλο σ ανιχνεύει ακμές μεγάλης κλίμακας
- μικρό σ ανιχνεύει λεπτές ακμές



Εύρεση ευθείας γραμμής

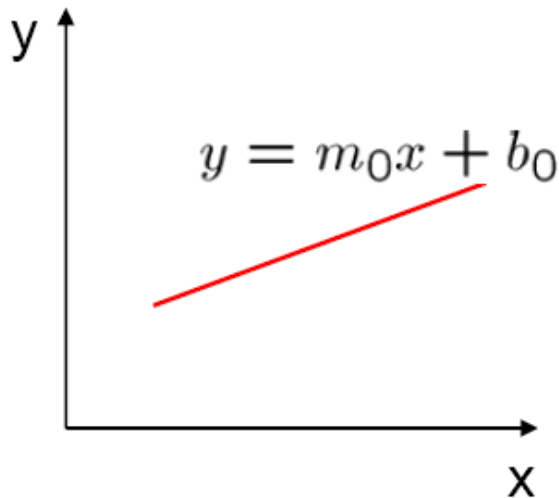
- Μία λύση: δοκιμή πολλών ευθειών και έλεγχος για το από πόσα σημεία διέρχεται
- Ο μετασχηματισμός Hough παρέχει ένα «γρήγορο» τρόπο για να γίνει αυτό



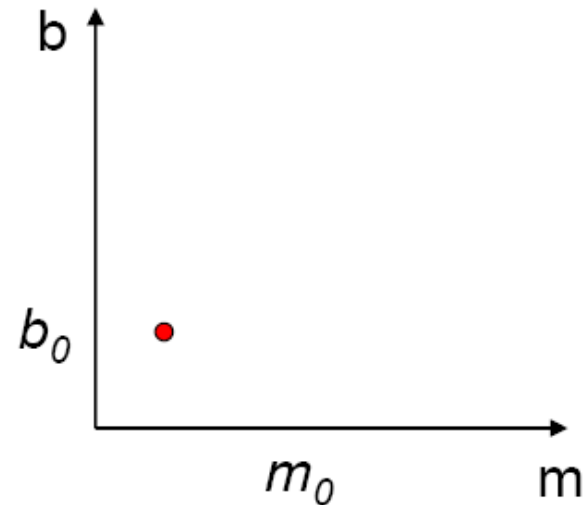
Αναπαράσταση παραμετρικού χώρου

- Μία ευθεία στην εικόνα αντιστοιχεί σε ένα σημείο στο χώρο Hough

Εικόνα

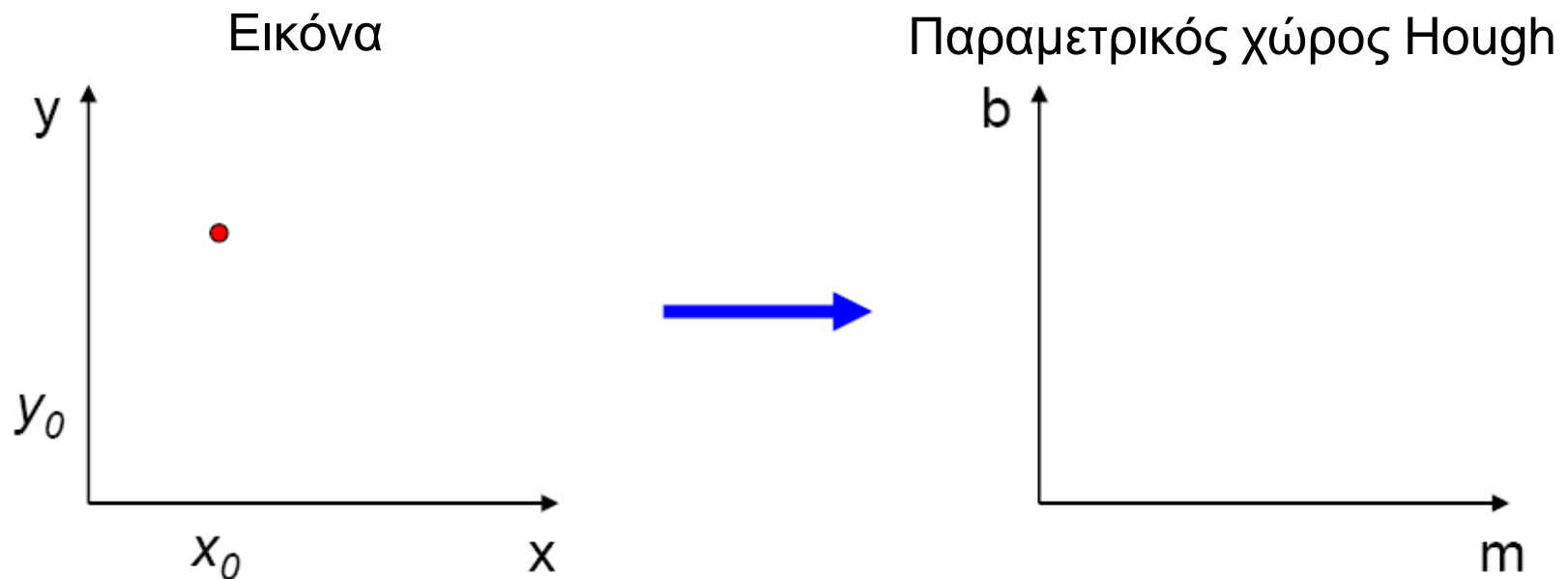


Παραμετρικός χώρος Hough



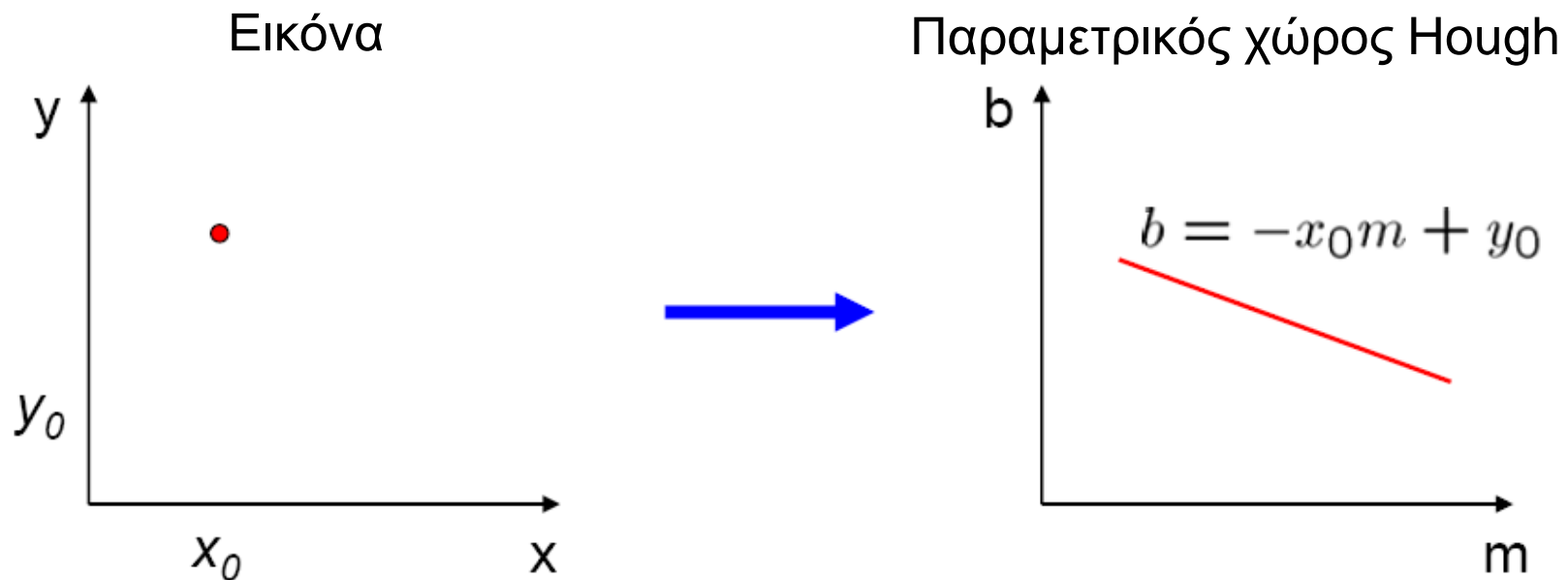
Αναπαράσταση παραμετρικού χώρου

- Σε τι αντιστοιχίζεται ένα σημείο (x_0, y_0) της εικόνας στο χώρο Hough;



Αναπαράσταση παραμετρικού χώρου

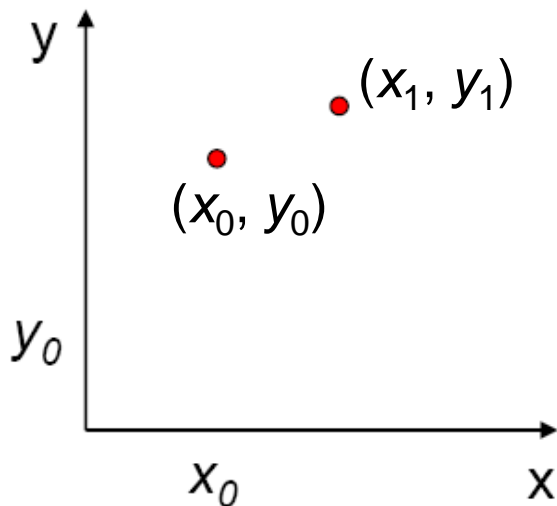
- Σε τι αντιστοιχίζεται ένα σημείο (x_0, y_0) της εικόνας στο χώρο Hough;
 - Απάντηση: οι λύσεις της $b = -x_0m + y_0$
 - Αυτό είναι μία ευθεία στο χώρο Hough



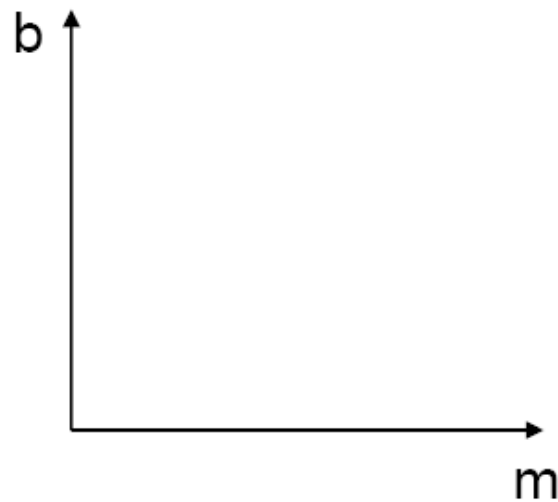
Αναπαράσταση παραμετρικού χώρου

- Που είναι η ευθεία που περιέχει τα σημεία (x_0, y_0) και (x_1, y_1) ;

Εικόνα



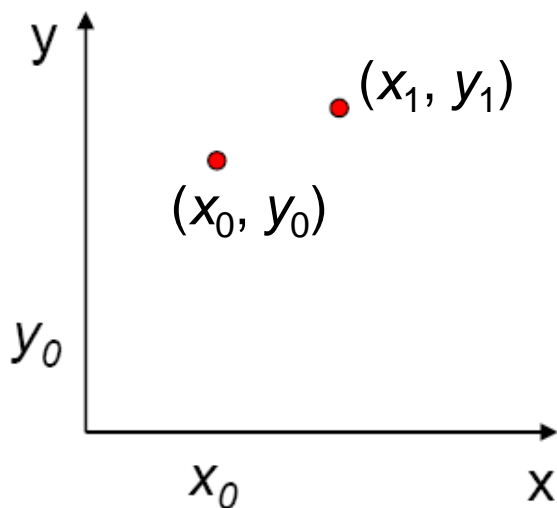
Παραμετρικός χώρος Hough



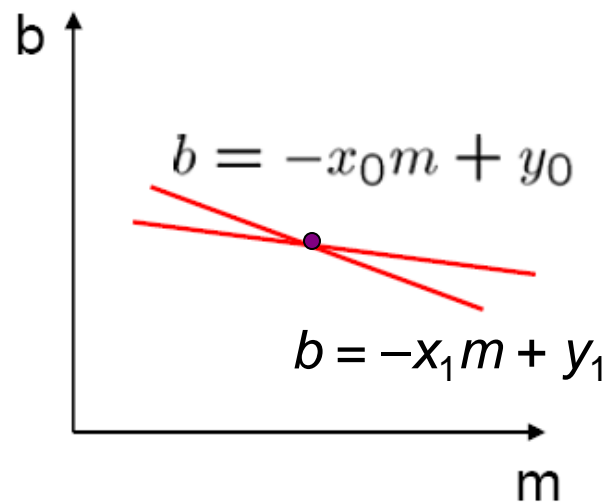
Αναπαράσταση παραμετρικού χώρου

- Που είναι η ευθεία που περιέχει τα σημεία (x_0, y_0) και (x_1, y_1) ;
 - Είναι η τομή των ευθειών $b = -x_0m + y_0$ και $b = -x_1m + y_1$

Εικόνα



Παραμετρικός χώρος Hough



Περιγραφή μετασχηματισμού Hough

1. Δημιουργία ενός πλέγματος παραμέτρων
2. Κάθε σημείο «ψηφίζει» για ένα σύνολο παραμέτρων αυξάνοντας τις τιμές τους στο πλέγμα
3. Εύρεση μεγίστου ή τοπικών μεγίστων στο πλέγμα



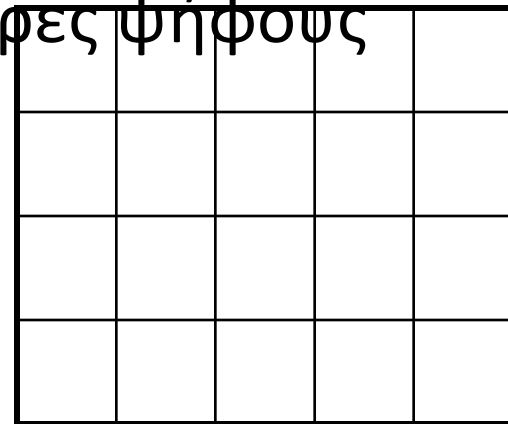
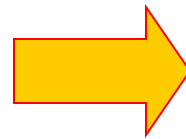
Μετασχηματισμός Hough

- Γενικά:

- Διακριτοποίηση παραμετρικού χώρου σε κελιά
- Για κάθε χαρακτηριστικό σημείο της εικόνας, αύξηση ψήφων κάθε κελιού κατά ένα, εάν θα μπορούσε να είχε δημιουργήσει το σημείο αυτό
- Εύρεση κελιών με τις περισσότερες ψήφους



Εικόνα



Παραμετρικός χώρος Hough

P.V.C. Hough, *Machine Analysis of Bubble Chamber Pictures*,
Proc. Int. Conf. High Energy Accelerators and
Instrumentation, 1959



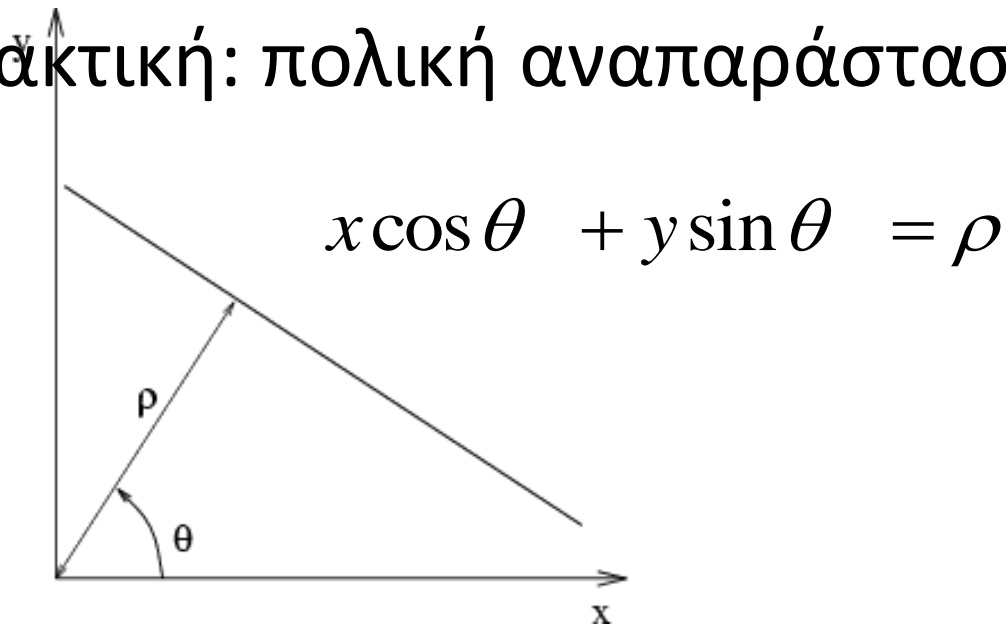
Αναπαράσταση παραμετρικού χώρου

- Προβλήματα με το χώρο (m,b) :
 - Μη φραγμένος παραμετρικός χώρος
 - Οι κάθετες ευθείες απαιτούν άπειρη τιμή για το m



Αναπαράσταση παραμετρικού χώρου

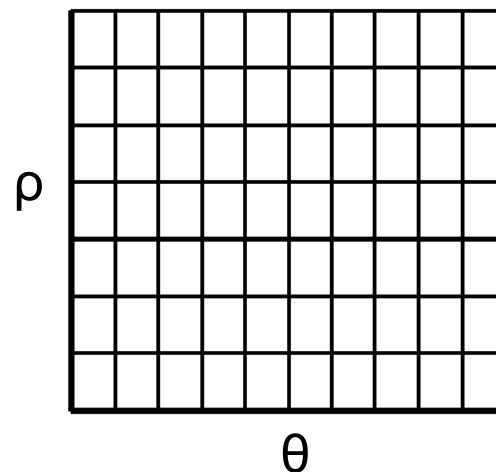
- Προβλήματα με το χώρο (m,b) :
 - Μη φραγμένος παραμετρικός χώρος
 - Οι κάθετες ευθείες απαιτούν άπειρη τιμή για το m
- Εναλλακτική: πολική αναπαράσταση



Κάθε σημείο προσθέτει ένα ημίτονο στον παραμετρικό χώρο (θ, ρ)

Περιγραφή αλγορίθμου

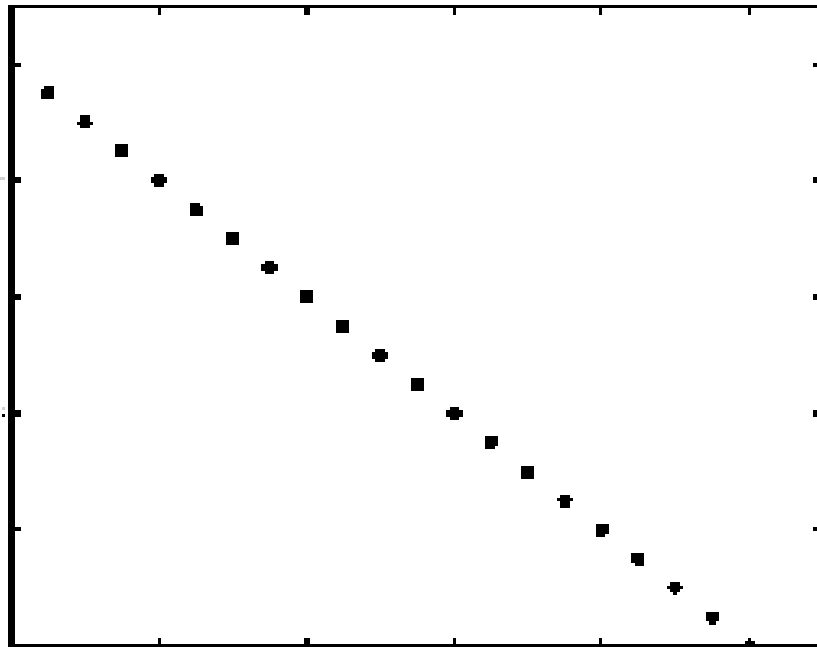
H: accumulator array (votes)



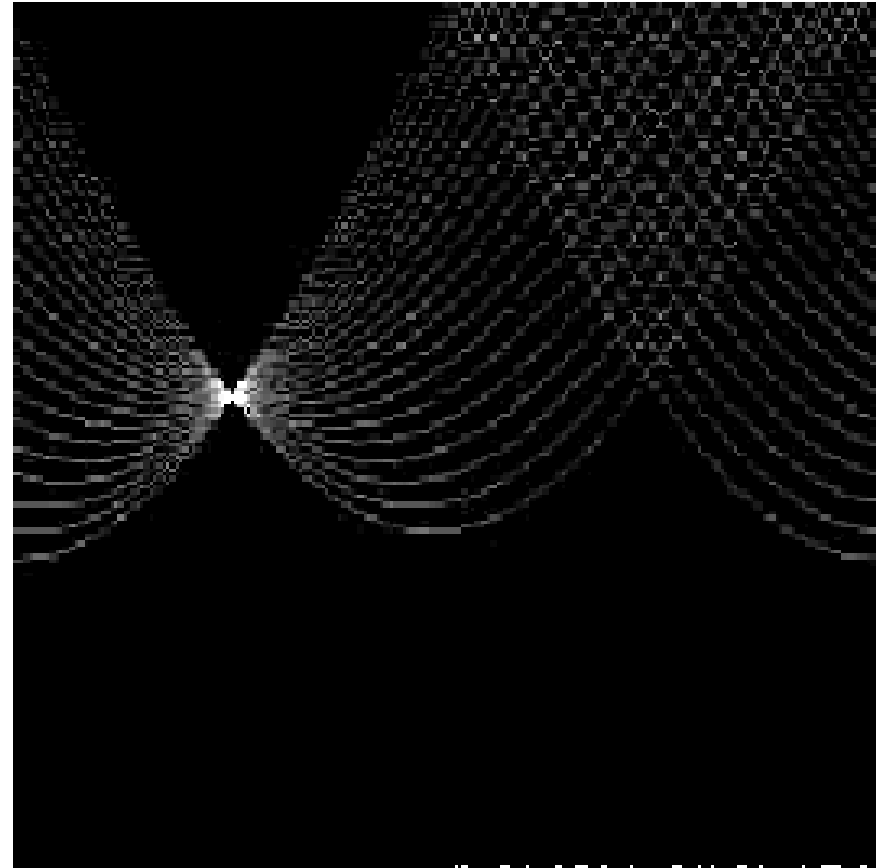
- Αρχικοποίηση συσσωρευτή H στο μηδέν
- for κάθε σημείο (x,y) της εικόνας
 - for $\theta = 0$ to 180
 - $\rho = x \cos \theta + y \sin \theta$
 - $H(\theta, \rho) = H(\theta, \rho) + 1$
- end
- end
- Εύρεση των τιμών του (θ, ρ) όπου ο $H(\theta, \rho)$ παρουσιάζει τοπικό μέγιστο
 - Η ανιχνευθείσα ευθεία στην εικόνα δίνεται από τον εξής τύπο
 - $\rho = x \cos \theta + y \sin \theta$



Απεικόνιση



χαρακτηριστικά

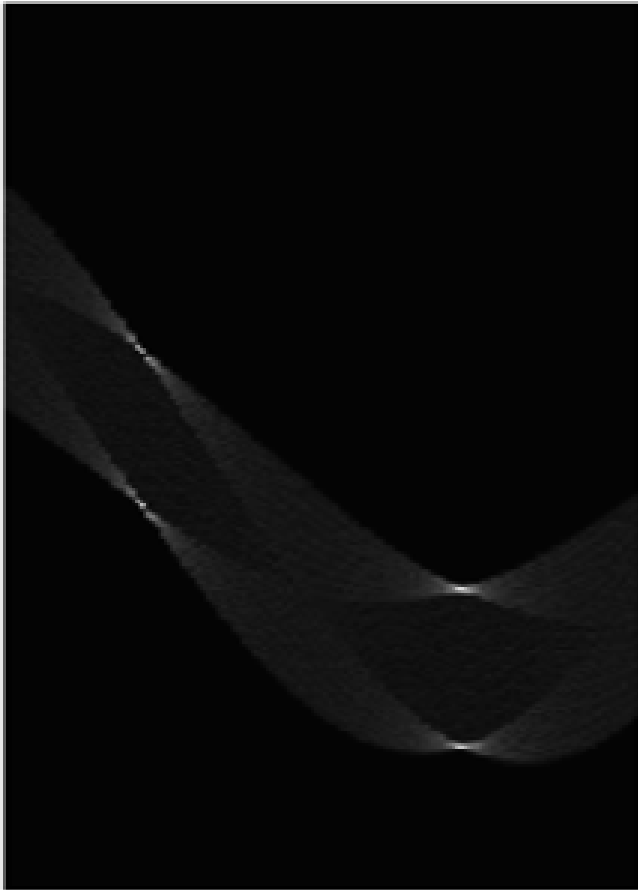


ψήφοι

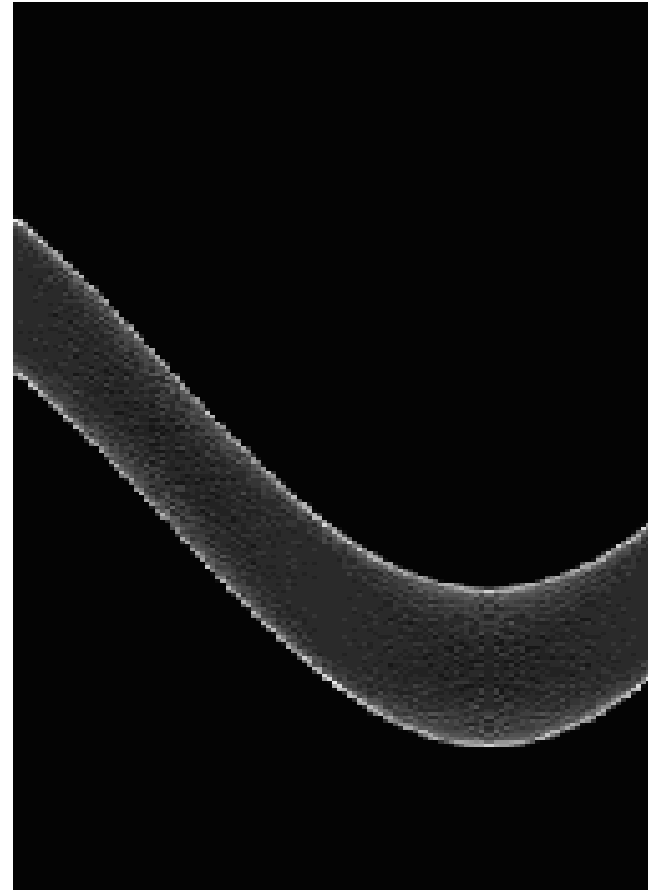


Άλλα σχήματα

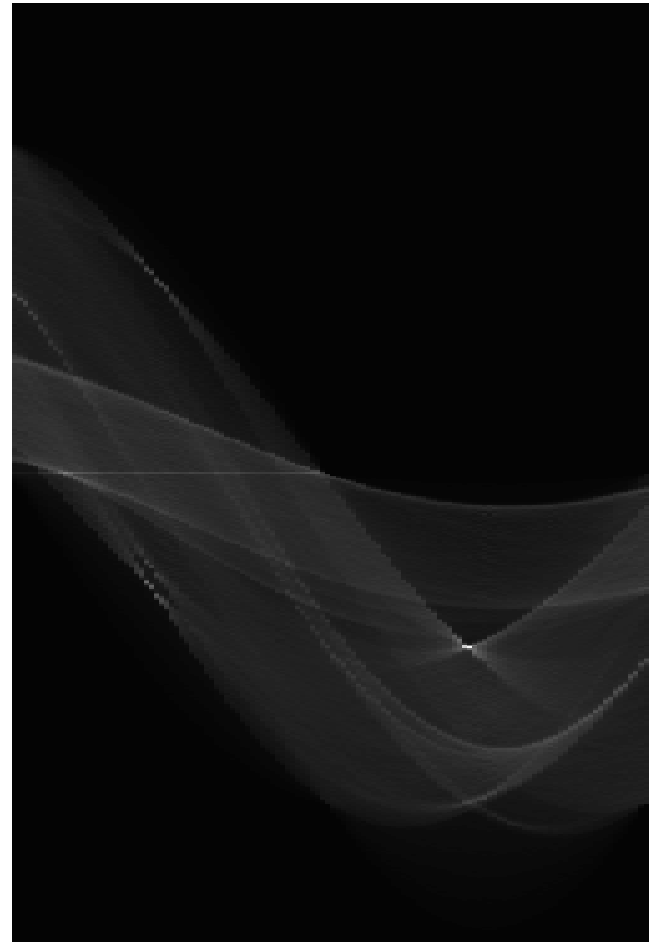
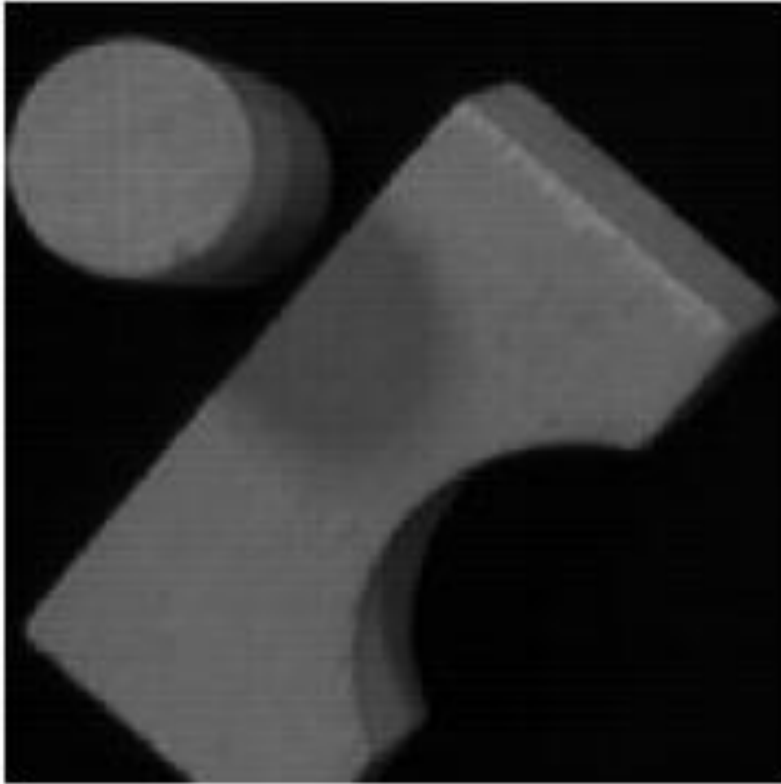
Τετράγωνο



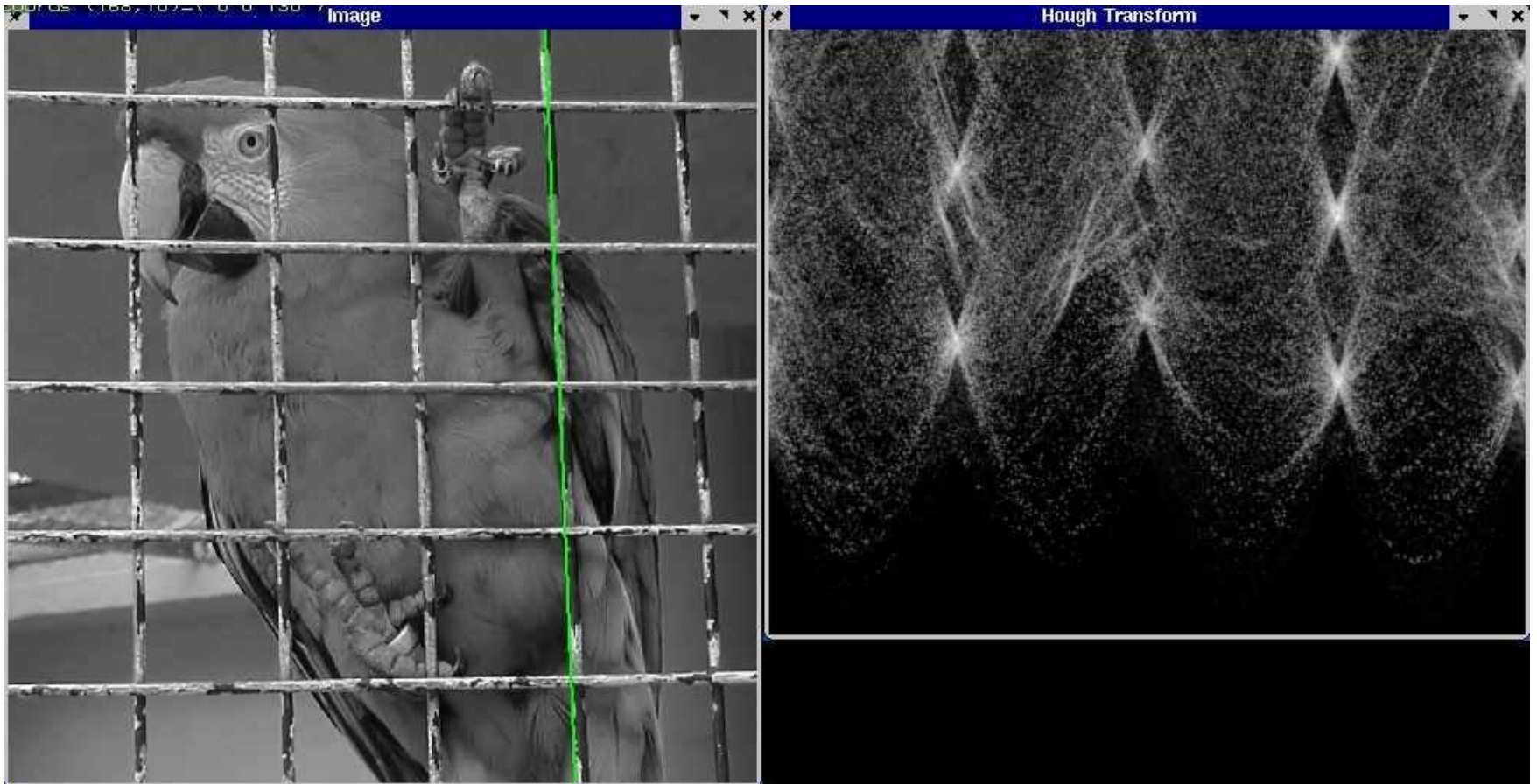
Κύκλος



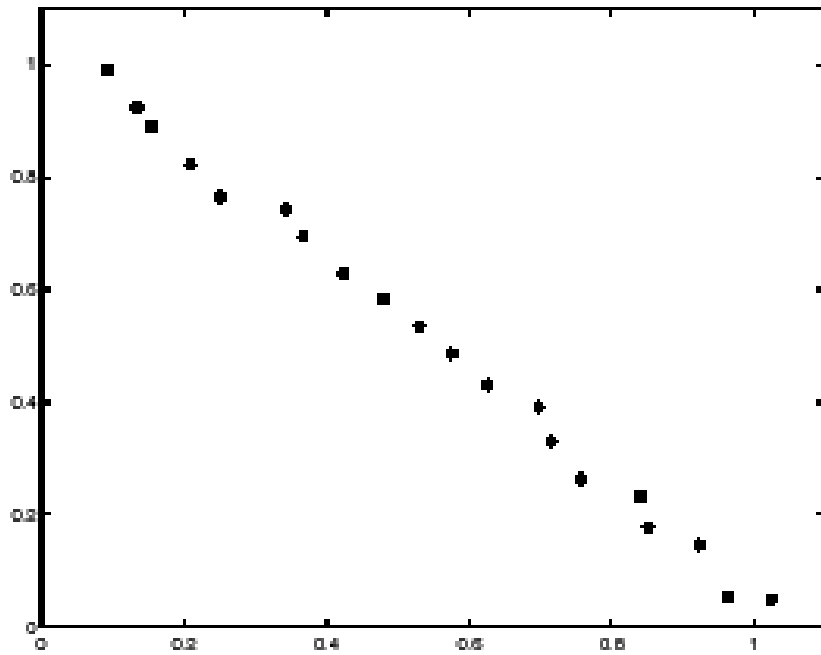
Πολλές ευθείες



Περίπλοκη εικόνα



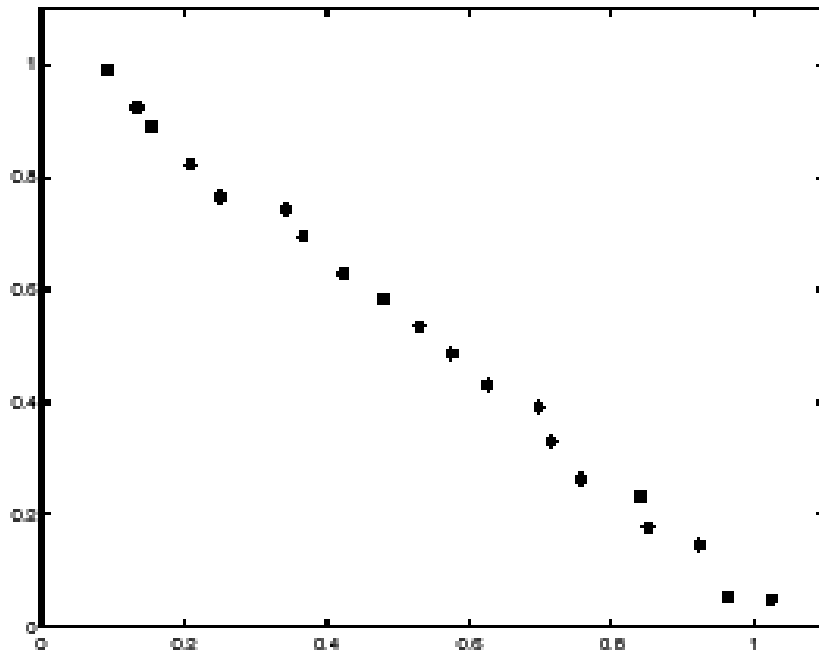
Αποτέλεσμα Θορύβου



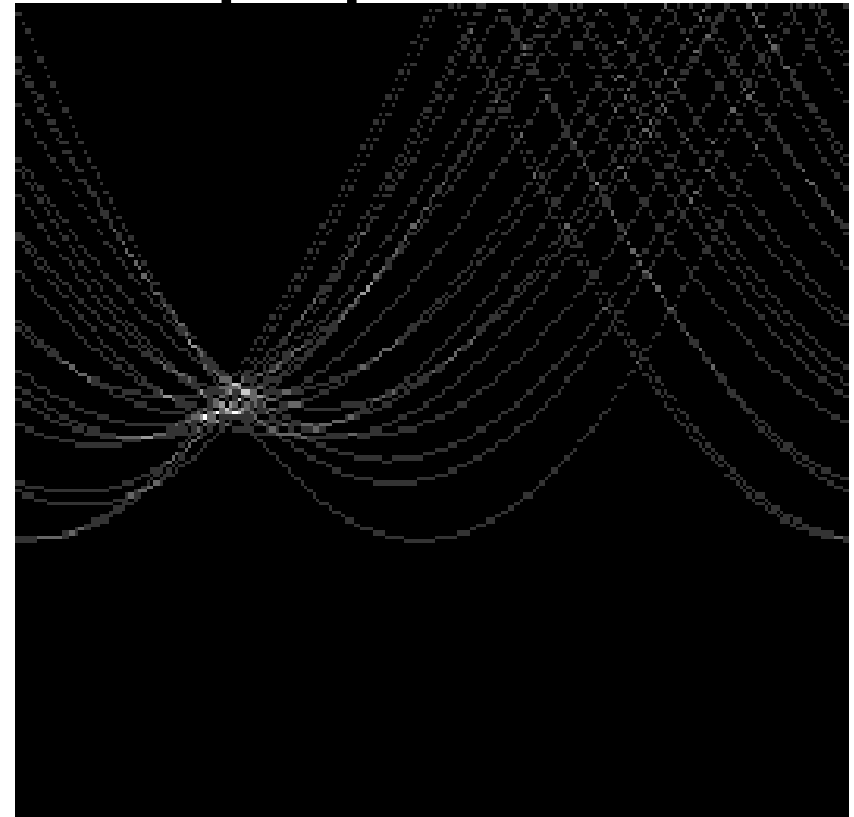
χαρακτηριστικά



Αποτέλεσμα θορύβου



χαρακτηριστικά

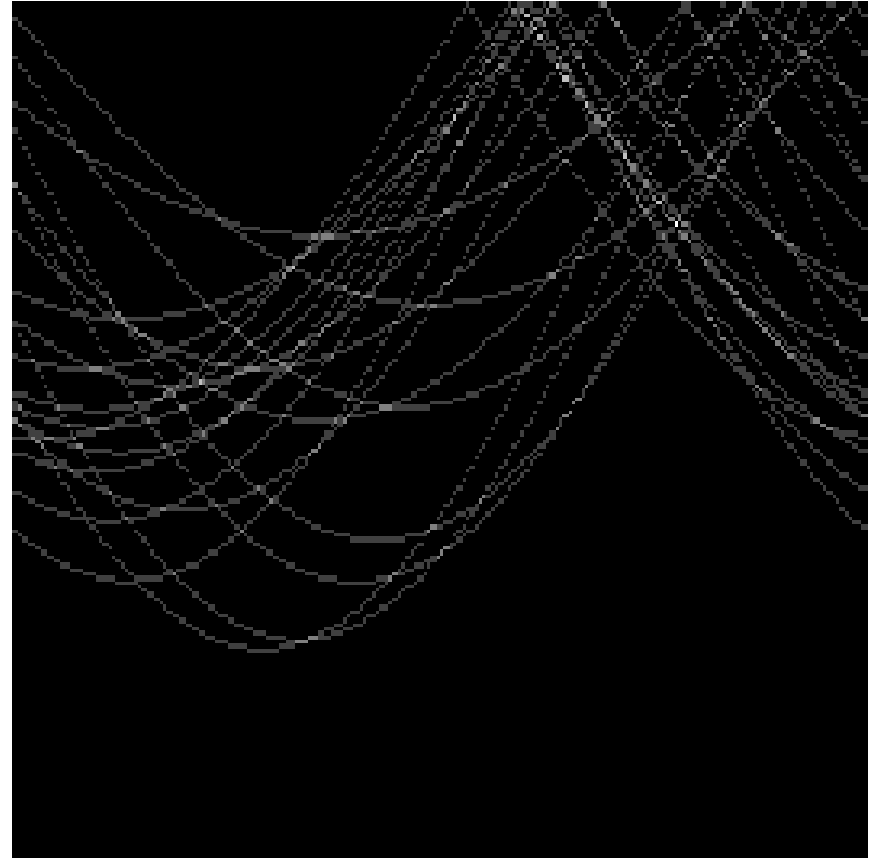
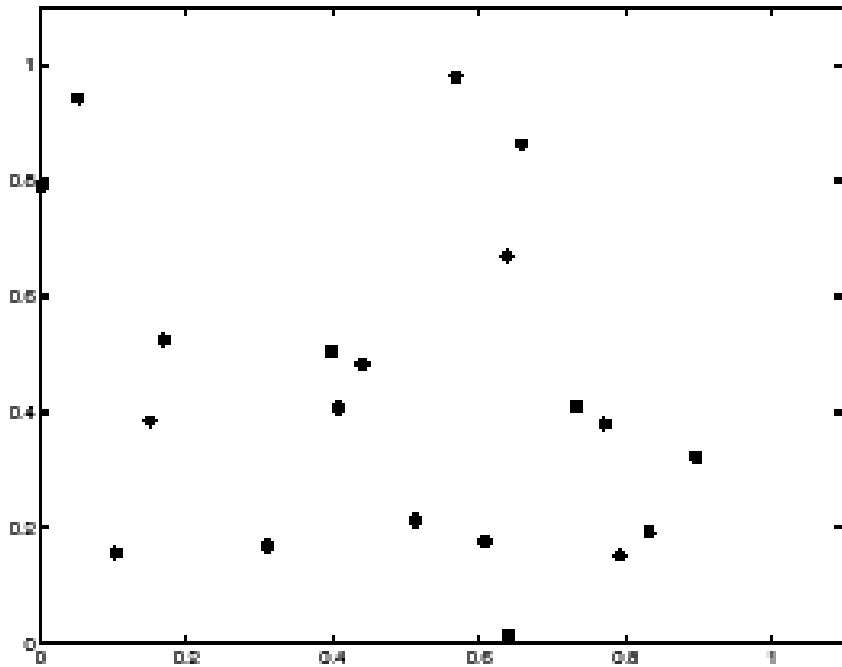


ψήφοι

- Ο εντοπισμός του μεγίστου γίνεται δύσκολος



Τυχαία σημεία



χαρακτηριστικά

ψήφοι

Ομοιόμορφος θόρυβος μπορεί να οδηγήσει σε τυχαία μέγιστα στον πίνακα H



Χειρισμός θορύβου

- Επιλογή ενός «σωστού» πλέγματος/διακριτοποίηση
 - Πολύ αραιό: πολλές ψήφοι παράγονται όταν πολλές διαφορετικές ευθείες αντιστοιχούν στο ίδιο κελί
 - Πολύ πυκνό: χάνονται ευθείες γιατί πολλά σημεία που δεν είναι ακριβώς συνευθειακά «ψηφίζουν» για διαφορετικά κελιά
- Συσχέτιση με γειτονικά κελιά (εξομάλυνση κελιών του συσσωρευτή)
- Αποφυγή χρήσης άσχετων χαρακτηριστικών
 - Χρήση μόνο των σημείων ακμών με σημαντικό πλάτος της κλίσης



Αναγνώριση αντικειμένων



Πόσες κατηγορίες αντικειμένων υπάρχουν;





~10,000 με 30,000



Αντικείμενα

Ζώα

Φυτά

Άψυχα

.....

Σπονδυλωτά

Φυσικά

Τεχνητά

Θυλαστικά

Πουλιά

Τίγρη

Σκύλος

Πέρδικα

Κάμερα



Corbis.com

Εργασίες αναγνώρισης



Κατηγοριοποίηση σκηνής

- Εξωτερική/εσωτερική
- Πόλη/δάσος...



Περιγραφή εικόνας / μεταδεδομένα

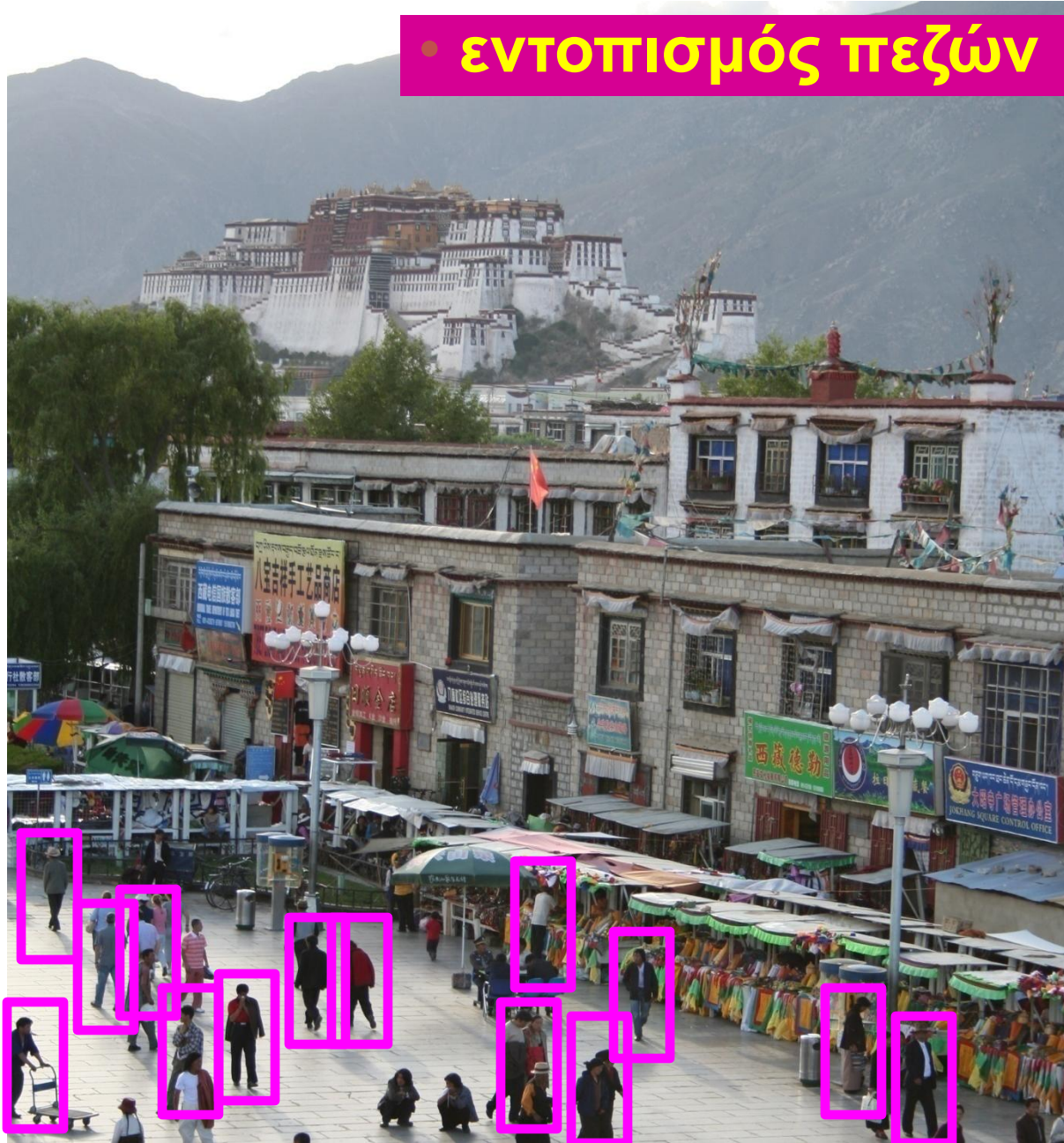


- δρόμος
- άνθρωποι
- κτίρια
- βουνό
- τουρισμός
- σύνεφα
- τούβλα
- ...



Αναγνώριση αντικειμένων

• ΕΝΤΟΠΙΣΜΟΣ ΠΕΖΩΝ



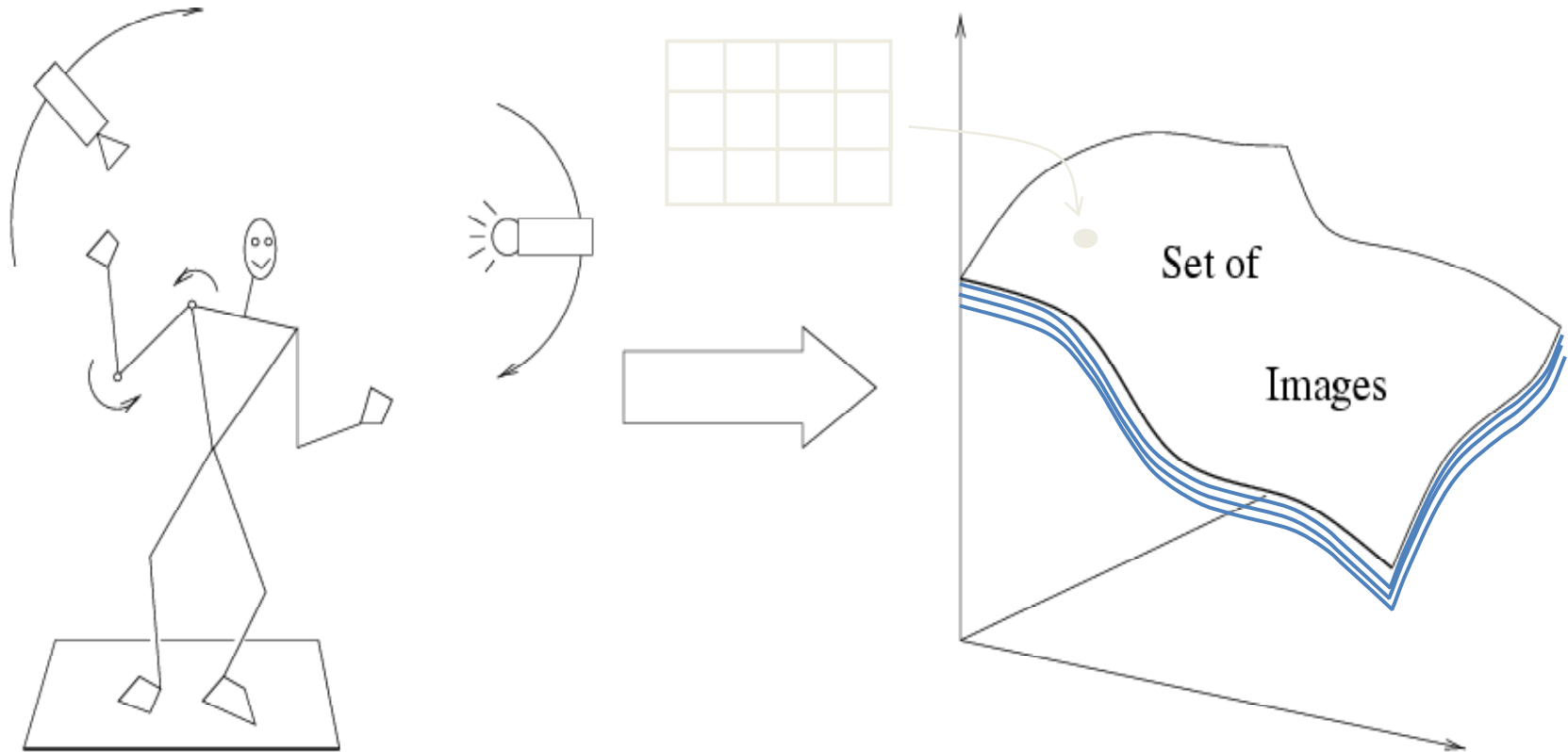
Αντικείμενα εικόνας



Κατανόηση σκηνής;



Μοντελοποίηση μεταβλητότητας



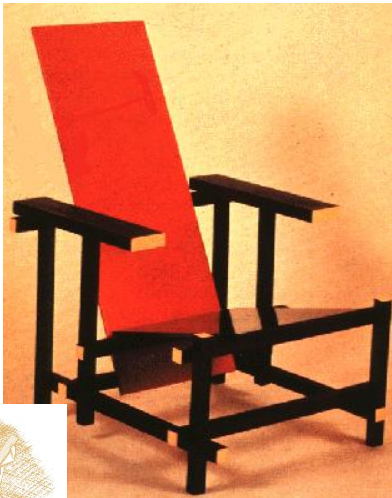
Μεταβλητότητα: Θέση κάμερας
Φωτισμός
Σχήμα



Αυτό-μεταβλητότητα;



Αυτο-μεταβλητότητα



Κατηγοριοποίηση εικόνας

Εκπαίδευση

Ground
truth

Εικόνες
εκπαίδευσης

Χαρακτηριστικά
Εικόνας

Εκπαίδευση
ταξινομητή

Ταξινο-
μητής



Κατηγοριοποίηση εικόνας

Εκπαίδευση

Ground truth

Εικόνες
εκπαίδευσης



Χαρακτηριστικά
Εικόνας

Εκπαίδευση
ταξινομητή

Ταξινομητής

Έλεγχος

Χαρακτηριστικά
Εικόνας

Ταξινομητής

Πρόβλεψη
Εξωτερικός



Εικόνα ελέγχου



Χαρακτηριστικά εικόνας

Εκπαίδευση

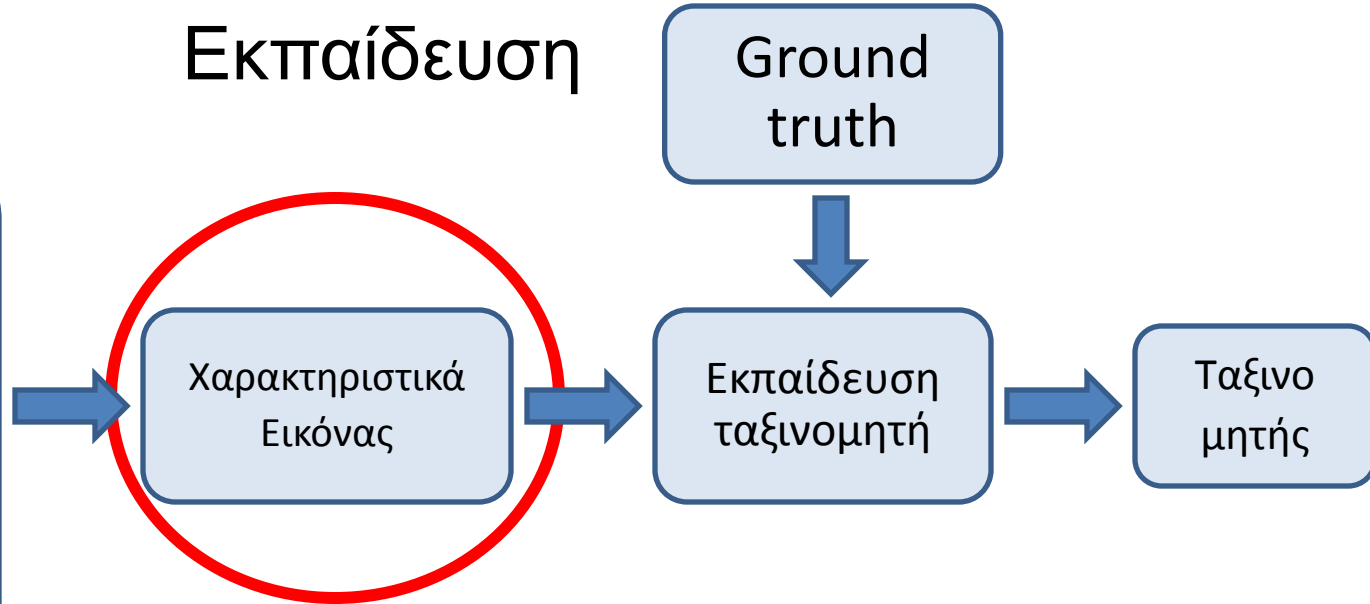
Ground truth

Εικόνες
εκπαίδευσης

Χαρακτηριστικά
Εικόνας

Εκπαίδευση
ταξινομητή

Ταξινομητής



Αναπαράσταση εικόνας

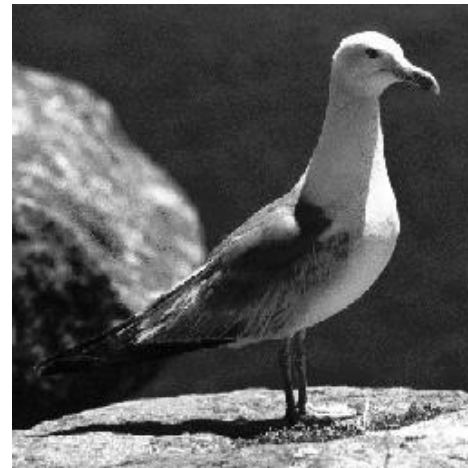
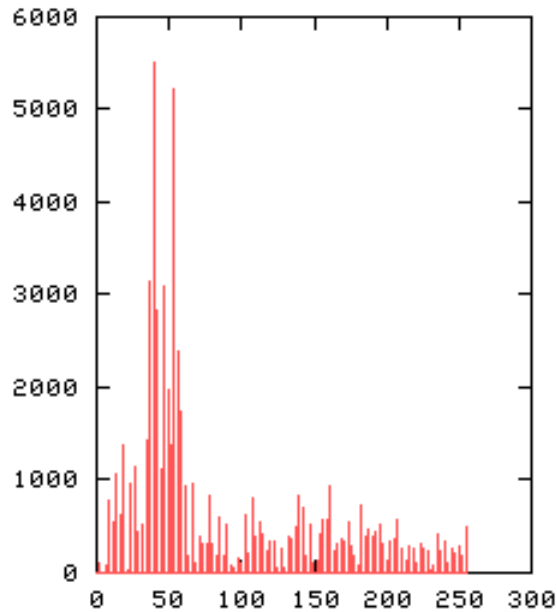
- Πρότυπα
 - Φωτεινότητα, κλίσεις, κλπ.



- Ιστογράμματα
 - Χρώμα, υφή, περιγραφείς SIFT, κλπ.



Αναπαράσταση εικόνας : Ιστόγραμμα



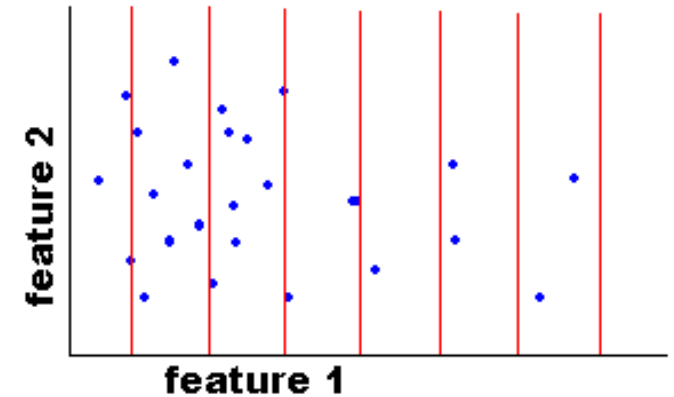
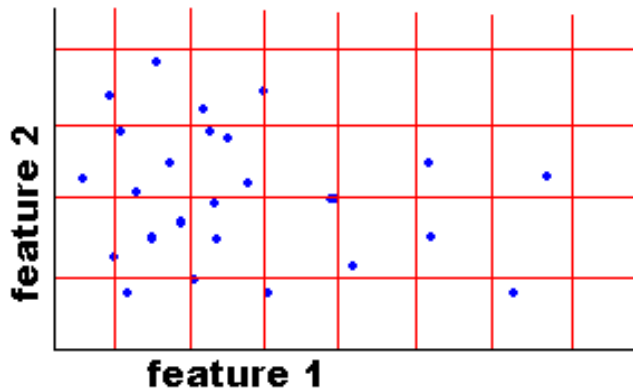
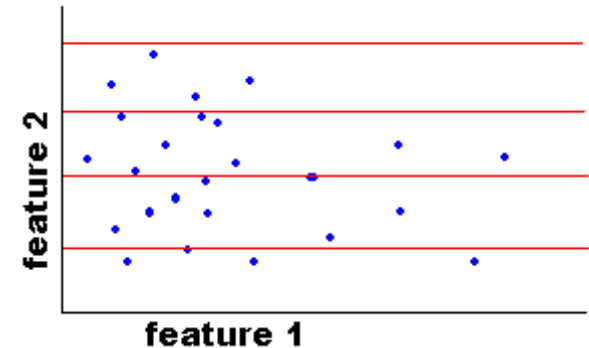
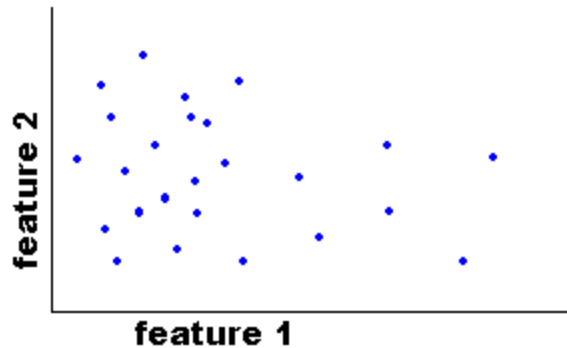
Ολικό ιστόγραμμα

- Αναπαριστά κατανομή χαρακτηριστικών
 - Χρώμα, υφή, βάθος, ...



Αναπαράσταση εικόνας : Ιστόγραμμα

Ιστόγραμμα: Πιθανότητα ή πλήθος σημείων σε κάθε κελί

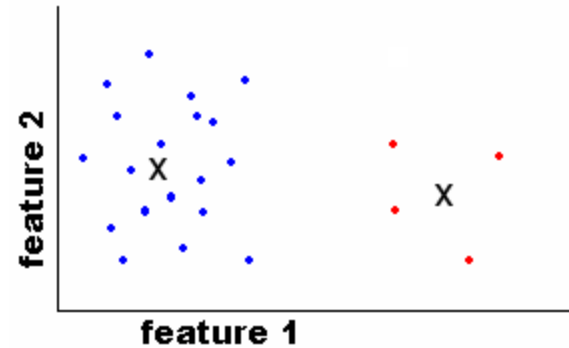
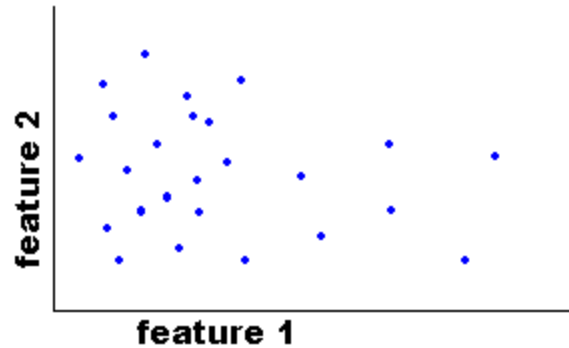


- Κοινό ιστόγραμμα
 - Απαιτεί πολλά δεδομένα
 - Απώλεια ανάλυσης για να μην υπάρχουν κενά κελιά



Αναπαράσταση εικόνας : Ιστόγραμμα

Ομαδοποίηση



Χρήση των ίδιων κέντρων ομάδων για όλες τις εικόνες



Αποστάσεις ιστογραμμάτων

$$\text{histint}(h_i, h_j) = 1 - \sum_{m=1}^K \min(h_i(m), h_j(m))$$

Τομή ιστογραμμάτων (απαιτεί κανονικοποίηση)

$$\chi^2(h_i, h_j) = \frac{1}{2} \sum_{m=1}^K \frac{[h_i(m) - h_j(m)]^2}{h_i(m) + h_j(m)}$$

Chi-squared απόσταση ιστογραμμάτων



Εκτός από όμοιο χρώμα
τι άλλο παρατηρείτε;;;

Αυτοκίνητα που βρέθηκαν με ιστογράμματα χρώματος και απόσταση chi-sq



Ιστογράμματα: Υλοποίηση

- Διακριτοποίηση/κβαντισμός
 - Πλέγμα: γρήγορο αλλά εφαρμόσιμο μόνο σε λίγες διαστάσεις
 - Ομαδοποίηση: αργό αλλά κατάλληλο και για πολλές διαστάσεις



Λίγα κελιά

Λίγα δεδομένα

Αραιή αναπαράσταση

- Απόσταση/Matching

- Η τομή ιστογραμμάτων ή η ευκλείδεια απόσταση είναι γρήγορες μέθοδοι
- Η Chi-squared πολλές φορές αποδίδει καλύτερα
- Η απόσταση Earth mover's είναι αποτελεσματική όταν γειτονικά κελιά αφορούν όμοιες τιμές (παράδειγμα)

Πολλά κελιά

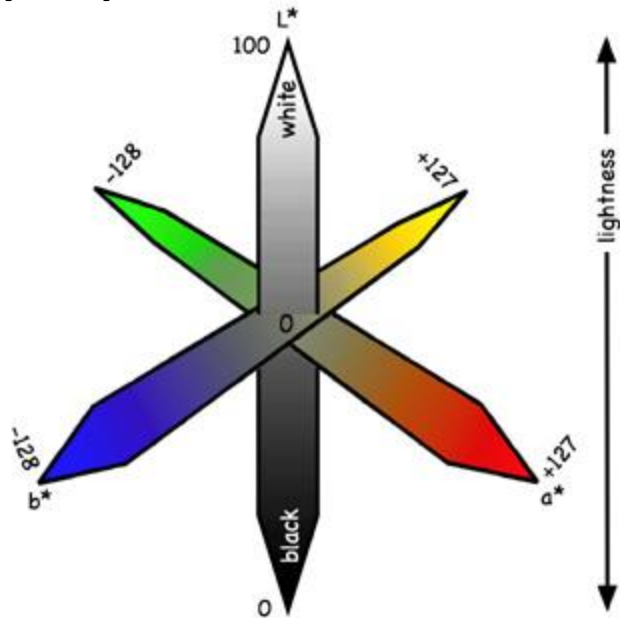
Πολλά δεδομένα

Λεπτομέρεια

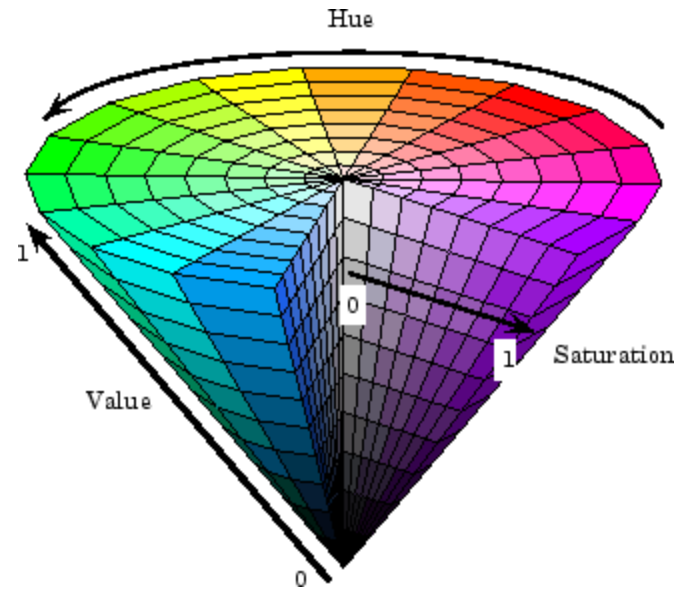


Ιστογράμματα σε τι;

- Χρώμα



L*a*b* χρωματικός χώρος



HSV χρωματικός χώρος

- Υφή (τράπεζες φίλτρων ή κλίσεις σε περιοχές)



Ιστογράμματα σε τι;

- Ιστογράμματα προσανατολισμένων κλίσεων

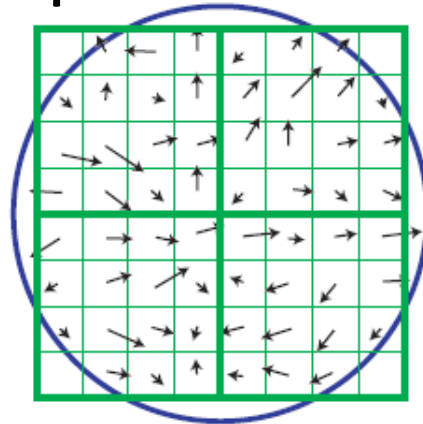
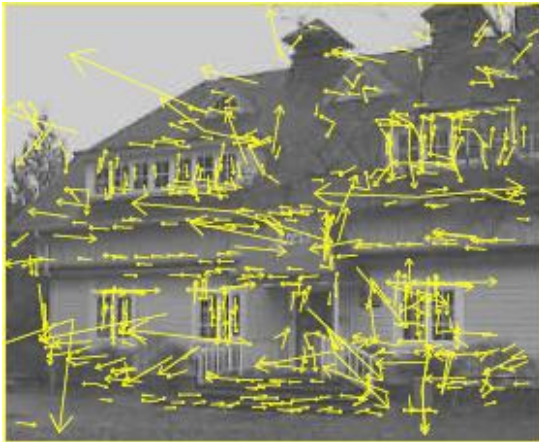
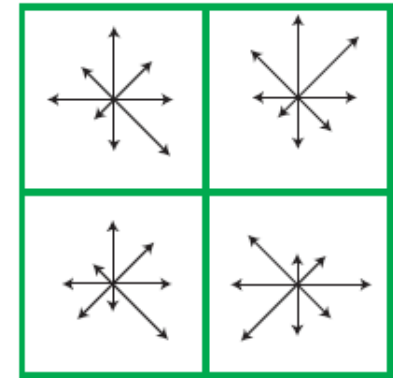


Image gradients



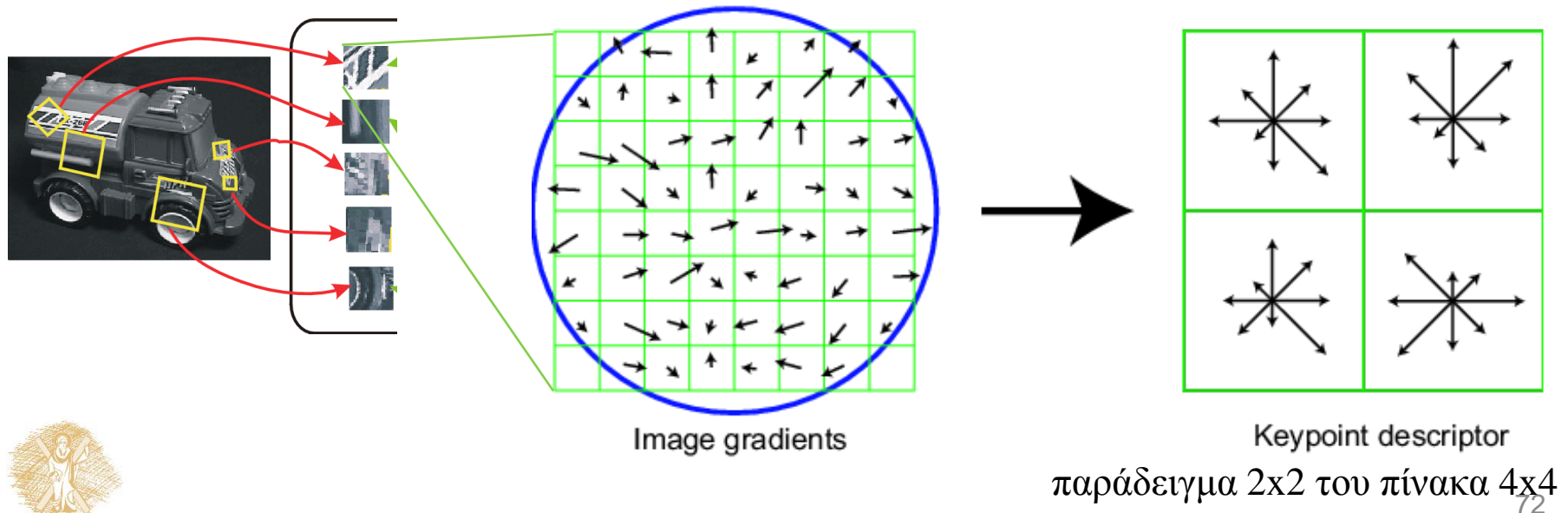
Keypoint descriptor

SIFT – Lowe IJCV 2004



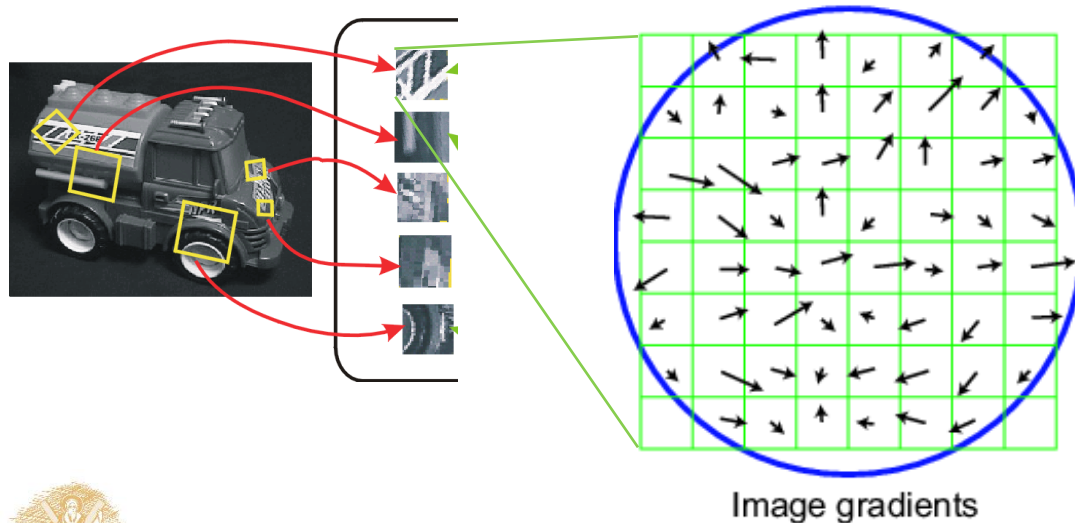
Διάνυσμα SIFT

- 4x4 πίνακας ιστογραμμάτων προσανατολισμού κλίσεων
 - Όχι ακριβώς ιστόγραμμα, στάθμιση βάσει πλάτους
- 8 κατευθύνσεις x 4x4 πίνακας = 128 διαστάσεις
- Κίνητρο: μικρή ευαισθησία σε θέση/ προσανατολισμό



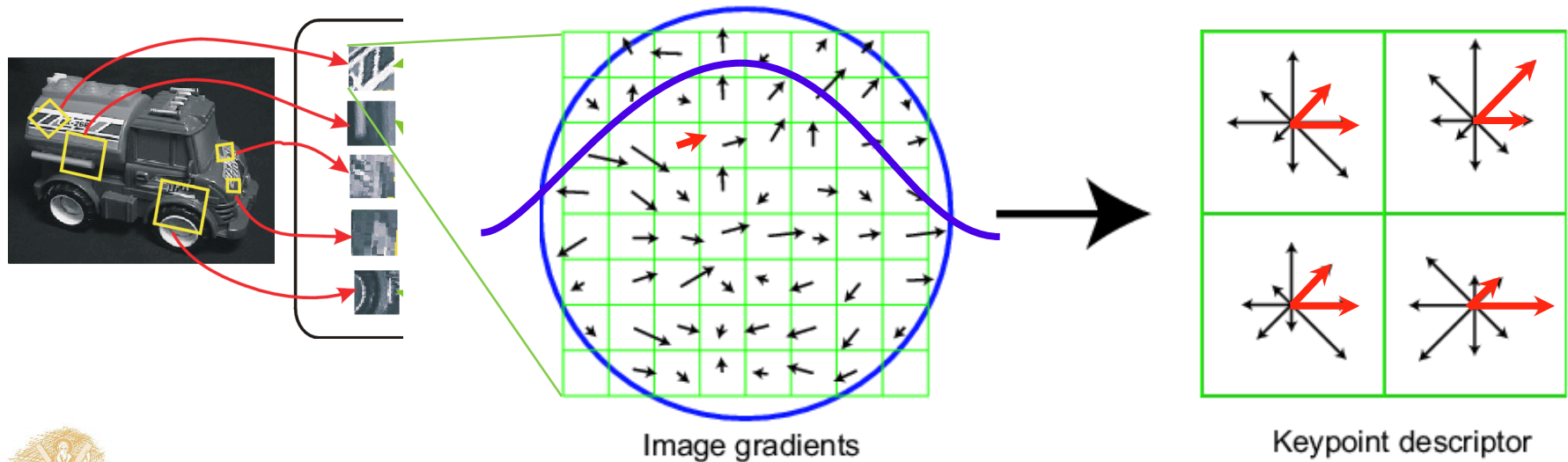
Διάνυσμα SIFT

- Υπολογίζεται σε περιστραμμένες και κλιμακωμένες εκδόσεις του παραθύρου ανάλογα με τον υπολογισμό προσανατολισμού και κλίμακας
 - Επαναδειγματοληψία του παραθύρου
- Βάσει κλίσεων σταθμισμένων μέσω Gaussian παραθύρου



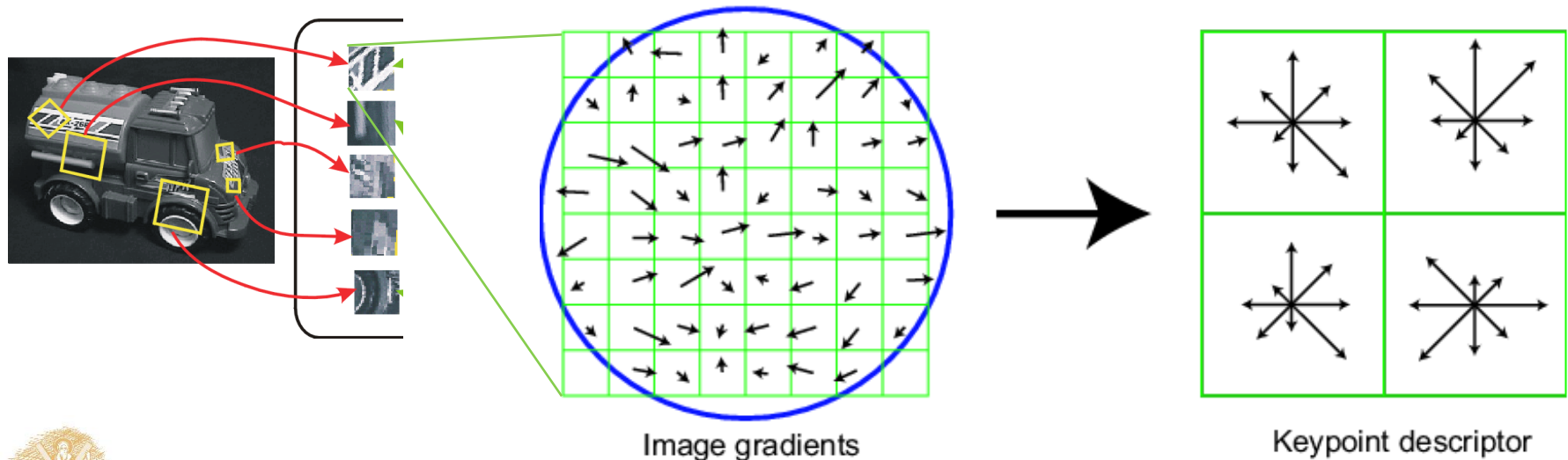
Ομαλότητα

- Gaussian στάθμιση
- Τρι-γραμμική παρεμβολή

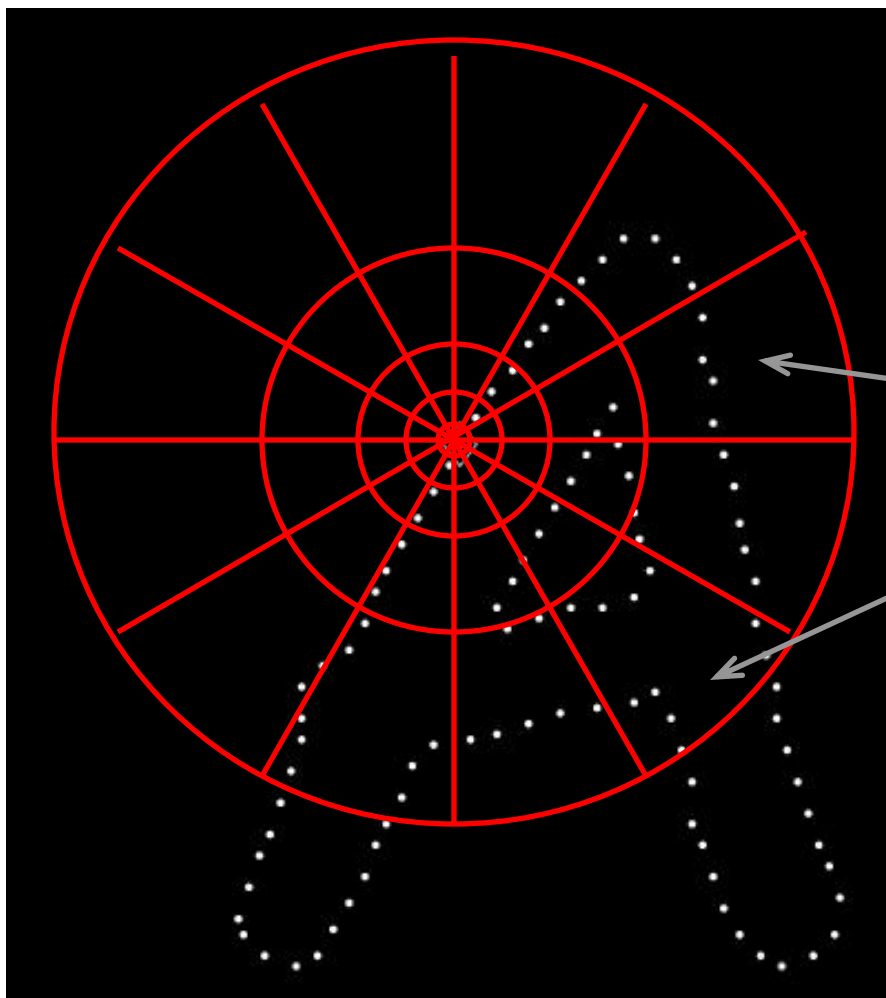


Μείωση επίδρασης φωτισμού

- 128-διάστατο διάνυσμα κανονικοποιημένο στο 1
- Φράγμα στο πλάτος της κλίσης για την αποφυγή μεγάλης επιροής μεγάλων κλίσεων



Τοπικοί περιγραφείς



Αρίθμηση του αριθμού σημείων σε κάθε κελί, π.χ.:

Count = 4

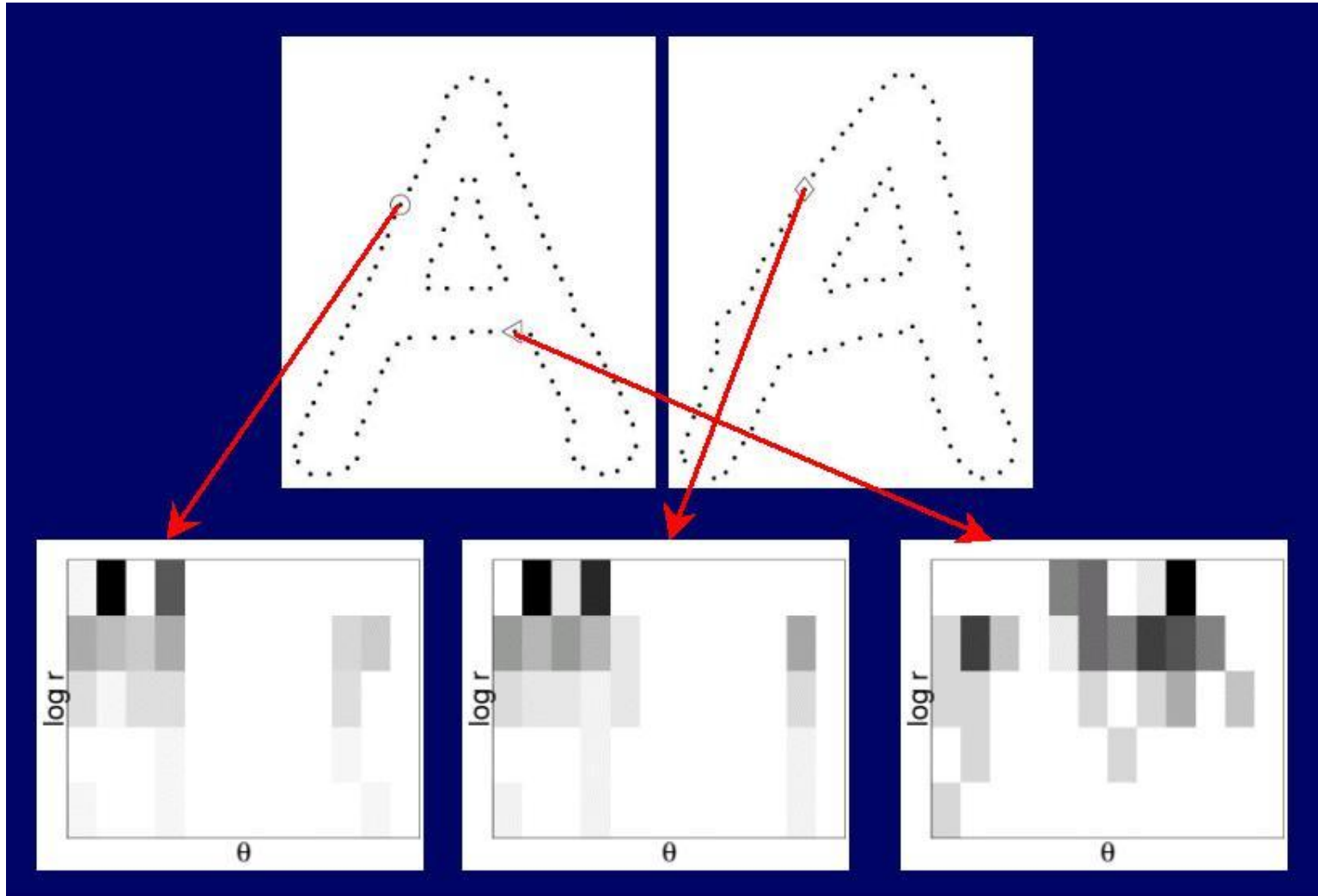
⋮

Count = 10

Log-polar διακριτοποίηση:
μεγάλη ακρίβεια σε
γειτονικά σημεία, ευελιξία
στα απομακρυσμένα.



Τοπικοί περιγραφείς: Σχήμα



Τέλος Ενότητας

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αθηνών**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



ΣΗΜΕΙΩΜΑΤΑ

Σημείωμα Ιστορικού Εκδόσεων Έργου

Το παρόν έργο αποτελεί την έκδοση **1.0**.

Έχουν προηγηθεί οι κάτωθι εκδόσεις:

- Έκδοση **1.0**. διαθέσιμη [εδώ](#).



Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών, Φακωτάκης Νίκος, Μουστάκας Κωνσταντίνος, Πέππας Παύλος, Σγάρμπας Κυριάκος. «Τεχνητή Νοημοσύνη II, Αντίληψη». Έκδοση: 1.0. Πάτρα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/modules/course_metadata/opencourses.php?fc=15



Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά, Μη Εμπορική Χρήση Παρόμοια Διανομή 4.0 [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



[1] <http://creativecommons.org/licenses/by-nc-sa/4.0/>

Ως **Μη Εμπορική** ορίζεται η χρήση:

- που δεν περιλαμβάνει άμεσο ή έμμεσο οικονομικό όφελος από την χρήση του έργου, για το διανομέα του έργου και αδειοδόχο
- που δεν περιλαμβάνει οικονομική συναλλαγή ως προϋπόθεση για τη χρήση ή πρόσβαση στο έργο
- που δεν προσπορίζει στο διανομέα του έργου και αδειοδόχο έμμεσο οικονομικό όφελος (π.χ. διαφημίσεις) από την προβολή του έργου σε διαδικτυακό τόπο

Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.



Σημείωμα Χρήσης Έργων Τρίτων

Το Έργο αυτό κάνει χρήση των ακόλουθων έργων:

Εικόνες/Σχήματα/Διαγράμματα/Φωτογραφίες

- Η διάλεξη είναι εμπνευσμένη και έχει χρησιμοποιηθεί εκπαιδευτικό και πολυμεσικό υλικό από τις διαλέξεις του James Hays του «Brown University, Computer Science Department» και του αντίστοιχου μαθήματος «CS 143 Introduction to Computer Vision» (<http://cs.brown.edu/courses/cs143/>).

