2.5 Ο μετασχηματισμός του Hough

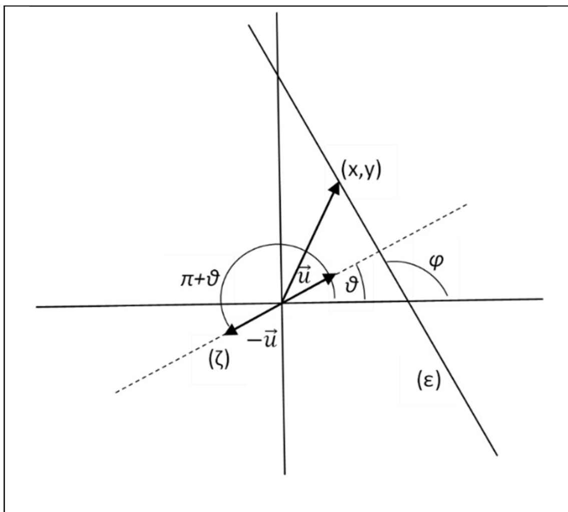
Συχνά οι ψηφιακές εικόνες περιέχουν γραμμές που ανήκουν σε σχήματα, τεχνικά σχέδια, γραφήματα, σειρές κειμένου ή άλλες αναπαραστάσεις. Σε πολλές εφαρμογές είναι επιθυμητή η εύρεση της θέσης και η αναγνώριση της μορφής των γραμμών (π.χ. ευθύγραμμα τμήματα, τόξα). Τέτοιες εφαρμογές είναι: η διανυσματική κωδικοποίηση τυπωμένων τεχνικών σχεδίων που ψηφιοποιήθηκαν από σαρωτές (scanners), η εύρεση περιοχών κειμένου σε έγγραφα, ο προσδιορισμός της υφής, βιομηχανικές εφαρμογές κατασκευών και ρομποτικής κ.α. Ο μετασχηματισμός του Hough (HT: Hough Transform) είναι μία μέθοδος για την εύρεση ευθειών σε μια εικόνα και αποτελεί την βάση τεχνικών για τον προσδιορισμό ευθυγράμμων τμημάτων, καμπυλών και μορφών που αναλύονται ανάλογα.

Ο HT προτάθηκε από τον Paul Hough το 1962 ως μέρος της κατασκευής μιας συσκευής ανίχνευσης της κίνησης σωματιδίων υψηλής ενέργειας και στόχευε στην αυτοματοποίηση και αντικατάσταση της οπτικής διαδικασίας που απαιτούσε πολλές ανθρωποώρες. Ο αρχικός αλγόριθμος εξελίχθηκε και η σημερινή διατύπωση του είναι η ακόλουθη.

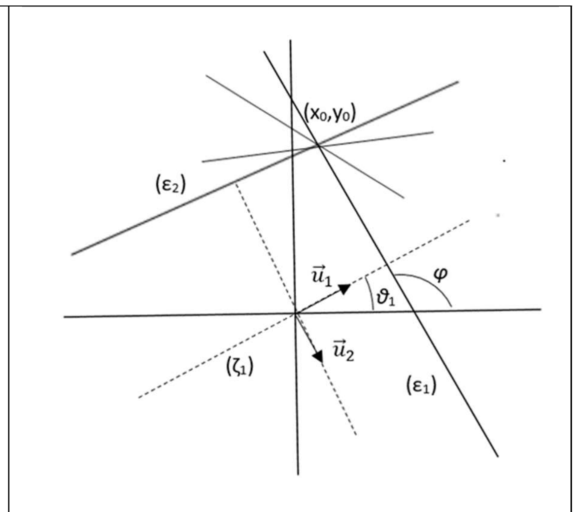
Κάθε ευθεία (ε) του καρτεσιανού επιπέδου περιγράφεται από την πολική της εξίσωση

$$\rho = x \text{ συν}\theta + y \text{ ημ}\theta$$

όπου (x,y) σημείο της ευθείας και ρ, θ οι πολικές της παράμετροι (Σχ.2.5.1).



Σχ.2.5.1.α



Σχ.2.5.1.β

Η σχέση προκύπτει ως ακολούθως. Έστω η ευθεία (ζ) που διέρχεται από το κέντρο των αξόνων και είναι κάθετος στην (ε) και τα μοναδιαία διανύσματα επάνω σ' αυτήν $\vec{u}$ και $-\vec{u}$.
Το ρ είναι η προσημασμένη προβολή του σημείου (x,y) στο $\vec{u}$ ή το $-\vec{u}$ που σχηματίζουν ευθεία γωνία π μεταξύ τους και ισούται αντίστοιχα:

$$\rho(\theta) = x \text{ συν}\theta + y \text{ ημ}\theta \text{ ή}$$

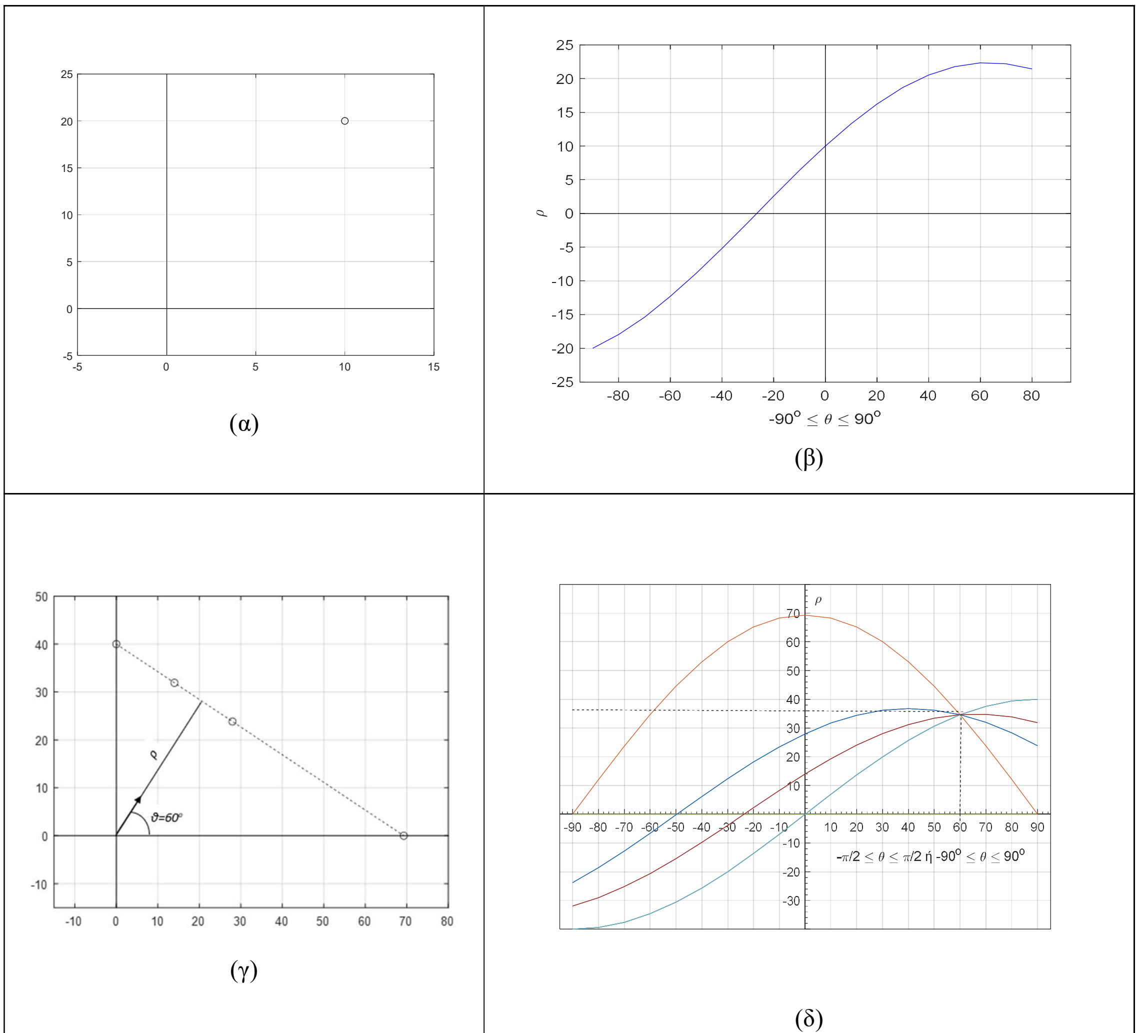
$$\rho(\pi+\theta) = x\ \sigma\upsilon\nu(\pi+\theta) + y\ \eta\mu(\pi+\theta) = -\rho(\theta)$$

Δηλαδή η προβολή ρ έχει πρόσημο ανάλογα με την κατεύθυνση του κάθετου μοναδιαίου διανύσματος. Αν η ευθεία (ε) διέρχεται από το κέντρο των αξόνων το $\rho=0$, εάν το μοναδιαίο διάνυσμα δείχνει την ευθεία το $\rho>0$ και εάν δείχνει αντίθετα $\rho<0$. Για δεδομένο σημείο του επιπέδου (x_0, y_0) οι εξισώσεις των ευθειών που διέρχονται από αυτό θα ικανοποιούν την σχέση

$$\rho(\theta) = x_0\ \sigma\upsilon\nu\theta + y_0\ \eta\mu\theta$$

Αν η γωνία φ της κλίσης μιας ευθείας (Σχ.2.5.1.β) παίρνει τιμές στο διάστημα $[0, \pi)$ μπορούν να περιγραφούν όλες οι ευθείες του επιπέδου. Επειδή ισχύει προφανώς ότι $\varphi=\pi/2+\theta$, αρκεί η γωνία θ του μοναδιαίου διανυσμάτος να παίρνει τιμές στο διάστημα $[-\pi/2,\pi/2)$. Σ' αυτήν την περίπτωση η τιμή του ρ θα είναι αρνητική για ευθείες όπως η ευθεία (ε2) του Σχ.2.5.1.β

Στο ακόλουθο Σχ.2.5.2 (β) δείχνεται η γραφική παράσταση της συνάρτησης $\rho(\theta)$ για $-\pi/2\leq\theta<\pi/2$ των ευθειών που διέρχονται από το σημείο (10,20).

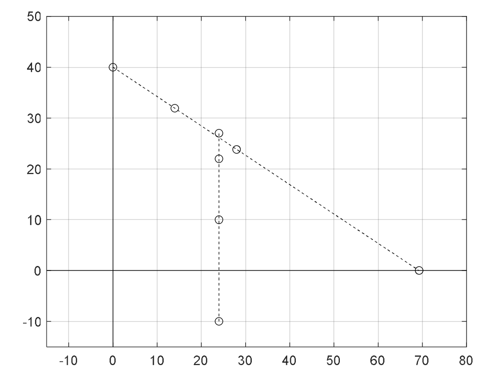


Σχήμα 2.5.2

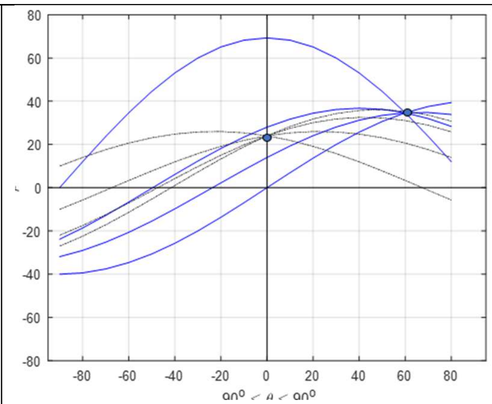
Οι τιμές ρ,θ αναπαρίστανται στους αντίστοιχους άξονες που ορίζουν το επίπεδο (ρ,θ) που αποκαλείται και πεδίο του Hough και οι τιμές αποτελούν τον μετασχηματισμό του Hough (Hough Transform, HT).

Στο Σχ.2.5.2.δ δείχνεται το πεδίο του HT που προκύπτει για τέσσερα συνευθειακά σημεία, Σχ. 2.5.2.γ.

Κάθε καμπύλη στο επίπεδο (ρ,θ) προκύπτει από ένα από τα τέσσερα σημεία του επιπέδου (x,y). Το σημείο (ρ,θ) της τομής των τεσσάρων καμπυλών στο πεδίο του Hough είναι το ρ και θ της ευθείας που διέρχεται από τα τέσσερα σημεία στο επίπεδο (x,y). Στο Σχ.2.5.3 βλέπουμε το πεδίο (x,y) με δύο ευθείες που τεσσάρων συνευθειακών σημείων και το πεδίο (ρ,θ) αυτών.



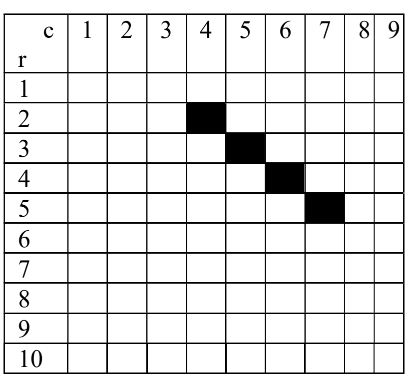
Σχ.2.5.3.α



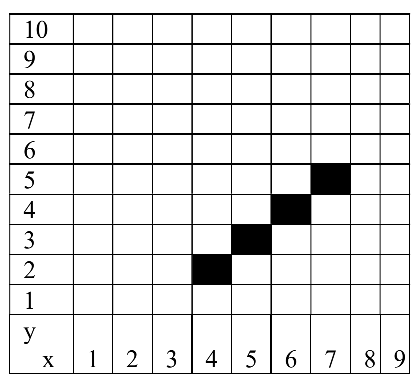
Σχ.2.5.3.β

Από τα παραπάνω συμπεραίνουμε πως το πρόβλημα μας ανάγεται στον εντοπισμό των σημείων τομής καμπυλών στο πεδίο του Hough (ρ,θ). Η αναλυτική επίλυση του προβλήματος είναι δύσκολη και η λύση δίνεται γεωμετρικά θεωρώντας ότι το επίπεδο (ρ,θ) σχεδιάζεται σε χαρτί μιλιμετρέ (πίνακας στην μνήμη του υπολογιστή). Έχουν προταθεί διαφορετικές προσεγγίσεις μιας τέτοιας επίλυσης που κατά βάση είναι παρόμοιες. Ακολούθως θα παρουσιάσουμε μία τέτοια ευρέως χρησιμοποιούμενη.

Η ψηφιακή (δυαδική) εικόνα περιγράφεται από ένα πίνακα R×C, και κάθε εικονοστοιχείο έχει θέση (r,c), r=1… R, c=1…C. Τοποθετούμε την εικόνα στο πρώτο τεταρτημόριο του επιπέδου (x,y), θέτωντας x=c και y=r, επειδή η αρίθμηση στο πίνακα της εικόνας αρχίζει από πάνω προς τα κάτω η εικόνα αναστρέφεται, Σχ.2.5.4. Θα εργασθούμε ακολούθως με την **ανεστραμμένη εικόνα**



Σχ.2.5.4.α: Αρχική εικόνα



Σχ.2.5.4.β: Ανεστραμμένη εικόνα

Η περιοχή του επιπέδου (θ,ρ) όπου θα απεικονισθούν οι καμπύλες ρ(θ) ορίζεται για

$$-90^o \leq \theta < 90^0 \text{ και}$$

$$-D \leq \rho < D,\ D=\sqrt{(R-1)^2+(C-1)^2}$$

Η περιοχή διακριτοποιείται με διαστήματα Δθ, και Δρ χωρίζοντας το διάστημα [-90,90) σε Ν τμήματα και το [-D,D) σε Μ. Θα ισχύει τότε

$$\Delta\theta = \frac{180^o}{N} \qquad (2.5.1)$$

$$\Delta\rho = \frac{2D}{M}, \qquad (2.5.2)$$

Τα Ν και Μ ορίζονται από τον χρήστη και καθορίζουν την ανάλυση στους άξονες θ και ρ. Με την διακριτοποίηση αυτή δημιουργούμε έναν πίνακα Η με διαστάσεις N × M και στοιχεία h(n,m) με
n=1…N και m=1…M. Σε κάθε τιμή του n αντιστοιχεί γωνία

$$\theta=-90^o+(n-1)\cdot\Delta\theta \qquad (2.5.3)$$

και σε κάθε τιμή του m, τιμή ρ

$$\rho=D-(m-1)\cdot\Delta\rho \qquad (2.5.4)$$

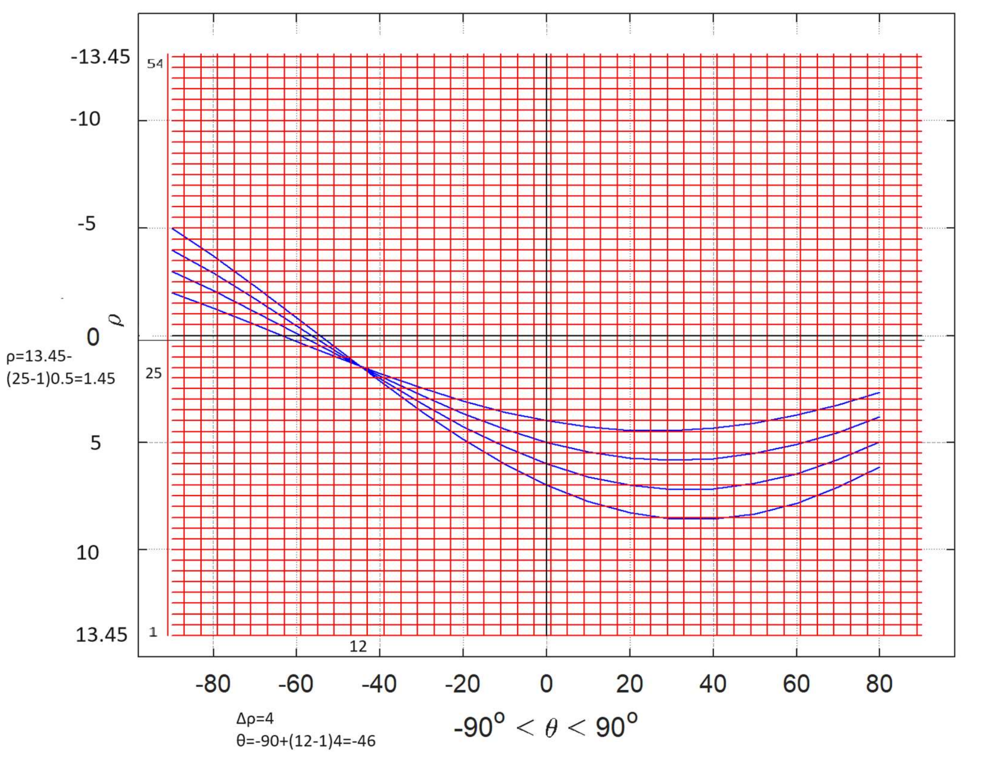
Αρχικά ο πίνακας Hough έχει μηδενικές τιμές h(n,m)=0, n=1…N και m=1…m. Για ένα σημείο $P_0(x_0,y_0)$ στο επίπεδο (x,y), υπολογίζονται οι τιμές $\rho(-90^0+(n-1)\cdot\Delta\theta)=\rho_n$. Υπολογίζεται για κάθ ρ η θέση γραμμής m στον πίνακα Η, επιλύοντας ως προς m την 2.5.4 από όπου προκύπτει ότι

$$m = round(\frac{M}{2} - \frac{\rho_n}{\Delta\rho} + 1)$$

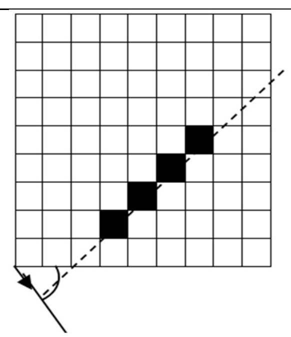
Η τιμή h(n,m) αυξάνεται κατά ένα. Το ίδιο θα γίνει και για τα σημεία π.χ. P_1, P_2, P_3 που περιλαμβάνονται στην εικόνα. Αν τα τέσσερα αυτά σημεία είναι συνευθειακά οι καμπύλες $\rho_0(\theta)$, $\rho_1(\theta)$, $\rho_2(\theta)$, $\rho_3(\theta)$ θα τέμνονται σε κάποια κοινή τιμή (ρ,θ) και η αντίστοιχη τιμή h(n,m) θα έχει την τιμή 4. Με τον τρόπο αυτό οι μεγάλες τιμές του πίνακα *H* προσδιορίζουν τις ευθείες στο επίπεδο (x,y). Αν h(n,m)=Λ μέγιστο τότε η ευθεία στο x,y με $\theta=-90^o+(n-1)\cdot\Delta\theta$ και $\rho=D-(m-1)\cdot\Delta\rho$ (σχέσεις 2.5.3, 2.5.4) θα διέρχεται από Λ το πλήθος σημεία για την συγκεκριμένη διακριτοποίηση που έγινε στο πεδίο του Hough. Πρέπει να τονισθεί ότι τα Μ, Ν με μεγάλες τιμές αυξάνουν την αυστηρότητα της συνευθειακότητας, αντίθετα με μικρές τιμές επιτρέπουν μια χαλαρότερη διέλευση ευθείας από σημεία. Στο παράδειγμα του Σχ.2.5.5.α δείχνεται ο πίνακας του Hough για την εικόνα του σχήματος 2.5.4 με $\Delta\theta=4^o$, $\Delta\rho=0.5$, $D=\sqrt{10^2+9^2}\approx 13.45$. σημείο που παρουσιάζεται μέγιστη τιμή στο πίνακα είναι για n=12 και m=25, ήτοι $\theta=-90^o+(12-1)\cdot 4=-46^o$ και $\rho=13.45-(25-1)\cdot 0.5=1.45$. Σημειώνεται ότι στο σχήμα αυτό ο άξονας των ρ είναι αντίθετα βαθμονομημένος για λόγους όμοιας παρουσίασης με αυτήν της Mathworks (MatLab). Στο Σχ.2.5.5.β φαίνεται η ευθεία που βρέθηκε στην ανεστραμμένη εικόνα και στο Σχ. η αρχική εικόνα και η ευθεία που διέρχεται από τα σημεία της. Στο παράδειγμα για λόγους απεικόνισης η διακριτοποίηση έγινε με μικρές τιμες των N,M γεγονός που δημιούργη σε αποκλίσεις από τις ακριβείς τιμές $\theta=-45^o$ και $\rho=\sqrt{2}\approx 1.4142$.

Προφανώς γειτονικά εικονοστοιχεία που ανήκουν σε μια ευθεία αποτελούν ευθύγραμμα τμήματα. Στα τμήματα αυτά μπορεί σε επόμενο βήμα να εφαρμοσθούν δύο διαδικασίες. Η μία αφορά το πλήθος τους ώστε να απορρίπτονται τα μικρά τμήματα (MinLength), η δεύτερη αφορά το πλήθος των εικονοστοιχείων του παρασκηνίου που παρεμβάλλονται σε διαδοχικά συνευθειακά τμήματα ώστε αυτά να συνενώνονται (FillGap). Στα Σχ.2.5.6 και Σχ.2.5.7 δείχνονται παραδείγματα με χρήση του MatLab.

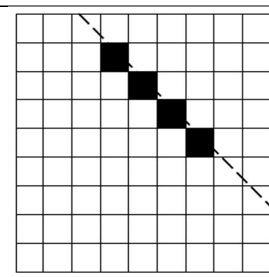
Ο Μετασχηματισμός του Hough επεκτείνεται και σε άλλα γεωμετρικά σχήματα που περιγράφονται αναλυτικά όπως κύκλοι και ελλείψεις. Τέλος ο Γενικευμένος Μετασχηματισμός του Hough στοχεύει στον εντοπισμό μορφών που δεν περιγράφονται αναλυτικά.



(α)

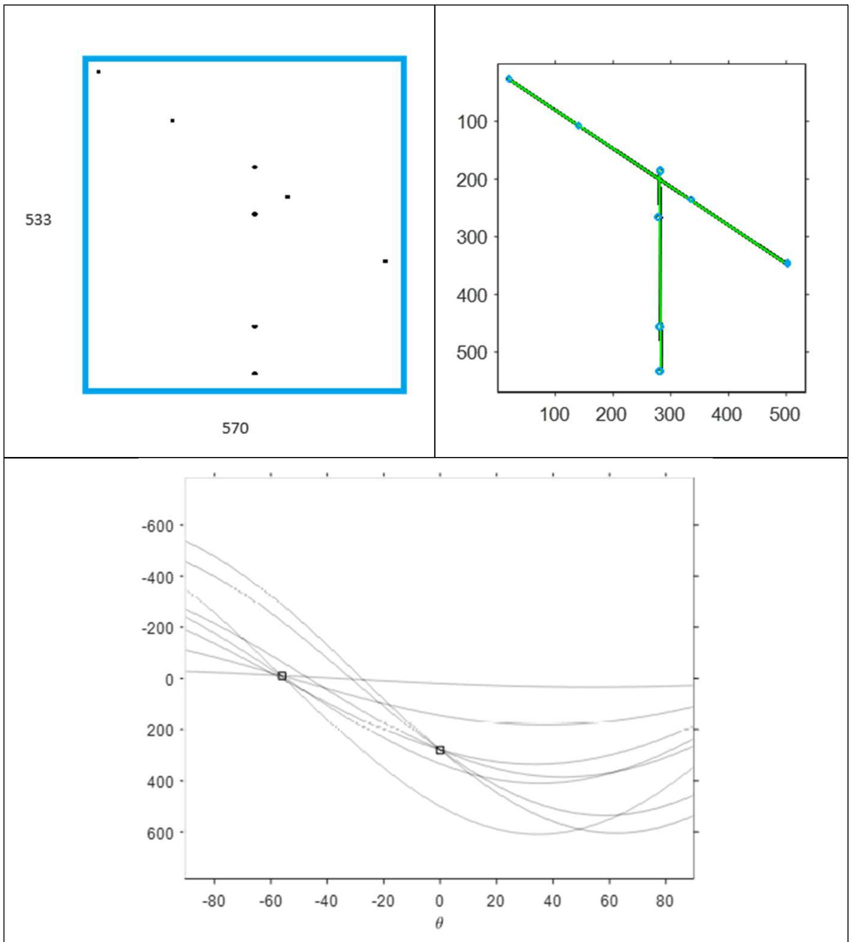


(α)

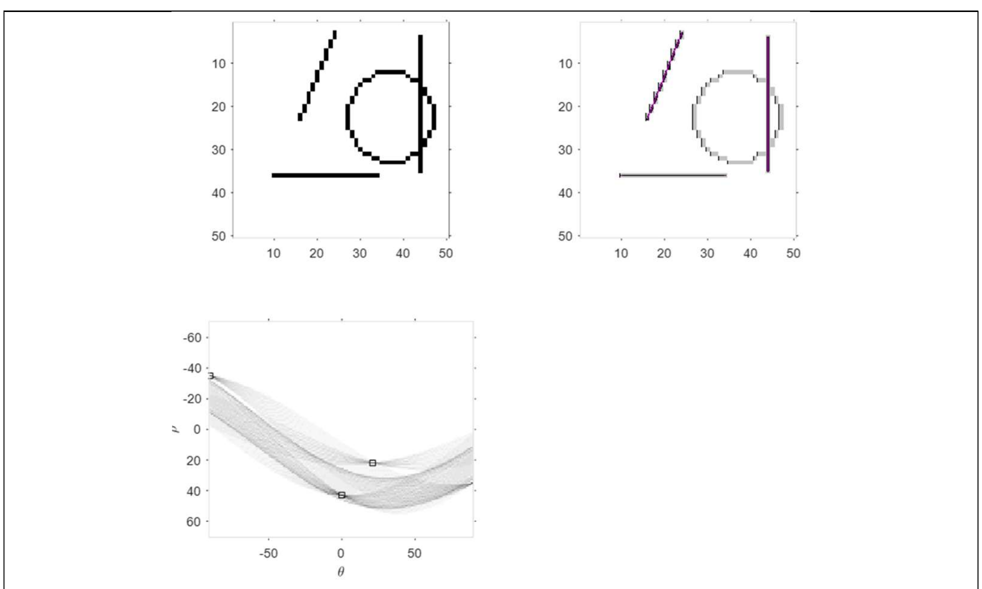


(β)

Σχήμα 2.5.5



Σχήμα 2.5.6



Σχήμα 2.5.7