ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΉΣ

ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ

Ενότητα 2.

Χαράλαμπος Π. Στρουθόπουλος
Καθηγητής

Ιούλιος 2015

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 2

ΕΠΕΞΕΡΓΑΣΙΕΣ ΣΕ ΔΥΑΔΙΚΕΣ ΕΙΚΟΝΕΣ

2.1 Κέντρο βάρους αντικειμένου.

Μία ψηφιακή δυαδική εικόνα με *J* γραμμές και *K* στήλες παριστάνεται με έναν πίνακα $I_{J\times K}$ και κάθε στοιχείο I_{jk} , j=0,...,J-1, k=0,...,K-1 παίρνει τιμή μηδέν (0) ή ένα (1). Αν η εικόνα αναπαριστά ένα αντικείμενο τότε το πλήθος των εικονοστοιχείων του αντικειμένου που έχουν τιμή ΄1΄ δίνεται από την σχέση

$$\mathrm{N} = \sum_{j=0}^{J-1}\sum_{k=0}^{K-1} \mathrm{I}_{jk}$$

Το κέντρο βάρους του αντικειμένου βρίσκεται στην θέση $(\bar{j},\bar{k})$ της εικόνας σύμφωνα με τις σχέσεις

$$\bar{k} = \frac{\sum_{j=0}^{J-1}\sum_{k=0}^{K-1} k\cdot \mathrm{I}_{jk}}{\mathrm{N}},$$

$$\bar{j} = \frac{\sum_{j=0}^{J-1}\sum_{k=0}^{K-1} j\cdot \mathrm{I}_{jk}}{\mathrm{N}}$$

2.2 Κωδικοποίηση δυαδικής εικόνας κατά μήκος διαδρομής

Μία δυαδική εικόνα μπορεί να κωδικοποιηθεί με κατάλληλο αλγόριθμο ώστε να ελαττωθεί η απαιτούμενη ποσότητα πληροφορίας για την αποθήκευση ή την μετάδοσή της και να διευκολυνθεί η αναγνώριση του περιεχομένου της. Ένας τέτοιος αλγόριθμος είναι η *κωδικοποίηση κατά μήκος διαδρομής* (*RLE:Run Lenght Ecoding*). Σύμφωνα με αυτόν ονομάζουμε *συστοιχία* μία ομάδα από διαδοχικά εικονοστοιχεία με την ίδια τιμή (όλα ΄1΄ ή όλα ΄0΄) και το πλήθος των εικονοστοιχείων ως *μήκος της συστοιχίας*. Διατρέχουμε κάθε σειρά της εικόνας και γράφουμε την θέση του πρώτου εικονοστοιχείου και το μήκος κάθε συστοιχίας από ΄1΄ ή εναλλακτικά την θέση του πρώτου και του τελευταίου εικονοστοιχείου κάθε συστοιχίας. Μία άλλη προσέγγιση είναι να γράψουμε το μήκος των διαδοχικών συστοιχειών από ΄1΄ και ΄0΄ δεχόμενοι ότι η πρώτη συστοιχία αποτελείται πάντα από ΄1΄ έστω και μηδενικού μήκους . Στο Σχ.2.2.1 που ακολουθεί φαίνεται η κωδικοποίηση των τριών πρώτων γραμμών μιας δυαδικής εικόνας με τους τρόπους που αναφέρθηκαν.

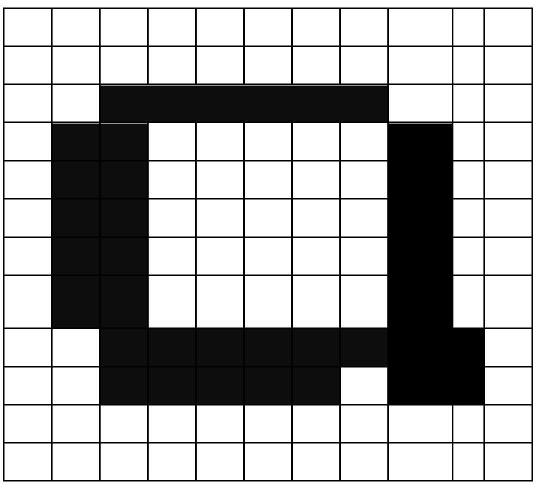
1	1	1	0	0	0	0	1	1	0	0
0	0	0	0	0	1	1	0	0	1	1
1	1	1	1	1	0	0	1	1	1	0

(0,3),(7,2)	(0,2),(8,9)	3,4,2,2
(5,2),(9,2)	(5,6),(9,10)	0,5,2,2,2
(0,5),(7,3)	(0,4),(7,9)	5,2,3,1

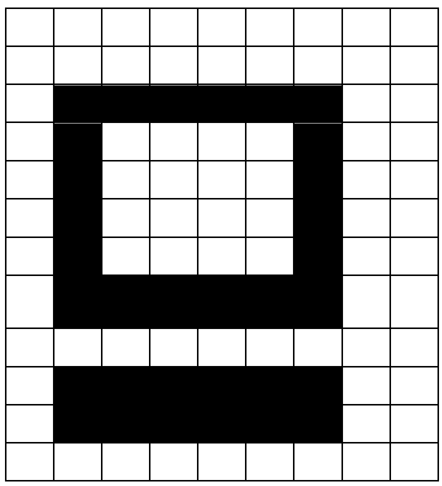
Σχήμα 2.2.1.

2.3 Συνδεδεμένα στοιχεία.

Θεωρούμε ένα σύνολο *S* εικονοστοιχείων μιας δυαδικής εικόνας με την ίδια τιμή (0 ή 1). Δύο εικονοστοιχεία ρ_1 και ρ_2 που ανήκουν στο *S* ονομάζονται *συνδεδεμένα* όταν υπάρχει διαδρομή από εικονοστοιχεία του *S* που οδηγεί από το ρ_1 στο ρ_2. Ένα σύνολο εικονοστοιχείων λέγεται *συνδεδεμένο συστατικό* (*connected component*) όταν όλα τα εικονοστοιχεία του είναι μεταξύ τους συνδεδμένα. Στο ακόλουθο Σχ.2.3.1 φαίνεται ένα συνδεδεμένο και ένα μη συνδεδεμένο συστατικό δυαδικής εικόνας.



(α) Συνδεμένο συστατικό

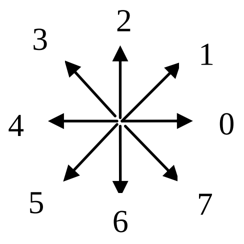


(β) Μη συνδεδεμένο συστατικό

Σχήμα 2.3.1

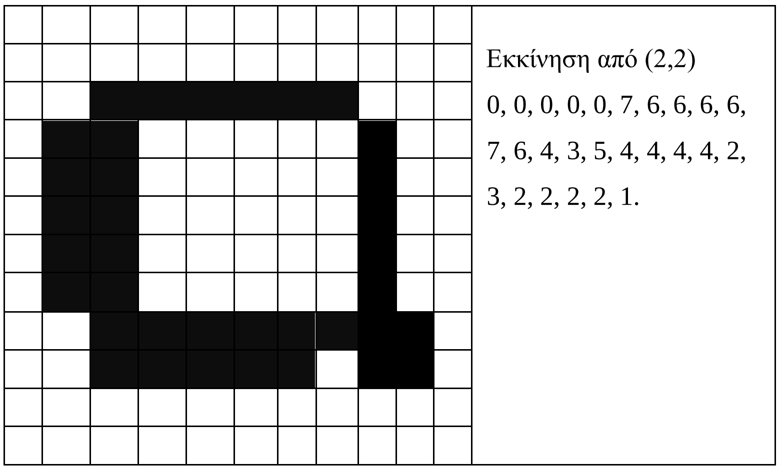
2.4 Κωδικοποίηση αλυσίδας.

Με την κωδικοποίηση αλυσίδας κωδικοποιούμε το περίγραμμα ενός αντικειμένου της εικόνας που είναι συνδεδεμένο. Για τον σκοπό αυτό καθορίζουμε και κωδικοποιούμε της διευθύνσεις που ξεκινούν από το κεντρικό εικονοστοιχείο και καταλήγουν στα γειτονικά του σε μια 3×3 γειτονιά. όπως φαίνεται στο Σχ.2.4.1



Σχήμα 2.4.1.

Ακολούθως ξεκινώντας από οποιοδήποτε εικονοστοιχείο του εξωτερικού περιγράμματος του αντικειμένου (συνήθως το πάνω αριστερό εικονοστοιχείο) διατρέχουμε το περίγραμμα γράφοντας τον κωδικό της σχετικής διεύθυνσης κάθε εικονοστοιχείου με το επόμενό του. Στο Σχ.2.4.2 δείχνεται η εφαρμογή της κωδικοποίησης αλυσίδας.



Σχήμα 2.4.2

2.5 Ο μετασχηματισμός του Hough

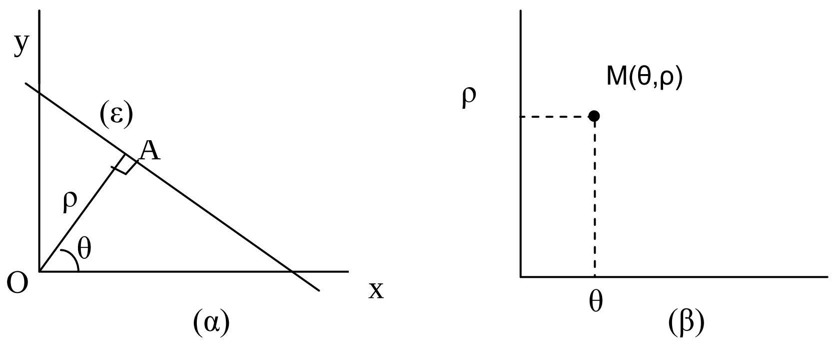
Συχνά οι ψηφιακές εικόνες περιέχουν γραμμές που ανήκουν σε σχήματα, τεχνικά σχέδια, γραφήματα, σειρές κειμένου ή άλλες αναπαραστάσεις. Σε πολλές εφαρμογές είναι επιθυμητή η εύρεση της θέσης και η αναγνώριση της μορφής των γραμμών (π.χ. ευθύγραμμα τμήματα, τόξα). Τέτοιες εφαρμογές είναι: η διανυσματική 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) καταδεικνύει την ύπαρξη ευθειών σε μια εικόνα και αποτελεί την βάση πολλών 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.

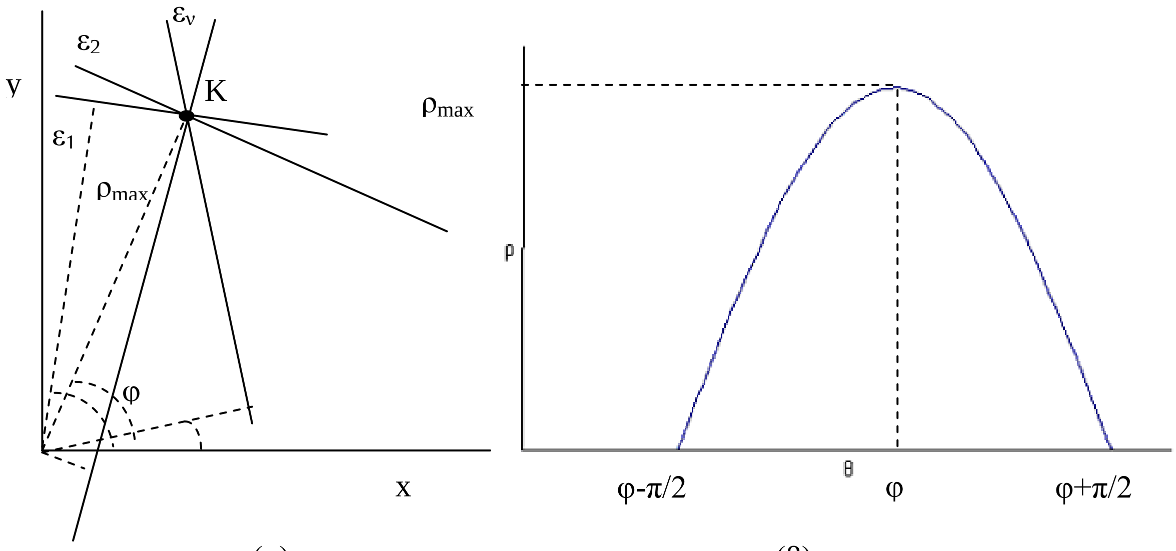
Κάθε ευθεία (ε) του καρτεσιανού επιπέδου Π_{xy} περιγράφεται από την πολική της εξίσωση

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

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



Σχ. 2.5.1



Σχήμα 2.5.2.

Το ρ είναι το μήκος του ευθύγραμμου τμήματος ΟΑ, $OA \perp (\varepsilon)$ και θ η γωνία που σχηματίζει το ΟΑ με τον άξονα Οχ. Ισχύει ότι $\rho \geq 0$ και $-\pi \leq \theta < \pi$. Θεωρούμε το επίπεδο $\Pi_{\theta\rho}$ όπου στον κάθετο άξονα σημειώνουμε τις τιμές του ρ και στον οριζόντιο τις τιμές του θ. Οι πολικές παράμετροι (θ,ρ) της ευθείας (ε) καθορίζουν ένα σημείο Μ(θ,ρ) στο επίπεδο $\Pi_{\theta\rho}$ (Σχ.2.5.1.(β))

Με τον τρόπο αυτό η ευθεία (ε) του επιπέδου Π_{xy} αντιστοιχίζεται (μετασχηματίζεται) με την χρήση της πολικής της εξίσωσης σε ένα σημείο του επιπέδου $\Pi_{\theta\rho}$. Ισοδύναμα κάθε σημείο Μ(θ,ρ) του επιπέδου $\Pi_{\theta\rho}$ ορίζει μια ευθεία $\varepsilon \rightarrow \rho = x\cos\theta + y\eta\mu\theta$ στο επίπεδο Π_{xy}. Από ένα σημείο $K(x_\kappa,y_\kappa)$ του επιπέδου Π_{xy} διέρχονται άπειρες ευθείες (ε_ν), ν=1,2… που ικανοποιούν την σχέση

$$\rho_\nu = x_\kappa\ \sigma\upsilon\nu\theta_\nu + y_\kappa\ \eta\mu\theta_\nu$$

(Σχ.2.5.2(α)). Οι τιμές ρ_ν και θ_ν των μεταβλητών ρ και θ ανήκουν στην καμπύλη $\rho = x_\kappa\ \sigma\upsilon\nu\theta + y_\kappa\eta\mu\theta$ του επιπέδου $\Pi_{\theta\rho}$. Η καμπύλη είναι ημιτονοειδής (Σχ.2.5.2(β)), όπως δείχνεται ακολούθως. Έστω φ η γωνία χΟΚ. Ισχύει ότι

$$\varepsilon\phi\phi = \frac{y_\kappa}{x_\kappa} \Leftrightarrow x_\kappa = y_\kappa\sigma\phi\phi = y_\kappa\varepsilon\phi(\frac{\pi}{2} - \phi) \Leftrightarrow$$

$$\rho = y_\kappa \frac{\eta\mu(\frac{\pi}{2} - \phi)}{\sigma\upsilon\nu(\frac{\pi}{2} - \phi)}\sigma\upsilon\nu\theta + y_\kappa\eta\mu\theta \Leftrightarrow$$

$$\rho = \frac{y_\kappa}{\eta\mu\phi}\eta\mu(\theta + \frac{\pi}{2} - \phi)$$

Αν $x_κ$ και $y_κ$ είναι θετικά ($1^ο$ τεταρτημόριο) και επειδή ρ≥0 έπεται ότι

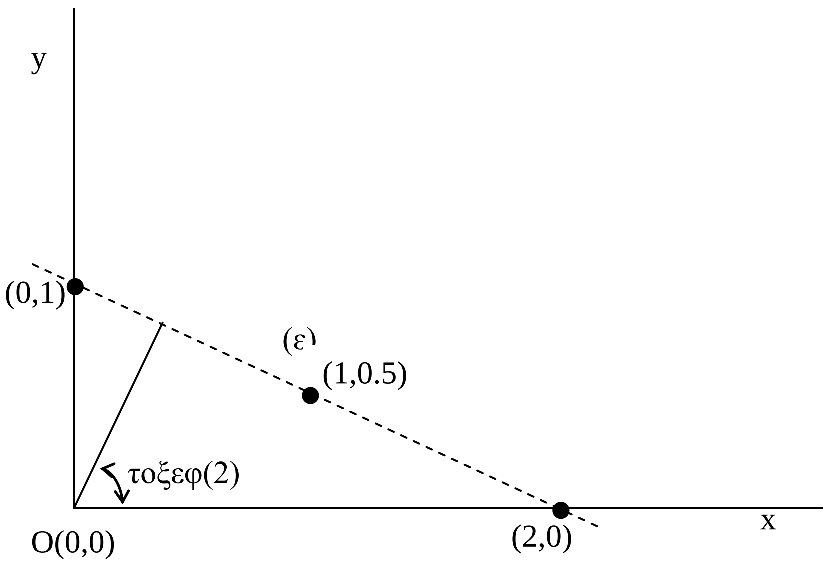
$$0 \le \phi \le \frac{\pi}{2} \qquad (α)$$

$$\phi - \frac{\pi}{2} \le \theta \le \phi + \frac{\pi}{2}, \qquad (β)$$

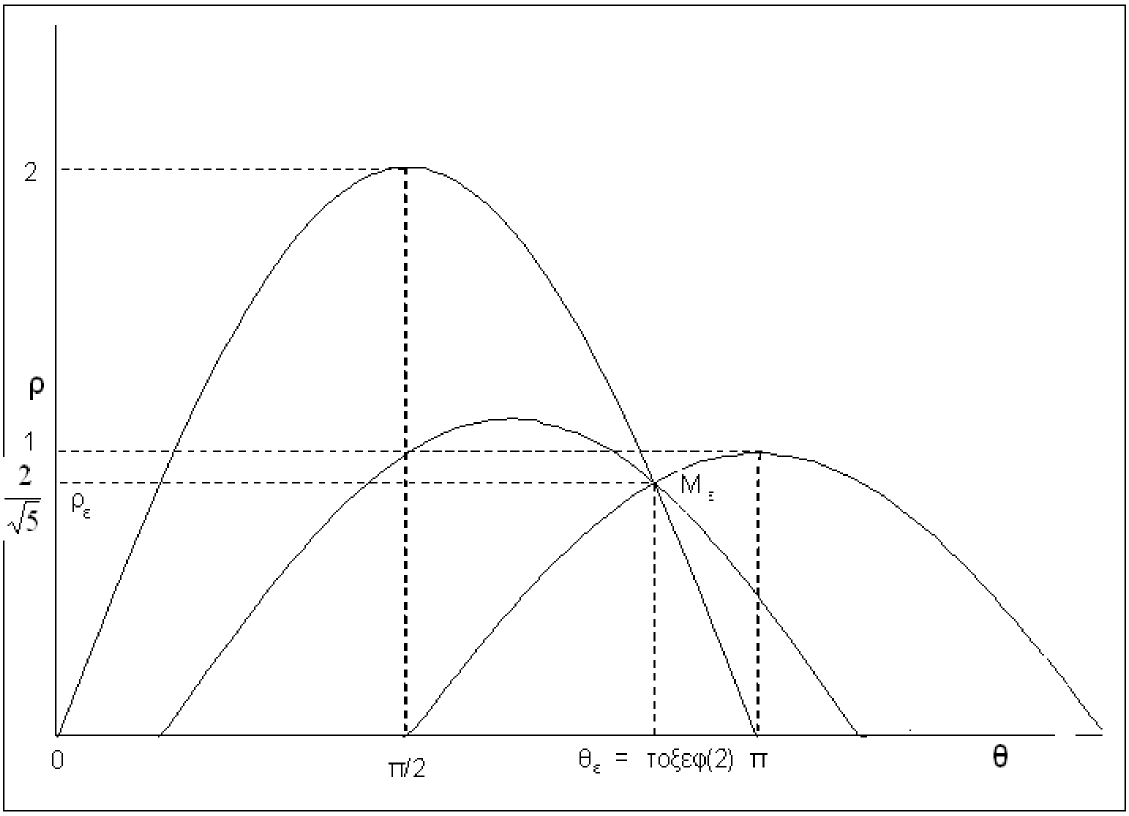
η μεγαλύτερη τιμή του ρ προκύπτει για θ=φ και είναι

$$\rho_{max} = \sqrt{x_κ^2 + y_κ^2} \qquad (γ)$$

Ο αναλυτικός υπολογισμός όλων των δυνατών τιμών των ρ και θ των ευθειών που ορίζουν ανά δύο τα σημεία της εικόνας και η εύρεση κοινών τιμών είναι επίπονος και πρακτικά ασύμφορος. Για αυτό ακολουθούμε το ακόλουθο σκεπτικό. Έστω τρία τουλάχιστον συνευθειακά σημεία στο επίπεδο $Π_{xy}$ με τις αντίστοιχες καμπύλες των ευθειών που διέρχονται από αυτά (Σχ.2.5.3(α)). Οι καμπύλες αυτές τέμνονται σε σημείο $Μ_ε(θ_ε,ρ_ε)$ με $θ_ε$ και $ρ_ε$ τις πολικές παραμέτρους της ευθείας που διέρχεται από αυτά (Σχ.2.5.3(β)). Άρα αν κατασκευάσουμε τις καμπύλες του επιπέδου $Π_{ρθ}$ για όλα τα σημεία του επιπέδου $Π_{xy}$ τα σημεία τομής τους έχουν συντεταγμένες τις πολικές παραμέτρους των ευθειών που ορίζουν τα σημεία αυτά. Προφανώς, όσες καμπύλες διέρχονται από ένα σημείο τομής στο $Π_{ρθ}$, τόσα σημεία του $Π_{xy}$ ανήκουν στην ευθεία που ορίζεται από το σημείο τομής.



Σχ.2.5.2(α).



Σχ.2.5.2(β).

Προγραμματιστικά και υπολογιστικά η μέθοδος μπορεί να εφαρμοσθεί με τεχνικές όπως η ακόλουθη. Ορίζεται ένας δισδιάστατος πίνακα $H_{M \times N}$ που κβαντίζει το επίπεδο $\Pi_{\theta\rho}$ σε M γραμμές και N στήλες με μηδενικές αρχικές τιμές. Η ψηφιακή εικόνα με πίνακα $I_{J \times K}$ τοποθετείται στο 1^o τεταρτημόριο. Οι συντεταγμένες ενός εικονοστοιχείου P(k,j) με I(k,j)=1 θα είναι x_P=k, y_P=J-j. Επειδή η εικόνα βρίσκεται στο 1^o τεταρτημόριο το διάστημα μεταβολής της γωνίας θ μπορεί να παίρνει τιμές από –π/2 έως π. Με βήμα

$$\Delta\theta = \frac{3\pi}{2N}$$

κάθε δείκτης ν=1,…,N αντιστοιχίζεται στο διάστημα $[\theta_\nu, \theta\nu+\Delta\theta)$, $\theta_\nu = -\pi/2+(\nu-1)\Delta\theta$.

Με βήμα

$$\Delta\rho = \frac{\sqrt{K^2 + J^2}}{M}$$

κάθε δείκτης μ=1,…,M αντιστοιχίζεται στο διάστημα [(μ-1)Δρ, μΔρ).

Για εκείνα τα ν που ισχύει

$$\phi - \frac{\pi}{2} \le \theta_\nu \le \phi + \frac{\pi}{2} \ , \ \phi = \tau o\xi \varepsilon \phi(\frac{y_P}{x_P})$$

υπολογίζεται η τιμή $\rho = x_P\ \sigma\upsilon\nu\theta_\nu + y_P\ \eta\mu\theta_\nu$ και ο δείκτης μ του διαστήματος που ανήκει το ρ. Η τιμή H(μ,ν) αυξάνεται κατά ένα. Η διαδικασία επαναλαμβάνεται για όλα τα εικονοστοιχεία της εικόνας. Μετά το τέλος της διαδικασίας οι υψηλές τιμές του πίνακα H προσδιορίζουν τις πολικές παραμέτρους ευθειών της εικόνας. Συγκεκριμένα αν H(μ,ν) έχει υψηλή τιμή ισχύει ότι

$$\theta = \theta_\nu + \frac{\Delta\theta}{2} = -\frac{\pi}{2} + \nu\Delta\theta - \frac{\Delta\theta}{2} \ ,$$

$$\rho = \rho_\mu + \frac{\Delta\rho}{2} = \mu\Delta\rho - \frac{\Delta\rho}{2}$$