



**ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ**

: μ μ

μ μ μ

Κεφάλαιο 10

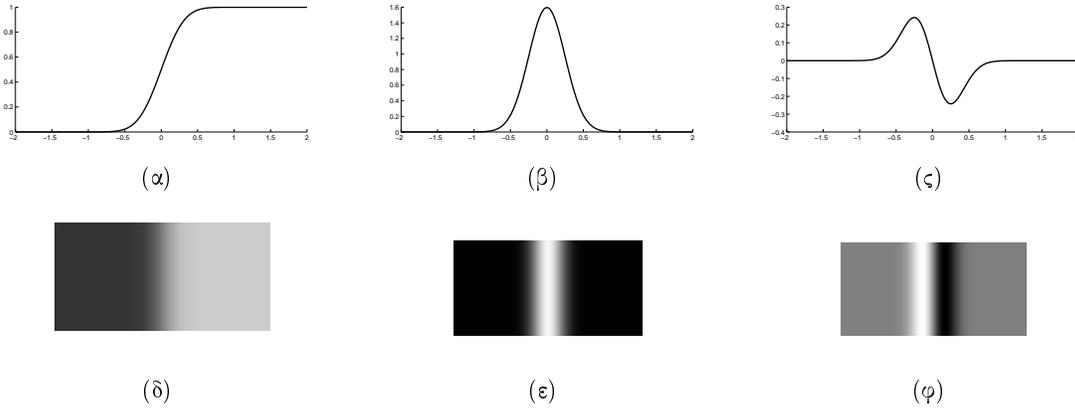
Τμηματοποίηση εικόνων

Η τμηματοποίηση των εικόνων είναι το πρώτο στάδιο στην ανάλυση των εικόνων. Η αναγνώριση αντικειμένων, η αναγνώριση γραπτών χαρακτήρων, η χαρτογράφηση, η ανάλυση μιας τρισδιάστατης σκηνής με τη στερεοσκοπία ή την ανάλυση της κίνησης βασίζονται κατά κανόνα στο αποτέλεσμα της τμηματοποίησης των εικόνων. Επομένως η ανάλυση των εικόνων θα είναι τόσο πιο ακριβής, όσο πιο επιτυχής είναι ο προσδιορισμός των τμημάτων της εικόνας, που αντιστοιχούν σε επιφάνειες ή τμήματα επιφανειών τρισδιάστατων αντικειμένων, ή ακόμα σε περιοχές με διαφορετική σύσταση στο εσωτερικό σωμάτων.

Δύο ισοδύναμοι και συμπληρωματικοί τρόποι προσδιορίζουν τα τμήματα μιας εικόνας: τα σύνορα και τα ομοιογενή χαρακτηριστικά του κάθε τμήματος. Σε αυτό το κεφάλαιο θα αναφερθούμε αποκλειστικά στην τμηματοποίηση με βάση τη φωτεινή ένταση. Επομένως τα σύνορα αντιστοιχούν σε ασυνέχειες της φωτεινής έντασης, ενώ τα τμήματα παρουσιάζουν ομοιογένεια της φωτεινής έντασης. Παρουσιάζονται στη συνέχεια τέσσερις προσεγγίσεις του προβλήματος της τμηματοποίησης των εικόνων. Στην ανίχνευση των ακμών αναζητούνται με τη χρήση αριθμητικής παραγωγίσις τα σημεία ασυνέχειας της φωτεινής έντασης. Στην ανάλυση του ιστογράμματος γίνεται κατάταξη των σημείων της εικόνας σε διαφορετικές κλάσεις ανάλογα με το διάστημα τιμών της φωτεινής έντασης. Η χρήση ενός μοντέλου Markov επιτρέπει να ληφθεί υπόψη η συνεκτικότητα των περιοχών μέσα από ένα κριτήριο στο σύνολο της εικόνας. Τέλος, η τμηματοποίηση μπορεί να γίνει χρησιμοποιώντας τοπικά ένα κριτήριο ομοιογένειας για την κατάτμηση μιας περιοχής ή τη συγχώνευση γειτονικών περιοχών.

10.1 Ανίχνευση ακμών

Ως σημείο ακμής ορίζεται εκείνο το σημείο όπου η μεταβολή της φωτεινής έντασης είναι τοπικά μέγιστη. Ένα σύνολο συνδεδεμένων σημείων ακμής αποτελεί ένα περίγραμμα. Σε κάθε σημείο ενός περιγράμματος μπορεί να ορισθεί ο προσανατολισμός του σε σχέση με το σύστημα συντεταγμένων του επιπέδου της εικόνας. Η φωτεινή ένταση παρουσιάζει τη μέγιστη μεταβολή στην κάθετη στο περίγραμμα κατεύθυνση. Με αυτή την έννοια η ανίχνευση ακμής ισοδυναμεί με τον εντοπισμό μιας μονοδιάστατης μεταβολής. Στο Σχήμα 10.1 δίδεται μια οριζόντια βαθμιαία μεταβολή και δείχνεται πώς η πρώτη ή η δεύτερη παράγωγος μπορούν να χρησιμοποιηθούν για την ανίχνευση και τον εντοπισμό της μεταβολής. Σε δισδιάστατες εικόνες η ανίχνευση των σημείων ακμής μπορεί να γίνει με τη βοήθεια είτε του διανύσματος των δύο πρώτων μερικών παραγώγων της φωτεινής έντασης, που ονομάζεται κλίση, είτε του αθροίσματος των δύο δεύτερων μερικών παραγώγων της φωτεινής έντασης, που καλείται λαπλασιανός (Laplacian) τελεστής. Ο τελεστής αυτός είναι ιστροπικός, δηλαδή η απόκρισή του είναι ανεξάρτητη από τον προσανατολισμό του



Σχήμα 10.1: Επάνω δίδεται η συνάρτηση μιας βαθμιαίας μεταβολής, η πρώτη και η δεύτερη παράγωγός της. Κάτω δίδεται μια εικόνα με οριζόντια βαθμιαία μεταβολή, η πρώτη οριζόντια παραγωγή και η δεύτερη οριζόντια παραγωγή.

περιγράμματος. Η ισοτροπικότητα είναι σημαντική ιδιότητα για την ακρίβεια του εντοπισμού των περιγραμμάτων.

10.1.1 Μέθοδος της κλίσης

Το διάνυσμα της κλίσης της φωτεινής έντασης, $I(x, y)$, έχει σα συνιστώσες τις δύο πρώτες μερικές παραγώγους

$$\nabla I = \begin{bmatrix} I_x \\ I_y \end{bmatrix}$$

Για την ανίχνευση των ακμών χρησιμοποιείται το μέτρο του διανύσματος της κλίσης

$$\|\nabla I\| = \sqrt{I_x^2 + I_y^2}.$$

Αποδεικνύεται ότι το μέτρο του διανύσματος της κλίσης για συνεχή σήματα ικανοποιεί την ιδιότητα της ισοτροπικότητας. Η κατεύθυνση του διανύσματος της κλίσης δίδεται από τη γωνία

$$\theta = \arctan \frac{I_y}{I_x}.$$

Η κλίση της φωτεινής έντασης σε μια δοσμένη κατεύθυνση, ϕ , είναι

$$I_\phi = I_x \cos \phi + I_y \sin \phi.$$

Η κλίση της φωτεινής έντασης είναι μέγιστη στην κατεύθυνση του διανύσματος της κλίσης, και μηδενική στην κάθετη σ' αυτήν κατεύθυνση. Σημεία ακμής ανιχνεύονται ακριβώς στην κατεύθυνση που η κλίση είναι μέγιστη, οπότε είναι ίση με το μέτρο του διανύσματος της κλίσης, $\|\nabla I\|$. Τελικά τα σημεία ακμής ανιχνεύονται ως τα τοπικά μέγιστα του μέτρου του διανύσματος της κλίσης, με τα τοπικά μέγιστα να ευρίσκονται στην κατεύθυνση αυτού του ίδιου διανύσματος.

Για την εφαρμογή της μεθόδου της κλίσης απαιτείται επομένως ο αριθμητικός υπολογισμός των πρώτων μερικών παραγώγων της φωτεινής έντασης. Εφόσον πρόκειται για μερική παραγωγή, το χρησιμοποιούμενο φίλτρο είναι μονοδιάστατο. Ωστόσο λόγω της ύπαρξης θορύβου, είναι αναγκαία επιπλέον η χρήση ενός βαθυπερατού φίλτρου στην άλλη κατεύθυνση, ως προς

αυτή της παραγωγίσης. Έτσι στην περίπτωση της οριζόντιας παραγωγίσης, πραγματοποιείται ολοκλήρωση στην κατακόρυφη κατεύθυνση, και αντίστροφα. Τα δύο φίλτρα χρησιμοποιούνται ξεχωριστά, ώστε το 2- Δ φίλτρο που προκύπτει να είναι διαχωρίσιμο. Ας είναι $g(\cdot)$ η κρουστική απόκριση του φίλτρου παραγωγίσης, και $h(\cdot)$ η κρουστική απόκριση του φίλτρου ολοκλήρωσης. Η ακολουθία $g(\cdot)$ είναι πάντοτε αντισυμμετρική

$$g(n) = -g(-n),$$

ενώ συνιστάται η ακολουθία $h(\cdot)$ να είναι συμμετρική

$$h(n) = h(-n).$$

Πολύ συχνά μεταξύ των δύο αυτών ακολουθιών υπάρχει ακριβώς η σχέση παραγώγου και ολοκλήρωματος

$$h(n) = \sum_{k=-\infty}^{n-1} g(k) + \frac{1}{2}g(n). \quad (10.1)$$

Παρουσιάζονται στη συνέχεια μερικά φίλτρα παραγωγίσης και ολοκλήρωσης με πεπερασμένη κρουστική απόκριση. Το απλούστερο αντισυμμετρικό φίλτρο παραγωγίσης είναι το ακόλουθο

$$g(n) = \begin{cases} 1/2 & n = -1 \\ -1/2 & n = 1 \\ 0 & |n| \neq 1 \end{cases} \quad (10.2)$$

Το αντίστοιχο φίλτρο ολοκλήρωσης, με βάση την Εξίσωση (10.1), είναι

$$h(m) = \begin{cases} 1/4 & |m| = 1 \\ 1/2 & m = 0 \\ 0 & |m| > 1 \end{cases} \quad (10.3)$$

Το 2- Δ φίλτρο που προκύπτει είναι γνωστό σαν το φίλτρο Sobel. Αν είναι



Σχήμα 10.2: Χρήση του φίλτρου Sobel για την ανίχνευση ακμών. Αριστερά δίδεται η οριζόντια παραγωγίση, δεξιά η κατακόρυφη παραγωγίση και στο κέντρο το μέτρο του διανύσματος της κλίσης.

$$h(m) = \begin{cases} 1/3 & |m| \leq 1 \\ 0 & |m| > 1 \end{cases} \quad (10.4)$$

πρόκειται για το φίλτρο Prewitt. Αν λόγω θορύβου απαιτείται ένα φίλτρο ολοκλήρωσης με πιο τονισμένο το βαθυπερατό χαρακτήρα, αρκεί να πάρουμε τη συνέλιξή του (10.3) με τον εαυτό του, που δίνει το εξής αποτέλεσμα

$$h(m) = \begin{cases} 1/16 & |m| = 2 \\ 1/4 & |m| = 1 \\ 6/16 & m = 0 \\ 0 & |m| > 2 \end{cases}$$

Το φίλτρο παραγωγίσης που ικανοποιεί την Εξίσωση (10.1), είναι

$$g(n) = \begin{cases} 1/8 & n = -2 \\ 1/4 & n = -1 \\ -1/4 & n = 1 \\ -1/8 & n = 2 \\ 0 & n = 0 \text{ είτε } |n| > 2 \end{cases}$$

Το πρόβλημα της σχεδίασης του φίλτρου παραγωγίσης μπορεί να τεθεί σαν ένα πρόβλημα βελτιστοποίησης ορισμένων κριτηρίων που να εκφράζουν την καλή ανίχνευση ακμών με παρουσία θορύβου. Το πρόβλημα τέθηκε αρχικά από τον Canny και λύθηκε για ένα συνεχές μονοδιάστατο σήμα, όπου το μοντέλο της ασυνέχειας δίδεται από μια βηματική συνάρτηση, στην οποία προστίθεται λευκός θόρυβος με μηδενική μέση τιμή και διασπορά σ^2 . Ας υποθέσουμε ότι το σήμα έχει ως εξής

$$f(x) = \begin{cases} A & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (10.5)$$

Ζητείται το φίλτρο παραγωγίσης που ικανοποιεί τρία κριτήρια: καλής ανίχνευσης, μικρής τοπικής απόκλισης και μικρής παρενόχλησης από πολλαπλά μέγιστα. Η καλή ανίχνευση μετριέται με τη σηματοθορυβική σχέση στην έξοδο του φίλτρου στο σημείο της ασυνέχειας

$$\mathcal{R} = \frac{|\int_{-\infty}^{\infty} g(-x)f(x)dx|}{\sigma\sqrt{\int_{-\infty}^{\infty} g^2(x)dx}}. \quad (10.6)$$

Η σηματοθορυβική σχέση πρέπει να είναι η μέγιστη δυνατή. Η ασυνέχεια πρέπει επίσης να ανιχνεύεται όσο γίνεται πλησιέστερα στο σημείο της πραγματικής ασυνέχειας του σήματος. Έχει προταθεί το ακόλουθο κριτήριο καλού εντοπισμού

$$\mathcal{L} = \frac{|\int_{-\infty}^{\infty} f'(-x)g'(x)dx|}{\sigma\sqrt{\int_{-\infty}^{\infty} g'^2(x)dx}}. \quad (10.7)$$

Τέλος, αποδεικνύεται ότι η μέση απόσταση μεταξύ δύο μεγίστων στην έξοδο του φίλτρου είναι

$$\mathcal{D} = 2\pi\sqrt{\frac{\int_{-\infty}^{\infty} g'^2(x)dx}{\int_{-\infty}^{\infty} g''^2(x)dx}}. \quad (10.8)$$

Με δοσμένη την απαίτηση ικανοποίησης πολλαπλών κριτηρίων, έχει προταθεί ο ορισμός του βέλτιστου φίλτρου παραγωγίσης ως εκείνου που μεγιστοποιεί το γινόμενο $\mathcal{R}\mathcal{L}$ με μία συνθήκη για το \mathcal{D} . Κάτω από αυτούς τους όρους είναι δυνατή η λύση του προβλήματος βελτιστοποίησης για την περίπτωση του σήματος (10.5) και για πεπερασμένη ή άπειρη κρουστική απόκριση. Για την περίπτωση μιας πεπερασμένης, και οπωσδήποτε αντισυμμετρικής, κρουστικής απόκρισης,

εκτεινόμενης στο διάστημα $[-M, M]$, ζητείται επομένως, για διάφορες τιμές του \mathcal{D}/M , η μεγιστοποίηση του

$$\frac{|g'(0)| \int_{-M}^0 g(x) dx}{\sqrt{\int_{-M}^M g^2(x) dx \int_{-M}^M g'^2(x) dx}}$$

Για χαμηλές τιμές του $\mathcal{D}/M (\approx 0, 1)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από

$$g(x) = \begin{cases} 1/M^2 & -M \leq x < 0 \\ -1/M^2 & 0 < x \leq M \\ 0 & x = 0 \text{ είτε } |x| > M \end{cases} \quad (10.9)$$

Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

$$h(x) = \begin{cases} \frac{M-|x|}{M^2} & |x| \leq M \\ 0 & |x| > M \end{cases} \quad (10.10)$$

Για υψηλές τιμές του $\mathcal{D}/M (\approx 1, 5)$, η λύση που προκύπτει μπορεί να προσεγγισθεί από την παράγωγο μιας συνάρτησης Gauss

$$g(x) = -\frac{x}{\sigma^3 \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}. \quad (10.11)$$

Προφανώς το φίλτρο αυτό έχει άπειρη κρουστική απόκριση. Αν το φίλτρο ζητηθεί κατευθείαν να έχει άπειρη κρουστική απόκριση, η λύση του προβλήματος βελτιστοποίησης δίδει

$$g(x) = -\frac{1}{4} \alpha^3 x e^{-\alpha|x|}. \quad (10.12)$$

Υψηλές τιμές του α αντιστοιχούν σε χαμηλές τιμές του \mathcal{D} , και αντίστροφα. Το αντίστοιχο φίλτρο ολοκλήρωσης είναι

$$h(x) = \frac{1}{4} \alpha (1 + \alpha|x|) e^{-\alpha|x|}. \quad (10.13)$$

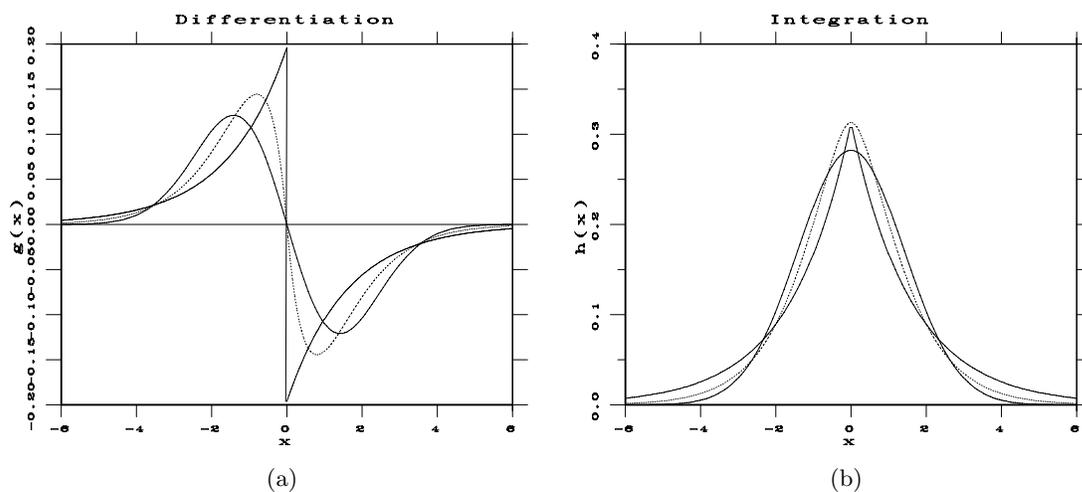
Άλλο χρησιμοποιούμενο ζεύγος φίλτρων με άπειρη κρουστική απόκριση είναι το ακόλουθο

$$g(x) = -\frac{\alpha^2}{2} \text{sign}(x) e^{-\alpha|x|} \quad \text{και} \quad h(x) = \frac{\alpha}{2} e^{-\alpha|x|} \quad (10.14)$$

με παρόμοιο ρόλο για το α . Όλα τα παραπάνω φίλτρα χρησιμοποιούνται φυσικά σε διακριτή μορφή, ανεξάρτητα αν είναι πεπερασμένης ή άπειρης κρουστικής απόκρισης. Για τα φίλτρα άπειρης κρουστικής απόκρισης υπάρχουν υλοποιήσεις μέσω αναδρομικών σχέσεων.

Στο Σχήμα 10.3(a) δίδονται τα παραπάνω φίλτρα παραγωγής, με τέτοιες τιμές για τις παραμέτρους, ώστε η σηματοθορυβική σχέση στην έξοδο του φίλτρου να είναι ίδια. Συγκεκριμένα θα είναι για την παράγωγο της συνάρτησης Gauss, $\sigma = \sqrt{2}$, για το φίλτρο της Εξίσωσης (10.12), που δίδεται με διακοπτόμενη γραμμή, $\alpha = \sqrt{\pi/2}$, και για το φίλτρο της Εξίσωσης (10.14), $\alpha = \frac{1}{2} \sqrt{\frac{\pi}{2}}$. Τα αντίστοιχα φίλτρα ολοκλήρωσης δίδονται στο Σχήμα 10.3(b).

Στο Σχήμα 10.4 δίδεται το αποτέλεσμα χρήσης της παραγωγού της συνάρτησης Gauss (φίλτρο Canny) με $\sigma = 1$ και με ανίχνευση που χρησιμοποιεί δύο κατώφλια και την τεχνική της υστέρησης. Με αυτή την τεχνική κατά την εφαρμογή του δεύτερου και μικρότερου κατωφλιού λαμβάνεται επίσης υπόψη η συνδεσιμότητα των σημείων ακμής.



Σχήμα 10.3: Φίλτρα παραγώγισης και ολοκλήρωσης.



Σχήμα 10.4: Αποτέλεσμα ανίχνευσης ακμών με τη μέθοδο του Canny.

10.1.2 Μέθοδος δευτέρων παραγώγων

Η διέλευση από το μηδέν των τιμών του αθροίσματος των δύο δευτέρων παραγώγων δίδει επίσης σημεία ακμών της φωτεινής έντασης. Ο τελεστής που δίδει αυτό το άθροισμα, ονομάζεται λαπλασιανός (Laplacian), και ορίζεται ως ακολούθως

$$\Delta I = I_{xx} + I_{yy}.$$

Αποδεικνύεται ότι ο λαπλασιανός τελεστής είναι ισοτροπικός, αρκεί να είναι: $I_{xy} = I_{yx}$.

Απλά φίλτρα πεπερασμένης κρουστικής απόκρισης που χρησιμοποιούνται για την υλοποίηση του λαπλασιανού τελεστή είναι τα ακόλουθα:

$$g(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 1 & m^2 + n^2 = 1 \\ 0 & m^2 + n^2 > 1 \end{cases} \quad (10.15)$$

και

$$g(m, n) = \begin{cases} -4 & m^2 + n^2 = 0 \\ 0,5 & m^2 + n^2 = 1 \text{ είτε } m^2 + n^2 = 2 \\ 0 & m^2 + n^2 > 2 \end{cases} \quad (10.16)$$

Ωστόσο αυτά τα φίλτρα χαρακτηρίζονται από ενίσχυση του θορύβου. Για τη μείωση του θορύβου προτείνεται η χρήση ενός βαθυπερατού φίλτρου Gauss, του οποίου η κρουστική απόκριση, σε συνεχείς συντεταγμένες είναι

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

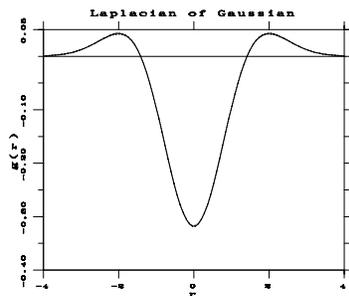
Η όλη επεξεργασία είναι ισοδύναμη με τη συνέλιξη του σήματος της φωτεινής έντασης με το φίλτρο που προκύπτει από την εφαρμογή του λαπλασιανού τελεστή στην κρουστική απόκριση Gauss, που δίδει το ακόλουθο αποτέλεσμα

$$g(r) = \frac{1}{\pi\sigma^4} \left(\frac{r^2}{2\sigma^2} - 1 \right) e^{-\frac{r^2}{2\sigma^2}} \quad (10.17)$$

όπου $r = \sqrt{x^2 + y^2}$. Αποδεικνύεται επίσης ότι η παραπάνω κρουστική απόκριση μπορεί να προσεγγισθεί από μία διαφορά δύο συναρτήσεων Gauss,

$$g(r) \propto \frac{1}{2\pi\sigma_1^2} e^{-\frac{r^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{r^2}{2\sigma_2^2}}, \quad (10.18)$$

όπου $\sigma_1^2 = 1,28\sigma^2$ και $\sigma_2^2 = 0,8\sigma^2$. Στο Σχήμα 10.5 δίδεται τόσο η λαπλασιανή της συνάρτησης Gauss, όσο και η προσέγγισή της, με διακοπτόμενη γραμμή, με τη διαφορά δύο συναρτήσεων Gauss. Επειδή η συνάρτηση Gauss δύο μεταβλητών είναι διαχωρίσιμη, η διαφορά δύο συναρτήσεων Gauss απαιτεί την υλοποίηση δύο διαχωρίσιμων φίλτρων. Η συνάρτηση Gauss μιας μεταβλητής προσεγγίζεται σε διακριτή μορφή, είτε από ένα φίλτρο πεπερασμένης κρουστικής απόκρισης, είτε από δύο αναδρομικά φίλτρα.



(a)



(b)

Σχήμα 10.5: Λαπλασιανή της συνάρτησης Gauss και αποτέλεσμα ανίχνευσης ακμών με εντοπισμό των διελεύσεων από το μηδέν.

10.2 Ανάλυση ιστογράμματος

Με την ανάλυση του ιστογράμματος η τμηματοποίηση καθίσταται ένα πρόβλημα κατάταξης των σημείων της εικόνας σε κλάσεις. Από το ιστόγραμμα, μετά από τη λείανση του, εφόσον είναι αναγκαία, προσδιορίζονται οι επικρατούσες τιμές της ολικής κατανομής. Η κατάταξη των σημείων της εικόνας γίνεται με βάση αυτές τις τιμές, ή ενδεχόμενα μέσω της κατανομής ανά κλάση που θα μπορούσε να υπολογισθεί. Η μέθοδος αυτή δεν εξασφαλίζει οπωσδήποτε τη συνεκτικότητα

των αντικειμένων που εντοπίζονται, αφού η κατάταξη κάθε σημείου βασίζεται αποκλειστικά στη φωτεινή ένταση του δοσμένου μόνο σημείου. Η αξιοπιστία της μεθόδου εξαρτάται από την ομοιογένεια του φωτισμού, και από το μέγεθος των αντικειμένων, που επιπλέον ως προς το είδος θα πρέπει να είναι ολίγα τον αριθμό.

Ο χωρισμός του ιστογράμματος σε διαστήματα μπορεί να βασισθεί σε κάποιο αρχικό μοντέλο για τη συνάρτηση πυκνότητας πιθανότητας της μεταβλητής που εκφράζει τη φωτεινή ένταση. Ας θεωρήσουμε την περίπτωση δύο επικρατουσών τιμών, ή ενός είδους αντικειμένου που ανιχνεύεται σε αντίθεση προς το βάθος της εικόνας. Ας υποθέσουμε ότι η τυχαία μεταβλητή του μοντέλου ακολουθεί και στις δύο περιπτώσεις την κανονική κατανομή, με την ίδια διασπορά, σ^2 , και μέση τιμή μ_0 για το βάθος της εικόνας, και μ_1 για το αντικείμενο. Με το κριτήριο της μέγιστης αληθοφάνειας το κατώφλι ανίχνευσης τοποθετείται στη μέση της απόστασης μεταξύ των θέσεων των δύο μεγίστων τιμών του ιστογράμματος

$$\kappa = \frac{\mu_0 + \mu_1}{2}. \quad (10.19)$$

Αν επιπλέον θεωρήσουμε ότι είναι γνωστές οι a priori πιθανότητες του αντικειμένου (P_1) και συμπληρωματικά του βάθους της εικόνας ($P_0 = 1 - P_1$), τότε η μεγιστοποίηση της a posteriori πιθανότητας δίνει το ακόλουθο κατώφλι για την ανίχνευση του αντικειμένου

$$\kappa = \frac{\mu_0 + \mu_1}{2} + \frac{\sigma^2}{\mu_1 - \mu_0} \ln \frac{P_0}{P_1}. \quad (10.20)$$

Ο προσδιορισμός των παραμέτρων $\mu_0, \mu_1, \sigma^2, P_0, P_1$ μπορεί να γίνει με τη βοήθεια του ιστογράμματος.

Ο προσδιορισμός του κατωφλιού μπορεί επίσης να βασισθεί σ' ένα κριτήριο αυτόματης ομαδοποίησης. Αν μείνουμε στην περίπτωση των δύο επικρατουσών τιμών, ένα τέτοιο κριτήριο προς ελαχιστοποίηση είναι το ακόλουθο

$$\sum_{i=0}^{k-1} p_i (i - \mu_0)^2 + \sum_{i=k}^{N-1} p_i (i - \mu_1)^2,$$

όπου p_i είναι η συχνότητα εμφάνισης, ή εμπειρική πιθανότητα, της τιμής i , για ένα σύνολο από N δυνατές τιμές. Ζητούνται οι τιμές του κατωφλιού k , και των παραμέτρων μ_0 και μ_1 που ελαχιστοποιούν το παραπάνω κριτήριο. Οι τιμές αυτές μπορούν να προσδιορισθούν χρησιμοποιώντας δύο αναγκαίες συνθήκες που ισχύουν στη θέση του ελάχιστου του κριτηρίου. Η πρώτη κατηγορία αναγκαίων συνθηκών προκύπτει για δοσμένο κατώφλι, και δίνει τις αντιπροσωπευτικές τιμές των δύο κλάσεων

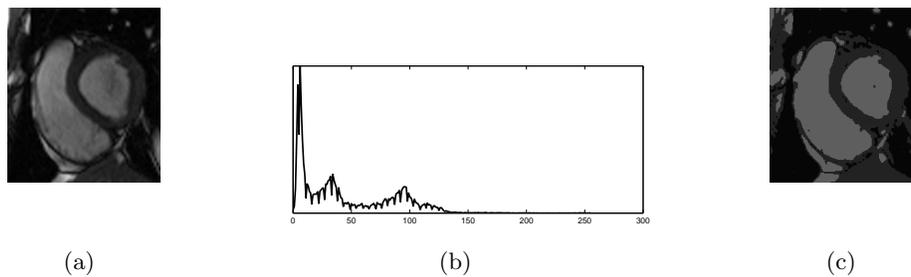
$$\hat{\mu}_0 = \frac{\sum_{i=0}^{k-1} i p_i}{\sum_{i=0}^{k-1} p_i} \quad \text{και} \quad \hat{\mu}_1 = \frac{\sum_{i=k}^{N-1} i p_i}{\sum_{i=k}^{N-1} p_i} \quad (10.21)$$

Η δεύτερη αναγκαία συνθήκη δίνει το κατώφλι για δοσμένες αντιπροσωπευτικές τιμές

$$k = \left\lfloor \frac{\mu_0 + \mu_1}{2} + 1 \right\rfloor \quad (10.22)$$

Η διαδοχική χρήση των παραπάνω συνθηκών σ' ένα επαναληπτικό αλγόριθμο, επιτρέπει την ελαχιστοποίηση του κριτηρίου και τον προσδιορισμό του κατωφλιού που οδηγεί στην τμηματοποίηση της εικόνας. Η μέθοδος αυτή μπορεί να επεκταθεί για περισσότερες των δύο τελικές αποχρώσεις. Ένα παράδειγμα εφαρμογής της με τρεις αποχρώσεις δίδεται στο Σχήμα 10.6.

Κάποια από τα μειονεκτήματα της ανάλυσης του ιστογράμματος, όπως η ύπαρξη μικρών στην έκταση αντικειμένων ή η ανομοιογένεια του φωτισμού, που αναφέρθηκαν εισαγωγικά, μπορούν να αντιμετωπισθούν με την εφαρμογή της μεθόδου τοπικά κατά τμήματα της εικόνας.

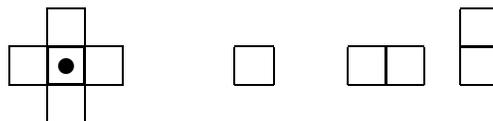


Σχήμα 10.6: Το ιστόγραμμα της εικόνας αριστερά και το αποτέλεσμα της τμηματοποίησης με χρήση τριών αποχρώσεων.

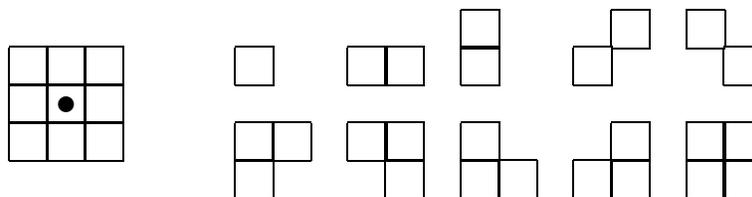
10.3 Ανάλυση ενός μοντέλου Markov ή Gibbs

Σε αντιδιαστολή με την απλή ανάλυση του ιστογράμματος, η ανάλυση με βάση ένα μοντέλο Markov για το πεδίο των αποφάσεων κατάταξης των σημείων σε τμήματα, χρησιμοποιεί τις αλληλεξαρτήσεις που υπάρχουν τοπικά σε γειτονικά σημεία. Για κάθε μοντέλο Markov που δίνει τις τοπικές αλληλεξαρτήσεις ορίζεται ισοδύναμα μία κατανομή Gibbs που δίνει συνολικά όλες τις εξαρτήσεις του πεδίου των αποφάσεων. Ο ορισμός της κατανομής Gibbs δίδεται στη συνέχεια, αφού προηγουμένα δοθούν οι σχέσεις γειτνίασης που είναι αναγκαίες για την περιγραφή των εξαρτήσεων του πεδίου.

Ας είναι $L = \{(m, n) : 0 \leq m < M, 0 \leq n < N\}$ το σύνολο των σημείων της εικόνας. Η γειτονιά ενός σημείου (m, n) ορίζεται ως ένα υποσύνολο $\gamma(m, n)$ του L , που δεν περιλαμβάνει το σημείο (m, n) , με τον όρο ότι, εάν $(k, l) \in \gamma(m, n)$, τότε $(m, n) \in \gamma(k, l)$. Το σύστημα γειτνίασης στο L ορίζεται ως το σύνολο των γειτονιών όλων των σημείων του L . Με τη βοήθεια του ορισμού της γειτονιάς ορίζεται η παρέα, c , ως ένα σύνολο, είτε ενός μόνο σημείου, είτε περισσότερων του ενός σημείων, αρκεί για κάθε σημείο μιας παρέας τα υπόλοιπα σημεία της παρέας να ανήκουν στη γειτονιά του. Στο Σχήμα 10.7 δίδεται ο ορισμός μιας γειτονιάς 4 σημείων και οι αντίστοιχες σ' αυτή παρέες. Παρόμοια δίδονται οι παρέες για μια γειτονιά 8 σημείων στο Σχήμα 10.8.



Σχήμα 10.7: Γειτονιά 4 σημείων και αντίστοιχες παρέες.



Σχήμα 10.8: Γειτονιά 8 σημείων και αντίστοιχες παρέες.

Ένα πεδίο Gibbs περιγράφει συνολικά το πεδίο των αποφάσεων σε όλη την εικόνα

$$Pr\{X(m, n) = x(m, n) : (m, n) \in L\} = \frac{e^{-U(x)}}{\Psi}, \quad (10.23)$$

όπου $U(x)$ ονομάζεται συνάρτηση ενέργειας, και είναι συνάρτηση MN μεταβλητών, όσα δηλαδή είναι τα σημεία της εικόνας. Ψ είναι μία σταθερά, $\Psi = \sum e^{-U(x)}$. Ιδιαίτερα ενδιαφέρουσα είναι η περίπτωση ενός πεδίου Gibbs, όπου η συνάρτηση ενέργειας ορίζεται με τη βοήθεια του συνόλου των παρεών, C , των σημείων της εικόνας

$$U(x) = \sum_{c \in C} \phi_c(x). \quad (10.24)$$

Το $\phi_c(x)$ ονομάζεται δυναμικό της παρέας c . Συχνά για τη γειτονιά 8 σημείων μόνο ζεύγη σημείων θεωρούνται ότι έχουν μη μηδενικό δυναμικό. Ένα παράδειγμα δυναμικού για διμελείς παρέες που μπορεί να χρησιμοποιηθεί για την τμηματοποίηση των εικόνων ακολουθεί

$$\phi_c(x) = \begin{cases} 0 & \text{εάν τα δύο σημεία ανήκουν στην ίδια κλάση} \\ \zeta > 0 & \text{διαφορετικά} \end{cases} \quad (10.25)$$

Ας θεωρήσουμε τώρα το πρόβλημα της τμηματοποίησης της εικόνας κατατάσσοντας τα σημεία της σε K κλάσεις. Υποθέτουμε ότι η παρατηρούμενη εικόνα για κάθε σημείο εξαρτάται μόνο από την κλάση που ανήκει αυτό το σημείο, με μία παρενόχληση από αθροιστικό θόρυβο, που παρακάτω θα θεωρήσουμε ότι ακολουθεί την κανονική κατανομή με διασπορά σ^2

$$p(y(m, n) | X(k, l) = x(k, l), (k, l) \in L) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y(m, n) - f(x(m, n)))^2}{2\sigma^2}}. \quad (10.26)$$

Η συνάρτηση $f(\cdot)$ δίδει τη μέση αντιπροσωπευτική τιμή για κάθε κλάση S_i

$$f(x(m, n)) = \mu_i, \quad \text{εάν } (m, n) \in S_i, \quad 0 \leq i \leq K - 1. \quad (10.27)$$

Η τμηματοποίηση της εικόνας μπορεί να γίνει μεγιστοποιώντας την a posteriori πιθανότητα του πεδίου κατάταξης των σημείων σε κλάσεις

$$\begin{aligned} & Pr\{X(m, n) = x(m, n) | Y(m, n) = y(m, n); (m, n) \in L\} \\ & \propto Pr\{X(m, n) = x(m, n) : (m, n) \in L\} \prod_{(m, n) \in L} p(y(m, n) | X(m, n) = x(m, n)) \end{aligned} \quad (10.28)$$

Αυτό ισοδυναμεί με την ελαχιστοποίηση της ακόλουθης ποσότητας

$$\sum_{(m, n) \in L} \phi_c(x(m, n)) + \frac{1}{2\sigma^2} \sum_{i=0}^{K-1} \sum_{(m, n) \in S_i} (y(m, n) - \mu_i)^2.$$

Ωστόσο το κόστος εύρεσης της βέλτιστης λύσης με διεξοδική αναζήτηση είναι εξαιρετικά υψηλό, γιατί το πλήθος των δυνατών καταστάσεων είναι K^{MN} , ενώ υπάρχουν πολλά τοπικά ελάχιστα. Η ελαχιστοποίηση μπορεί να επιτευχθεί με μία πιθανοκρατική αναζήτηση του ελάχιστου προσομοιώνοντας το πεδίο Gibbs. Ωστόσο και σε αυτή την περίπτωση το υπολογιστικό κόστος είναι απαγορευτικό για πολλές εφαρμογές.

Θα πρέπει όμως να διακρίνουμε την περίπτωση όπου $K = 2$, γιατί τότε υπάρχει αλγόριθμος που ευρίσκει το ελάχιστο σε πολυωνυμικό χρόνο. Πράγματι θέτοντας

$$\lambda(m, n) = \frac{1}{\sigma^2} ((y(m, n) - \mu_0)^2 - (y(m, n) - \mu_1)^2) = \frac{1}{\sigma^2} (2y(m, n) - \mu_0 - \mu_1)(\mu_1 - \mu_0), \quad (10.29)$$

μπορούμε να γράψουμε ότι η προς ελαχιστοποίηση ποσότητα ισούται με

$$- \sum_{(m,n) \in L} x(m,n) \lambda(m,n) + \zeta \sum_{\substack{(m,n) \in L \\ (k,l) \in \gamma(m,n)}} (x(m,n) - x(k,l))^2.$$

Αποδεικνύεται ότι το ανωτέρω πρόβλημα ελαχίστου ταυτίζεται με αυτό της εύρεσης της ελάχιστης τομής σε διμερή γράφο, και ισοδύναμα της επίλυσης του προβλήματος της μέγιστης ροής. Πράγματι η ανωτέρω ποσότητα διαφέρει κατά μία σταθερή ποσότητα από την ακόλουθη

$$\begin{aligned} C(x) = & \sum_{(m,n) \in L} x(m,n) \max(0, -\lambda(m,n)) + \sum_{(m,n) \in L} (1 - x(m,n)) \max(0, \lambda(m,n)) \\ & + \zeta \sum_{\substack{(m,n) \in L \\ (k,l) \in \gamma(m,n)}} (x(m,n) - x(k,l))^2. \end{aligned} \quad (10.30)$$

Άρα μπορούμε να θεωρήσουμε ένα γράφο με κόμβους τα σημεία της εικόνας και επιπλέον ένα κόμβο ‘πηγή’ (s) και ένα κόμβο ‘δεξαμενή’ (t). Ένα οποιοδήποτε σημείο της εικόνας είναι συνδεδεμένο είτε με την ‘πηγή’ είτε με τη ‘δεξαμενή’. Εφόσον $\lambda(m,n) > 0$, το σημείο (m,n) συνδέεται με την ‘πηγή’ και η χωρητικότητα της σύνδεσης είναι $\lambda(m,n)$. Ενώ εάν $\lambda(m,n) < 0$, το σημείο (m,n) συνδέεται με τη ‘δεξαμενή’ με χωρητικότητα σύνδεσης $-\lambda(m,n)$. Εάν ορίσουμε τα σύνολα $\{s\} \cup \{(m,n) : x(m,n) = 1\}$ και $\{t\} \cup \{(m,n) : x(m,n) = 0\}$, το ζητούμενο είναι η διαμέριση του συνόλου των σημείων της εικόνας, ώστε η τομή να έχει την ελάχιστη χωρητικότητα $C(x)$.

Εναλλακτικά, μπορεί να χρησιμοποιηθεί, για οποιονδήποτε αριθμό κλάσεων K , η μέθοδος των επαναλαμβανόμενων επικρατούσων τιμών. Από μία αρχική κατάταξη, που συχνά είναι αυτή που προκύπτει από απλή ανάλυση του ιστογράμματος, γίνεται μία επανεκτίμηση, ώστε να μεγιστοποιείται η ακόλουθη πιθανότητα

$$Pr\{X(m,n) = x(m,n) | Y(m,n) = y(m,n), \hat{x}(k,l); (k,l) \in L - \{(m,n)\}\}.$$

Το αποτέλεσμα της μεγιστοποίησης είναι η επικρατούσα τιμή του πεδίου για δοσμένη παρατήρηση, και με δοσμένη την αρχική εκτίμηση. Αποδεικνύεται ότι αρκεί να ελαχιστοποιηθεί η ακόλουθη ποσότητα σε κάθε σημείο

$$\frac{1}{2\sigma^2} (y(m,n) - \mu_i)^2 + \sum \phi_c(x(m,n)),$$

όπου αθροίζονται μόνο τα δυναμικά των παρεών που περιλαμβάνουν το σημείο (m,n) , κι όπου οι τιμές του πεδίου για τα γειτονικά σημεία είναι αυτές της προηγούμενης εκτίμησης. Αν επομένως από τα 8 γειτονικά σημεία, κατά την προηγούμενη εκτίμηση κατατάχθηκαν $\alpha_i(m,n)$ στην κλάση i , τότε η τρέχουσα κατάταξη βασίζεται στην τοπική ελαχιστοποίηση του κριτηρίου

$$(y(m,n) - \mu_i)^2 - \sigma^2 \zeta \alpha_i(m,n).$$

Όταν ευρεθούν οι επικρατούσες τιμές σε όλα τα σημεία του πεδίου η διαδικασία επαναλαμβάνεται με τον ίδιο ακριβώς τρόπο, και σταματά όταν μειωθεί σημαντικά ο αριθμός των σημείων που από μία επανάληψη στην επόμενη αλλάζουν κλάση, ή ιδανικά όταν ουδένα σημείο αλλάζει κλάση. Ο αλγόριθμος επαναλαμβανόμενων τοπικά επικρατούσων τιμών δίδεται κατωτέρω σε ψευδοκώδικα.

Initialization:

```
for each pixel p
{
  current_label(p) := init_label(p)
  next_state(p) := active
}
change_pixels := number_of_pixels
```

Iterations:

```
while (change_pixels != 0)
{
  change_pixels := 0
  for each pixel p
  {
    previous_label(p) := current_label(p)
    current_state(p) := next_state(p)
    next_state(p) := inactive
  }
  for each pixel p
  {
    if (current_state(p) == active)
    {
      current_label(p) := find_optimal_label(p)
      if (current_label(p) != previous_label(p))
      {
        change_pixels++
        next_state(neighbour(p)) := active
      }
    }
  }
}
```

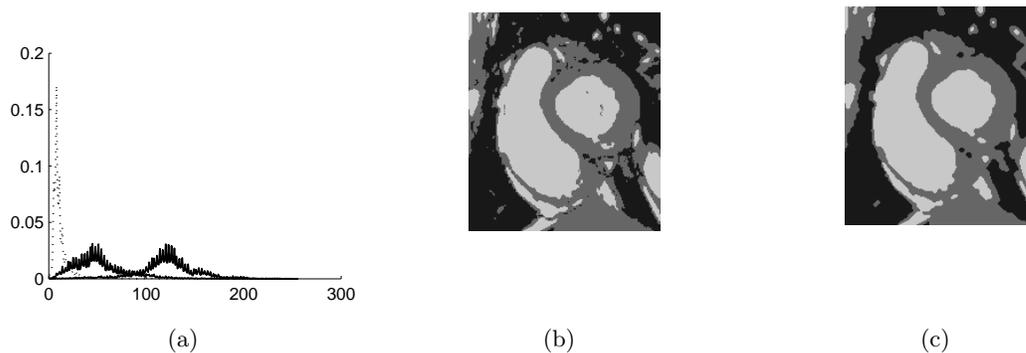
Στο Σχήμα 10.9 δίδονται εκτιμήσεις των πυκνοτήτων πιθανότητας για τρεις κλάσεις που προέκυψαν από ανάλυση της εικόνας του Σχήματος 10.6, η αρχική κατάταξη με βάση τις πιθανοφάνειες των τριών κλάσεων και το αποτέλεσμα της τμηματοποίησης με τη μέθοδο των επαναλαμβανόμενων επικρατουσών τιμών.

10.4 Τμηματοποίηση σε τοπολογικά συνεκτικές περιοχές

Σ' αυτή τη μέθοδο η συνεκτικότητα και η ομοιογένεια των περιοχών απαιτείται με άμεσο τρόπο. Ο αριθμός των κατηγοριών που κατατάσσονται τα σημεία της εικόνας θα μπορούσε να είναι άγνωστος και να προκύψει ως αποτέλεσμα της τμηματοποίησης, σε αντίθεση με τις προηγούμενες μεθόδους, όπου ο αριθμός αυτός ήταν δοσμένος.

Αν R είναι όλη η εικόνα σα σύνολο σημείων, ζητείται η τμηματοποίηση σε τοπολογικά συνεκτικές περιοχές R_1, R_2, \dots, R_K , ώστε

$$\bullet \bigcup_{k=1}^K R_k = R$$



Σχήμα 10.9: Τμηματοποίηση με τον αλγόριθμο επαναλαμβανόμενων επικρατουσών τιμών.

- $R_i \cap R_j = \emptyset, i \neq j$

Η τμηματοποίηση βασίζεται σ' ένα κριτήριο ομοιογένειας, που θα πρέπει να ικανοποιείται από όλα τα τμήματα της εικόνας, και να μην ικανοποιείται από την ένωση δύο γειτονικών τμημάτων, δηλαδή τμημάτων με κοινό σύνορο. Σαν κριτήριο ομοιογένειας μπορεί να χρησιμοποιηθεί η αντίθεση της φωτεινής έντασης σε μια περιοχή, μετρημένη είτε με τη μέγιστη αντίθεση, είτε με την τυπική απόκλιση της φωτεινής έντασης.

Με συσσωρεύσεις γύρω από κάποιο ή κάποια αφετηριακά σημεία μπορεί να υλοποιηθεί μια τμηματοποίηση που να ικανοποιεί τις παραπάνω απαιτήσεις. Δίδεται κατωτέρω ένας αλγόριθμος διαδοχικού προσεταιρισμού των σημείων της εικόνας με βάση ένα προκαθορισμένο κριτήριο ομοιότητας, $\rho(\cdot, \cdot)$, μέσω του οποίου προσδιορίζεται και η σειρά των κατατάξεων.

B1 Ονοματοθεσία των σημείων των αρχικών συνόλων.

B2 Εκτίμηση των χαρακτηριστικών των συνόλων (π.χ. μέση φωτεινότητα).

B3 Εισαγωγή στη διαταγμένη λίστα όλων των γειτονικών σημείων των αρχικών συνόλων.

B4 Όσο η λίστα δεν είναι κενή:

B4.1 Κατάταξη του πρώτου σημείου της λίστας α και αφαίρεσή του απ' αυτή.

B4.2 Έλεγχος των γειτονικών σημείων του α και ενημέρωση της λίστας:

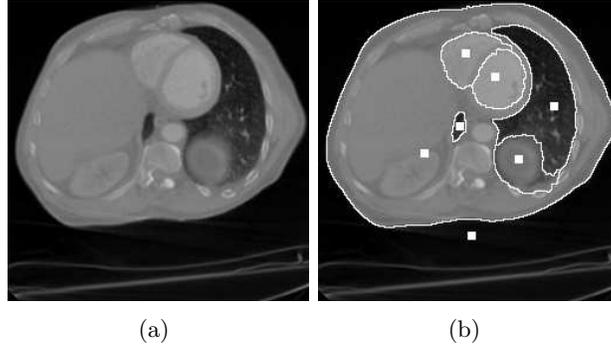
B4.2.1 Πρόσθεση των γειτόνων του α που

- δεν έχουν ήδη όνομα
- δεν είναι ήδη στη λίστα

σύμφωνα με την ομοιότητα $\rho(\cdot, \cdot)$.

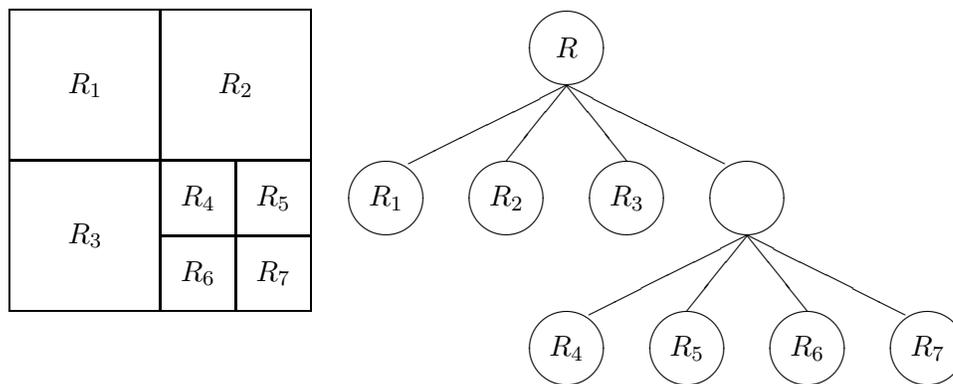
B4.2.2 Έλεγχος για γείτονες που είναι ήδη στη λίστα και που λόγω της κατάταξης του α συνορεύουν με ένα νέο σύνολο. Αυτά σημαδεύονται ως σημεία συνόρου. Επιπλέον, αν η ομοιότητα μειώνεται λόγω της νέας γειτνίασης, ανελίσσονται αντίστοιχα στη λίστα.

Ένα παράδειγμα εφαρμογής αυτής της μεθόδου δίδεται στο Σχήμα 10.10, όπου ορίζονται αφετηριακά μπλοκ σημείων και τηρείται σειρά προτεραιότητας για τις συνενώσεις, που γίνονται με κριτήριο την κατ' απόλυτο τιμή απόκλιση της φωτεινότητας του δοσμένου σημείου από τη μέση τιμή της περιοχής στην οποία γειτονεύει.



Σχήμα 10.10: Τομογραφική εικόνα της καρδιάς και τμηματοποίησή της με συσσωρεύσεις γύρω από τα αφετηριακά μπλοκ σημείων που ορίσθησαν από το χρήστη.

Ωστόσο το αποτέλεσμα επηρεάζεται από την επιλογή των αφετηριακών σημείων και επίσης απαιτείται μεγάλος αριθμός ελέγχων για την υπόθεση της ομοιογένειας. Το αποτέλεσμα είναι πιο σταθερό, αν η τμηματοποίηση επιτυγχάνεται με ένα συνδυασμό διαιρέσεων ανομοιογενών τμημάτων και συγχωνεύσεων γειτονικών ομοιογενών τμημάτων. Η υλοποίηση ενός τέτοιου αλγορίθμου διευκολύνεται με τη χρήση μιας πυραμιδωτής δομής για την παράσταση των τμημάτων της εικόνας. Στην κορυφή της πυραμίδας βρίσκεται ολόκληρη η εικόνα, και στη βάση της τα σημεία της εικόνας. Ισοδύναμα η τμηματοποίηση μπορεί να παρασταθεί με ένα δένδρο, του οποίου η ρίζα είναι ολόκληρη η εικόνα, και τα φύλλα του τα σημεία της εικόνας. Πιο συχνά το δένδρο είναι τετράκλωνο, που αντιστοιχεί σε κατάτμηση κάθε τμήματος της εικόνας σε τέσσερα ίσα τμήματα. Στο Σχήμα 10.11 δίδεται ένα παράδειγμα τέτοιας παράστασης των τμημάτων της εικόνας. Το αποτέλεσμα είναι πιο ενδιαφέρον, και η πολυπλοκότητα περιορισμένη, αν η εκκίνηση του αλγορίθμου γίνει σε ένα ενδιάμεσο επίπεδο της πυραμίδας.



Σχήμα 10.11: Τμήματα της εικόνας και αντίστοιχο τετράκλωνο δένδρο.

