



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ  
ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΒΙΝΤΕΟ ΚΑΙ ΠΟΛΥΜΕΣΩΝ

Κατάτμηση Φυσικών Εικόνων  
με Τεχνικές Αποσύνθεσης Διάμεσου Άξονα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Σπυρίδωνος Μ. Λεονάρδου

Επιβλέπων: Στέφανος Κόλλιας  
Καθηγητής Ε.Μ.Π.

Αθήνα, Μάρτιος 2012





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ  
ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΒΙΝΤΕΟ ΚΑΙ ΠΟΛΥΜΕΣΩΝ

Κατάτμηση Φυσικών Εικόνων  
με Τεχνικές Αποσύνθεσης Διάμεσου Άξονα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Σπυρίδωνος Μ. Λεονάρδου

Επιβλέπων: Στέφανος Κόλλιας  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 19η Μαρτίου 2012.

(Υπογραφή)

.....  
Στέφανος Κόλλιας  
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....  
Σταφυλοπάτης Ανδρέας-Γεώργιος  
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....  
Στάμου Γεώργιος  
Λέκτορας Ε.Μ.Π.

Αθήνα, Μάρτιος 2012

.....  
**Σπυρίδων Μ. Λεονάρδος**

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Σπυρίδων Μ. Λεονάρδος (2012) Εθνικό Μετσόβιο Πολυτεχνείο.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

# Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2011-2012 στο Εργαστήριο Ψηφιακής Επεξεργασίας Εικόνας, Βίντεο και Πολυμέσων του Εθνικού Μετσόβιου Πολυτεχνείου. Θα ήθελα να ευχαριστήσω τον επιβλέποντα Καθηγητή κ. Στέφανο Κόλλια για την εμπιστοσύνη που μου έδειξε αναθέτοντάς μου την εργασία αυτή και για τη δυνατότητα που μου έδωσε να ασχοληθώ με το συγκεκριμένο ενδιαφέρον θέμα. Επίσης, ευχαριστώ ιδιαίτερα τον Δρ. Ιωάννη Αβρίθη για την καθοδήγησή του και για την εξαιρετική συνεργασία που είχαμε. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου για την καθοδήγηση και την ηθική συμπαράσταση που μου προσέφεραν όλα αυτά τα χρόνια.



# Περίληψη

Στην παρούσα διπλωματική εργασία παρουσιάζονται τεχνικές κατάτμησης εικόνων βασισμένες σε μία διαδικασία αποσύνθεσης διάμεσου άξονα. Ξεκινώντας από μια gray-scale εικόνα ακμών υπολογίζουμε ένα σταθμισμένο μετασχηματισμό απόστασης και τον αντίστοιχο διάμεσο άξονα. Εφαρμόζοντας, τώρα, τον ίδιο μετασχηματισμό απόστασης από το διάμεσο άξονα ανάποδα, λαμβάνουμε μια αρχική κατάτμηση της εικόνας και ένα γράφο που αναπαριστά τη δομή της εικόνας. Αναπαριστώντας την εικόνα ως disjoint-set forest δομή, ενώνουμε γειτονικές περιοχές βάσει κάποιων κριτηρίων. Διάφορα κριτήρια υπολοποιήθηκαν και αξιολογήθηκαν. Κατά το πρώτο από αυτά, χρησιμοποιούμε το ύψος των σημείων σέλλας του διάμεσου άξονα για να εκφράσουμε την ομοιομορφία μεταξύ γειτονικών περιοχών. Μια δεύτερη διαφορετική κατεύθυνση που ακολουθούμε είναι η χρήση ενός μέτρου κλειστότητας των περιοχών της εικόνας για να αποφασίσουμε εάν πρέπει να ενώσουμε δύο γειτονικές περιοχές. Τρίτη και τελευταία κατεύθυνση είναι η χρήση ενός μέτρου ανομοιομορφίας μεταξύ περιοχών που ικανοποιεί την υπερμετρική ιδιότητα. Με αυτό τον τρόπο, υλοποιούμε την κατάτμηση ιεραρχικά. Όλες οι παραπάνω τεχνικές αξιολογούνται με βάση τη συλλογή δεδομένων του πανεπιστημίου Berkeley και συγκρίνονται με γνωστές τεχνικές της βιβλιογραφίας. Χωρίς μάθηση, πετυχαίνουμε πολύ καλά αποτελέσματα κοντά στη στάθμη της τεχνικής και με πρακτικούς χρόνους εκτέλεσης.

## Λέξεις Κλειδιά

Όραση υπολογιστών, επεξεργασία εικόνας, κατάτμηση, διάμεσος άξονας, σκελετός, αποσύνθεση διάμεσου άξονα, γράφοι γειτνίασης περιοχών, μέθοδοι ταχείας πόρευσης, σταθμισμένος μετασχηματισμός απόστασης.





# Abstract

In the framework of this thesis, we present new image segmentation techniques based on a weighted medial axis decomposition procedure. Starting from image contour map, we compute a weighted distance map and its weighted medial axis. Applying the same distance propagation from the medial axis backwards, we dually obtain an initial image partition and a graph representing image structure. Using a disjoint-set data structure, we then merge adjacent regions according to some criteria. Several different criteria were examined and tested. First, we use medial axis saddle point height to express similarity between adjacent regions and merge correspondingly. A second distinct direction we follow is to merge adjacent regions according to how fragmented they are. Last but not least, we use ultrametric contour map representation to implement hierarchical segmentation. As inter-region ultrametric dissimilarities, we use mean boundary strength on the common boundary between adjacent regions and inter-region fragmentation. All the above mentioned techniques are evaluated using the Berkeley Segmentation Dataset and compared with some state of the art algorithms. Without learning, we achieve performance near the state of the art with very practical running times.

## Keywords

Computer vision, image processing, segmentation, partition, grouping, medial axis, skeleton, medial axis decomposition, weighted distance transform, region adjacency graphs, exact group marching, fast marching methods, contour detection.



# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>19</b>
1.1	Ορίζοντας την κατάτμηση εικόνων . . . . .	19
1.2	Κίνητρα . . . . .	20
1.3	Σύνοψη . . . . .	22
<b>2</b>	<b>Βιβλιογραφική ανασκόπηση</b>	<b>25</b>
2.1	Γενικές τεχνικές κατάτμησης εικόνων . . . . .	25
2.2	Σχετικές εργασίες . . . . .	27
2.2.1	Μετασχηματισμός Watershed . . . . .	27
2.2.2	Κατευθυντικός μετασχηματισμός watershed και υπερμετρική εικόνα ορίων . . . . .	30
2.2.3	Αποδοτική κατάτμηση εικόνας με χρήση γράφων . . . . .	32
<b>3</b>	<b>Ανίχνευση ακμών</b>	<b>35</b>
3.1	Canny ανίχνευση ακμών . . . . .	35
3.2	Ανιχνευτής συνολικής πιθανότητας ακμών . . . . .	36
<b>4</b>	<b>Μετασχηματισμός απόστασης</b>	<b>41</b>
4.1	Θεωρητικό υπόβαθρο . . . . .	41
4.1.1	Ορισμοί . . . . .	41
4.1.2	Μετασχηματισμός απόστασης ως κυματική διάδοση . . . . .	42
4.1.3	Μετασχηματισμός απόστασης ως infimal συνέλιξη . . . . .	43
4.2	Αλγόριθμοι αποδοτικού υπολογισμού μετασχηματισμών απόστασης . . . . .	44
4.2.1	Λύνοντας την eikonal ΜΔΕ με τεχνικές ταχείας πόρευσης . . . . .	44

4.2.2	Σταθμισμένος μετασχηματισμός απόστασης και αλγόριθμος exact group marching . . . . .	46
<b>5</b>	<b>Μετασχηματισμός διάμεσου άξονα</b>	<b>51</b>
5.1	Εισαγωγή . . . . .	51
5.2	Υπολογισμός διάμεσου άξονα . . . . .	51
<b>6</b>	<b>Αρχική κατάτμηση με αποσύνθεση του διάμεσου άξονα</b>	<b>57</b>
6.1	Αποσύνθεση του διάμεσου άξονα . . . . .	57
6.2	Δυϊκός μετασχηματισμός απόστασης . . . . .	58
<b>7</b>	<b>Τεχνικές συνένωσης γειτονικών περιοχών</b>	<b>61</b>
7.1	Αποδοτική συνένωση περιοχών με βάση την μεταξύ τους ομοιομορφία . . . . .	61
7.2	Συνένωση περιοχών με βάση τον παράγοντα κατακερματισμού σχήματος . . . . .	63
7.3	Ιεραρχική κατάτμηση εικόνων και υπερμετρική εικόνα ορίων . . . . .	64
7.3.1	Υπερμετρική ανομοιομορφία με βάση των παράγοντα κατακερματισμού κοινού συνόρου . . . . .	66
7.3.2	Υπερμετρική ανομοιομορφία με βάση την ισχύ των ορίων . . . . .	67
<b>8</b>	<b>Αξιολόγηση</b>	<b>69</b>
8.1	Συλλογή δεδομένων . . . . .	69
8.2	Κριτήρια αξιολόγησης . . . . .	69
8.2.1	Precision-recall . . . . .	69
8.2.2	Variation of information . . . . .	71
8.2.3	Δείκτης Rand . . . . .	71
8.2.4	Επικάλυψη . . . . .	72
8.3	Αποτελέσματα . . . . .	72
8.4	Αποτελέσματα αξιολόγησης . . . . .	79
8.5	Σχολιασμός . . . . .	80
<b>9</b>	<b>Συμπεράσματα</b>	<b>83</b>
9.1	Συμπεράσματα . . . . .	83
9.2	Μελλοντικές κατευθύνσεις . . . . .	83





## Κατάλογος Πινάκων

- 8.1 Συνολικά αποτελέσματα αξιολόγησης με βάση τα όρια των ανιχνευθέντων περιοχών. . . . . 80
- 8.2 Συνολικά αποτελέσματα αξιολόγησης με βάση τις ανιχνευθείσες περιοχές. 81





# Κατάλογος Σχημάτων

1.1	Προτεινόμενη μεθοδολογία κατάτμησης . . . . .	23
2.1	Οπτικοποίηση του μετασχηματισμού watershed . . . . .	28
2.2	Διαδικασία πλημμυρισμού στον μετασχηματισμό watershed . . . . .	29
2.3	Υπερκατάτμηση που προκύπτει από εφαρμογή μετασχηματισμού watershed απευθείας στο μέτρο της κλίσης . . . . .	30
2.4	Προσδιορισμός σημαδευτών μέσω της εξαγωγής μεγίστων . . . . .	30
2.5	Κατάτμηση βιοϊατρικής εικόνας με χρήση εργαλείων σήμανσης . . . . .	31
3.1	Canny ανίχνευση ακμών . . . . .	36
3.2	Κατευθυντική κλίση με διαφορές ιστογράμματος . . . . .	37
3.3	Ανιχνευτής γκριζων ακμών $Pb$ . . . . .	38
3.4	Ιδιοδιανύσματα της εικόνας και πληροφορία σχετική με τις ακμές . . . . .	39
3.5	Ανιχνευτής ακμών $gPb$ . . . . .	40
4.1	Μετασχηματισμός απόστασης με ευκλείδιο μετρικό . . . . .	42
4.2	Μετασχηματισμός απόστασης με ευκλείδιο μετρικό και οπτικοποιημένος ως κυματική διάδοση . . . . .	43
4.3	Urwind κατασκευή αποδεκτών τιμών της λύσης . . . . .	45
4.4	Σταθμισμένος μετασχηματισμός απόστασης υπολογισμένος με τον αλγόριθμο exact group marching. . . . .	49
5.1	Διαδικασίες σχήμα και ο αντίστοιχος διάμεσος άξονας . . . . .	52
5.2	Διαδικασίες σχήματα και αντίστοιχοι διάμεσοι άξονες . . . . .	52
5.3	Διάμεσος άξονας για πολλαπλές τιμές της παραμέτρου scale . . . . .	54
5.4	Υπολογισμός διάμεσου άξονα με τον αλγόριθμο WMA . . . . .	55

6.1	Παράδειγμα δυϊκότητας πηγών και διάμεσου άξονα στη μία διάσταση . . .	59
6.2	Διαδικασία αρχικής κατάτμησης . . . . .	60
6.3	Επίδραση της παραμέτρου κλίμακας $\sigma$ της συνάρτησης $f$ στην αρχική κατάτμηση . . . . .	60
8.1	Ανθρώπινες κατατμήσεις εικόνων του Berkeley Segmentation Dataset .	70
8.2	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm . . . . .	73
8.3	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm . . . . .	74
8.4	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm . . . . .	75
8.5	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-sfm . . . . .	76
8.6	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-ucm . . . . .	77
8.7	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-ucm . . . . .	78
8.8	Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-ucm . . . . .	79

# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Ορίζοντας την κατάτμηση εικόνων

Το πρόβλημα της κατάτμησης εικόνων είναι ένα από τα πιο θεμελιώδη προβλήματα στον τομέα της όρασης υπολογιστών. Αν και έχει μελετηθεί από τα πρώτα χρόνια της όρασης υπολογιστών, η κατάτμηση εικόνων παραμένει ακόμα μια μεγάλη πρόκληση. Από την εποχή του Gestalt κινήματος στην ψυχολογία, ήταν γνωστό ότι η αντιληπτική ομαδοποίηση διαδραματίζει έναν ισχυρό ρόλο στην ανθρώπινη οπτική αντίληψη.

Πιο συγκεκριμένα, η κατάτμηση περιγράφεται ως η διαδικασία χωρισμού μιας εικόνας σε διακριτές περιοχές, καθεμία από τις οποίες είναι ομοιογενής όσον αφορά ένα ή πολλαπλά χαρακτηριστικά όπως η φωτεινότητα, η υφή, το χρώμα, η κίνηση και άλλα. Ο στόχος της κατάτμησης είναι να συσταδοποιηθούν τα εικονοστοιχεία (pixels) σε προεξέχουσες περιοχές εικόνας, δηλ. περιοχές που αντιστοιχούν σε επιφάνειες, αντικείμενα ή φυσικά μέρη αντικειμένων. Ως εκ τούτου, κάποιος μπορεί εύκολα να δει ότι δεν υπάρχει μοναδικός ορισμός της κατάτμησης. Στην πραγματικότητα, είναι αδύνατο να διατυπωθούν γενικά οι ακριβείς στόχοι της κατάτμησης, όπως ο Marr [25] έχει υπογραμμίσει. Κατά συνέπεια, μπορεί να συναχθεί το συμπέρασμα ότι η κατάτμηση είναι εξαρτημένη από την εκάστοτε εφαρμογή και ότι η εύρεση μιας μοναδικής λύσης είναι, γενικά, διφορούμενη.

Ένας πιο τυπικός ορισμός της κατάτμησης [16] παρουσιάζεται στη συνέχεια. Έστω  $X$  το χωρίο που καταλαμβάνεται από μια εικόνα. Σε αυτήν την περίπτωση, η διαδικασία κατάτμησης εικόνας μπορεί να θεωρηθεί ως η διαδικασία που χωρίζει το χωρίο  $X$  σε  $n$  στις διακριτές περιοχές  $R_1, R_2, \dots, R_n$  με έναν τέτοιο τρόπο ώστε οι ακόλουθες ιδιότητες να ικανοποιούνται:

(α) Το σύνολο  $R_i$  είναι συνεκτικό για κάθε  $i = 1, 2, 3, \dots, n$

(β)  $\bigcup_{i=1}^n R_i = X$

(γ)  $R_i \cap R_j = \emptyset$  για κάθε  $i, j$  με  $i \neq j$

Η συνθήκη (α) δηλώνει πως όλα τα στοιχεία μιας περιοχής  $R_i$  πρέπει να συνδέονται σύμφωνα με έναν προκαθορισμένο τρόπο. Η επόμενη συνθήκη απαιτεί η διαδικασία της κατάτμησης να είναι πλήρης, δηλαδή κάθε εικονοστοιχείο πρέπει να ανήκει σε κάποια περιοχή  $R_i$ . Σύμφωνα με την τρίτη συνθήκη, όλες οι περιοχές στις οποίες χωρίζεται μια εικόνα πρέπει να είναι διακριτές ή ισοδύναμα αμοιβαία αποκλειόμενες.

Επιπροσθέτως, για τον τερματισμό της διαδικασίας της κατάτμησης πρέπει να ισχύουν δύο ακόμα συνθήκες. Δοθέντος ενός κατηγορήματος  $Q(R_i)$  ορισμένου για όλες τις διακριτές περιοχές  $R_i$ , η διαδικασία της κατάτμησης πρέπει να τερματίζει όταν:

(δ)  $Q(R_i) = \text{αληθές}$  για κάθε  $i = 1, 2, 3, \dots, n$

(ε)  $Q(R_i \cap R_j) = \text{ψευδές}$  για κάθε ζεύγος γειτονικών περιοχών  $R_i$  και  $R_j$

όπου δύο περιοχές  $R_i$  και  $R_j$  είναι γειτονικές αν και μόνο αν η ένωσή τους είναι συνεκτικό σύνολο. Η συνθήκη (δ) δηλώνει πως το κατηγορήμα  $Q$  πρέπει να είναι αληθές για κάθε περιοχή  $R_i$  ενώ δύο γειτονικές περιοχές  $R_i$  και  $R_j$  πρέπει να διαχωρίζονται σύμφωνα με το  $Q$ .

Η κατάτμηση εικόνων είναι ένα από τα πιο δημοφιλή προβλήματα στην όραση υπολογιστών επειδή συνδέεται με ένα μεγάλο εύρος εφαρμογών. Είναι στενά συνδεδεμένη με την ανίχνευση και αναγνώριση αντικειμένων καθώς υπάρχει εκτενής βιβλιογραφία που συνδυάζει τα δύο αυτά προβλήματα [14, 15, 21, 40]. Άλλες δημοφιλείς εφαρμογές της κατάτμησης εικόνων είναι η συμπίεση εικόνων με βάση αντικείμενα [39], ανάκτηση εικόνων με βάση το περιεχόμενο [6], ανάλυση βιοϊατρικών εικόνων [33], ανίχνευση προσώπου [22, 43] και αναγνώριση δακτυλικών αποτυπωμάτων [28].

## 1.2 Κίνητρα

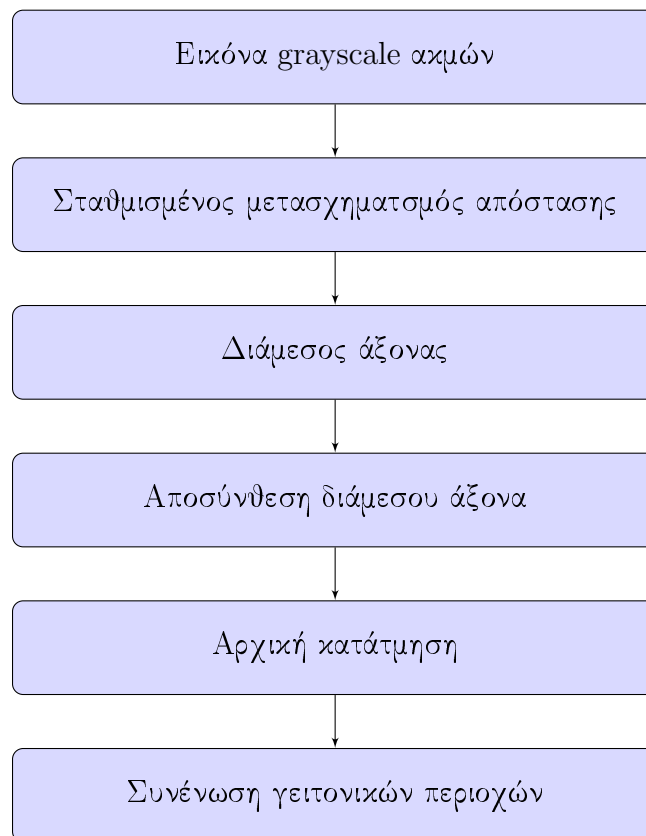
Σύμφωνα με σημαντικά επιστημονικά ευρήματα της ψυχοφυσικής προκύπτει πως κοιτάζοντας μια φυσική σκηνή ο εγκέφαλος διαφοροποιεί τη σκηνή αυτή σε περιοχές και σε περιγράμματα που σχηματίζουν τα πιθανά σημαντικά μέρη των αντικειμένων. Κατ' αρχάς, οι Kovacs και Julesz [18] έδειξαν ότι η ανίχνευση κλειστών περιγραμμάτων σε μια σκηνή από το ανθρώπινο σύστημα όρασης είναι πιο εύκολη από την ανίχνευση ανοιχτών καμπυλών. Επιπλέον, κατέληξαν στο ότι η ανίχνευση των κλειστών περιγραμμάτων υποβοηθείται από την ανίχνευση χαρακτηριστικών στο εσωτερικό των περιοχών τα οποία συνέδεσαν με το διάμεσο άξονα των σχημάτων και αντικειμένων της σκηνής. Παρόμοια ευρήματα υπάρχουν και σε άλλες παρόμοιες έρευνες [19, 20]. Συμπερασματικά, υπάρχει ισχυρή επιστημονική ένδειξη ότι ο ανθρώπινος εγκέφαλος χρησιμοποιεί το διάμεσο άξονα για να βοηθήσει στην κατάτμηση σκηνής και την οπτική ομαδοποίηση. Το κίνητρο για μία μέθοδο κατάτμησης βασισμένη στο διάμεσο άξονα προέρχεται από αυτά τα στοιχεία από κοινού με το γεγονός ότι ο διάμεσος άξονας δεν έχει μελετηθεί εκτενώς ως μέσο για κατάτμηση εικόνων.

Η μέθοδος κατάτμησης εικόνων που παρουσιάζεται στα πλαίσια της παρούσας δι-

πλωματικής εργασίας βασίζεται στον *medial feature detector*. Οι Αβρίθης και Ραπαντζίκος [5] ανέπτυξαν μία μέθοδο ανίχνευσης περιοχών ενδιαφέροντος. Παρόλο που η εργασία αυτή έχει διαφορετικό στόχο, παράγει μια αρχική κατάτμηση (*oversegmentation*) της εκάστοτε εικόνας. Εισήγαγαν μία τεχνική αποσύνθεσης του σκελετού σε συνιστώσες για να λάβουν την αρχική κατάτμηση. Η αναπαράσταση αυτή είναι πιο περιγραφική από τον κλασικό μετασχηματισμό *watershed* [7]. Χρησιμοποιώντας την αναπαράσταση αυτή μπορούμε να μετρήσουμε πόσο κατακερματισμένα ή συμπαγή είναι τα όρια των περιοχών και οι ανιχνευθείσες κλειστές καμπύλες.

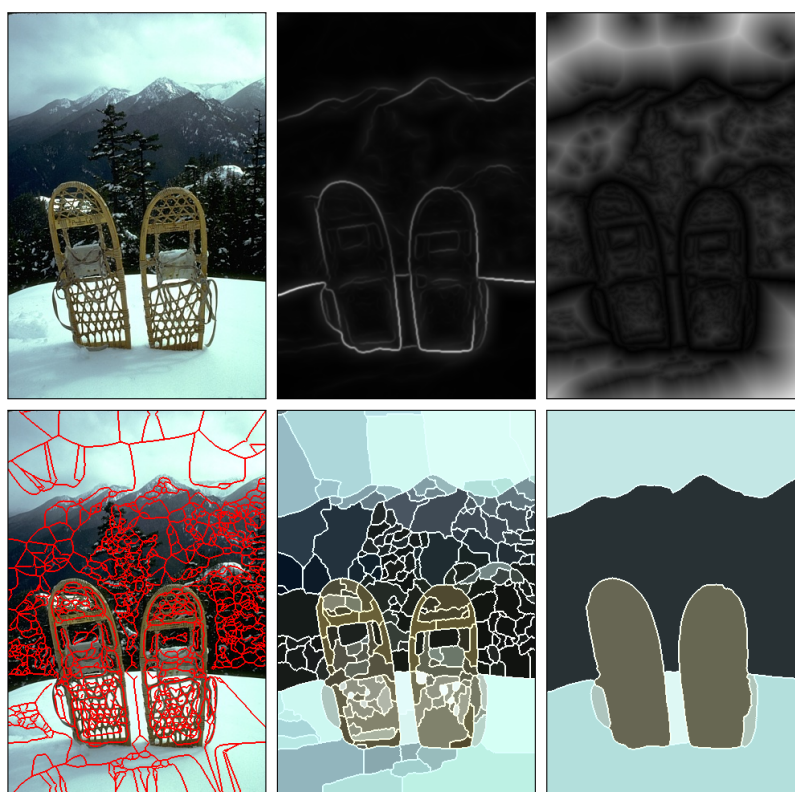
Η συνεισφορά της παρούσας εργασίας είναι η υλοποίηση, η εξέταση και η αξιολόγηση διαφόρων κριτηρίων και τεχνικών συνένωσης περιοχών με στόχο να μετατρέψουμε την αρχική κατάτμηση σε μία πιο “καλή” κατάτμηση. Επειδή ο όρος “καλή” κατάτμηση είναι ασαφής, ορίζουμε ως “καλή” κατάτμηση στα πλαίσια αυτής της διπλωματικής εργασίας την κατάτμησης που είναι όσο το δυνατόν κοντινότερη στις ανθρώπινες κατατμήσεις όπως αυτές δίνονται από το *Berkeley Segmentation Dataset* [3, 26].

Συνολικά, η προτεινόμενη μέθοδος κατάτμησης εικόνων συνοψίζεται στα ακόλουθα βήματα:



## 1.3 Σύνοψη

Η παρούσα εργασία έχει δομηθεί βάσει των διακριτών βημάτων της προτεινόμενης μεθόδου. Το **Κεφάλαιο 2** περιέχει μια βιβλιογραφική ανασκόπηση διαφόρων μεθοδολογιών κατάτμησης εικόνων. Εκτενέστερη αναφορά έχει γίνει σε μεθόδους σχετικές με την προτεινόμενη. Στο **Κεφάλαιο 3** περιγράφονται οι δύο τεχνικές ανίχνευσης grayscale ακμών που χρησιμοποιήθηκαν. Στο **Κεφάλαιο 4** περιέχονται ορισμοί και θεωρία σχετικά με μετασχηματισμούς απόστασης καθώς και βασικοί αλγόριθμοι για αποδοτικό υπολογισμό διαφορετικών μετασχηματισμών απόστασης. Στο **Κεφάλαιο 5** παρουσιάζεται η έννοια του διάμεσου άξονα ή σκελετού καθώς και ο αντίστοιχος αλγόριθμος υπολογισμού του. Η τεχνική αποσύνθεσης του σκελετού και η μέθοδος που προσδίδει μια αρχική κατάτμηση της εικόνας εισόδου μέσω ενός δυϊκού μετασχηματισμού απόστασης περιγράφονται αναλυτικά στο **Κεφάλαιο 6**. Στο **Κεφάλαιο 7** γίνεται η παρουσίαση των διαφόρων τεχνικών συνένωσης περιοχών που υλοποιήθηκαν. Ακολουθεί η αξιολόγηση των προηγούμενων τεχνικών και η σύγκρισή τους με γνωστές τεχνικές της βιβλιογραφίας χρησιμοποιώντας τη συλλογή δεδομένων του πανεπιστημίου Berkeley στο **Κεφάλαιο 8** καθώς και η παράθεση κάποιων αντιπροσωπευτικών αποτελεσμάτων. Συμπεράσματα και πιθανές μελλοντικές κατευθύνσεις αναλύονται στο **Κεφάλαιο 9**. Το **Παράρτημα Α'** περιέχει βασικές αποδείξεις διαφόρων λημμάτων και ιδιοτήτων που αναφέρθηκαν στα προηγούμενα κεφάλαια.



Σχήμα 1.1: Πάνω, από αριστερά προς τα δεξιά: εικόνα εισόδου, grayscale εικόνα ακμών και μετασχηματισμός απόστασης. Κάτω, από αριστερά προς τα δεξιά: διάμεσος άξονας (κόκκινο χρώμα), αρχική και τελική κατάτμηση.





## Κεφάλαιο 2

# Βιβλιογραφική ανασκόπηση

### 2.1 Γενικές τεχνικές κατάτμησης εικόνων

Η απλούστερη τεχνική κατάτμησης εικόνων είναι η *κατωφλίωση*. Ορίζεται ένα κατώφλι αρχικά και εν συνεχεία κάθε εικονοστοιχείο συγκρίνεται με το κατώφλι αυτό. Εάν η τιμή της φωτεινότητας του εκάστοτε εικονοστοιχείου είναι μεγαλύτερη από το κατώφλι τότε το εικονοστοιχείο αυτό σημειώνεται ως *foreground* ενώ στην αντίθετη περίπτωση ως *background*. Το κατώφλι λαμβάνει συνήθως τιμές φωτεινότητας ή χρώματος. Άλλες πιο σύνθετες τεχνικές κατωφλίωσης επιτρέπουν στο κατώφλι να λαμβάνει πολλαπλές τιμές για την ίδια εικόνα. Το κρίσιμο σημείο όλων των τεχνικών κατωφλίωσης είναι η επιλογή της τιμής του κατωφλίου. Για το σκοπό αυτό διάφορες τεχνικές έχουν αναπτυχθεί όπως η μέθοδος της μέγιστης εντροπίας και η μέθοδος του Otsu [32] (*maximum variance*). Η τεχνική της κατωφλίωσης θεωρείται ιδιαίτερα απλοϊκή και μπορεί να χρησιμοποιηθεί μόνο για πολύ απλές περιπτώσεις.

Μία άλλη κατηγορία τεχνικών κατάτμησης εικόνων αποτελείται από αυτές που βασίζονται στη *συμπύεση* δεδομένων. Σύμφωνα με αυτές τις τεχνικές, η βέλτιστη κατάτμηση είναι αυτή που ελαχιστοποιεί το μήκος κωδικοποίησης των δεδομένων. Οι Mobahi et al. [29] περιγράφουν κάθε περιοχή της εικόνας με βάση την υφή του και το σχήμα του συνόρου του. Για κάθε δοθείσα κατάτμηση μιας εικόνας, η εν λόγω μέθοδος δίνει έναν αριθμό από bits που απαιτούνται για να κωδικοποιηθεί η εικόνα. Με αυτόν τον τρόπο, για κάθε πιθανή κατάτμηση της εικόνας ο στόχος είναι η εύρεση της κατάτμησης με τον ελάχιστο αριθμό bits. Αυτό επιτυγχάνεται με μια απλή συσταδοποίηση. Η αλλοίωση (*distortion*) στη συμπύεση με απώλειες ορίζει πόσο χονδροειδής είναι και η τελική κατάτμηση. Έχει το αρνητικό πως η βέλτιστη τιμή για την αλλοίωση μπορεί να διαφέρει από εικόνα σε εικόνα.

Μία μεγάλη κατηγορία τεχνικών κατάτμησης περιλαμβάνει μεθόδους *συσταδοποίησης* (*clustering*). Μία κύρια αντιπροσωπευτική τεχνική είναι ο αλγόριθμος *k-means*, μια επαναληπτική τεχνική που χρησιμοποιείται για τον χωρισμό της εικόνας σε  $K$  κέντρα. Τα βασικά βήματα του εν λόγω αλγορίθμου έχουν ως εξής:

1. Αρχικοποίηση των  $K$  κέντρων.
2. Κάθε εικονοστοιχείο αντιστοιχίζεται στο κοντινότερο κέντρο.
3. Επανυπολογισμός των κέντρων.
4. Επανάληψη των βημάτων 2 και 3 έως να συγκλίνει ο αλγόριθμος.

Στην περίπτωση αυτή, η απόσταση είναι η τετραγωνισμένη ή απόλυτη διαφορά μεταξύ ενός εικονοστοιχείου και ενός κέντρου μιας συστάδας. Η διαφορά συνήθως βασίζεται στο χρώμα, στην ένταση, στην υφή, στην χωρική απόσταση ή σε ένα σταθμισμένο συνδυασμό αυτών. Η μέθοδος θα συγκλίνει εγγυημένα αλλά μπορεί να μην επιστρέψει τη βέλτιστη λύση. Ένα μειονέκτημα του  $k$ -means αλγορίθμου είναι ότι ο αριθμός των συστάδων είναι μια παράμετρος εισόδου. Μια ακατάλληλη επιλογή του  $K$  μπορεί να αποφέρει πενιχρά αποτελέσματα. Ο αλγόριθμος υποθέτει επίσης ότι η διακύμανση αποτελεί το κατάλληλο μέτρο της διασποράς των συστάδων.

Οι μέθοδοι που βασίζονται στα *ιστογράμματα* μιας εικόνας είναι ιδιαίτερα αποδοτικές συγκριτικά με άλλες τεχνικές κατάτμησης εικόνων επειδή πρακτικά χωρίζουν μια εικόνα σε περιοχές με ένα μοναδικό πέρασμα από κάθε εικονοστοιχείο. Σε αυτήν την οικογένεια τεχνικών, ένα ιστόγραμμα υπολογίζεται από όλα τα εικονοστοιχεία της εικόνας και οι κορυφές και οι κοιλάδες του ιστογράμματος χρησιμοποιούνται για την ανίχνευση των συστάδων της εικόνας [37]. Ως μέτρο για τον υπολογισμό του ιστογράμματος χρησιμοποιούνται συνήθως η φωτεινότητα της εικόνας και το χρώμα. Υπάρχει το εμφανές μειονέκτημα στην κατηγορία αυτή πως είναι δύσκολο να βρεθούν οι κορυφές και οι κοιλάδες των ιστογραμμάτων.

Οι τεχνικές *κατάτμησης γράφων* μπορούν να χρησιμοποιηθούν αποτελεσματικά για την κατάτμηση εικόνων. Σύμφωνα με τις μεθόδους αυτές, η εικόνα αναπαρίσταται ως ένας μη κατευθυνόμενος γράφος με βάρη. Οι κορυφές του γράφου αναπαριστούν είτε μεμονωμένα εικονοστοιχεία είτε ομάδες αυτών. Οι ακμές του γράφου συνδέουν γειτονικά εικονοστοιχεία ή περιοχές της εικόνας και τα βάρη των ακμών εκφράζουν ένα μέτρο ομοιομορφίας ή ανομοιομορφίας μεταξύ των γειτονικών εικονοστοιχείων ή περιοχών. Δημοφιλείς αλγόριθμοι της εν λόγω κατηγορίας είναι ο αλγόριθμος *normalized cuts* [36], ο αλγόριθμος *minimum cut* [41] και ο αλγόριθμος αποδοτικής συνένωσης με βάση το ελάχιστο συνεκτικό δέντρο [13].

Η *φασματική θεωρία γράφων* [10] και συγκεκριμένα το κριτήριο *normalized cuts* [36] παρέχουν μία τεχνική που συμπεριλαμβάνει συνολική πληροφορία από ολόκληρη την εικόνα στη διαδικασία της κατάτμησης. Στον αλγόριθμο *normalized cuts*, η εικόνα αναπαρίσταται ως μη κατευθυνόμενος γράφος με βάρη. Οι κορυφές αναπαριστούν τα εικονοστοιχεία και ακμές υπάρχουν μεταξύ όλων των πιθανών ζευγαριών. Τα βάρη των ακμών αποτελούν ένα μέτρο ομοιομορφίας μεταξύ γειτονικών κορυφών. Ο γράφος μπορεί να χωριστεί σε δύο διακριτά σύνολα  $A, B$  απλά αφαιρώντας τις μεταξύ τους ακμές. Ο βαθμός ομοιομορφίας μεταξύ των δύο αυτών συνόλων υπολογίζεται ως το άθροισμα των βαρών των μεταξύ τους ακμών. Αυτή η ποσοτητα ονομάζεται *τομή (cut)*. Η βέλτιστη κατάτμηση του γράφου σε δύο σύνολα είναι αυτή που ελαχιστοποιεί την τιμή της τομής. Ωστόσο, χρησιμοποιώντας το κριτήριο ελαχίστης τομής προκύπτουν μικρά σύνολα απομονωμένων κορυφών του γράφου. Οι Wu και Leahy [41] πρότειναν μια

μέθοδο κατάτμησης βασισμένη στο κριτήριο της ελάχιστης τομής και παρατήρησαν το προηγούμενο πρόβλημα στην εργασία τους. Για την επίλυση του προβλήματος αυτού, οι Shi και Malik [36] εισήγαγαν ένα κανονικοποιημένο μέτρο ομοιομορφίας μεταξύ δύο συνόλων λαμβάνοντας υπόψιν όλες τις ακμές του γράφου. Το μέτρο αυτο ονομάστηκε *normalized cut* ( $Ncut$ ) και ορίζεται ως εξής:

$$Ncut(A, B) = \frac{cut(A, B)}{assoc(A, V)} + \frac{cut(A, B)}{assoc(B, V)} \quad (2.1)$$

όπου  $assoc(A, V)$  είναι η συνολική σύνδεση των κορυφών του  $A$  προς όλες τις κορυφές του γράφου. Οι Shi και Malik απέδειξαν πως η εύρεση της βέλτιστης κατάτμησης του γράφου σε δύο σύνολα ελαχιστοποιώντας την ποσότητα  $Ncut$  ανάγεται σε ένα πρόβλημα γενικευμένων ιδιοτιμών. Ο γράφος (εικόνα) κατατμείται σε δύο σύνολα βάση του ιδιοδιανύσματος με τη δεύτερη μικρότερη ιδιοτιμή και στη συνέχεια κάθε σύνολο μπορεί να χωριστεί επαναληπτικά σε περισσότερα υποσύνολα. Ωστόσο, η προσέγγιση αυτή μπορεί να σπάσει ομοιόμορφες περιοχές της εικόνας σε περισσότερες υποπεριοχές. Πρόσφατα, μια πολυκλιμακωτή παράλλαξη του αλγορίθμου *normalized cuts* Cour et al. [12] βελτίωσε το πρόβλημα αυτό καθώς και κατέστησε πιο αποδοτικό υπολογιστικά τον αρχικό αλγόριθμο.

Πολλές προσεγγίσεις του προβλήματος της κατάτμησης εικόνων θέτουν ένα πρόβλημα λογισμού μεταβολών. Ένα τέτοιο μοντέλο προταθηκε από τους Mumford και Shah [30], όπου η κατάτμηση της εικόνας εισόδου  $u_0$  δίνεται από την ελαχιστοποίηση του συναρτησιακού:

$$\mathcal{F}(u, C) = \int_{\Omega} (u - u_0)^2 dx + \mu \int_{\Omega \setminus C} |\nabla u|^2 dx + \nu \oint_C ds \quad (2.2)$$

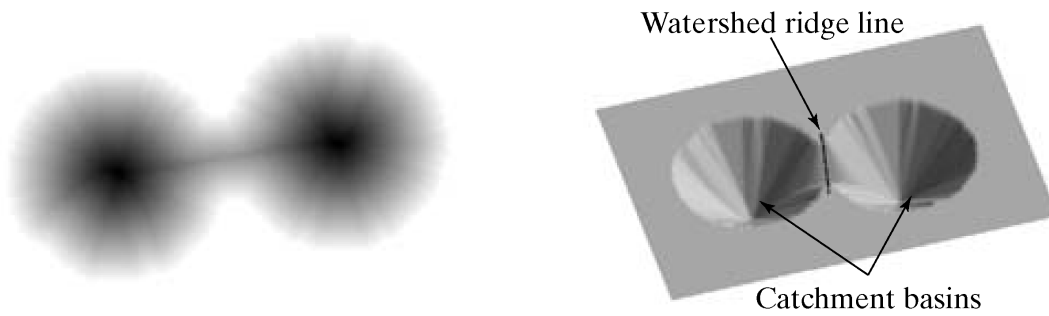
όπου  $C$  είναι τα κλειστά και ομαλά περιγράμματα της κατάτμησης,  $u$  η τμηματικά ομαλή προσέγγιση του  $u_0$  με ασυνέχειες μόνο κατα μήκος της καμπύλης  $C$ ,  $\Omega$  το χωρίο της εικόνας και  $\mu, \nu$  παράγοντες στάθμισης. Διάφοροι αλγόριθμοι έχουν αναπτυχθεί για την ελαχιστοποίηση του (2.2) καθώς και της απλουστευμένης έκδοσής του όπου η  $u$  είναι τμηματικά σταθερή στο  $\Omega \setminus C$ .

## 2.2 Σχετικές εργασίες

### 2.2.1 Μετασχηματισμός Watershed

Μία ευρέως διαδεδομένη μεθοδολογία κατάτμησης εικόνων είναι ο *μετασχηματισμός watershed*. Καθώς η μεθοδολογία αυτή είναι σχετική με την προτεινόμενη τεχνική της διπλωματικής αυτής εργασίας θα παρουσιαστεί εκτενώς. Η ανάλυση που ακολουθεί βασίζεται στα [7, 16].

Προκειμένου να κατανοήσει κάποιος τη λογική του μετασχηματισμού αυτού θα πρέπει να συλλάβει με το μυαλό του μια εικόνα ορισμένη στο χώρο των τριών διαστάσεων: δύο χωρικών συντεταγμένων και μίας έντασης. Σε μία τέτοια τοπολογία



Σχήμα 2.1: Εικόνα εισόδου και αντίστοιχη τοπογραφική επιφάνεια.

θεωρούμε τρεις κατηγορίες σημείων: (α) σημεία τοπικών ελαχίστων, (β) σημεία που εάν τοποθετήσουμε μία σταγόνα νερού αυτή θα κυλήσει προς ένα μοναδικό τοπικό ελάχιστο και (γ) σημεία που εάν τοποθετήσουμε μία σταγόνα νερού, αυτή διαθέτει ίσες πιθανότητες να κυλήσει προς περισσότερα τοπικά ελάχιστα. Για κάθε τοπικό ελάχιστο, το σύνολο των σημείων που ικανοποιούν τη συνθήκη (β) είναι γνωστό ως *λεκάνη απορροής ύδατος*. Τα σημεία που ανήκουν στην τρίτη κατηγορία είναι γνωστά ως *γραμμές διαχωρισμού* και αποτελούν τα όρια των περιοχών της κατάτμησης.

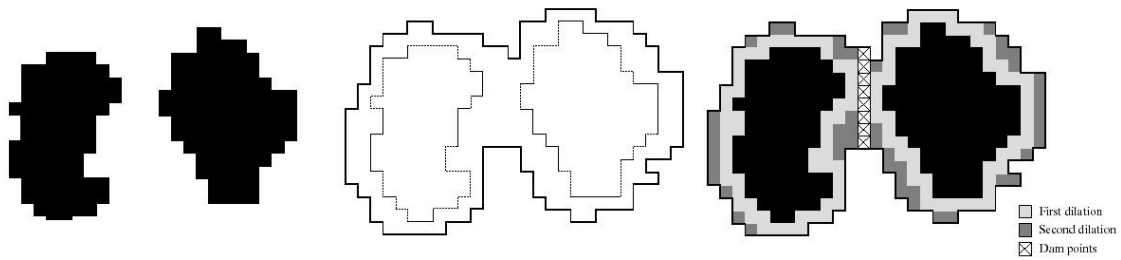
Ο βασικός στόχος των αλγορίθμων κατάτμησης που βασίζονται σε αυτές τις έννοιες είναι ο προσδιορισμός των γραμμών διαχωρισμού. Η βασική ιδέα είναι απλή. Υποθέστε πως σε κάθε ελάχιστο περιοχής ανοίγουμε μια οπή και στη συνέχεια πλημμυρίζουμε την τοπογραφία στο σύνολό της από κάτω προς τα πάνω, αφήνοντας το νερό να ανέβει με σταθερό ρυθμό μέσα από αυτές τις οπές. Όταν το νερό που ανεβαίνει μέσα στις διακριτές λεκάνες απορροής είναι έτοιμο να συνενωθεί, κατασκευάζεται ένα φράγμα προκειμένου να αποτραπεί αυτή η συνένωση. Επομένως, το πλημμύρισμα της τοπολογίας θα φτάσει τελικά σε ένα στάδιο όπου μόνο τα φράγματα θα είναι ορατά πάνω από το νερό. Τα φράγματα αυτά είναι ακριβώς οι γραμμές διαχωρισμού. Συνεπώς, αυτές θα αποτελούν τα συνδεδεμένα περιγράμματα που θα έχουν εξαχθεί από το συγκεκριμένο αλγόριθμο κατάτμησης.

Μία από τις κύριες εφαρμογές αυτού του τύπου κατάτμησης είναι η εξαγωγή σχεδόν ομοιόμορφων αντικειμένων από το background της εικόνας. Οι περιοχές που χαρακτηρίζονται από μικρές μεταβολές στην ένταση έχουν και μικρές τιμές κλίσης. Στην πράξη, λοιπόν, αυτός ο τύπος κατάτμησης εφαρμόζεται στην κλίση μιας εικόνας και όχι στην ίδια την εικόνα.

Πιο αναλυτικά, έστω  $M_1, M_2, \dots, M_R$  τα τοπικά ελάχιστα μιας εικόνας  $g(x, y)$ ,  $C(M_i)$  οι αντίστοιχες λεκάνες απορροής και  $minlevel$  και  $maxlevel$  το ελάχιστο και το μέγιστο αντίστοιχα επίπεδο της  $g(x, y)$ . Το σύνολο  $T[n]$  ορίζεται ως:

$$T[n] = \{(s, t) : g(s, t) < n\} \quad (2.3)$$

Ας υποθέσουμε τώρα πως η τοπογραφία πλημμυρίζει με νερό, με το ύψος της πλημμύρας να αυξάνει σε ακέραια βήματα από την τιμή  $n = minlevel + 1$  έως την τιμή  $n = maxlevel + 1$ . Εάν συμβολίσουμε με  $C_n(M_i)$  τις λεκάνες απορροής του ελαχίστου  $M_i$



Σχήμα 2.2: **Αριστερά:** δύο μερικώς πλημμυρισμένες λεκάνες απορροής στο βήμα  $n - 1$ . **Κέντρο:** πλημμύρα στο βήμα  $n$  που μας δείχνει πως το νερό έχει ξεχυθεί ανάμεσα στις λεκάνες. **Δεξιά:** αποτέλεσμα της διαστολής και κατασκευή φράγματος (από [16])

στο βήμα  $n$  έχουμε:

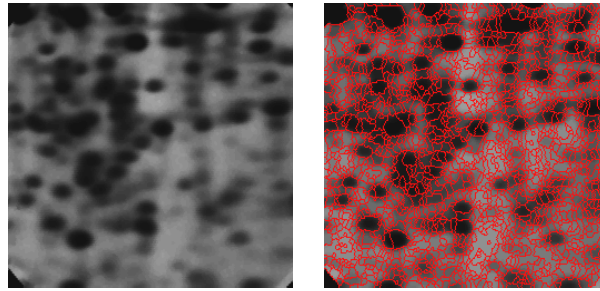
$$C_n(M_i) = C(M_i) \cap T[n] \quad (2.4)$$

Στη συνέχεια, ας συμβολίσουμε με  $C[n]$  την ένωση των πλημμυρισμένων λεκανών απορροής στο βήμα  $n$ . Θα ισχύει πως  $C[n] = \bigcup_{i=1}^R C_n(M_i)$  και προφανώς  $C[\maxlevel + 1] = \bigcup_{i=1}^R C(M_i)$ . Ο αλγόριθμος αρχικοποιείται με  $C[\minlevel + 1] = T[\minlevel + 1]$ . Ο αλγόριθμος σε κάθε βήμα κατασκευάζει το  $C[n]$  με δεδομένο το  $C[n - 1]$ . Το σύνολο των συνεκτικών συνιστώσων του  $T[n]$  συμβολίζεται με  $Q[n]$ . Για κάθε συνιστώσα  $q \in Q[n]$  υπάρχουν τρεις πιθανότητες:

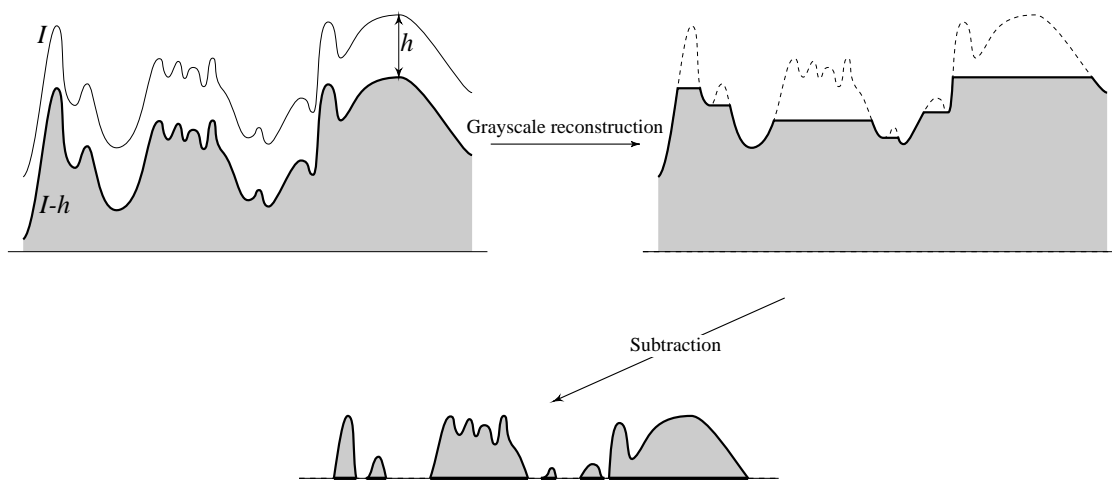
1. Το  $q \cap C[n - 1]$  είναι κενό. Τότε ένα καινούργιο ελάχιστο έχει συναντηθεί και η συνιστώσα  $q$  ενσωματώνεται στο σύνολο  $C[n - 1]$  για να σχηματιστεί το  $C[n]$ .
2. Το  $q \cap C[n - 1]$  περιέχει μόνο μία συνεκτική συνιστώσα του  $C[n - 1]$ . Τότε, η συνιστώσα  $q$  ενσωματώνεται στο σύνολο  $C[n - 1]$  για να σχηματιστεί το  $C[n]$ .
3. Το  $q \cap C[n - 1]$  περιέχει πάνω από μία συνεκτικές συνιστώσες του  $C[n - 1]$ . Τότε, θα πρέπει να κατασκευάσουμε ένα ή περισσότερα φράγματα προκειμένου να αποτρέψουμε την υπερχείλιση του νερού.

Η άνωθι διαδικασία επαναλαμβάνεται έως ότου  $n = \maxlevel + 1$

Η απευθείας εφαρμογή μετασχηματισμού watershed στο μέτρο της κλίσης της εικόνας έχει ως αποτέλεσμα μια υπερχατάττηση της εικόνας όπως αυτή που παρουσιάζεται στο Σχ. 2.3. Για την αποφυγή του προβλήματος αυτού, έχει εισαχθεί η έννοια των εργαλείων σήμανσης. Τα εργαλεία σήμανσης αποτελούν συνδεδεμένες συνιστώσες της εικόνας που ελέγχουν τη διαδικασία πλημμυρισμού με στόχο την αποφυγή της υπερχατάττησης. Ένα βασικό παράδειγμα εξαγωγής τέτοιων εργαλείων σήμανσης είναι η γενικευμένη εξαγωγή μεγίστων-ελαχίστων (*maxima-minima extraction*) [38]. Για τα μέγιστα, αφαιρείται μια τυχαία σταθερά  $h$  από την αρχική εικόνα  $I$  και πραγματοποιείται ένα μορφολογικό άνοιγμα ανακατασκευής της  $I$  από την  $I - h$ . Η ανασκευασμένη



Σχήμα 2.3: Εικόνα εισόδου και αντίστοιχη υπερκατάτμηση που προκύπτει από εφαρμογή μετασχηματισμού watershed απευθείας στο μέτρο της κλίσης.

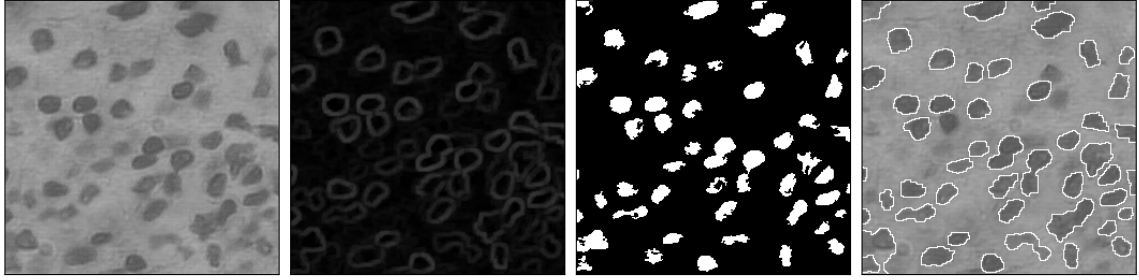


Σχήμα 2.4: Προσδιορισμός σημαδευτών μέσω της εξαγωγής μεγίστων (από [38]).

εικόνα αφαιρείται στη συνέχεια από την αρχική προσδίδοντας μια grayscale εικόνα  $J$ . Από την εικόνα  $J$  είναι εύκολη η εξαγωγή των πιο σημαντικών μεγίστων μέσω μιας απλής κατωφλίωσης. Μια δυϊκή διαδικασία μπορεί να χρησιμοποιηθεί για την εξαγωγή των ελαχίστων της εικόνας  $I$ .

## 2.2.2 Κατευθυντικός μετασχηματισμός watershed και υπερμετρική εικόνα ορίων

Μια πρόσφατη τροποποίηση του κλασικού μετασχηματισμού watershed που αξιολογεί πληροφορία για την ισχύ αλλά και την κατεύθυνση των συνόρων των περιοχών της κατάτμησης είναι ο κατευθυντικός μετασχηματισμός watershed (*oriented watershed transform*) [3]. Ως σήμα εισόδου στον άνωθι αλγόριθμο μπορεί να χρησιμοποιηθεί οποιοσδήποτε grayscale εικόνα ακμών  $E(x, y, \theta)$  που εκφράζει την πιθανότητα ύπαρξης ακμής στο σημείο  $(x, y)$  με κατεύθυνση  $\theta$ .



Σχήμα 2.5: Από αριστερά προς τα δεξιά: αρχική εικόνα, μορφολογική κλίση, εργαλεία σήμανσης μεγίστων, τελική κατάτμηση.

Αρχικά, εφαρμόζεται ο μετασχηματισμός watershed στην εικόνα:

$$E(x, y) = \max_{\theta} E(x, y, \theta) \quad (2.5)$$

Οι λεκάνες απορροής ( $\mathcal{P}_0$ ) των τοπικών ελαχίστων της  $E(x, y)$  αποτελούν τις περιοχές της αρχικής υπερκατάτμησης και οι γραμμές διαχωρισμού μεταξύ των περιοχών ( $\mathcal{K}_0$ ) τα πιθανά σύνορα των περιοχών της εικόνας.

Στη συνέχεια, ορίζεται ένα μέτρο ανομοιομορφίας μεταξύ γειτονικών περιοχών βασισμένο στην ισχύ των μεταξύ τους ορίων. Η χρήση της μέσης τιμής της  $E(x, y)$  ως βάρος του κάθε ορίου μπορεί να εισάγει θόρυβο. Για τη διόρθωση του προβλήματος αυτού προτάθηκε η ακόλουθη τροποποίηση: τα όρια των περιοχών προσεγγίζονται αρχικά από ευθύγραμμα τμήματα. Στη συνέχεια, σε κάθε εικονοστοιχείο  $(x, y)$  αντιστοιχίζεται η κατεύθυνση  $\alpha(x, y) \in [0, \pi)$  του ευθυγράμμου τμήματος στο οποίο ανήκει. Η ισχύς του ορίου στο  $(x, y)$  γίνεται  $E(x, y, \alpha(x, y))$  αντί για  $E(x, y)$ .

Επιπροσθέτως, το παραπάνω μέτρο ανομοιομορφίας μεταξύ γειτονικών περιοχών ικανοποιεί την υπερμετρική ιδιότητα [1]. Το επόμενο βήμα είναι η χρήση της υπερμετρικής εικόνας ορίων (ultrametric contour map (UCM)). Μια αναλυτική περιγραφή των υπερμετρικών εικόνων ορίων γίνεται στην παράγραφο 7.3. Η ιεραρχία των περιοχών κατασκευάζεται από έναν άπληστο αλγόριθμο συνένωσης περιοχών που βασίζεται σε γράφους. Ορίζεται ένας αρχικός γράφος  $G = (\mathcal{P}_0, \mathcal{K}_0)$  όπου οι κορυφές αντιστοιχούν στις περιοχές  $\mathcal{P}_0$ , οι ακμές στα όρια  $\mathcal{K}_0$  μεταξύ γειτονικών περιοχών και τα βάρη  $W : \mathcal{K}_0 \rightarrow \mathbb{R}_+$  αποτελούν ένα μέτρο ανομοιομορφίας μεταξύ περιοχών. Ο αλγόριθμος αρχικά ταξινομεί τις ακμές του γράφου κατά αύξουσα σειρά βαρών και επαναληπτική συνενώνει τις πιο όμοιες περιοχές:

- 1) Επιλογή ορίου ελαχίστου βάρους:  $C^* = \arg \min_{C \in \mathcal{K}_0} W(C)$
- 2) Έστω  $R_1, R_2 \in \mathcal{P}_0$  οι περιοχές που χωρίζονται από το  $C^*$
- 3)  $R = R_1 \cup R_2$  και ανανέωση του γράφου:  
 $\mathcal{P}_0 \leftarrow \mathcal{P}_0 \setminus \{R_1, R_2\} \cup R$  ανδ  $\mathcal{K}_0 \leftarrow \mathcal{K}_0 \setminus \{C^*\}$
- 4) Τέλος εάν το σύνολο  $\mathcal{K}_0$  είναι κενό. Αλλιώς, ανανέωση των βαρών  $W(\mathcal{K}_0)$  και επανάληψη από το βήμα 1.

### 2.2.3 Αποδοτική κατάτμηση εικόνας με χρήση γράφων

Ο αλγόριθμος κατάτμησης εικόνας που προτείνουν οι Felzenswalb και Huttenlocher [13] επιχειρεί να χωρίσει την εικόνα σε περιοχές έτσι ώστε η τελική κατάτμηση να μην είναι ούτε πολύ λεπτομερής ούτε πολύ χονδροειδής. Στα πλαίσια της άνωθι εργασίας, οι όροι “πολύ λεπτομερής”, “σωστά λεπτομερέστερη κατάτμηση” “πολύ χονδροειδής” έχουν ως εξής:

**Ορισμός 2.2.1** Μια κατάτμηση  $S$  είναι πολύ λεπτομερής εάν υπάρχει ένα ζεύγος περιοχών  $R_1, R_2 \in S$  για τις οποίες δεν υπάρχει ένδειξη για ύπαρξη συνόρου μεταξύ τους.

**Ορισμός 2.2.2** Δοθέντων δύο κατατμήσεων  $S$  και  $T$  της ίδιας εικόνας, η  $T$  αποτελεί μια σωστά λεπτομερέστερη κατάτμηση της  $S$  εάν κάθε περιοχή της  $T$  περιέχεται σε κάποια περιοχή της  $S$ . Επιπροσθέτως, για να αποτελεί η κατάτμηση  $T$  μια σωστά λεπτομερέστερη κατάτμηση της  $S$  πρέπει  $T \neq S$ .

**Ορισμός 2.2.3** Μια κατάτμηση  $S$  είναι πολύ χονδροειδής εάν υπάρχει σωστά λεπτομερέστερη κατάτμηση της  $S$  που δεν είναι πολύ λεπτομερής.

Σε αυτήν την προσέγγιση, η εικόνα αναπαρίσταται ως ένας μη κατευθυνόμενος γραφος  $G = (V, E)$ . Οι κορυφές  $v_i \in V$  αντιστοιχούν στα εικονοστοιχεία και οι ακμές  $(v_i, v_j) \in E$  αντιστοιχούν σε ζεύγη γειτονικών κορυφών. Κάθε ακμή  $(v_i, v_j) \in E$  έχει ένα βάρος  $w(v_i, v_j)$ , που εκφράζει ένα μη αρνητικό μέτρο ανομοιομορφίας μεταξύ γειτονικών κορυφών (π.χ. διαφορά φωτεινότητας ή χρώματος).

Η εσωτερική διαφορά μιας περιοχής (συνιστώσας)  $R \subseteq V$  ορίζεται ως το μέγιστο βάρος στο ελάχιστο συνεκτικό δέντρο της περιοχής αυτής  $MST(R, E)$ :

$$\text{Int}(R) = \max_{e \in MST(R, E)} w(e) \quad (2.6)$$

Η διαφορά μεταξύ δύο περιοχών  $R_1, R_2 \subseteq V$  ορίζεται ως το ελάχιστο βάρος ακμής που συνδέει τις δύο αυτές περιοχές:

$$\text{Dif}(R_1, R_2) = \min \{w(v_i, v_j) : v_i \in R_1, v_j \in R_2, (v_i, v_j) \in E\} \quad (2.7)$$

Για την ύπαρξη ορίου μεταξύ δύο περιοχών έχει εισαχθεί ένα λογικό κατηγορημα που ελέγχει εάν η διαφορά των δύο εν λόγω περιοχών είναι μεγάλη συγκρινόμενη με την εσωτερική διαφορά τουλάχιστον μίας εκ των δύο περιοχών. Για μικρές περιοχές, η εσωτερική διαφορά μιας περιοχής δεν αποτελεί μια καλή εκτίμηση. Για το λόγο αυτόν, γίνεται χρήση μιας συνάρτησης κατωφλίου  $\tau(R)$  η οποία είναι βασισμένη στο μέγεθος της περιοχής. Για μικρές περιοχές απαιτείται μεγαλύτερη ένδειξη για ύπαρξη συνόρου. Το κατηγορημα για δύο περιοχές ορίζεται ως εξής:

$$D(R_1, R_2) = \begin{cases} \text{αληθ'ες} , & \text{ε'αν } \text{Dif}(R_1, R_2) > \text{MIInt}(R_1, R_2) \\ \text{ψευδ'ες} , & \text{αλλι'ως} \end{cases} \quad (2.8)$$



όπου η ελάχιστη εσωτερική διαφορά  $MInt$  ορίζεται ως:

$$MInt(R_1, R_2) = \min(\text{Int}(R_1) + \tau(R_1), \text{Int}(R_2) + \tau(R_2)) \quad (2.9)$$

Σύμφωνα με τον κύριο αλγόριθμο, κάθε κορυφή αντιστοιχεί αρχικά σε μία ξεχωριστή περιοχή. Η επεξεργασία των ακμών γίνεται κατά αύξουσα σειρά βαρών. Σε κάθε βήμα ο αλγόριθμος συνενώνει τις δύο περιοχές που συνδέει η εκάστοτε ακμή εάν το αντίστοιχο κατηγορημα είναι ψευδές. Η πολυπλοκότητα του αλγορίθμου αυτού είναι  $O(n \log n)$ , όπου  $n$  είναι το πλήθος των εικονοστοιχείων. Τα βάρη των ακμών είναι συνάρτηση της απόλυτης διαφοράς φωτεινότητας μεταξύ εικονοστοιχείων. Ένα σημαντικό χαρακτηριστικό της τεχνικής αυτής είναι πως διατηρεί λεπτομέρειες σε περιοχές μικρής μεταβλητότητας ενώ αγνοεί λεπτομέρειες σε περιοχές μεγάλης μεταβλητότητας.



# Κεφάλαιο 3

## Ανίχνευση ακμών

### 3.1 Canny ανίχνευση ακμών

Αρχικά, πρέπει να αναφέρουμε πως ο όρος grayscale ακμές αναφέρεται σε μια εικόνα γκριζου τόνου (grayscale) που εκφράζει την πιθανότητα ύπαρξης ορίου περιοχής σε κάθε εικονοστοιχείο. Ως βασικό ανιχνευτή γκριζων ακμών χρησιμοποιήθηκε μια grayscale έκδοση του ανιχνευτή ακμών Canny [9].

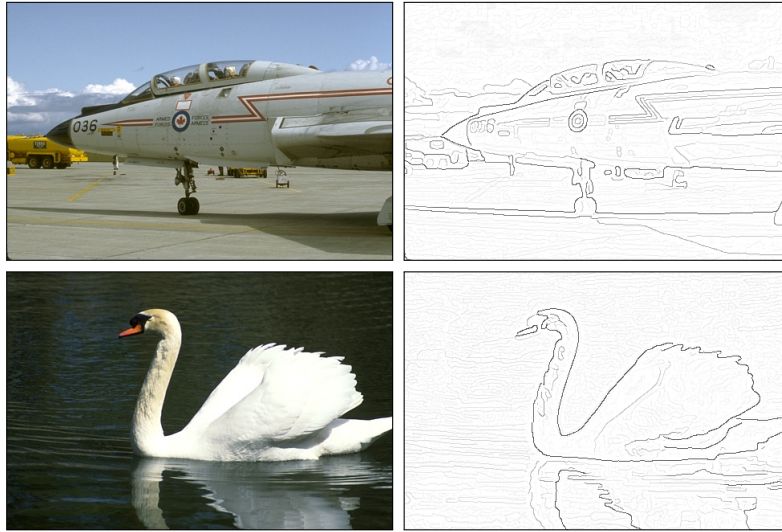
Το πρώτο βήμα της ανίχνευσης ακμών με τη μέθοδο του Canny είναι η συνέλιξη της εικόνας εισόδου με ένα γκαουσιανό φίλτρο για αποθορυβοποίηση. Στη συνέχεια, πρώτης τάξης γκαουσιανές παράγωγοι χρησιμοποιούνται για τον υπολογισμό της διαστάτης κλίσης της εικόνας. Η διαδικασία αυτή δίνει παράγωγο στην οριζοντια διεύθυνση ( $G_x$ ) και στην κάθετη διεύθυνση ( $G_y$ ). Από αυτές υπολογίζεται το μέτρο της κλίσης  $G$  και η διεύθυνση  $\Theta$  ως εξής:

$$G = \sqrt{G_x^2 + G_y^2} \quad (3.1)$$

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (3.2)$$

Η διεύθυνση της κλίσης στρογγυλοποιείται σε 4 κατευθύνσεις (0, 45, 90 και 135 μοίρες). Στη συνέχεια, ακολουθεί η διαδικασία της καταπίεσης των μη μεγίστων (non-maximum suppression) κατά την οποία διατηρούνται μόνο τα εικονοστοιχεία που αποτελούν τοπικά μέγιστα στην κατεύθυνση της κλίσης ενώ στα υπόλοιπα θέτουμε την τιμή 0.

Το τελευταίο βήμα είναι η καταφλίωση με υστέρηση (hysteresis thresholding). Η λειτουργία αυτή χρησιμοποιεί δύο τιμές καταφλίου: μία υψηλή και μία χαμηλή για να ανιχνεύσει αντιστοίχα τις “ισχυρές” και τις “αδύναμες” ακμές. Οι αδύναμες ακμές υπολογίζονται ως ακμές μόνο όταν συνδέονται με ισχυρές ακμές. Οπότε η μέθοδος Canny είναι λιγότερο πιθανό, σε σχέση με άλλες τεχνικές, να είναι ευαίσθητη στο θόρυβο. Συνήθως, η χαμηλή τιμή καταφλίου είναι ένα σταθερό κλάσμα της υψη-



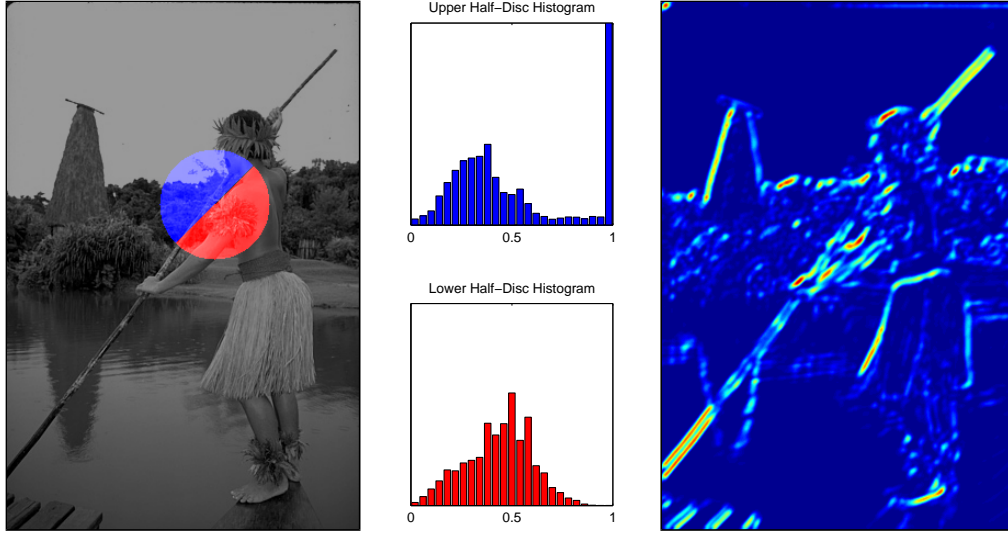
Σχήμα 3.1: Εικόνες εισόδου και οι αντίστοιχες grayscale εικόνες ακμών που προκύπτουν από τον τροποποιημένο αλγόριθμο Canny

λής τιμής (π.χ. 1/3). Χρησιμοποιώντας πολλαπλές τιμές για την υψηλή τιμή κατωφλίου προκύπτει η επιθυμητή εικόνα ακμών γκριζου τόνου. Παραδείγματα παρουσιάζονται στο Σχ. 3.1. Η υλοποίηση που χρησιμοποιήθηκε στα πλαίσια της παρούσας διπλωματικής εργασίας βασίζεται στην υλοποίηση του D.R. Martin η οποία είναι διαθέσιμη στον σύνδεσμο <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/code/Detectors/>.

## 3.2 Ανιχνευτής συνολικής πιθανότητας ακμών

Στο σημείο αυτό θα γίνει μια σύντομη περιγραφή του δεύτερου ανιχνευτή ακμών που χρησιμοποιήθηκε ως είσοδος στην προτεινόμενη μεθοδολογία κατάτμησης εικόνων. Αυτός είναι ο ανιχνευτής συνολικής πιθανότητας ακμών (globalized probability of boundary (*gPb*)) [2]. Η τεχνική αυτή συνδυάζει πολλαπλά τοπικά στοιχεία όπως φωτεινότητα, χρώμα, υφή σε πολλαπλές κλίμακες και τα συνδυάζει με πληροφορία προερχόμενη από το σύνολο της εικόνας. Είναι ο αλγόριθμος στη στάθμη της τεχνικής και παράγει αποτελέσματα πλησιέστερα στην ανθρώπινη όραση από οποιοδήποτε άλλο γνωστό σήμερα αλγόριθμο.

Αρχικά, ο Martin et al. [27] ορίζουν τη συνάρτηση  $Pb(x, y, \theta)$  για να περιγράψουν την πιθανότητα ύπαρξης ακμής στο σημείο  $(x, y)$  με κατεύθυνση  $\theta$  μετρώντας τοπικές διαφορές στα κανάλια της φωτεινότητας, των χρωμάτων και της υφής. Ο Arbelaez et al. [2] εισάγουν τον ανιχνευτή συνολικής πιθανότητας ακμών (globalized probability of boundary (*gPb*)). Η τεχνική αυτή περιέχει μια πολυκλιμακωτή έκδοση του προηγούμενου ανιχνευτή  $Pb$  και ένα επιπρόσθετο στάδιο φασματικής συσταδοποίησης για την εισαγωγή πληροφορίας από το σύνολο της εικόνας.



Σχήμα 3.2: Από αριστερά προς τα δεξιά: κανάλι φωτεινότητας με κυκλικό δίσκο κατεύθυνσης  $\theta = \pi/4$  με κέντρο ένα τυχαίο εικονοστοιχείο, ιστογράμματα των δύο ημικυκλικών δίσκων, τελικό αποτέλεσμα στο κανάλι φωτεινότητας για τη δεδομένη κατεύθυνση.

Το βασικό στάδιο του  $Pb$  ανιχνευτή ακμών είναι ο υπολογισμός ενός σήματος  $G(x, y, \theta)$  από μια εικόνα έντασης  $I$ . Σε κάθε εικονοστοιχείο  $(x, y)$  τοποθετείται ένας κυκλικός δίσκος και χωρίζεται σε δύο ίσα μέρη από μια διάμετρο στη διεύθυνση  $\theta$ . Το μέτρο  $G$  στο εκάστοτε σημείο ορίζεται από την  $\chi^2$  μεταξύ των δύο ιστογραμμάτων  $g$  και  $h$  που αντιστοιχούν στους δύο ημικυκλικούς δίσκους:

$$\chi^2 = \frac{1}{2} \sum_i \frac{(g(i) - h(i))^2}{g(i) + h(i)} \quad (3.3)$$

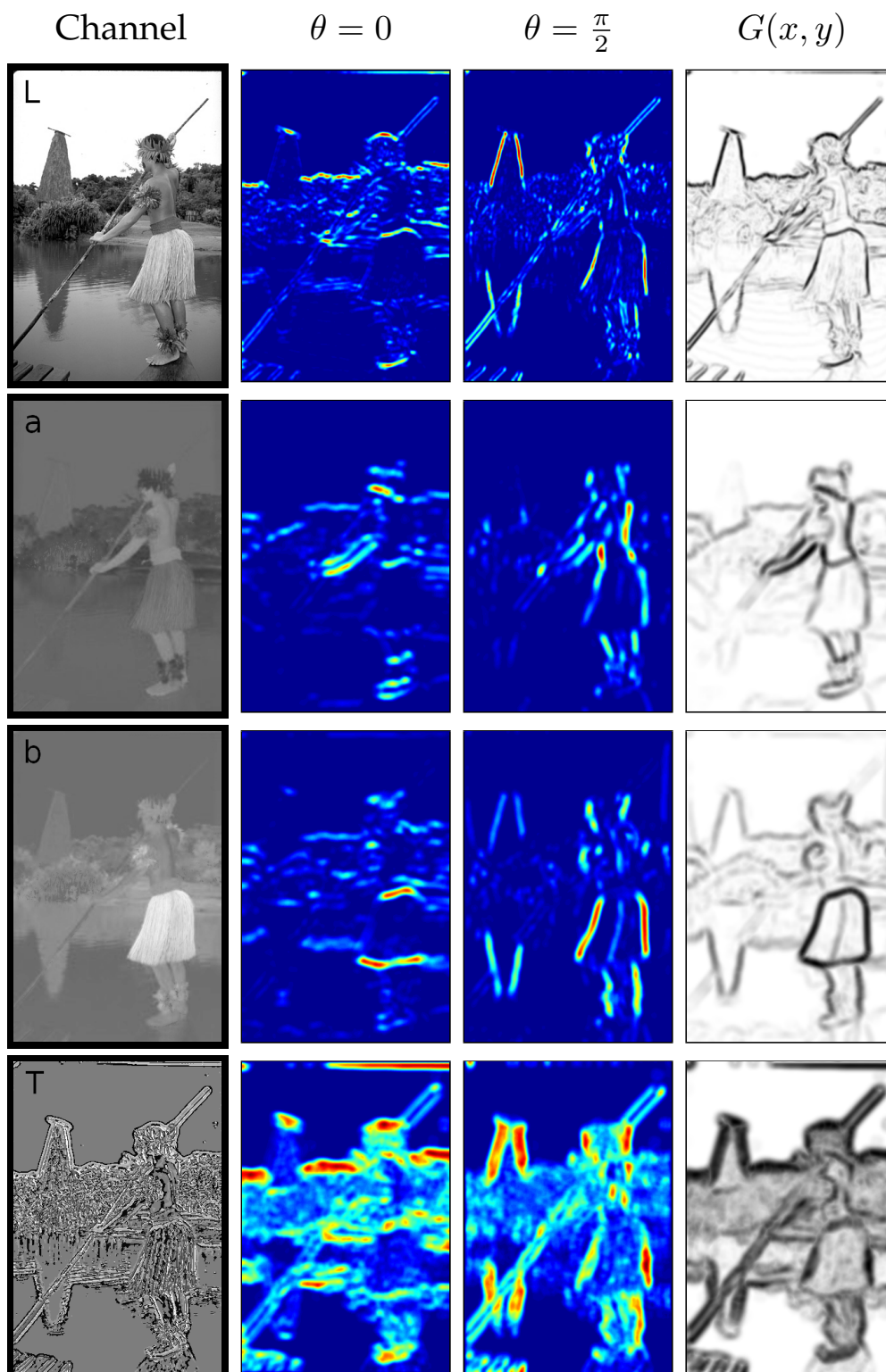
Ένα παράδειγμα της άνωθι διαδικασίας παρουσιάζεται στο Σχ. 3.2. Συνολικά τέσσερα κανάλια χρησιμοποιηθήκαν, τρία του χρωματικού χώρου CIELab ( $L, a, b$ ) και το τέταρτο είναι το κανάλι υψής. Το πρώτο αυτό στάδιο της τεχνικής  $gPb$  παρουσιάζεται στο Σχ. 3.2.

Όλα τα παραπάνω αποτελούν τον  $Pb$  ανιχνευτή. Για την ανίχνευση μικρών και μεγάλων δομών στην εικόνα οι υπολογισμοί λαμβάνουν χώρα σε τρεις κλίμακας (ακτίνες των δίσκων):  $[\frac{\sigma}{2}, \sigma, 2\sigma]$ . Το πολυκλιμακωτό κατευθυντικό σήμα υπολογίζεται τώρα ως εξής:

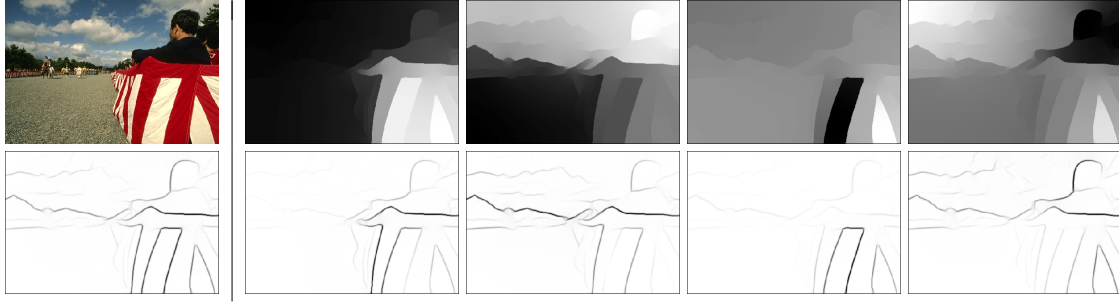
$$mPb(x, y, \theta) = \sum_s \sum_i a_{i,s} G_{i,\sigma(i,s)}(x, y, \theta) \quad (3.4)$$

όπου η κατεύθυνση  $\theta$  δειγματοληπτείται σε 8 ισαπέχοντες γωνίες στο διάστημα  $[0, \pi)$  και τα βάρη  $a_{i,s}$  υπολογίζονται με τεχνικές μάθησης με χρήση της συλλογής δεδομένων του πανεπιστημίου Berkeley.

Επιπροσθέτως, υπάρχει και ένα στάδιο εισαγωγής συνολικής πληροφορίας στην τε-



Σχήμα 3.3: Αριστερά, από πάνω προς τα κάτω: φωτεινότητα, χρωματικά κανάλια a και b του χρωματικού χώρου Lab και κανάλι υψής. Γραμμές, από αριστερά προς τα δεξιά: δίπλα σε κάθε κανάλι είναι το αντίστοιχο σήμα  $G$  για κατευθύνσεις  $\theta = 0$  και  $\theta = \pi/2$  καθώς και η μέγιστη απόκριση ως προς τις  $\delta$  κατευθύνσεις στο διάστημα  $[0, \pi)$ .



Σχήμα 3.4: **Αριστερά:** εικόνα εισόδου και σήμα  $sPb(x, y) = \max_{\theta} \{sPb(x, y, \theta)\}$ . **Πάνω και δεξιά:** τέσσερα πρώτα γενικευμένα ιδιοδιανύσματα  $\mathbf{v}_1, \dots, \mathbf{v}_4$ . **Κάτω δεξιά:** μέγιστη απόκριση ως προς όλες τις κατευθύνσεις των γκαουσιανών παραγωγών  $\max_{\theta} \{\nabla_{\theta} \mathbf{v}_{\mathbf{k}}(x, y)\}$  για κάθε ιδιοδιάνυσμα.

χνική αυτή. Ως είσοδος στο στάδιο αυτό χρησιμοποιείται ένας αραιός πίνακας συνάφειας  $\mathbf{W}$  μεταξύ όλων των εικονοστοιχείων της εικόνας. Ο πίνακας αυτός κατασκευάζεται χρησιμοποιώντας τη μέγιστη τιμή του σήματος  $mPb$  πάνω στο ευθύγραμμο τμήμα  $\bar{ij}$  που συνδέει δύο εικονοστοιχεία  $i, j$ . Όλα τα εικονοστοιχεία  $i$  και  $j$  που απέχουν λιγότερο από μια σταθερά  $r$  συνδέονται με το ακόλουθο μέτρο συνάφειας  $W_{ij}$ :

$$\mathbf{W}_{ij} = \exp(-\max_{p \in \bar{ij}} \{mPb(p)\} / \rho) \quad (3.5)$$

όπου  $\rho$  μια σταθερά. Για την εισαγωγή πληροφορίας από ολόκληρη την εικόνα, εφαρμόζεται μία τεχνικά παρόμοια με αυτήν της μεθόδου Normalized Cuts[36]. Ορίζεται ένας διαγώνιος πίνακας  $\mathbf{D}_{ii} = \sum_j \mathbf{W}_{ij}$  και επιλύεται το ακόλουθο πρόβλημα γενικευμένων ιδιοτιμών:

$$(\mathbf{D} - \mathbf{W})\mathbf{v} = \lambda \mathbf{D}\mathbf{v} \quad (3.6)$$

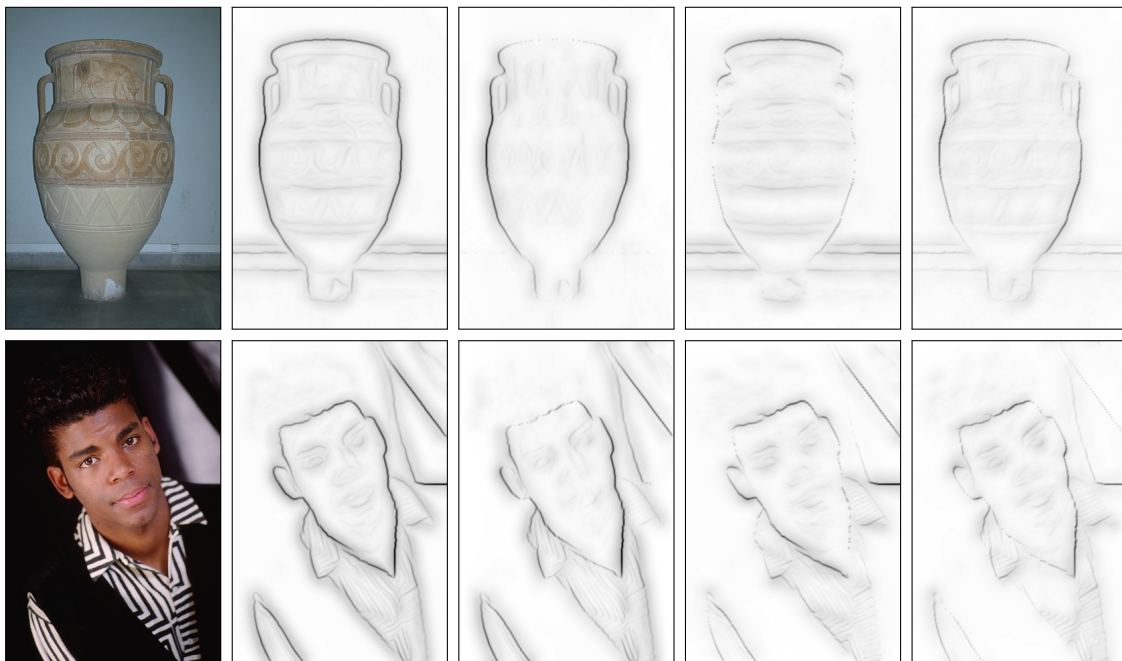
Τα ιδιοδιανύσματα περιέχουν πληροφορία σχετική με τις ακμές τις εικόνας όπως μπορεί να δει κανείς στο Σχ. 3.4. Με βάση την παρατήρηση αυτή, κάθε ιδιοδιάνυσμα  $\mathbf{v}_{\mathbf{k}}$  χειρίζεται ως εικόνα και συνελίσσεται με γκαουσιανά κατευθυντικά φίλτρα παραγωγού σε πολλαπλές κατευθύνσεις  $\theta$ . Με αυτόν τον τρόπο λαμβάνονται τα σήματα  $\{\nabla_{\theta} \mathbf{v}_{\mathbf{k}}(x, y)\}$  και η φασματική συνιστώσα του ανιχνευτή  $gPb$  προκύπτει ως ακολούθως:

$$sPb(x, y, \theta) = \sum_{k=1}^n \frac{1}{\sqrt{\lambda_k}} \mathbf{v}_{\mathbf{k}}(x, y) \quad (3.7)$$

Τελικά, η συνολική πιθανότητα ακμών ( $gPb$ ) είναι ένας γραμμικός συνδυασμός των  $mPb$  και  $sPb$ :

$$gPb(x, y, \theta) = \sum_s \sum_i \beta_{i,s} G_{i,\sigma(i,s)}(x, y, \theta) + \gamma \cdot sPb(x, y, \theta) \quad (3.8)$$

Τα βάρη  $\beta_{i,s}$  ανδ  $\gamma$  υπολογίζονται με τεχνικές μάθησης χρησιμοποιώντας τη συλλογή δεδομένων του πανεπιστημίου Berkeley. Παραδείγματα του τελικού ανιχνευτή γκρίζων ακμών  $gPb$  παρουσιάζονται στο Σχ. 3.5.



Σχήμα 3.5: Από αριστερά προς τα δεξιά: εικόνα εισόδου, σήμα  $gPb(x, y) = \max_{\theta} \{gPb(x, y, \theta)\}$ ,  $gPb$  για  $\theta = \pi/2$ ,  $\theta = 0$  και  $\theta = \pi/4$  αντίστοιχα.



# Κεφάλαιο 4

## Μετασχηματισμός απόστασης

### 4.1 Θεωρητικό υπόβαθρο

#### 4.1.1 Ορισμοί

Πριν την ανάλυση των μετασχηματισμών απόστασης, θα επεξηγηθούν οι συμβολισμοί που χρησιμοποιούνται στη συνέχεια της εργασίας αυτής. Οι συμβολισμοί αυτοί είναι ίδιοι με τους συμβολισμούς των Αβρίθη και Ραπαντζίκου [5] καθώς πρέπει να υπάρχει συνέπεια με την εργασία στην οποία βασίζεται η παρούσα διπλωματική. Αρχικά, οι δισδιάστατες εικόνες αναπαριστώνται από συναρτήσεις  $f : \mathbb{X} \rightarrow \mathbb{V}$ . Ως εύρος  $\mathbb{V}$  χρησιμοποιείται το σύνολο  $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$  και ως χωρίο  $\mathbb{X}$  το συνεχές (διακριτό) σύνολο  $\mathbb{R}^2$  ( $\mathbb{Z}^2$ ). Στην πράξη χρησιμοποιείται ένα πεπερασμένο υποσύνολο  $X \subseteq \mathbb{X}$ . Για τη σύνδεση γειτονικών pixels της εικόνας χρησιμοποιούμε συνεκτικότητα κατά 4 ή κατά 8 και γράφουμε αντίστοιχα  $u \diamond v$  ή  $u * v$ .

Μια συνάρτηση  $g$  από το  $\mathbb{X} \times \mathbb{X} \rightarrow \bar{\mathbb{R}}_+$  ονομάζεται:

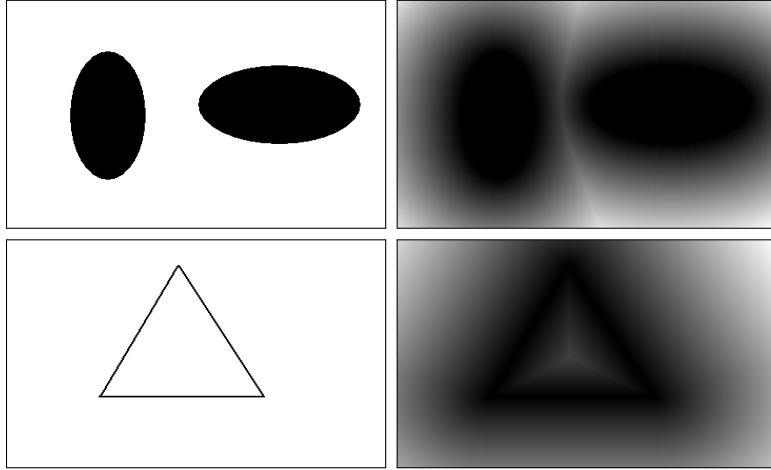
(α) *Θετικά ορισμένη*: εάν  $g(u, v) = 0 \Leftrightarrow u = v$ ,  $\forall u, v \in V$

(β) *Συμμετρική*: εάν  $g(u, v) = g(v, u)$ ,  $\forall u, v \in V$

(γ) *Τριγωνική*: εάν  $g(u, w) \leq g(u, v) + g(v, w)$ ,  $\forall u, v, w \in V$

Εάν μία συνάρτηση  $g$  ικανοποιεί τις (α)-(γ), ονομάζεται *συνάρτηση απόστασης* ή *μετρικό* [34].

Ο μετασχηματισμός απόστασης είναι μια ειδική συνάρτηση απόστασης. Αρχικά, έστω μια δυαδική εικόνα που αποτελείται από foreground και background εικονοστοιχεία τα οποία μπορεί να αντιστοιχούν σε μεμονωμένα σημεία, σε ακμές τις εικόνες ή και σε ολόκληρα αντικείμενα. Ο μετασχηματισμός απόστασης είναι μία διαδικασία που μετατρέπει μία δυαδική εικόνα σε μία εικόνα γκριζου τόνου όπου όλα τα σημεία έχουν μια τιμή που αντιστοιχεί στην απόσταση από το κοντινότερο σημείο του foreground.



Σχήμα 4.1: Από αριστερά προς τα δεξιά: σύνολο πηγών (μαύρο χρώμα), εξωτερικός μετασχηματισμός απόστασης με ευκλείδιο μετρικό.

Το σύνολο του foreground της εικόνας συμβολίζεται με  $S \subseteq X$  και το αντίστοιχο background συμβολίζεται ως  $S^c \subseteq X$ , όπου το  $S^c$  δηλώνει το συμπλήρωμα  $X \setminus S$  του  $S$ . Δοθέντος ενός μετρικού  $d$ , ο μετασχηματισμός απόστασης του  $S$  ορίζεται ως εξής:

$$D_d(S)(x) \triangleq \bigwedge_{y \in S} d(x - y) \quad (4.1)$$

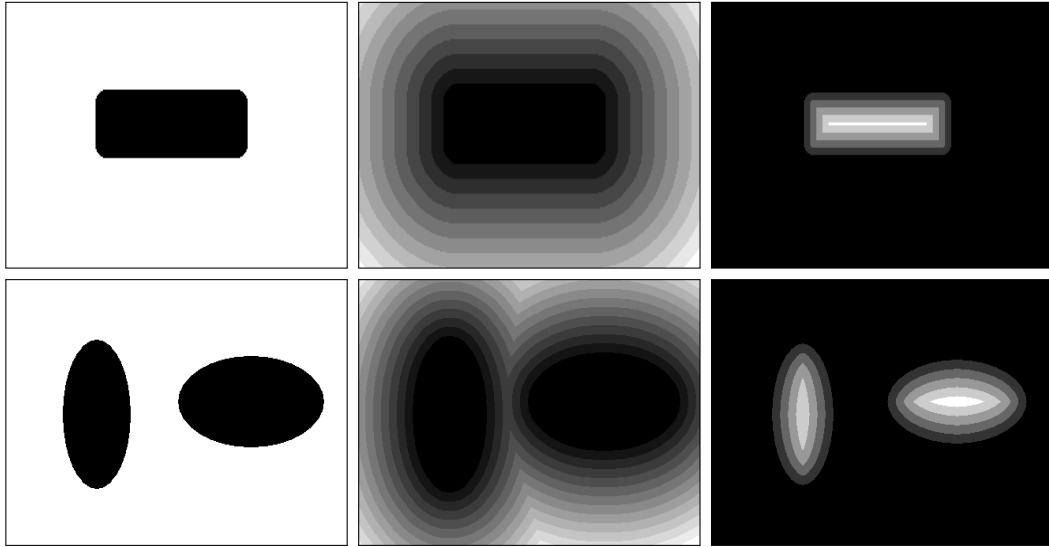
όπου με  $\bigwedge$  συμβολίζεται το infimum. Προφανώς, η συνάρτηση  $D_d(S)(x)$  λαμβάνει μηδενικές τιμές για όλα τα  $x \in S$ .

#### 4.1.2 Μετασχηματισμός απόστασης ως κυματική διάδοση

Καθ' αυτόν τον τρόπο, το σύνολο  $S$  μπορεί να θεωρηθεί ως ένα σύνολο πηγών που διαδίδει ένα κύμα απόστασης προς το background της εικόνας. Από αυτήν την οπτική γωνία, ο ορισμός (4.1) αποτελεί έναν εξωτερικό μετασχηματισμό απόστασης και είναι χρήσιμος όταν μας ενδιαφέρει η γεωμετρία και η τοπολογία του background  $S^c$ . Σε αντίθετη περίπτωση, είναι προτιμώτερη η χρήση του εσωτερικού μετασχηματισμού απόστασης  $D_d(S^c)$ . Συνολικά, η ευκλείδια συνάρτηση απόστασης  $D_2(S)$  αποτελεί ασθενή λύση της ακόλουθης μη γραμμικής μερικής διαφορικής εξίσωσης:

$$\begin{aligned} \|\nabla u(x)\| &= 1 & x \in S^c \\ u(x) &= 0 & x \in \partial S \end{aligned} \quad (4.2)$$

Η άνωθεν εξίσωση αποτελεί ειδική περίπτωση της *eikonal MΔE* που αντιστοιχεί σε κυματική διάδοση σε ετερογενή μέσα και η λύση της  $u$  αποτελεί ένα σταθμισμένο μετασχηματισμό απόστασης, όπου τα βάρη  $F(x)$  είναι αντιστρόφως ανάλογα της μεταβλητής



Σχήμα 4.2: Από αριστερά προς τα δεξιά: σύνολο πηγών, εξωτερικός και εσωτερικός μετασχηματισμός απόστασης με ευκλείδιο μετρικό και οπτικοποιημένος ως κυματική διάδοση.

ταχύτητας διάδοσης:

$$\begin{aligned} \|\nabla u(x)\|_2 &= F(x) \text{ εν } \Omega, & F(x) > 0 \\ u(x) &= g(x) \text{ ον } \Gamma \end{aligned} \quad (4.3)$$

όπου  $\Omega$  είναι του  $\mathbb{R}^2$  ή του  $\mathbb{R}^3$  και η συνάρτηση  $g(x)$  είναι δεδομένη πάνω σε μια καμπύλη ή επιφάνεια  $\Gamma$  εν  $\Omega$ .

### 4.1.3 Μετασχηματισμός απόστασης ως infimal συνέλιξη

Η *infimal* συνέλιξη  $f \star g$  δύο διασδιάστατων σημάτων  $f, g$  ορίζεται ως:

$$(f \star g)(x) \triangleq \bigwedge_{y \in X} g(x - y) + f(y), \quad x \in X \quad (4.4)$$

Ο μετασχηματισμός απόστασης είναι στενά συνδεδεμένος με την πράξη της *infimal* συνέλιξης. Μπορεί εύκολα να παρατηρηθεί πως όταν  $g(x - y) = \|x - y\|_d$  για ένα μετρικό  $d$ , ο μετασχηματισμός απόστασης μιας συνάρτησης  $f$  είναι ακριβώς η *infimal* συνέλιξη της  $f$  με την  $g$ . Αναλυτικά έχουμε:

$$\begin{aligned} D_d(S)(x) &= \bigwedge_{y \in S} d(x - y), \quad x \in X \Rightarrow \\ D_d(S)(x) &= \bigwedge_{y \in X} d(x - y) + 1_S(y), \quad x \in X \Rightarrow \\ D_d(S)(x) &= (1_S \star d)(x) \end{aligned}$$

όπου  $1_S(x)$  η  $0/\infty$  συνάρτηση δείκτη του συνόλου  $S$ :

$$1_S(x) = \begin{cases} 0, & x \in S \\ +\infty, & \text{οτηρωρισε} \end{cases}$$

## 4.2 Αλγόριθμοι αποδοτικού υπολογισμού μετασχηματισμών απόστασης

### 4.2.1 Λύνοντας την eikonal ΜΔΕ με τεχνικές ταχείας πόρευσης

Οι τεχνικές ταχείας πόρευσης (fast marching methods) [35] είναι υπολογιστικές τεχνικές που προσεγγίζουν τη λύση μη γραμμικών μερικών διαφορικών εξισώσεων της μορφής:

$$\begin{aligned} \|\nabla u(x)\| &= F(x) \text{ εν } \Omega, & F(x) > 0 \\ u &= g(x) \text{ ον } \Gamma \end{aligned} \quad (4.5)$$

όπου  $\Omega$  είναι του  $\mathbb{R}^2$  ή του  $\mathbb{R}^3$  και η συνάρτηση  $g(x)$  είναι δεδομένη πάνω σε μια καμπύλη ή επιφάνεια  $\Gamma$  εν  $\Omega$ .

Η εξίσωση (4.5) είναι ανήκει στην κατηγορία των εξισώσεων Hamilton-Jacobi που έχουν την ακόλουθη μορφή:

$$H(u_x, u_y, u_z, x, y, z) = 0 \quad (4.6)$$

Στην περίπτωση της eikonal ΜΔΕ, η συνάρτηση  $H$  γίνεται  $H = |\nabla u(x)| - F(x)$ .

Αρχικά, πρέπει να αναφέρουμε την προσέγγιση πεπερασμένων διαφορών που ο Sethian [35] χρησιμοποιεί στην προσέγγισή του:

$$u_x^2 \approx \left( \max(D_i^{+x}u, 0)^2 + \min(D_i^{-x}u, 0)^2 \right) \quad (4.7)$$

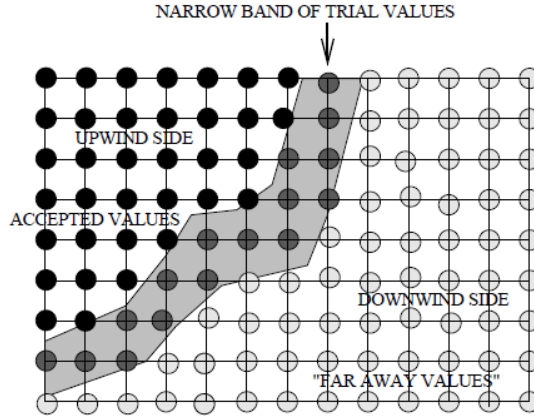
όπου

$$D_i^{-x}u = \frac{u_i - u_{i-1}}{h}, \quad D_i^{+x}u = \frac{u_{i+1} - u_i}{h} \quad (4.8)$$

Εδώ,  $u_i$  είναι η τιμή της  $u$  σε ένα πλέγμα στο σημείο  $ih$  με σταθερή απόσταση  $h$  μεταξύ των σημείων του πλέγματος.

Επεκτείνοντας τις προηγούμενες προσεγγίσεις για την κλίση σε πολλαπλές διαστάσεις προκύπτουν τα ακόλουθα σχήματα:

$$|\nabla u| \approx \left[ \begin{aligned} &\max(D_{ij}^{-x}u, 0)^2 + \min(D_{ij}^{+x}u, 0)^2 \\ &+ \max(D_{ij}^{-y}u, 0)^2 + \min(D_{ij}^{+y}u, 0)^2 \end{aligned} \right]^{1/2} = F_{ij} \quad (4.9)$$



Σχήμα 4.3: Upwind κατασκευή αποδεκτών τιμών της λύσης (από [35])

ή ισοδύναμα:

$$\left[ \begin{array}{l} \max (D_{ij}^{-x}u, -D_{ij}^{+x}u, 0)^2 \\ + \max (D_{ij}^{-y}u, -D_{ij}^{+y}u, 0)^2 \end{array} \right]^{1/2} = F_{ij} \quad (4.10)$$

Μία πρόταση για την επίλυση της εξίσωσης (4.10) είναι η χρήση επαναληπτικής μεθόδου. Θεωρώντας πλέγμα  $N \times N$  σημείων και  $N$  επαναλήψεις μέχρι τη σύγκλιση, η υπολογιστική πολυπλοκότητα της επαναληπτικής μεθόδου αυτής προκύπτει της τάξης  $O(N^3)$ . Οι τεχνικές ταχείας πόρευσης μειώνουν την πολυπλοκότητα αυτή σε  $O(N^2 \log N)$ .

Η κεντρική ιδέα των τεχνικών ταχείας πόρευσης είναι η συστηματική κατασκευή της λύσης σε μία κατεύθυνση για την παραγωγή της λύσης  $u$ . Η ιδέα αυτή βασίζεται στην παρατήρηση πως η πληροφορία διαδίδεται προς μία κατεύθυνση, δηλαδή από χαμηλές τιμές της λύσης  $u$  προς υψηλότερες. Ο αλγόριθμος ταχείας πόρευσης λύνει την εξίσωση (4.10) χτίζοντας τη λύση προς τα έξω από τη μικρότερη τιμή της συνάρτησης  $u$ . Ο αλγόριθμος επιταχύνεται περιορίζοντας τη “ζώνη χτισίματος” σε μια στενή ζώνη (narrow band) πάνω στο μέτωπο της κυματικής διάδοσης (βλέπε Σχ. 4.2.1).

Ο αλγόριθμος αρχικοποιείται ως εξής: αρχικά, όλα τα σημεία των αρχικών συνθηκών σημειώνονται ως *Alive*. Στη συνέχεια, όλα τα γειτονικά σημεία των προηγούμενων σημειώνονται ως *Close* και όλα τα υπόλοιπα ως *Far*. Το κύριο κομμάτι του αλγορίθμου συνοψίζεται στα κάτωθι βήματα:

1. Αρχή βρόχου: έστω *Trial* το σημείο του συνόλου *Close* με τη μικρότερη τιμή  $u$ .
2. Όλοι οι γείτονες του *Trial* που δεν είναι *Alive* σημειώνονται ως *Close*.
3. Επαναυπολογισμός όλων των τιμών της λύσης  $u$  σε όλους τους γείτονες του *Trial* που είναι *Close* σύμφωνα με την (4.10).
4. Το σημείο *Trial* γίνεται *Alive*.
5. Επανάληψη του βρόχου εάν το σύνολο *Close* δεν είναι άδειο.

Το κλειδί για μια αποδοτική υλοποίηση της παραπάνω τεχνικής βασίζεται σε έναν

γρήγορο τρόπο εύρεσης εκείνου του σημείου του πλέγματος που βρίσκεται στη στενή ζώνη και έχει την μικρότερη τιμή του  $u$ . Αυτό επιτυγχάνεται με χρήση μιας `min-heap` δομής δεδομένων. Το συνολικό υπολογιστικό κόστος για την αλλαγή της τιμής ενός στοιχείου της δομής αυτής και προώθηση της τιμής του προς τα πάνω είναι της τάξης  $O(\log M)$ , όπου  $M$  είναι το συνολικό πλήθος των στοιχείων της δομής. Άρα συνολικά, για  $M$  σημεία πλέγματος η πολυπλοκότητα του αλγορίθμου ταχείας πόρευσης προκύπτει  $O(M \log M)$  αφού κάθε σημείο επεξεργάζεται μία και μόνο φορά.

#### 4.2.2 Σταθμισμένος μετασχηματισμός απόστασης και αλγόριθμος `exact group marching`

Στο σημείο αυτό, αναλύεται ο προσθετικά σταθμισμένος μετασχηματισμός απόστασης μαζί με έναν αποδοτικό αλγόριθμο γραμμικού χρόνου για τον υπολογισμό του. Ο προσθετικά σταθμισμένος μετασχηματισμός απόστασης  $D_d(f)$  της συνάρτησης  $f$  δοθέντος ενός μετρικού  $d$  ορίζεται ως:

$$D_d(f)(x) = \bigwedge_{y \in X} d(x, y) + f(y), \quad x \in X, \quad (4.11)$$

Πιο συχνά, χρησιμοποιούμε ένα μετρικό που ορίζεται από μια νόρμα  $\|\cdot\|$ , δηλαδή  $d(x, y) = \|x - y\|$  για  $x, y \in \mathbb{X}$ . Ακόμη, μπορούμε να αγνοήσουμε το  $d$  στον συμβολισμό και να γράψουμε απλά  $\mathcal{D}(f)$ . Παρόλο που ο προηγούμενος ορισμός είναι για τυχαία συνάρτηση  $f$ , εάν το πρόβλημα προς επίλυση είναι η κατάτμηση εικόνας τότε η συνάρτηση  $f$  πρέπει να σχετίζεται με τα όρια των περιοχών όπως π.χ. η κλίση της εικόνας ή η εικόνα ακμών.

Ο αλγόριθμος `exact group marching` είναι μια παραλλαγή του αλγορίθμου `group marching` (GMM) [17], μιας γραμμικού χρόνου τεχνικής ταχείας πόρευσης που επιλέγει έναν αριθμό σημείων στο μέτωπο της διάδοσης και τα κινεί ως ομάδα γλιτώνοντας έτσι το κόστος της ταξινόμησης. Όλα τα σημεία μετακινούνται από το μέτωπο ως ομάδα με χρήση μιας ουράς προτεραιότητας κβαντισμένων επιπέδων απόστασης. Αυτό είναι παρόμοιο με το [42] και στη περίπτωση δυαδικών εικόνων είναι ισοδύναμο με το σχήμα δύο ουρών του [24]. Ωστόσο, ο όλος υπολογισμός είναι ακριβής.

Έχουμε, λοιπόν, πως δοθείσας μιας εικόνας  $f$ , το ελαχιστοσύνολο  $\hat{S}(x)$  ορίζεται για κάθε σημείο  $x \in X$  ως το σύνολο των σημείων  $y \in X$  για τα οποία ελαχιστοποιείται η ποσότητα  $d(x, y) + f(y)$ :

$$\hat{S}(x) = \{y \in X : d(x, y) + f(y) = \mathcal{D}(f)(x)\} \quad (4.12)$$

για  $x \in X$ . Εάν  $y \in \hat{S}(x)$  γράφουμε ισοδύναμα  $y \succ x$ . Το σύνολο πηγών  $S(x)$  του  $x$  ορίζεται ως το υποσύνολο του ελαχιστοσυνόλου έτσι ώστε δεν υπάρχουν δύο σημεία  $y, z \in S(x)$  για τα οποία  $y \succ z$ :

$$S(x) = \{y \in \hat{S}(x) : \nexists z(y \succ z \succ x)\}. \quad (4.13)$$

Ένα  $y$  είναι πηγή του  $x$  ή ισοδύναμα  $y \succ x$  αν  $y \in S(x)$ . Γενικότερα, το σημείο  $y \in X$  είναι πηγή αν  $y \succ x$  για κάποιο  $x \in X$ , ακόμα και για το ίδιο. Στο διακριτό χωρίο  $X$ , κάθε σημείο  $x \in X$  έχει τουλάχιστον μία πηγή.

**Λήμμα 4.2.1** Δοθέντος  $y \in X$ , τα παρακάτω είναι ισοδύναμα:

(α)  $y$  είναι πηγή.

(β)  $y \succ y$ .

(γ)  $\mathcal{D}(f)(y) = f(y)$ .

(δ)  $S(y) = \{y\}$ .

(ε)  $y \succ y$ .

Το σύνολο πηγών  $S(f)$  της  $f$  ορίζεται ως το σύνολο όλων των πηγών  $y \in X$ . Προκύπτει πως  $S(f) = \{x \in X : x \succ x\}$  και με αυτόν τον τρόπο καθίσταται εύκολη η ανίχνευση των πηγών. Με  $s(x)$  συμβολίζεται η πηγή του  $x$  εάν είναι μοναδική, αλλιώς κάποιος αντιπρόσωπος του  $S(x)$ . Το σύνολο πηγών συνδέεται άμεσα με τα όρια των περιοχών της εικόνας. Αντίστοιχα, ορίζεται το εσωτερικό σύνολο της  $f$  ως  $I(f) = X \setminus S(f)$ .

**Λήμμα 4.2.2** Ο μετασχηματισμός απόστασης  $\mathcal{D}_d(f)$  είναι μονοσήμαντα ορισμένος από τον περιορισμό  $f|_{S(f)}$  της  $f$  στο αντίστοιχο σύνολο πηγών.

Για τον υπολογισμό του σταθμισμένου μετασχηματισμού απόστασης  $h = \mathcal{D}(f)$  σύμφωνα με την (4.11) χρησιμοποιείται ο αλγόριθμος *exact group marching* (EGM) οποίος παρουσιάζεται αναλυτικά στη συνέχεια. Η διάδοση ξεκινά από το αρχικό σύνολο πηγών  $S_+(f)$ :

$$S_+(f) = \{x \in X : f(x) < \min_{y \diamond x} f(y) + 1\}. \quad (4.14)$$

Επειδή  $d(x, y) = 1$  για  $y \diamond x$ , μπορεί εύκολα να δειχθεί πως το  $S_+(f)$  είναι υπερσύνολο του  $S(f)$ .

Η διάδοση της απόστασης βασίζεται σε μια ουρά προτεραιότητας με διακριτά επίπεδα προτεραιότητας, υλοποιημένη ως πίνακας FIFO ουρών. Τα σημεία παίρνουν ετικέτες **far**, **near**, και **done**. Η ουρά προτεραιότητας κρατάει σημεία που είναι **near**, δηλαδή σημεία που ανήκουν στο μέτωπο της διάδοσης. Τα σημεία επεξεργάζονται κατά ομάδες: κάθε σημείο επεξεργάζεται σύμφωνα με το επίπεδό του  $\lfloor h(x) \rfloor$  και σημεία με ίδιο επίπεδο επεξεργάζονται σε τυχαία σειρά. Η διάδοση πραγματοποιείται σε γείτονες του σημείου που επεξεργάζομαστε. Ο υπολογισμός είναι ακριβής παρόλο που σημεία του ίδιου επιπέδου επεξεργάζονται κατά τυχαία σειρά.

---

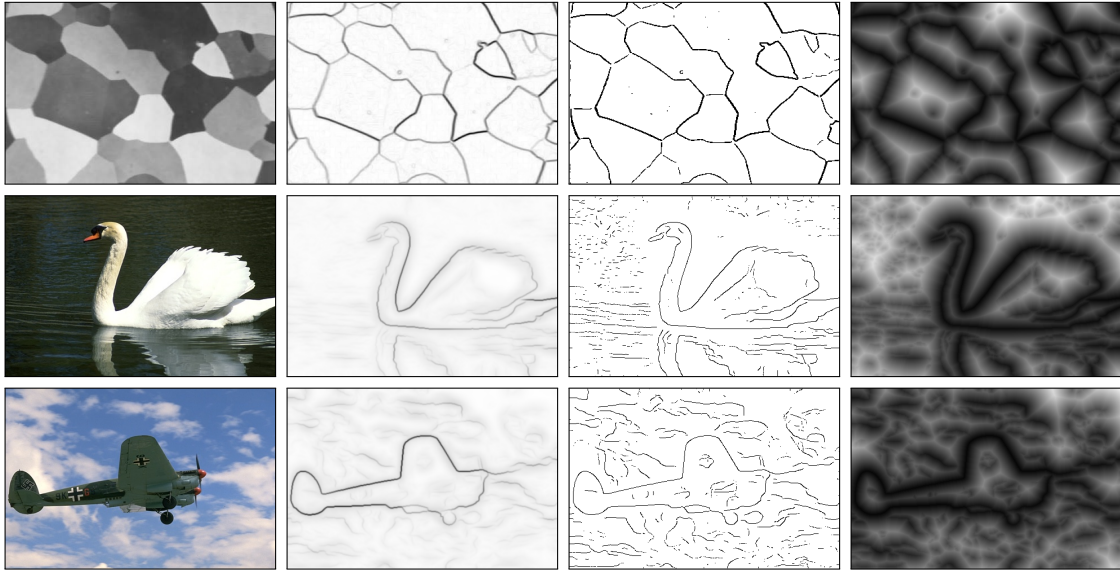
**Algorithm 1** Exact Group Marching

---

```
1: procedure EGM(image  $f$ )
2:   initialize  $q, h, s$ ; construct seed  $S_+$  as in (4.14)
3:   for all  $x \in S_+$  do  $\{s(x) \leftarrow x; \text{PROP}(x, x); \}$ 
4:   for all  $x \in X \setminus S_+$  do  $\{\text{label } x \text{ as } \textit{far}; \}$ 
5:   while  $\neg q.\text{EMPTY}(\ )$  do
6:      $x \leftarrow q.\text{POP}(\ )$ ; label  $x$  as done
7:     for  $y \diamond x, y$  near do  $\text{UPDATE}(y, x)$  ▷ incoming
8:     for  $y \diamond x, y$  near do  $\text{UPDATE}(x, y)$  ▷ outgoing
9:     for  $y \diamond x, y$  far do  $\text{PROP}(x, y)$ 
10:  end while
11:  return distance map  $h$ , source map  $s$ 
12: end procedure
13:
14: procedure PROP(point  $x$ , point  $y$ )
15:    $h(y) \leftarrow d(y, s(x)) + f(s(x))$ ;
16:    $s(y) \leftarrow s(x)$ ;
17:    $q.\text{PUSH}(y, \lfloor h(y) \rfloor)$ ;
18:   label  $y$  as near;
19: end procedure
20:
21: procedure UPDATE(point  $x$ , point  $y$ )
22:    $h_0 \leftarrow d(y, s(x)) + f(s(x))$ ;
23:   if  $h_0 \geq h(y)$  return
24:    $h(y) \leftarrow h_0$ ;
25:    $s(y) \leftarrow s(x)$ ;
26: end procedure
```

---





Σχήμα 4.4: Από αριστερά προς τα δεξιά: εικόνες εισόδου, αντίστοιχες grayscale εικόνες ακμών, σύνολα πηγών (μαύρο χρώμα) και σταθμισμένος μετασχηματισμός απόστασης με ευκλείδια νόρμα υπολογισμένος με τον αλγόριθμο exact group marching.

**Λήμμα 4.2.3** (α) Ο αλγόριθμος EGM υπολογίζει την ακριβή απόσταση  $\mathcal{D}(f)(x)$  που ορίζεται στην εξίσωση (4.11) και τη σωστή πηγή  $s(x)$  για κάθε  $x \in X$ . (β) Ο βρόχος *while* επεξεργάζεται κάθε σημείο ακριβώς μία φορά. (γ) Η υπολογιστική πολυπλοκότητα είναι  $O(n)$ , όπου  $n = |X|$ .

Το μόνο που παραμένει αδιευκρίνιστο είναι η συνάρτηση  $f$ . Στα πλαίσια της εργασίας αυτής, χρησιμοποιούμε την έξοδο  $g$  ενός ανιχνευτή γκριζων ακμών και ορίζουμε  $f(x) = \sigma/g(x)$  για  $x \in X$ , όπου  $\sigma$  είναι μια παράμετρος κλίμακας. Αυτή είναι μια γενίκευση της  $0/\infty$  συνάρτησης δείκτη που χρησιμοποιείται στον απλό μετασχηματισμό απόστασης.



# Κεφάλαιο 5

## Μετασχηματισμός διάμεσου άξονα

### 5.1 Εισαγωγή

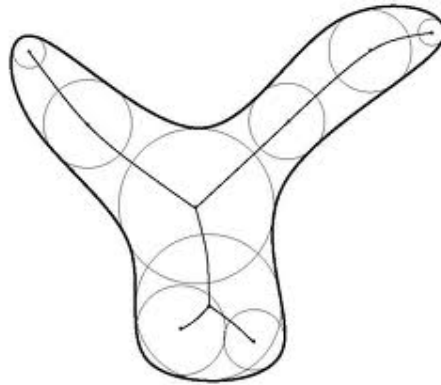
Ο μετασχηματισμός διάμεσου άξονα ή σκελετού είναι μια τεχνική που προτάθηκε από τον Blum [8] ως μέσο περιγραφής και αναπαράστασης ενός σχήματος. Ο μετασχηματισμός αυτός ορίζεται επίσημα ως εξής: δοθέντος ενός αντικειμένου, έστω ενός απλού πολυγώνου  $G$ , ο διάμεσος άξονας ή σκελετός  $A(G)$  είναι το σύνολο των σημείων  $q$ , εσωτερικών του  $G$ , που ισαπέχουν από δύο σημεία των ορίων του αντικειμένου. Άρρηκτα συνδεδεμένη με το διάμεσο άξονα είναι μια συνάρτηση απόστασης  $R$  που ορίζει την απόσταση κάθε σημείου του διάμεσου άξονα από τα όρια του αντικειμένου (βλέπε Σχ. 5.1). Με δεδομένα τον διάμεσο άξονα και τη συνάρτηση απόστασης καθίσταται εφικτή η ακριβής ανακατασκευή του αρχικού σχήματος.

### 5.2 Υπολογισμός διάμεσου άξονα

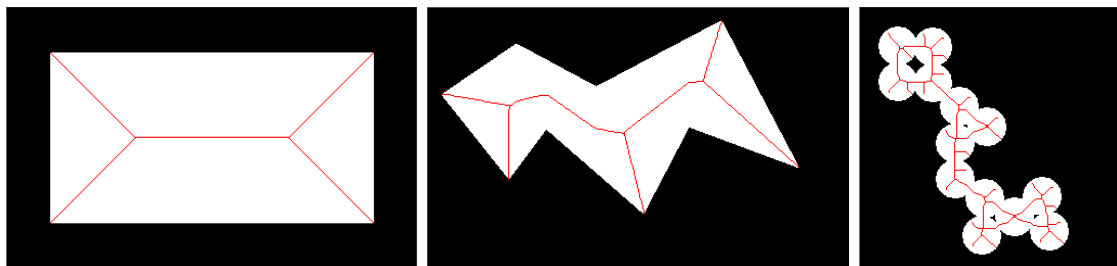
Πρόσφατα, οι Αβρίθης και Ραπαντζίκος [5] μελέτησαν το διάμεσο άξονα ενός σταθμισμένου μετασχηματισμού απόστασης. Αντί να δουλέψουν με ΜΔΕ, χρησιμοποίησαν ένα κριτήριο υπολοίπου βασισμένο στην εγγύτητα των πηγών του μετασχηματισμού απόστασης. Επέκτειναν το κριτήριο υπολοίπου από δυαδικά σχήματα σε τυχαίες συναρτήσεις του επιπέδου. Αυτή η προσέγγιση υιοθετείται στην παρούσα διπλωματική εργασία και περιγράφεται αναλυτικά στη συνέχεια.

Δοθέντων των ορισμών των πηγών του σταθμισμένου μετασχηματισμού απόστασης που παρουσιάστηκε στην παράγραφο 4.2.2, έχουμε πως ένα σημείο  $x \in X$  ανήκει στο διάμεσο άξονα της συνάρτησης (εικόνας)  $f$  εάν έχει τουλάχιστον δύο διακριτές πηγές. Ο διάμεσος άξονας  $A(f)$  ορίζεται ως το σύνολο όλων αυτών των σημείων:

$$A(f) = \{x \in X : |S_f(x)| > 1\} \quad (5.1)$$



Σχήμα 5.1: Δυαδικό σχήμα και ο αντίστοιχος σκελετός. Οι ακτίνες των κύκλων αντιστοιχούν στην απόσταση κάθε σημείου του διάμεσου άξονα από το σύνορο του σχήματος.



Σχήμα 5.2: Δυαδικά σχήματα (άσπρο χρώμα) και αντίστοιχοι διάμεσοι άξονες (κόκκινο χρώμα).

**Λήμμα 5.2.1** Το σύνολο πηγών και ο διάμεσος άξονας  $f$  αμοιβαίως αποκλειόμενα σύνολα:  $S(f) \cap A(f) = \emptyset$ . Άρα ο διάμεσος άξονας είναι υποσύνολο του εσωτερικού συνόλου, δηλαδή  $A(f) \subseteq I(f)$ .

Στη διακριτή περίπτωση έχουμε τις ακόλουθες ιδιότητες:

**Λήμμα 5.2.2** Έστω  $A$  ο διάμεσος άξονας της  $f$  σε ευκλείδιο χώρο και δύο σημεία  $x \in A$  και  $y \in S(x)$ .

(α) Έστω ένα ευθύγραμμο τμήμα από το  $x$  στο  $y$ . Κάθε σημείο  $z$  του ευθυγράμμου τμήματος έχει μοναδική πηγή  $s(z) = y$ .

(β) Ο διάμεσος άξονας είναι μηδενικού πάχους  $A$ , δηλαδή  $A \subseteq \partial A$ .

Δοθέντων δύο γειτονικών σημείων  $x, y$  με  $s(x) \neq s(y)$ , το λήμμα 5.2.2(β) προτείνει την ύπαρξη ενός σημείου  $m$  που ανήκει στο διάμεσο άξονα με  $S(m) = \{s(x), s(y)\}$  στο ευθύγραμμο τμήμα μεταξύ  $x$  και  $y$ . Για το λόγο αυτό, το εύγος  $(x, y)$  θεωρείται πως ανήκει στο διάμεσο άξονα.

Ο αλγόριθμος weighted medial axis (WMA) υπολογίζει το διάμεσο άξονα  $A(f)$  της εικόνας  $f$  δοθέντος του μετασχηματισμού απόστασης  $h = \mathcal{D}(f)$  και της αντιστοιχης συνάρτησης πηγών  $s$ . Η διάδοση ξεκινάει από ένα υπερσύνολο του διαμέσου άξονα που ορίζεται ως:

$$A_+(f) = \{x \in X : h(x) \geq \max_{y \circ x} h(y)\}, \quad (5.2)$$

και συνεχίζει προς χαμηλότερες τιμές απόστασης πάντα πάνω στο διάμεσο άξονα  $A(f)$ . Για κάθε σημείο  $x$  που επεξεργάζεται ο αλγόριθμος, ελέγχει τους γείτονές του για να αποφασίσει εάν το ζεύγος  $(x, y)$  ανήκει στο διάμεσο άξονα. Η απόφαση για το εάν τελικά ανήκει εάν σημείο στο διάμεσο άξονα βασίζεται στη συνάρτηση υπολοίπου  $r(x) = \max_{y \circ x} \text{res}(x, y)$  για  $x \in X$ . Η συνάρτηση υπολοίπου  $\text{res}$  συζητείται στη συνέχεια.

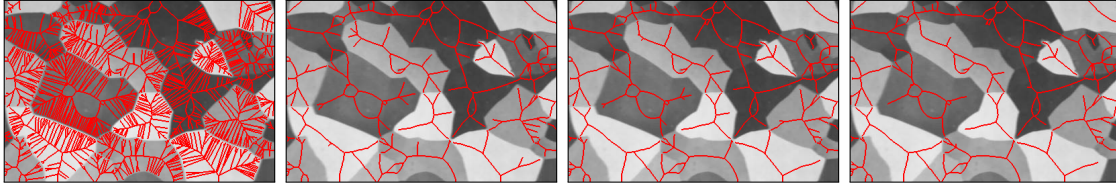
$$A(f) = \{x \in X : r(x) > 0\} \quad (5.3)$$

Το υπόλοιπο χορδής [31] ορίζεται, για δυαδικά σχήματα μόνο, ως τη διαφορά μεταξύ του μήκους ενός τμήματος μιας καμπύλης και του αντίστοιχου μήκους της χορδής ενός κύκλου που περιέχεται στο σχήμα και εφάπτεται στην καμπύλη στα δύο άκρα. Για τη γενίκευση του υπολοίπου χορδής στις τρεις διαστάσεις γίνεται χρήση του μετασχηματισμού απόστασης και η τιμή της απόστασης (ύψους) αντιμετωπίζεται ως τρίτη διάσταση. Τα διςδιάστατα τμήματα καμπυλών αντιστοιχούν σε τριςδιάστατα γεωδαιτικά μονοπάτια κατά μήκος της επιφάνειας των πηγών  $S(f)$ . Για τη μέτρηση των τριςδιάστατων αποστάσεων χρησιμοποιείται το μετρικό  $\delta$  το οποίο αποτελείται από το ευκλείδιο μετρικό  $d$  και την απόλυτη διαφορά:

$$\delta(u, v) = d(u, v) + |h(u) - h(v)|, \quad u, v \in \mathbb{X}. \quad (5.4)$$

Τώρα, δοθέντων δύο σημείων  $x, y \in X$  με πηγές  $u = s(x), v = s(y)$ , το υπόλοιπο χορδής γεικεύεται ως εξής:

$$\text{res}(x, y) = \ell(u, v) - \delta(u, v) \quad (5.5)$$



Σχήμα 5.3: Από αριστερά προς τα δεξιά: διάμεσος άξονας υπολογισμένος με τον αλγόριθμο WMA για  $scale = 0, 1, 2, 5$ .

όπου συνάρτηση μήκους  $\ell$  ορίζεται ως το μήκος του συντομότερου γεωδαιτικού μονοπατιού που συνδέει τα σημεία  $(u, f(u))$  και  $(v, f(v))$  κατά μήκος της επιφάνειας των πηγών  $\mathcal{S}(f)$ .

**Λήμμα 5.2.3** (α) Δοθέντων ζευγών σημείων  $(x, y), (x', y')$  της ίδιας συνιστώσας του εσωτερικού συνόλου  $I(f)$  με αντίστοιχα ζεύγη πηγών  $(u, v), (u', v')$ , ορίζονται μονοπάτια  $\pi = (u, \dots, v), \pi' = (u', \dots, v')$ . Εάν  $\pi \subset \pi'$ , τότε  $res(x, y) < res(x', y')$ . (β) Ο αλγόριθμος WMA παράγει ακριβώς μία συνιστώσα του διάμεσου άξονα  $A(f)$  για κάθε συνιστώσα του εσωτερικού συνόλου  $I(f)$ . (γ) Η πολυπλοκότητά του είναι  $O(k)$ , όπου  $k = |A(f)|$ .

Παρατηρούμε, δηλαδή, πως η συνάρτηση υπολοίπου αυξάνει όσο πιο εσωτερικά είμαστε στο  $I(f)$  και η απλοποίηση (pruning) του διάμεσου άξονα ανάγεται σε μία απλή διαδικασία κατωφλίωσης με μία παράμετρο κλίμακας  $scale$ . Συνήθως,  $scale = 2$ . Στο Σχ. 5.3 αναπαρίσταται ο διάμεσος άξονας για πολλαπλές τιμές της παραμέτρου  $scale$ . Ελάχιστες μεταβολές παρατηρούνται για τιμές της παραμέτρου μεγαλύτερες ή ίσες με την τυπική τιμή  $scale = 2$ .

Επιπροσθέτως, ορίζουμε για κάθε σημείο  $x \in X$  το σκελετικό σύνολο ζευγών  $P(x)$  ως το σύνολο των σημείων που μεγιστοποιούν τη συνάρτηση υπολοίπου  $r(x)$ . Ισοδύναμα:

$$P(x) = \{y \diamond x : r(x) = res(x, y)\} \quad (5.6)$$

και ο υπολογισμός του γίνεται κατά τη διαδικασία SCAN του αλγορίθμου WMA.

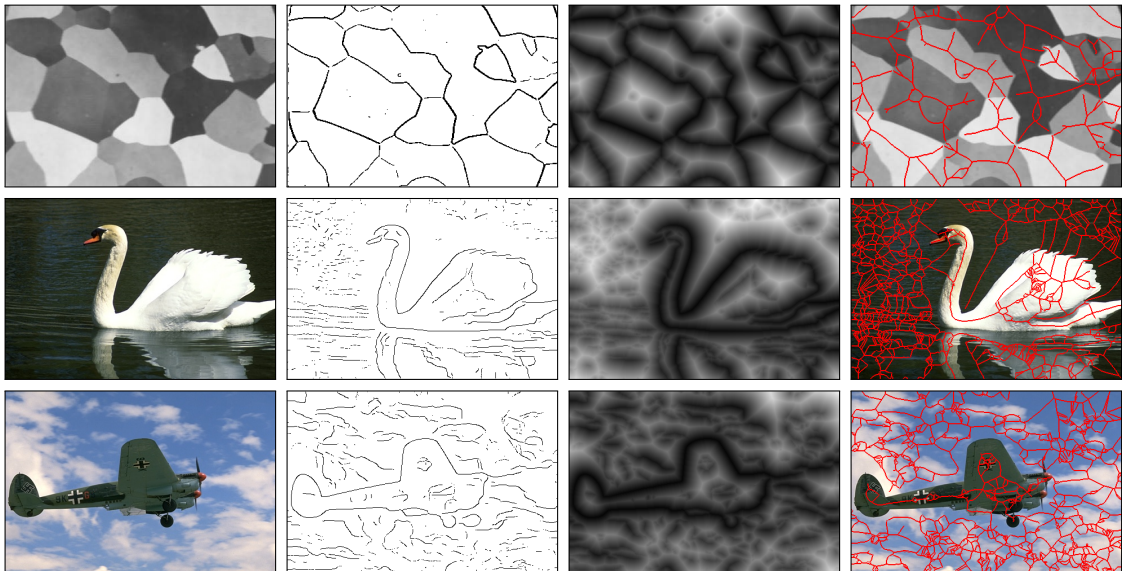
---

**Algorithm 2** Weighted Medial Axis

---

```
1: procedure MEDIAL(distance map  $h$ , source map  $s$ )
2:   initialize  $q, r$ ; construct  $A_+$ ;
3:   for  $x \in X$  do  $r(x) \leftarrow \emptyset$ ; label  $x$  as far;
4:   for  $x \in X$  do if  $x \succ x$  then label  $x$  as done;
5:   for  $x \in A_+$  do PROP( $x$ )
6:   while  $\neg q$ .EMPTY( ) do
7:      $x \leftarrow q$ .POP( );
8:     label  $x$  as done;
9:     for  $y \diamond x, \neg y$  done do SCAN( $x, y$ )
10:    if  $r(x) \neq 0$  then for  $y * x, y$  far do PROP( $y$ );
11:  end while
12:  return residue  $r$ 
13: end procedure
14:
15: procedure PROP(point  $x$ )
16:   $q$ .PUSH( $x$ );
17:  label  $x$  as near;
18: end procedure
19:
20: procedure SCAN(point  $x$ , point  $y$ )
21:   $\rho \leftarrow \text{res}(x, y)$ ;
22:  if  $s(x) = s(y) \vee \rho < \text{scale}$  then return
23:  if  $\rho > r(y) \wedge y$  far then PROP( $y$ );
24:   $r(x) \leftarrow \max(r(x), \rho)$ ;
25:   $r(y) \leftarrow \max(r(y), \rho)$ ;
26: end procedure
```

---



Σχήμα 5.4: Από αριστερά προς τα δεξιά:εικόνες εισόδου, σύνολα πηγών (μαύρο χρώμα), σταθμισμένοι μετασχηματισμοί απόστασης και αντίστοιχοι διάμεσοι άξονες (κόκκινο χρώμα).





# Κεφάλαιο 6

## Αρχική κατάτμηση με αποσύνθεση του διάμεσου άξονα

### 6.1 Αποσύνθεση του διάμεσου άξονα

Τεχνικές αποσύνθεσης διάμεσου άξονα χρησιμοποιούνται συχνά σε προβλήματα υπολογιστικής γεωμετρίας όπως η αποσύνθεση χωρίου σε δυαδικές εικόνες [23]. Η προσέγγιση που χρησιμοποιείται στα πλαίσια της εργασίας αυτής για την αποσύνθεση του διάμεσου άξονα είναι η ίδια που χρησιμοποιείται και από τους Αβρίθη και Ραπαντζίκο [5].

Ενώ οι περισσότερες εργασίες στη βιβλιογραφία χρησιμοποιούν τον διάμεσο άξονα για την περιγραφή ενός σχήματος, η χρήση του εδώ είναι για την περιγραφή ολόκληρης της εικόνας. Ο διάμεσος άξονας αποσυντίθεται σε συνιστώσες και ένας μη κατευθυνόμενος γράφος με βάρη  $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$  κατασκευάζεται ως εξής:

- (α) Οι κορυφές  $\mathcal{V}$  αντιστοιχούν στα τοπικά μέγιστα του μετασχηματισμού απόστασης.
- (β) Οι ακμές  $\mathcal{E}$  αντιστοιχούν στα τοπικά ελάχιστα κατά μήκος του σκελετού, δηλαδή, στα σημεία σέλας του διάμεσου άξονα.
- (γ) Τα βάρη των ακμών  $w(\mathcal{G}) : \mathcal{E} \rightarrow \mathbb{R}$  αποτελούν συνάρτηση του ύψους στα σημεία σέλας.

Ο αλγόριθμος medial axis decomposition (MAD) [4] κατασκευάζει τον άνωθι γράφο  $\mathcal{G}$  με δεδομένα το μετασχηματισμό απόστασης  $h = \mathcal{D}(f)$  και τον αντίστοιχο διάμεσο άξονα  $A(f)$ . Σημεία εκκίνησης του αλγορίθμου είναι τα τοπικά μέγιστα του διάμεσου άξονα:

$$\hat{A}_+(f) = A_+(f) \cap A(f) \quad (6.1)$$

και η διάδοση πραγματοποιείται προς χαμηλότερες τιμές ύψους όπως αναφέρεται και στον αλγόριθμο 3. Γίνεται και πάλι χρήση μιας ουράς προτεραιότητας αλλά με αντίθετη προτεραιότητα σε σχέση με τους προηγούμενους αλγορίθμους λόγω της κατεύθυνσης διάδοσης. Κάθε  $x \in X$  λαμβάνει μία ετικέτα  $\kappa(x)$  που αντιστοιχεί σε μία κορυφή του

γράφου  $\mathcal{G}$ . Ο γράφος  $\mathcal{G}$  κατασκευάζεται σταδιακά με εισάγοντας μια καινούργια κορυφή όποτε εξετάζεται ένα τοπικό μέγιστο για πρώτη φορά και μια ακμή όταν εξετάζονται δύο γειτονικά σημεία με διαφορετικές ετικέτες. Η διαδικασία αυτή είναι ισοδύναμη με εφαρμογή μετασχηματισμού watershed στον ανεστραμμένο μετασχηματισμό απόστασης αλλά μόνο πάνω στον διάμεσο άξονα.

---

### Algorithm 3 Medial Axis Decomposition

---

```

1: procedure MAD(distance map  $h$ , medial axis  $A$ )
2:   initialize  $q, \mathcal{G}$ ; construct  $\hat{A}_+$ 
3:   for  $x \in A$  do  $\kappa(x) \leftarrow \emptyset$ ; label  $x$  as far
4:   for  $x \in \hat{A}_+$  do PROP( $x$ )
5:   while  $\neg q.$ EMPTY( ) do
6:      $x \leftarrow q.$ POP( );
7:     label  $x$  as done;
8:     for  $y * x, y \in A$  do SCAN( $x, y$ )
9:     if  $\kappa(x) = \emptyset$  then  $\kappa(x) \leftarrow \mathcal{G}.$ VERTEX( $x$ )
10:  end while
11:  return graph  $\mathcal{G}$ 
12: end procedure
13:
14: procedure PROP(point  $x$ )
15:    $q.$ PUSH( $x, \lfloor -h(y) \rfloor$ );
16:   label  $x$  as near;
17: end procedure
18:
19: procedure SCAN(point  $x$ , point  $y$ )
20:   if  $y$  far then PROP( $y$ )
21:   if  $\kappa(y) = \emptyset$  return
22:   if  $\kappa(x) = \emptyset$  then  $\kappa(x) \leftarrow \kappa(y)$ ; return
23:   if  $\kappa(x) \neq \kappa(y)$  then  $\mathcal{G}.$ EDGE( $\kappa(x), \kappa(y), w(x)$ )
24: end procedure

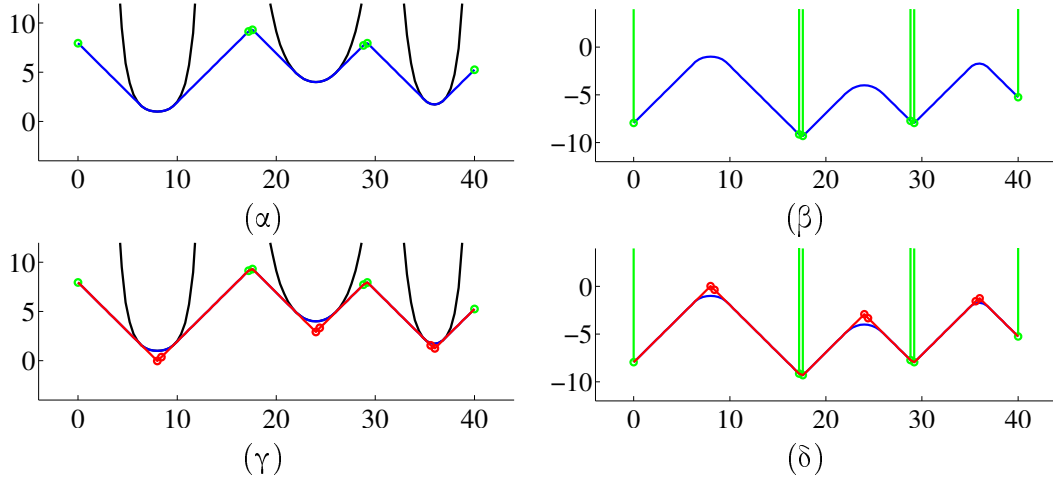
```

---

## 6.2 Δυϊκός μετασχηματισμός απόστασης

Στη συνέχεια, πραγματοποιείται η κατάτμηση της εικόνας μέσω μιας διαδικασίας ανακατασκευής. Μια ιδιότητα δυϊκότητας επιτρέπει την εφαρμογή της ανακατασκευής μέσω του αλγορίθμου EGM. Για οποιαδήποτε συνάρτηση  $f : U \rightarrow \mathbb{V}$ , ο τελεστής επέκτασης ορίζεται ως  $f|^X = f \cup ((X \setminus U) \times \{-\infty\})$ . Ο επεκταμένος τελεστής διάμεσου άξονα  $\mathcal{M}$  ορίζεται ως  $\mathcal{M}(f) = \mathcal{A}(f)|^X$  για  $f \in \mathbb{F}$ . Αφού ο τελεστής  $\mathcal{M}(f)$  ορίζεται στο  $X$  έχουμε:

**Πρόταση 6.2.1** Έστω συνάρτηση  $f$  και  $g = \mathcal{M}(f)$  σε έναν ευκλείδιο χώρο. Ορίζουμε  $f' = -\mathcal{M}(-g)$ ,  $g' = \mathcal{M}(f')$ . Τότε η συνάρτηση πηγών  $\mathcal{S}$  και ο διάμεσος άξονας  $\mathcal{A}$  είναι δυϊκά:



Σχήμα 6.1: Παράδειγμα δυϊκότητας πηγών και διάμεσου άξονα στη μία διάσταση. Οι συναρτήσεις των (β),(δ) είναι ανεστραμμένες παραλλαγές των (α), (γ). (α) Μαύρο χρώμα:  $f$ , μπλε χρώμα:  $\mathcal{D}(f)$ , πράσινες τελείες:  $\mathcal{A}(f)$ . Η  $f$  παίρνει χαμηλές τιμές στα όρια των περιοχών και υψηλές στο εσωτερικό των περιοχών. (β) μπλε χρώμα:  $-\mathcal{D}(f)$ , πράσινη γραμμή:  $-g$  φoor  $g = \mathcal{M}(f)$ , πράσινες τελείες:  $-\mathcal{A}(f) = \mathcal{S}(-g)$ . (γ) Κόκκινες τελείες:  $\mathcal{S}(f') = -\mathcal{A}(-g)$ . (δ) Κόκκινη γραμμή:  $\mathcal{D}(-g)$ , κόκκινες τελείες:  $\mathcal{A}(-g)$ .

$$(a) -\mathcal{S}(-g) = \mathcal{A}(f)$$

$$(β) \mathcal{S}(f') = -\mathcal{A}(-g) \subseteq \mathcal{S}(f)$$

$$(ς) g' = g.$$

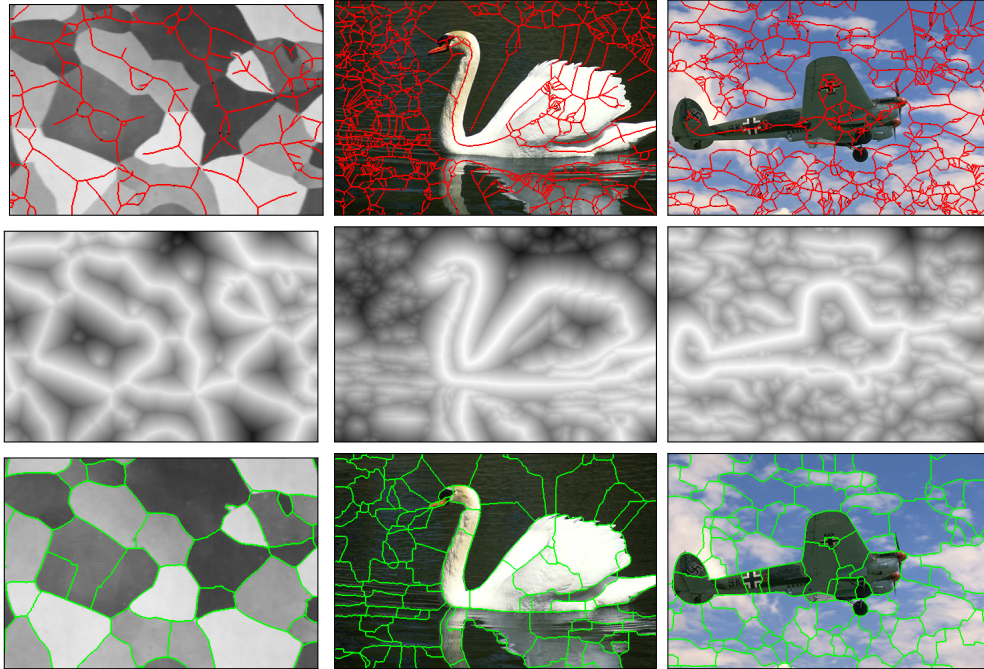
Στην πράξη, δεδομένης της συνάρτησης απόστασης  $h = \mathcal{D}(f)$  και του διάμεσου άξονα  $\mathcal{A}(f)$ , ο αλγόριθμος EGM με είσοδο  $g$ :

$$g(x) = \begin{cases} -h(x), & x \in \mathcal{A}(f) \\ +\infty, & \text{αλλιώς} \end{cases} \quad (6.2)$$

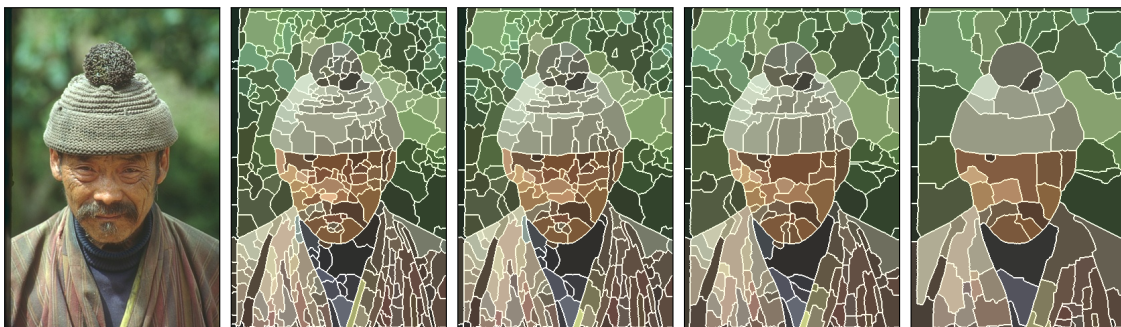
Οι ετικέτες  $\kappa$  από τον αλγόριθμο MAD χρησιμοποιούνται για την κατασκευή ετικετών περιοχών  $\kappa(x)$  για όλα τα  $x \in X$ . Η αρχική κατάτμηση  $\mathcal{P}_0$  λαμβάνεται ως εξής:

$$\kappa(x) = \kappa(\hat{s}(x)), \quad \text{για κάθε } x \in X \quad (6.3)$$

όπου με  $\hat{s}(x)$  συμβολίζεται η πηγή του  $x \in X$  όπως προέκυψε από τον δυϊκό μετασχηματισμό απόστασης. Οι ετικέτες  $\kappa$  που αντιστοιχούν στην αρχική κατάτμηση  $\mathcal{P}_0$  θα τις συμβολίζουμε ως  $\kappa_0$ . Η εικόνα αναπαρίσταται ως disjoint-set forest δομή όπου κάθε pixel αντιστοιχεί σε έναν κόμβο.



Σχήμα 6.2: Από πάνω προς τα κάτω: εικόνες εισόδου και διάμεσος άξονας (κόκκινο), δυϊκός μετασχηματισμός απόστασης και αρχική κατάτμηση. Με πράσινο χρώμα είναι τα όρια περιοχών της αρχικής κατάτμησης



Σχήμα 6.3: Επίδραση της παραμέτρου κλίμακας  $\sigma$  της συνάρτησης  $f$  στην αρχική κατάτμηση. Αρχική κατάτμηση για  $\sigma = 0.25, 0.5, 1.0$  ανδ  $2.0$  αντίστοιχα.

# Κεφάλαιο 7

## Τεχνικές συνένωσης γειτονικών περιοχών

### 7.1 Αποδοτική συνένωση περιοχών με βάση την μεταξύ τους ομοιομορφία

Η πρώτη τεχνική που προτείνουμε είναι μια τεχνική αποδοτικής συνένωσης περιοχών με βάση ένα μέτρο ομοιομορφίας που εισάγουμε. Επιστρέφοντας στην γραμμή 23 του αλγορίθμου 3, ορίζουμε το βάρος κάθε ακμής  $e \in \mathcal{E}$ , για την τεχνική αυτή, ως ακολούθως:

$$w(e) = h(x(e)) \quad (7.1)$$

όπου  $x(e)$  το αντίστοιχο σημείο σέλας.

Το άνωθι μέγεθος εκφράζει ομοιομορφία μεταξύ δύο γειτονικών κορυφών του γράφου ή ισοδύναμα μεταξύ δύο περιοχών της εικόνας. Το γιατί εξηγείται στη συνέχεια. Αρχικά, υπενθυμίζουμε τον ορισμό του μετασχηματισμού απόστασης στο σημείο  $x$ :

$$h(x) = d(x, s(x)) + f(s(x)) \quad (7.2)$$

Μεγάλη τιμή του ύψους στο σημείο σέλας καταδεικνύει την ύπαρξη μιας αντίστοιχα μεγάλης ασυνέχειας στο σύνολο των πηγών. Αυτή η ασυνέχεια εκφράζεται από τον όρο  $d(x, s(x))$ . Επιπροσθέτως, θα μπορούσε να καταδεικνύει ύπαρξη ασθενούς ορίου, δηλαδή μικρής τιμής κλίσης, σε περίπτωση υψηλής τιμής του παράγοντα  $f(s(x))$ . Και στις δύο περιπτώσεις, το ύψος του σημείου σέλας περιγράφει ένα μέτρο ομοιομορφίας μεταξύ γειτονικών περιοχών.

Μεταξύ δύο περιοχών είναι δυνατή η ύπαρξη πολλαπλών σημείων σέλας ή ισοδύναμα να ενώνονται με πολλαπλές ακμές. Ορίζουμε ως τελικό μέτρο ομοιομορφίας μεταξύ δύο περιοχών  $R, R' \subseteq X$  μέγιστο βάρος ακμής που τις συνδέει, δηλαδή:

$$\text{Sim}(R, R') = \max \{w(u, v) : u \in R, v \in R', (u, v) \in \mathcal{E}\} \quad (7.3)$$

Εάν οι  $R$  ανδ  $R'$  δεν συνδέονται από κάποια ακμή τότε θεωρούμε πως  $\text{Sim}(R, R') = 0$ .

Εισάγουμε ένα κατηγορήμα  $Q(R, R')$  για να ελέγχουμε εάν υπάρχει όριο μεταξύ δύο περιοχών  $R, R'$ . Εάν αυτό είναι αληθές για δύο περιοχές τότε δεν θα πρέπει να συνενώνονται. Το εν λόγω κατηγορήμα ελέγχει εάν η ομοιότητα μεταξύ δύο περιοχών,  $\text{Sim}(R, R')$ , είναι μικρότερη ή ίση από μια συνάρτηση κατωφλίου  $\tau(R, R')$ :

$$Q(R, R') = \begin{cases} \text{αληθ'ες} , & \text{ε'αν } \text{Sim}(R, R') < \tau(R, R') \\ \text{ψευδ'ες} , & \text{αλλι'ως} \end{cases} \quad (7.4)$$

Ο αλγόριθμος *efficient similarity merging* (ESM) που υλοποιήσαμε έχει τα ίδια βασικά βήματα με τον [13]. Αρχικά, ταξινομούμε τις ακμές κατά φθίνουσα σειρά βαρών. Στη συνέχεια, περνάμε από όλες τις ακμές με αυτή τη σειρά. Συνενώνουμε γειτονικές περιοχές εάν το αντίστοιχο κατηγορήμα είναι ψευδές. Ο αλγόριθμος συνοψίζεται παρακάτω.

---

#### Algorithm 4 Efficient Similarity Merging

---

```

1: procedure ESM(graph  $\mathcal{G}$ , initial partition  $\mathcal{P}_0$ )
2:   Sort  $\mathcal{E} = (e_1, \dots, e_m)$  by non increasing edge weight.
3:   for  $q = 1, \dots, m$  do
4:     Let  $e_q = (u, v)$ .
5:      $R_u \leftarrow \text{FIND}(u)$ 
6:      $R_v \leftarrow \text{FIND}(v)$ 
7:     if  $((-Q(R_u, R_v)) \wedge (R_u \neq R_v))$  then UNION( $R_u, R_v$ )
8:   end for
9:   return final partition  $\mathcal{P}_m$ 
10: end procedure

```

---

Αυτό που δεν έχουμε ακόμα διευκρινίσει είναι η συνάρτηση κατωφλίου. Ορίζουμε τη συνάρτηση με ορίσματα δύο περιοχές  $R, R'$  ως:

$$\tau(R, R') = \min(\tau(R), \tau(R')) \quad (7.5)$$

Για τη συνάρτηση κατωφλίου μίας περιοχής δοκιμάσαμε δύο επιλογές. Η πρώτη επιλογή αποτελεί η χρήση μίας απλής σταθεράς:

$$\tau(R) = \tau \quad (7.6)$$

όπου  $\tau$  μια σταθερά. Η δεύτερη επιλογή που δοκιμάσαμε εξαρτάται από το εμβαδό  $|R|$  της εκάστοτε περιοχής  $R$ . Δηλαδή για μικρότερη περιοχή απαιτείται και ανάλογα μικρότερο μέτρο ομοιότητας:

$$\tau(R) = |R|/k \quad (7.7)$$

όπου  $k$  ένας παράγοντας κλίμακας.

## 7.2 Συνένωση περιοχών με βάση τον παράγοντα κατακερματισμού σχήματος

Στην παράγραφο αυτή εισάγουμε μία εναλλακτική τεχνική συνένωσης περιοχών με βάση τον παράγοντα κατακερματισμού σχήματος. Η τεχνική αποσύνθεσης του διάμεσου άξονα που έχουμε χρησιμοποιήσει επιτρέπει την ακριβή μέτρηση του κατακερματισμού ή της κλειστότητας των ορίων μιας περιοχής. Αντιμετωπίζοντας την κατάτμηση εικόνων ως αναζήτηση προς ολοένα και πιο κλειστές περιοχές, εισάγουμε μια καινούργια τεχνική συνένωσης περιοχών και αντίστοιχα ένα διαφορετικό κατηγορήμα που βασίζονται στον παράγοντα κατακερματισμού σχήματος που θα αναλύσουμε στη συνέχεια.

Η χρησιμοποιηθείσα τεχνική αποσύνθεσης του διάμεσου άξονα μας βοηθά όχι μόνο να υπερβούμε τον κατακερματισμό του συνόλου των πηγών αλλά και να τον μετρήσουμε διότι σε κάθε ασυνέχεια των πηγών αντιστοιχεί ένα τοπικό ελάχιστο του μετασχηματισμού απόστασης πάνω στο διάμεσο άξονα, δηλαδή ένα σημείο σέλας. Επιστρέφοντας πάλι στη γραμμή 23 του αλγορίθμου 3, ορίζουμε το βάρος  $w(e)$  κάθε ακμής  $e \in \mathcal{E}$  ως το πλάτος της αντίστοιχης ασυνέχειας του συνόλου των πηγών. Το βάρος αυτό είναι ανάλογο της απόστασης  $d(x(e), s(x(e)))$ , όπου  $x(e)$  είναι το αντίστοιχο σημείο σέλας.

Ο παράγοντας κατακερματισμού σχήματος ορίζεται ως για μία περιοχή  $R$  εμβαδού  $|R|$  και αντίστοιχου συνόλου ακμών  $E(R)$ , ορίζεται ως [5]:

$$\phi(R) = \frac{1}{|R|} \sum_{e \in E(R)} w^2(e) \quad (7.8)$$

όπου  $\phi(R) = 0$  αν  $E(R) = \emptyset$ . Ο παράγοντας αυτός είναι αδιάστατος και αμετάβλητος σε αλλαγές κλίμακας.

Στη συνέχεια, ορίζουμε ένα διαφορετικό κατηγορήμα για δύο περιοχές  $R, R'$  :

$$Q(R, R') = \max(\phi(R'), \phi(R)) > \tau \quad (7.9)$$

όπου  $\tau$  μια σταθερά. Στο σημείο αυτό πρέπει να υπογραμμίσουμε πως όταν συνενώνονται δύο γειτονικές περιοχές  $(R, R' \cap R)$ , ο παράγοντας κατακερματισμού σχήματος της καινούργιας περιοχής  $R \cup R'$  υπολογίζεται ως εξής:

$$\phi(R \cup R') = \frac{\sum_{e \in E(R)} w^2(e) + \sum_{e \in E(R')} w^2(e) - 2 \cdot \sum_{e \in E(R) \cap E(R')} w^2(e)}{a(R) + a(R')} \quad (7.10)$$

Ο αλγόριθμος συνένωσης *shape fragmentation based merging* (SFM) είναι σε γενικές γραμμές παρόμοιος με τον αλγόριθμο της προηγούμενης παραγράφου εκτός από δύο διαφορές: συνενώνουμε γειτονικές περιοχές εάν το αντίστοιχο κατηγορήμα είναι αληθές και κάθε φορά που προκύπτει μια συνένωση γειτονικών περιοχών πρέπει να ανανεώσουμε τον παράγοντα κατακερματισμού σχήματος.

---

**Algorithm 5** SFM

---

```
1: procedure MERGING(graph  $\mathcal{G}$ , initial label map  $\kappa_0$ )
2:   Sort  $\mathcal{E} = (e_1, \dots, e_m)$  by non increasing edge weight.
3:   for  $q = 1, \dots, m$  do
4:     Let  $e_q = (u, v)$ ,  $u \in R$  and  $v \in R'$ 
5:      $\kappa_u \leftarrow \text{FIND}(u)$ 
6:      $\kappa_v \leftarrow \text{FIND}(v)$ 
7:     if  $Q(R, R')$  then
8:       UNION( $\kappa_u, \kappa_v$ )
9:       UPDATE( $\phi(R \cup R')$ ) according to (7.10)
10:    end if
11:  end for
12:  return label map  $\kappa$ 
13: end procedure
14:
```

---

### 7.3 Ιεραρχική κατάτμηση εικόνων και υπερμετρική εικόνα ορίων

Ο Arbelaez [1] ορίζει υπερμετρική εικόνα ορίων (ultrametric contour map (UCM)) ως μέσο για επίτευξη ιεραρχικής κατάτμησης εικόνων. Η βασική ιδέα είναι η αναπαράσταση ολόκληρης της διαδικασίας κατάτμησης εικόνων, από την αρχική κατάτμηση μέχρι την τελική κατάτμηση όπου η εικόνα αποτελείται από ένα τμήμα, σε μία εικόνα γκριζου τόνου. Η εικόνα αυτή ονομάζεται υπερμετρική εικόνα ορίων και η λήψη μιας κατάτμησης από την εικόνα αυτή σε οποιαδήποτε κλίμακα ανάγεται σε μια απλή λειτουργία καταφλίσωσης. Οι ορισμοί και η ανάλυση που παραθέτουμε στη συνέχεια παρουσιάστηκαν στο [1].

Έστω  $\mathcal{P}_0$  μια αρχική κατάτμηση του χώρου  $X$  και  $\lambda \in \mathbb{R}$  μία παράμετρος κλίμακας. Ο τελεστής ιεραρχικής κατάτμησης (hierarchical segmentation operator (HSO)) αποτελεί μια αντιστοίχιση της κατάτμησης  $\mathcal{P}_\lambda$  στο ζεύγος  $(\mathcal{P}_0, \lambda)$  έτσι ώστε:

$$\mathcal{P}_\lambda = \mathcal{P}_0, \quad \forall \lambda \leq 0 \quad (7.11)$$

$$\exists \lambda_1 \in \mathbb{R}^+ : \mathcal{P}_\lambda = \{X\}, \quad \forall \lambda \geq \lambda_1 \quad (7.12)$$

$$\lambda \leq \lambda' \Rightarrow \mathcal{P}_\lambda \sqsubseteq \mathcal{P}_{\lambda'} \quad (7.13)$$

όπου το σύμβολο  $\sqsubseteq$  δηλώνει τη σειρά των κατατμήσεων, έτσι ώστε  $\mathcal{P} \sqsubseteq \mathcal{P}'$  αν:

$$\forall R_i \in \mathcal{P}, \exists R'_i \in \mathcal{P}' : R_i \subseteq R'_i \quad (7.14)$$

Σύμφωνα με την (7.14), κατατμήσεις σε διαφορετικές κλίμακες είναι φωλιασμένες και δημιουργούν μια ιεραρχική δομή:

$$\mathcal{H} = \{R \subseteq X \mid \exists \lambda : R \in \mathcal{P}_\lambda\} \quad (7.15)$$



Ο δείκτης διαστρωμάτωσης μιας περιοχής  $R$  ορίζεται ως η κλίμακα  $\lambda$  στην οποία εμφανίζεται μια περιοχή στην ιεραρχία  $\mathcal{H}$ :

$$I(R) = \inf \{ \lambda \in [0, \lambda_1] : R \in \mathcal{P}_\lambda \} \quad (7.16)$$

Το ζεύγος  $(\mathcal{H}, I)$  ονομάζεται δεικτιοδοτημένη ιεραρχία των κατατμήσεων του  $X$ . Αυτή η ιεραρχία αναπαρίσταται από ένα δενδόγραμμα, όπου το ύψος κάθε περιοχής αντιστοιχεί στο δείκτη διαστρωμάτωσης της  $I(R)$ . Η κατασκευή της παραπάνω ιεραρχίας είναι ισοδύναμη με τον ορισμό ενός μετρικού ή μιας απόστασης μεταξύ δύο στοιχείων  $x, y \in X$ :

$$\mathcal{Y}(x, y) = \inf \{ I(R) \mid x \in R \wedge y \in R \wedge R \in \mathcal{H} \} \quad (7.17)$$

Η απόσταση  $\mathcal{Y}$  πρέπει να ανήκει σε μια ειδική κατηγορία αποστάσεων που ονομάζονται *υπερμετρικές* και εκτός της τριγωνικής ιδιότητας ικανοποιούν και την παρακάτω ιδιότητα:

$$\mathcal{Y}(x, y) \leq \max \{ \mathcal{Y}(x, z), \mathcal{Y}(y, z) \}, \quad x, y, z \in X \quad (7.18)$$

Εναλλακτικά, η κατάτμηση εικόνων μπορεί να οριστεί βάση των ορίων των περιοχών. Μια κατάτμηση  $K$  του  $X$  μπορεί να οριστεί ως ένα σύνολο καμπυλών Jordan, τα όρια των περιοχών. Οι περιοχές  $R_i$  της  $K$  είναι οι συνεκτικές συνιστώσες του  $X \setminus K$ . Το όριο μεταξύ των περιοχών  $R_i$  και  $R_j$  συμβολίζεται ως  $\partial_{ij}$ .

Η ιεραρχική κατάτμηση εικόνων μπορεί δυϊκά να εκφραστεί βάση των ορίων των περιοχών:

$$\mathcal{K}_\lambda = \mathcal{K}_0, \quad \forall \lambda \leq 0 \quad (7.19)$$

$$\exists \lambda_1 \in \mathbb{R}^+ : \mathcal{K}_\lambda = \partial X, \quad \forall \lambda \geq \lambda_1 \quad (7.20)$$

$$\lambda \leq \lambda' \Rightarrow \mathcal{K}_\lambda \supseteq \mathcal{K}_{\lambda'} \quad (7.21)$$

Έστω  $\mathcal{Y}$  μια υπερμετρική απόσταση. Η αντίστοιχη υπερμετρική εικόνα ορίων (ultrametric contour map (UCM)) ορίζεται ως:

$$\mathcal{C}(\mathcal{Y})(\partial) = \inf \{ \lambda \in [0, \lambda_1] \mid \partial \not\subseteq K_\lambda \}, \quad \forall \partial \in K_0 \quad (7.22)$$

Κάθε όριο δύο περιοχών στην υπερμετρική εικόνα ορίων λαμβάνει ως τιμή την κλίμακα στην οποία εξαφανίστηκε από την ιεραρχία των ορίων. Η υπερμετρική εικόνα ορίων αποτελεί μια αναπαράσταση μια ιεραρχικής κατάτμησης σε μία και μόνο εικόνα γκριζου τόνου. Εξ ορισμού, η κατωφλίωση της υπερμετρικής εικόνας ορίων σε μία κλίμακα  $\lambda$  μας παρέχει ένα σύνολο κλειστών καμπυλών, την κατάτμηση  $K$ .

Μία ανομοιομορφία  $\delta$  μεταξύ δύο περιοχών ονομάζεται υπερμετρική εάν το ζεύγος  $(\mathcal{H}, I)$  είναι δεικτιοδοτημένη ιεραρχία, όπου ο δείκτης  $I$  ορίζεται ως:

$$I(R_i \cup R_j) = \delta(R_i, R_j) \quad (7.23)$$

για όλες τις γειτονικές περιοχές  $(R_i, R_j) \in \mathcal{H}^2$ .

Η ανομοιομορφία  $\delta$  αποδεικνύεται πως είναι υπερμετρική αν και μόνο αν:

$$\delta(R_i, R_j) \leq \delta(R_i \cup R_j, R_k) \quad (7.24)$$

όπου  $(R_i, R_j)$  είναι ζεύγος περιοχών που ελαχιστοποιούν την ανομοιομορφία  $\delta$  και  $R_k$  είναι μια γειτονική περιοχή της  $R_i \cup R_j$  και εμφανίζεται στην κατάτμηση και μετά τη συνένωση των  $(R_i, R_j)$ .

Ο βασικός αλγόριθμος κατάτμησης ορίζει, αρχικά, ένα γράφο  $G = (\mathcal{P}_0, \mathcal{K}_0)$  όπου οι κορυφές αντιστοιχούν στις περιοχές  $\mathcal{P}_0$ , οι ακμές στα σύνορα  $\mathcal{K}_0$  μεταξύ γειτονικών περιοχών και τα βάρη  $W : \mathcal{K}_0 \rightarrow \mathbb{R}_+$  αποτελούν ένα μέτρο υπερμετρικής ανομοιομορφίας μεταξύ περιοχών. Ο αλγόριθμος αρχικά ταξινομεί τις ακμές του γράφου κατά αύξουσα σειρά βαρών και επαναληπτικά συνενώνει τις πιο όμοιες περιοχές:

- 1) Επιλογή ορίου ελαχίστου βάρους:  $C^* = \arg \min_{C \in \mathcal{K}_0} W(C)$
- 2) Έστω  $R_1, R_2 \in \mathcal{P}_0$  οι περιοχές που χωρίζονται από το  $C^*$
- 3)  $R = R_1 \cup R_2$  και ανανέωση του γράφου:  
 $\mathcal{P}_0 \leftarrow \mathcal{P}_0 \setminus \{R_1, R_2\} \cup R$  ανδ  $\mathcal{K}_0 \leftarrow \mathcal{K}_0 \setminus \{C^*\}$
- 4) Τέλος εάν το σύνολο  $\mathcal{K}_0$  είναι κενό. Αλλιώς, ανανέωση των βαρών  $W(\mathcal{K}_0)$  και επανάληψη από το βήμα 1.

Καθώς έχει υπολογιστεί η υπερμετρικός εικόνα ορίων, η τελική κατάτμηση λαμβάνεται με κατωφλίωσή της στην επιθυμητή κλίμακα.

### 7.3.1 Υπερμετρική ανομοιομορφία με βάση των παράγοντα κατακερματισμού κοινού συνόρου

Στην παράγραφο αυτή, ορίζουμε ένα μέτρο ανομοιομορφίας μεταξύ γειτονικών περιοχών που βασίζεται στον παράγοντα κατακερματισμού του ορίου μεταξύ γειτονικών περιοχών. Ο παράγοντας κατακερματισμού του ορίου μεταξύ των περιοχών  $R_i$  και  $R_j$  ορίζεται ως το πλάτος των ασυνεχειών του συνόλου των πηγών που σχετίζονται με το όριο των περιοχών αυτών δια το μήκος του ορίου:

$$\Phi(R_i, R_j) = \frac{\sum_{e \in E(R_i) \cap E(R_j)} w(e)}{L(\partial_{ij})} \quad (7.25)$$

Μεγάλη τιμή του παράγοντα κατακερματισμού δείχνει την απουσία ισχυρού ορίου μεταξύ των εν λόγω περιοχών. Δηλαδή, αποτελεί μέτρο ομοιομορφίας και όχι ανομοιομορφίας. Για το λόγο αυτό, ορίζουμε την υπερμετρική ανομοιομορφία  $\delta_\Phi$  ως εξής:

$$\delta_\Phi(R_i, R_j) = \exp(-\Phi(R_i, R_j)) \quad (7.26)$$

**Λήμμα 7.3.1** Η ανομοιομορφία  $\delta_\Phi$  είναι υπερμετρική.

### 7.3.2 Υπερμετρική ανομοιομορφία με βάση την ισχύ των ορίων

Η δεύτερη υπερμετρική ανομοιομορφία που υλοποιήθηκε στην παρούσα διπλωματική εργασία είναι η ισχύ του ορίου δύο γειτονικών περιοχών. Η ισχύς αυτή ορίζεται ως η μέση τιμή της εικόνας ακμών γκρίζου τόνου πάνω στο όριο των περιοχών. Αυτή η ανομοιομορφία έχει χρησιμοποιηθεί αρκετά στη βιβλιογραφία [1, 3] και ορίζεται ως:

$$\delta_g(R_i, R_j) = \frac{\sum_g(\partial_{ij})}{L(\partial_{ij})} \quad (7.27)$$

όπου  $\sum_g(\partial_{ij}) = \int_{\partial_{ij}} g(x(s))ds$ , δηλαδή το άθροισμα της εικόνας ακμών γκρίζου τόνου  $g$  πάνω στο όριο των δύο περιοχών.

**Λήμμα 7.3.2** Η ανομοιομορφία  $\delta_g$  της εξίσωσης (7.27) είναι υπερμετρική.



# Κεφάλαιο 8

## Αξιολόγηση

### 8.1 Συλλογή δεδομένων

Για την αξιολόγηση των προτεινόμενων τεχνικών κατάτμησης χρησιμοποιήθηκε η συλλογή δεδομένων του πανεπιστημίου Berkeley η οποία αποτελεί την κύρια συλλογή δεδομένων με στόχο την αξιολόγηση τεχνικών κατάτμησης εικόνων. Η συλλογή δεδομένων του πανεπιστημίου Berkeley (BSDS500) αποτελείται από 500 φυσικές και φωτογραφικές εικόνες οι οποίες έχουν κατατμηθεί από έναν αριθμό πέντε ανθρώπων κατά μέσο όρο. Η συλλογή δεδομένων χωρίζεται σε 300 εικόνες για εκπαίδευση της μεθόδου (training set) και 200 εικόνες για αξιολόγηση της μεθόδου (test set).

### 8.2 Κριτήρια αξιολόγησης

Στη συνέχεια, αναλύονται κάποια κριτήρια αξιολόγησης αυτομάτων κατατμήσεων όπως παρουσιάζονται στην τεκμηρίωση της συλλογής δεδομένων BSDS500 [3].

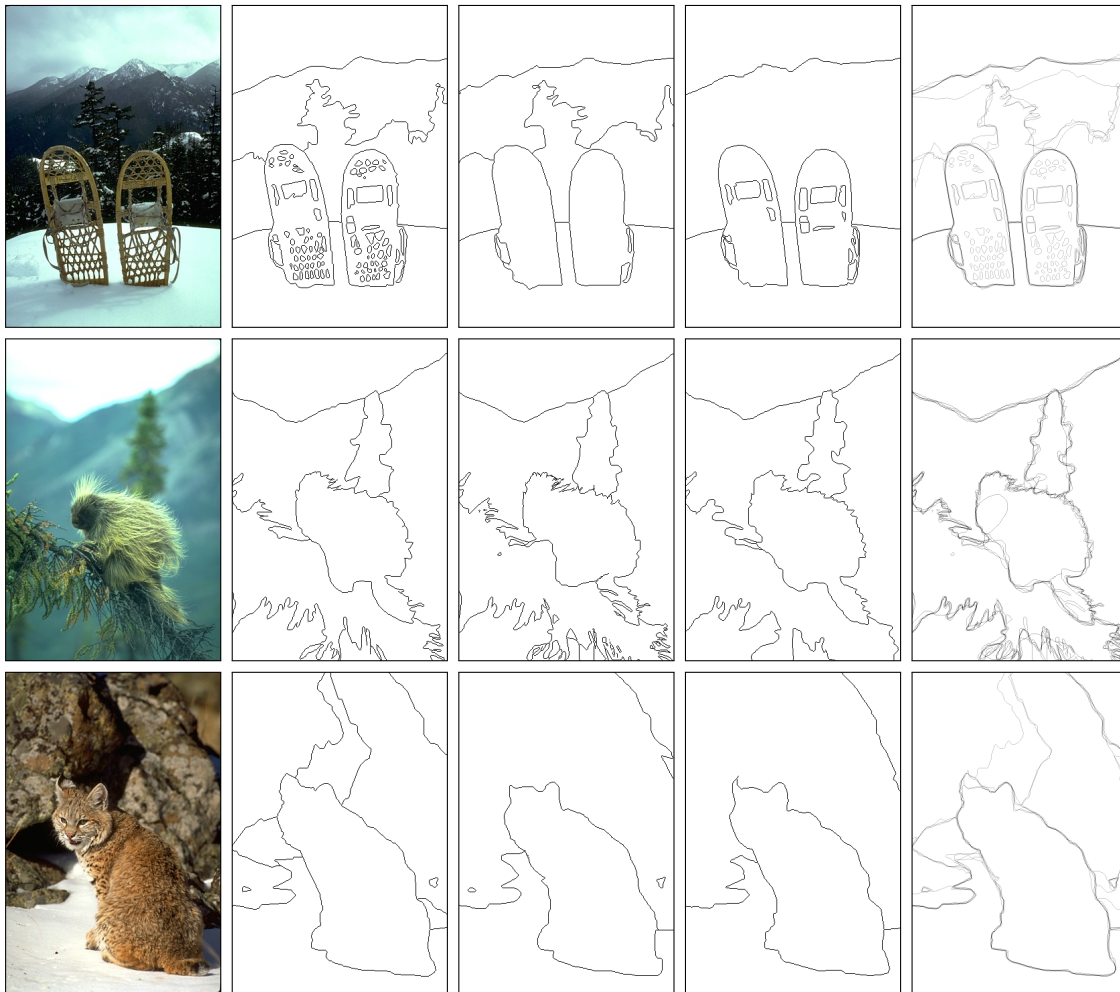
#### 8.2.1 Precision-recall

Δύο γνωστά μέτρα, δανεισμένα από την ανάκτηση πληροφοριών, εφαρμόσιμα σε προβλήματα όπως η κατάτμηση εικόνων, είναι το μέτρο ακρίβειας (*precision*) και το μέτρο ανάκτησης (*recall*). Ορίζοντας σαν  $|\cdot|$  το πλήθος των στοιχείων ενός συνόλου, τα μέτρα *precision* και *recall* δίνονται από τους παρακάτω τύπους:

$$Precision = \frac{| \{ \text{Relevant items} \} \cap \{ \text{Retrieved items} \} |}{| \{ \text{Retrieved items} \} |} \quad (8.1)$$

και αντίστοιχα:

$$Recall = \frac{| \{ \text{Relevant items} \} \cap \{ \text{Retrieved items} \} |}{| \{ \text{Relevant items} \} |} \quad (8.2)$$



Σχήμα 8.1: Ανθρώπινες κατατιμήσεις εικόνων του Berkeley Segmentation Dataset

Τα μέτρα precision και recall χρησιμοποιούνται για την αξιολόγηση τεχνικών ανίχνευσης ακμών αλλά και τεχνικών κατάτμησης εικόνων με τον ίδιο τρόπο καθώς και οι δύο κατηγορίες τεχνικών παράγουν ως αποτέλεσμα όρια περιοχών. Όσον αφορά αυτές τις κατηγορίες προβλημάτων, το μέτρο precision (P) εκφράζει το ποσοστό των πραγματικών ορίων, δηλαδή ορίων των ανθρώπινων κατατμήσεων, που ανιχνεύτηκαν επί του συνόλου των ανιχνευθέντων ενώ το μέτρο recall (R) εκφράζει το ποσοστό των πραγματικών ορίων που ανιχνεύτηκαν επί του συνόλου των πραγματικών ορίων της εικόνας. Για ανιχνευτές ακμών που παράγουν ως έξοδο grayscale εικόνες ακμών προκύπτει μια παραμετροποιημένη καμπύλη precision και recall διαφορετικές τιμές κατωφλίου. Ένα μέγεθος που συνοψίζει την παραμετροποιημένη αυτή καμπύλη είναι το μέτρο  $F$  που ορίζεται ως ο αρμονικός μέσο των μεγεθών precision και recall:

$$F = \frac{2PR}{P + R} \quad (8.3)$$

## 8.2.2 Variation of information

Το μέτρο Variation of Information (VI) χρησιμοποιείται για σύγκριση μεταξύ συσταδοποιήσεων (clusterings). Μετρά την απόσταση μεταξύ δύο συσταδοποιήσεων  $C$  και  $C'$  ως τη μέση υπό συνθήκη εντροπία τους:

$$VI(C, C') = H(C) + H(C') - 2I(C, C') \quad (8.4)$$

όπου με  $H$  και  $I$  συμβολίζονται η εντροπία και η κοινή πληροφορία αντίστοιχα. Για το πρόβλημα της κατάτμησης εικόνων, οι συσταδοποιήσεις που συγκρίνονται είναι η αυτόματη κατάτμηση του αλγορίθμου προς αξιολόγηση με τις ανθρώπινες κατατμήσεις.

## 8.2.3 Δείκτης Rand

Ο δείκτης Rand λειτουργεί ελέγχοντας τις αντιστοιχίες μεταξύ ζευγών στοιχείων σε δύο συστάδες. Ο δείκτης Rand μεταξύ μιας αυτόματης κατάτμησης  $S$  και μιας ανθρώπινης κατάτμησης  $G$  ορίζεται ως το πλήθος των ζευγών pixels που ανήκουν είτε στην ίδια περιοχή και στις κατατμήσεις είτε που ανήκουν σε διαφορετική περιοχή και στις δύο κατατμήσεις. Ο πιθανωτικός δείκτης Rand (probabilistic Rand index (PRI)) αποτελεί μια παραλλαγή του δείκτη Rand για να μπορούν να συγκριθούν πολλαπλές ανθρώπινες κατατμήσεις με μία αυτόματη και ορίζεται ως εξής:

$$PRI(S, \{G_k\}) = \frac{1}{T} \sum_{i < j} [c_{ij}p_{ij} + (1 - c_{ij})(1 - p_{ij})] \quad (8.5)$$

όπου  $T$  είναι το πλήθος των ζευγών των εικονοστοιχείων της εικόνας,  $c_{ij}$  είναι η λογική μεταβλητή που εκφράζει το γεγονός τα εικονοστοιχεία  $i$  και  $j$  να ανήκουν στην ίδια περιοχή και  $p_{ij}$  η αντίστοιχη πιθανότητα.

### 8.2.4 Επικάλυψη

Η επικάλυψη (covering) μεταξύ δύο περιοχών  $R_1$  και  $R_2$  ορίζεται ως:

$$\mathcal{O}(R_1, R_2) = \frac{|R_1 \cap R_2|}{|R_1 \cup R_2|} \quad (8.6)$$

Η επικάλυψη μιας κατάτμησης  $S_1$  από μια κατάτμηση  $S_2$  ορίζεται ως:

$$\mathcal{C}(S_1 \rightarrow S_2) = \frac{1}{N} \sum_{R_1 \in S_1} |R_1| \cdot \max_{R_2 \in S_2} \mathcal{O}(R_1, R_2) \quad (8.7)$$

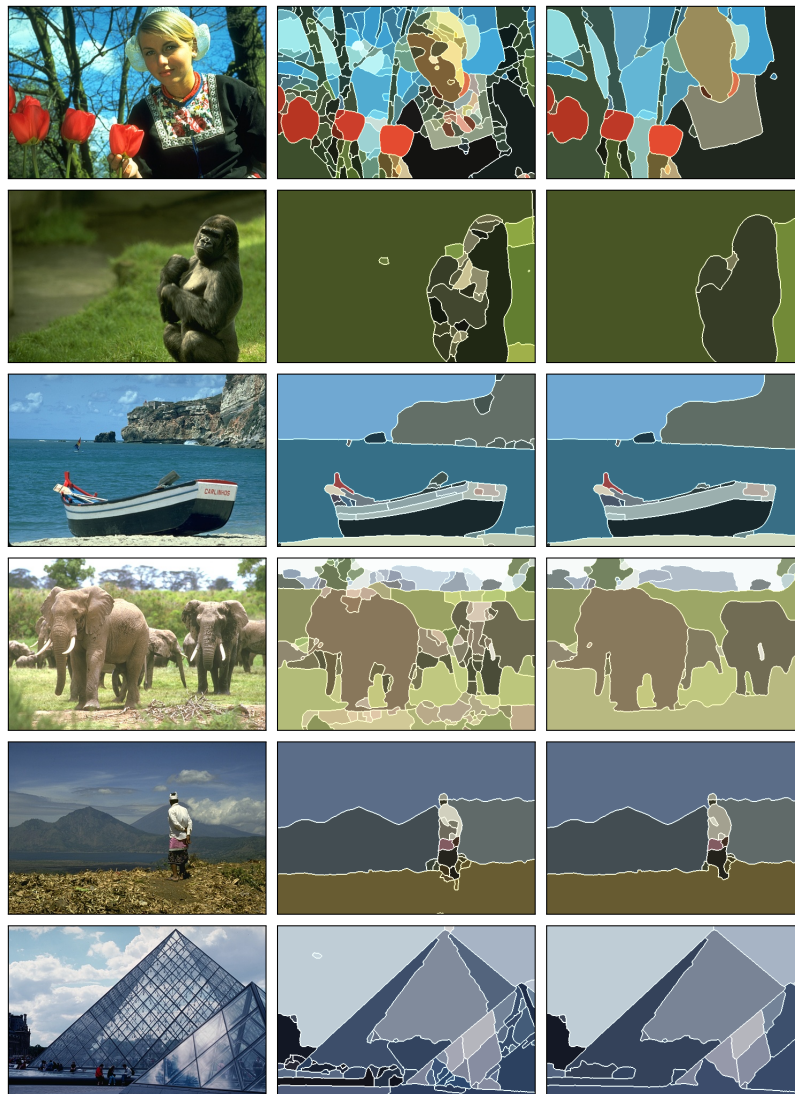
όπου  $N$  το πλήθος των pixels της εικόνας.

Με παρόμοια λογική, η επικάλυψη μιας αυτόματης κατάτμησης  $S$  από μια ομάδα πολλαπλών ανθρώπινων κατατμήσεων  $\{G_k\}$  υπολογίζεται από τη μέση τιμή της επικάλυψης της αυτόματης κατάτμησης  $S$  ξεχωριστά με κάθε ανθρώπινη  $\{G_k\}$ .

## 8.3 Αποτελέσματα

Στην παράγραφο αυτή παραθέτουμε κάποια αντιπροσωπευτικά αποτελέσματα των προτεινόμενων τεχνικών με είσοδο από τον *gPb* ανιχνευτή ακμών. Όλες οι προτεινόμενες τεχνικές βασίζονται σε μία και μόνο τιμή κατωφλίου. Η τιμή κατωφλίου που μεγιστοποιεί το μέτρο F για ολόκληρη τη συλλογή δεδομένων θα αναφέρεται ως *optimal dataset scale* (ODS) ενώ η τιμή κατωφλίου που μεγιστοποιεί το μέτρο F για μία συγκεκριμένη εικόνα θα αναφέρεται στη συνέχεια ως *optimal image scale* (OIS).

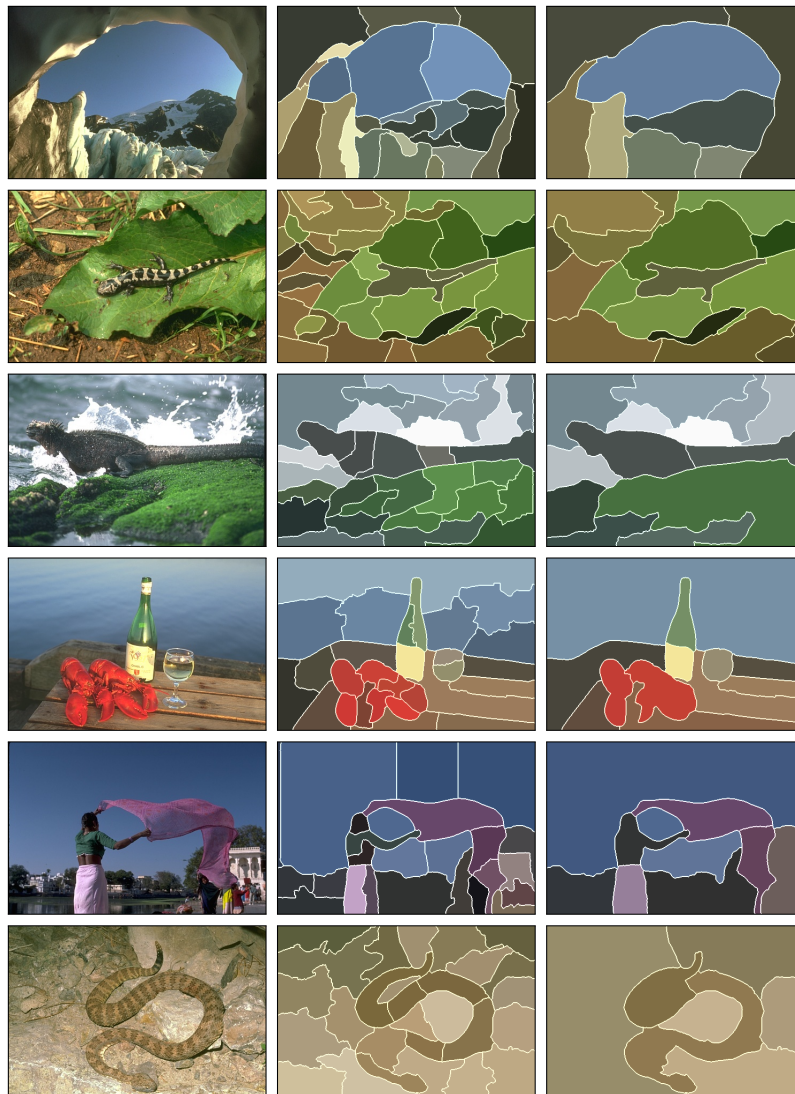




Σχήμα 8.2: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm με σταθερο κατώφλι. Από αριστερά προς τα δεξιά: εικόνα εισόδου και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



Σχήμα 8.3: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm με σταθερο κατώφλι. Από πάνω προς τα κάτω: εικόνα εισόδου και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



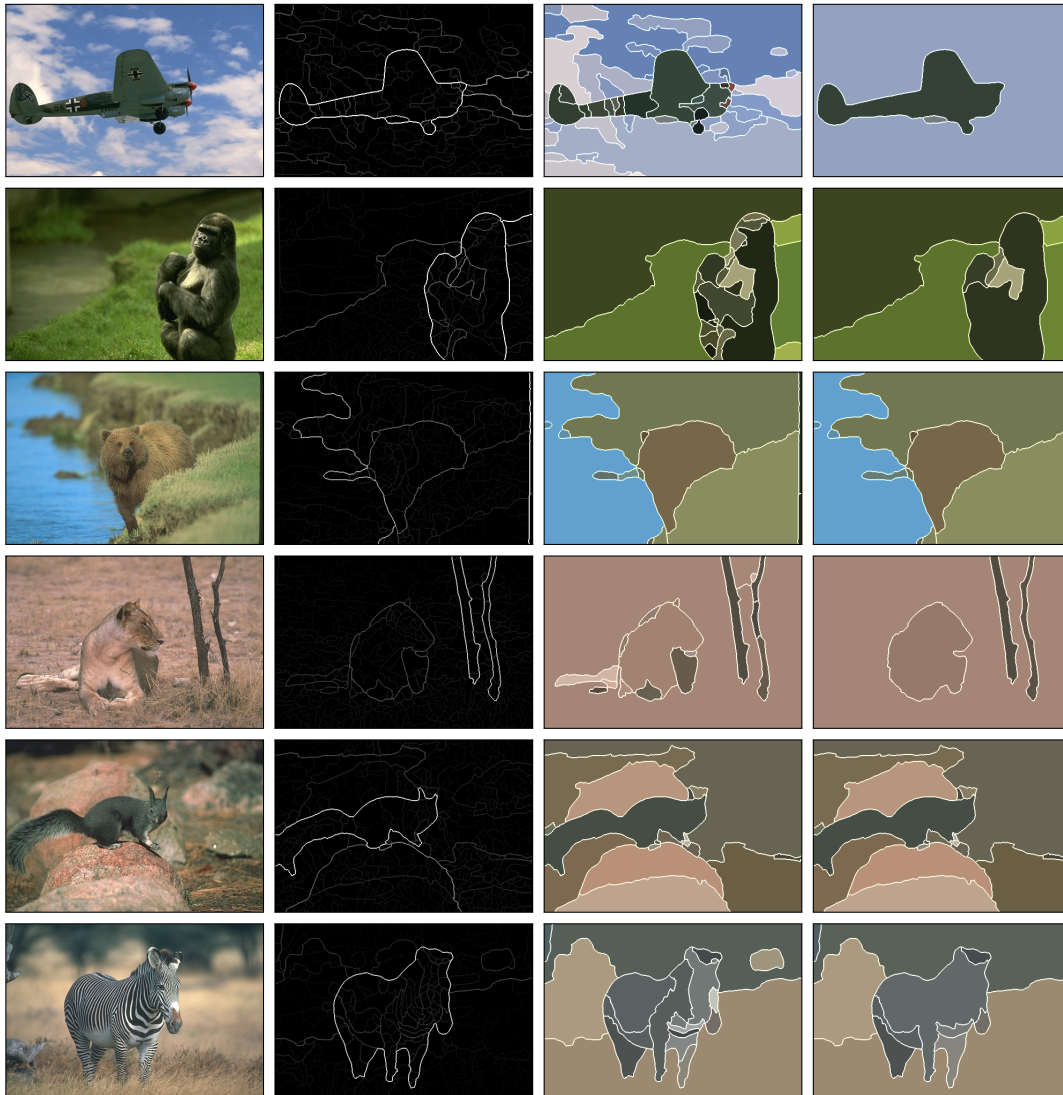
Σχήμα 8.4: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-esm με κατώφλι ανάλογο του εμβαδού της εκάστοτε περιοχής. Από αριστερά προς τα δεξιά: εικόνα εισόδου και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



Σχήμα 8.5: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-sfm. Από πάνω προς τα κάτω: εικόνα εισόδου και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



Σχήμα 8.6: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-ucm με μέτρο ανομοιομορφίας βασισμένο στον παράγοντα κατακερματισμού κοινού συνόρου. Από αριστερά προς τα δεξιά: εικόνα εισόδου και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



Σχήμα 8.7: Αποτελέσματα κατάτμησης με την τεχνική gPb-mad-ucm με μέτρο ανομοιομορφίας βασισμένο στον ισχύ των ορίων των περιοχών. Από αριστερά προς τα δεξιά: εικόνα εισόδου, υπερμετρική εικόνα ορίων και αυτόματες κατατμήσεις με ODS και IDS αντίστοιχα.



Σχήμα 8.8: Αποτελέσματα κατάρτησης με την τεχνική gPb-mad-ucm με μέτρο ανομοιομορφίας βασισμένο στον ισχύ των ορίων των περιοχών. Από πάνω προς τα κάτω: εικόνα εισόδου, υπερμετρική εικόνα ορίων και αυτόματες κατάρτησεις με ODS και IDS αντίστοιχα.

## 8.4 Αποτελέσματα αξιολόγησης

Στην παράγραφο αυτή παρέθετουμε τα αποτελέσματα αξιολόγησης των διαφόρων προτεινόμενων τεχνικών καθώς και σύγκρισή τους με γνωστούς αλγορίθμους της βιβλιογραφίας. Οι αλγόριθμοι αυτοί είναι συγκεκριμένα: ο αλγόριθμος gPb-owt-ucm [2] και η απλουστευμένη έκδοσή του Canny-owt-ucm, ο αλγόριθμος αποδοτικής κατάρτησης με χρήση γράφων [13] (Felz-Hutt), ο αλγόριθμος Mean Shift [11], η τεχνική Multiscale Normalized Cuts [12].

Όσον αφορά τις προτεινόμενες τεχνικές της παρούσας εργασίας, συμβολίζουμε με  $ucm_g$  την τεχνική ιεραρχικής κατάτμησης με βάση την ισχύ των ορίων ενώ με  $ucm_\phi$  την τεχνική ιεραρχικής κατάτμησης με βάση τον παράγοντα κατακερματισμού ορίου. Με  $esm_c$  και  $esm_a$  συμβολίζουμε αντίστοιχα τον αλγόριθμο αποδοτικής με βάση την ομοιότητα με χρήση σταθερού κατωφλίου και κατωφλίου ανάλογου του εμβαδού των περιοχών αντίστοιχα. Τέλος, συμβολίζουμε με  $gCanny$  την παραλλαγή του γνωστού ανιχνευτή ακμών Canny που υλοποιήθηκε.

Τα συνολικά αποτελέσματα αξιολόγησης με βάση τα όρια των ανιχνευθέντων περιοχών παρουσιάζονται στον Πίνακα 8.4 που ακολουθεί:

	BSDS500	
	ODS	OIS
Human	0.80	0.80
gPb-owt-ucm [2]	<b>0.73</b>	<b>0.76</b>
gPb-mad-ucm <sub>g</sub>	0.72	0.75
gPb-mad-esm <sub>c</sub>	0.69	0.72
gPb-mad-ucm <sub>ϕ</sub>	0.69	0.71
gPb-mad-sfm	0.67	0.70
gPb-mad-esm <sub>a</sub>	0.64	0.68
Mean Shift [11]	0.64	0.68
NCuts [12]	0.64	0.68
gCanny-mad-ucm <sub>g</sub>	0.61	0.65
Felz-Hutt [13]	0.61	0.64
Canny-owt-ucm [2]	0.60	0.64
gCanny-mad-esm <sub>c</sub>	0.58	0.61
gCanny-mad-esm <sub>a</sub>	0.58	0.61
gCanny-mad-sfm	0.56	0.59
gCanny-mad-ucm <sub>ϕ</sub>	0.55	0.58
gPb [2]	0.71	0.74
gCanny	0.61	0.64
Canny [2]	0.60	0.63

Πίνακας 8.1: Συνολικά αποτελέσματα αξιολόγησης με βάση τα όρια των ανιχνευθέντων περιοχών.

Τα συνολικά αποτελέσματα αξιολόγησης με βάση τις ανιχνευθείσες περιοχές της αυτόματης κατάτμησης παρουσιάζονται στον Πίνακα 8.4 που ακολουθεί:

## 8.5 Σχολιασμός

Από το σύνολο των προτεινόμενων τεχνικών της παρούσας διπλωματικής εργασίας, η τεχνική  $gPb-mad-ucm_g$  εμφανίζει την καλύτερη απόδοση σύμφωνα όλα τα κριτήρια αξιολόγησης. Εμφανίζει παραπλήσια ή ακόμα και καλύτερη απόδοση από τον state of



	BSDS500						
	Covering			PRI		VI	
	ODS	OIS	Best	ODS	OIS	ODS	OIS
Human	0.72	0.72	–	0.88	0.88	1.17	1.17
gPb-owt-ucm[2]	<b>0.59</b>	<b>0.65</b>	<b>0.74</b>	<b>0.83</b>	<b>0.86</b>	1.69	1.48
gPb-mad-ucm <sub>g</sub>	0.58	0.64	<b>0.74</b>	<b>0.83</b>	<b>0.86</b>	<b>1.62</b>	<b>1.39</b>
gPb-mad-esm <sub>c</sub>	0.55	0.62	0.71	0.82	<b>0.86</b>	1.83	1.51
Mean Shift [11]	0.54	0.58	0.66	0.79	0.81	1.85	1.64
Felz-Hutt [13]	0.52	0.57	0.69	0.80	0.82	2.21	1.87
gPb-mad-sfm	0.52	0.56	0.62	0.79	0.82	1.83	1.70
gPb-mad-esm <sub>a</sub>	0.51	0.54	0.60	0.79	0.80	1.86	1.82
Canny-owt-ucm [2]	0.49	0.55	0.66	0.79	0.83	2.19	1.89
myCanny-mad-ucm <sub>g</sub>	0.48	0.55	0.65	0.79	0.83	2.10	1.77
gPb-mad-ucm <sub>ϕ</sub>	0.46	0.54	0.63	0.77	0.80	2.07	1.81
NCuts [12]	0.45	0.53	0.67	0.78	0.80	2.23	1.89
gCanny-mad-esm <sub>c</sub>	0.45	0.53	0.63	0.78	0.83	2.31	1.91
gCanny-mad-sfm	0.42	0.49	0.58	0.77	0.80	2.18	1.95
gCanny-mad-esm <sub>a</sub>	0.40	0.47	0.53	0.76	0.77	2.41	2.27
gCanny-mad-ucm <sub>ϕ</sub>	0.35	0.42	0.50	0.74	0.77	2.43	2.29

Πίνακας 8.2: Συνολικά αποτελέσματα αξιολόγησης με βάση τις ανιχνευθείσες περιοχές.

art αλγόριθμο gPb-owt-ucm . Η χρήση της υπερμετρικής ανομοιομορφίας με βάση τον παράγοντα κατεκερματισμού του ορίου γειτονικών περιοχών εμφανίζει χειρότερη απόδοση καθώς αποτελεί μια πιο χονδροειδή προσέγγιση της ισχύος ενός ορίου σε σχέση με τη μέση τιμή της grayscale εικόνας ακμών. Είναι πιθανή ύπαρξης ασυνέχειας στο σύνολο των πηγών της εικόνας ενώ υπάρχει, έστω και τοπικά ασθενές, όριο περιοχής.

Η τεχνική αποδοτικής συνένωσης, μολονότι εμφανίζουν χειρότερη απόδοση από την τεχνική gPb-mad-ucm<sub>g</sub>, έχουν το πλεονέκτημα πως είναι υπολογιστικά πιο αποδοτικές. Με ένα και μόνο πέρασμα από κάθε ακμή του γράφου που προκύπτει από την αρχική κατάτμηση και χωρίς καμία διαδικασία ανανέωσης του γράφου ολοκληρώνεται η διαδικασία συνένωσης γειτονικών περιοχών.

Επιπροσθέτως, η τεχνική συνένωσης γειτονικών περιοχών με βάση των παράγοντα κατακερματισμού σχήματος έχει το πλεονέκτημα πως λαμβάνει υπόψιν πληροφορία από το σύνολο των περιοχών και όχι μόνο από το κοινό τους σύνορο. Από τη μία πλευρά, μπορούν να συνενωθούν επιτυχώς μικρές μη συμπαγείς περιοχές που δεν θα συνενώνονταν από άλλες τεχνικές. Από την άλλη, μπορεί μία ομοιογενής περιοχή να παραμείνει διασπασμένη σε περίπτωση που οι επιμέρους περιοχές έχουν μεγάλο εμβαδό.



# Κεφάλαιο 9

## Συμπεράσματα

### 9.1 Συμπεράσματα

Στα πλαίσια της διπλωματικής αυτής εργασίας, παρουσιάσαμε διάφορες τεχνικές κατάτμησης φυσικών και φωτογραφικών εικόνων με βάση μια πρόσφατη τεχνική αποσύνθεσης διάμεσου άξονα που παράγει μια αρχική κατάτμηση και μια αναπαράσταση της εκάστοτε εικόνας. Ερευνήσαμε πως μπορεί να αξιοποιηθεί η αναπαράσταση αυτή για την επίλυση του προβλήματος της κατάτμησης από μόνη της αλλά και σε συνδυασμό με κριτήρια συνένωσης περιοχών της βιβλιογραφίας. Όλες οι προτεινόμενες τεχνικές αξιολογήθηκαν με βάση τη συλλογή δεδομένων του πανεπιστημίου Berkeley και συγκρίθηκαν με γνωστές τεχνικές της βιβλιογραφίας. Παρόλο που ο μετασχηματισμός διάμεσου άξονα θεωρείται γενικά ασταθής, αποδείχτηκε πως μπορεί να χρησιμοποιηθεί με επιτυχία για το σκοπό της κατάτμησης φυσικών εικόνων που αποτελεί μέχρι σήμερα ένα ανοικτό πρόβλημα και από τις μεγαλύτερες προκλήσεις της όρασης υπολογιστών.

### 9.2 Μελλοντικές κατευθύνσεις

Οι τεχνικές αποσύνθεσης διάμεσου άξονα για το στόχο της κατάτμησης φυσικών εικόνων δεν έχουν πλήρως ερευνηθεί στα στενά πλαίσια μιας διπλωματικής εργασίας. Κάποια παραδείγματα μελλοντικών κατευθύνσεων είναι τα ακόλουθα:

- Ένας τροποποιημένος μετασχηματισμός απόστασης ώστε εκτός από το μέτρο της εικόνας ακμών να περιέχει και τον προσανατολισμό των ακμών.
- Μία τεχνική κατάτμησης που εκμεταλλεύομενη την αναπαράσταση της εικόνας μέσω του διάμεσου άξονα να περιέχει πληροφορία από όλη την εικόνα.
- Συνδυασμός κάποιας από τις προτεινόμενες τεχνικές κατάτμησης εικόνων μαζί με χρήση του διάμεσου άξονα για περιγραφή σχημάτων με στόχο την αναγνώριση αντικειμένων.

- Μία αποδοτικότερη τεχνική υπολογισμού του μετασχηματισμού απόστασης καθώς αποτελεί το πιο χρονοβόρο κομμάτι των τεχνικών μετά την ανίχνευση ακμών.
- Συνδυασμός κάποιων εκ των προτεινόμενων τεχνικών για βελτίωση της απόδοσης.

# Παράρτημα Α'

## Αποδείξεις

**Απόδειξη του λήμματος 4.2.1** (Από [4])

(α)  $\rightarrow$  (β). Με βάση τους ορισμούς (4.13), (4.12) και (4.11), έχουμε πως εάν το  $y$  είναι πηγή τότε:

$$d(x, y) + f(y) \leq d(x, z) + f(z) \quad \forall z \in X \quad (A'.1)$$

για κάποιο  $x \in X$ , ή ισοδύναμα:

$$f(y) \leq d(x, z) - d(x, y) + f(z) \quad \forall z \in X. \quad (A'.2)$$

Χρησιμοποιώντας την τριγωνική ιδιότητα έχουμε πως:

$$d(y, y) + f(y) \leq d(y, z) + f(z) \quad \forall z \in X, \quad (A'.3)$$

η οποία συνεπάγεται πως  $y \succcurlyeq y$ .

(β)  $\rightarrow$  (γ). Εάν  $y \succcurlyeq y$ , τότε από τον ορισμό (4.12),  $\mathcal{D}(f)(y) = f(y)$ .

(γ)  $\rightarrow$  (δ). Εάν  $\mathcal{D}(f)(y) = f(y)$ , τότε από τον ορισμό (4.12) προκύπτει πως  $y \in \hat{S}(y)$  ή  $y \succcurlyeq y$ . Έστω πως κάποιο άλλο σημείο  $z \in S(y)$ , τότε  $z \in \hat{S}(y)$  ή ισοδύναμα  $z \succcurlyeq y \succcurlyeq y$ . Από τον ορισμό (4.13) έχουμε πως  $z \notin S(y)$ , το οποίο αποτελεί αντίφαση. Έτσι  $\hat{S}(y) = \{y\}$  καταλήγοντας στο ότι  $S(y) = \{y\}$ .

□

**Απόδειξη του λήμματος 5.2.1** (Από [4])

Έστω  $y \in S(f)$ . Από το λήμμα 4.2.1 έχουμε πως  $S(y) = \{y\}$  και συνεπώς  $|S(y)| = 1$ . Έτσι, το  $y$  δεν δύναται να ανήκει στο διάμεσο άξονα:  $y \notin A(f)$ . □

**Απόδειξη του λήμματος 5.2.3** (Από [4])

(α) Προφανώς:

$$\ell(u', v') = \ell(u', u) + \ell(u, v) + \ell(v, v'), \quad (A'.4)$$

Από την τριγωνική ανισότητα:

$$\delta(u', v') < \ell(u', u) + \delta(u, v) + \ell(v, v'). \quad (A'.5)$$

Αντικαθιστώντας την ποσότητα  $\ell(u', u) + \ell(v, v')$  από την ποσότητα  $\ell(u', v') - \ell(u, v)$  προκύπτει:

$$\delta(u', v') < \delta(u, v) + \ell(u', v') - \ell(u, v), \quad (A'.6)$$

και από τον ορισμό (;;) καταλήγουμε στο ζητούμενο:

$$\text{res}(x, y) < \text{res}(x', y') \quad (A'.7)$$

□

**Απόδειξη του λήμματος 7.3.1** Η απόδειξη πως η ανομοιομορφία  $\delta_\Phi$  είναι υπερμετρική ανάγεται στο να αποδείξουμε πως η  $\delta_\Phi$  ικανοποιεί την ιδιότητα (7.24). Έστω  $(R_i, R_j)$  το ζεύγος περιοχών που ελαχιστοποιούν την  $\delta_\Phi$  και  $R_k$  οποιαδήποτε περιοχή που γειτνοεύει με την  $R_i \cup R_j$  και ανήκει στην κατάτμηση της εικόνας μετά την συνένωση των  $R_i, R_j$ . Ακόμη, έστω  $\mathcal{A}_{ij} = \sum_{e \in E(R_i) \cap E(R_j)} w(e)$ . Έχουμε:

$$\begin{aligned} \delta_\Phi(R_i, R_j) &\leq \delta_\Phi(R_i, R_k) \Rightarrow \\ \exp(-\Phi(R_i, R_j)) &\leq \exp(-\Phi(R_i, R_k)) \Rightarrow \\ \Phi(R_i, R_j) &\geq \Phi(R_i, R_k) \Rightarrow \\ \frac{\mathcal{A}_{ij}}{L(\partial_{ij})} &\geq \frac{\mathcal{A}_{ik}}{L(\partial_{ik})} \Rightarrow \\ L(\partial_{ik}) \cdot \mathcal{A}_{ij} &\geq L(\partial_{ij}) \cdot \mathcal{A}_{ik} \end{aligned} \quad (A'.8)$$

Ομοίως,

$$L(\partial_{jk}) \cdot \mathcal{A}_{ij} \geq L(\partial_{ij}) \cdot \mathcal{A}_{jk} \quad (A'.9)$$

Συνδυάζοντας τις εξισώσεις (A'.8) και (A'.9) έχουμε:

$$\begin{aligned} L(\partial_{jk}) \cdot \mathcal{A}_{ij} + L(\partial_{ik}) \cdot \mathcal{A}_{ij} &\geq L(\partial_{ij}) \cdot \mathcal{A}_{jk} + L(\partial_{ij}) \cdot \mathcal{A}_{ik} \Rightarrow \\ (L(\partial_{jk}) + L(\partial_{ik})) \cdot \mathcal{A}_{ij} &\geq L(\partial_{ij}) \cdot (\mathcal{A}_{jk} + \mathcal{A}_{ik}) \Rightarrow \\ \frac{\mathcal{A}_{ij}}{L(\partial_{ij})} &\geq \frac{\mathcal{A}_{jk} + \mathcal{A}_{ik}}{L(\partial_{jk}) + L(\partial_{ik})} \Rightarrow \\ \frac{\sum_{e \in E(R_i) \cap E(R_j)} w(e)}{L(\partial_{ij})} &\geq \frac{\sum_{e \in E(R_j) \cap E(R_k)} w(e) + \sum_{e \in E(R_i) \cap E(R_k)} w(e)}{L(\partial_{jk}) + L(\partial_{ik})} \Rightarrow \\ \frac{\sum_{e \in E(R_i) \cap E(R_j)} w(e)}{L(\partial_{ij})} &\geq \frac{\sum_{e \in E(R_i \cup R_j) \cap E(R_k)} w(e)}{L(\partial_{i \cup j, k})} \Rightarrow \\ \Phi(R_i, R_j) &\geq \Phi(R_i \cup R_j, R_k) \Rightarrow \\ \exp(-\Phi(R_i, R_j)) &\leq \exp(-\Phi(R_i \cup R_j, R_k)) \Rightarrow \\ \delta_\Phi(R_i, R_j) &\leq \delta_\Phi(R_i \cup R_j, R_k) \end{aligned} \quad (A'.10)$$

□

**Απόδειξη του λήμματος 7.3.2** Απόδειξη ανάλογη με την απόδειξη του λήμματος 7.3.1. Έστω  $(R_i, R_j)$  το ζεύγος περιοχών που ελαχιστοποιούν την  $\delta_\Phi$  και  $R_k$  οποιαδήποτε περιοχή που γειτονεύει με την  $R_i \cup R_j$  και ανήκει στην κατάτμηση της εικόνας μετά την συνένωση των  $R_i, R_j$ . Έχουμε:

$$\begin{aligned} \delta_g(R_i, R_j) &\leq \delta_g(R_i, R_k) \Rightarrow \\ \frac{\sum_g(\partial_{ij})}{L(\partial_{ij})} &\leq \frac{\sum_g(\partial_{ik})}{L(\partial_{ik})} \Rightarrow \\ L(\partial_{ik}) \cdot \sum_g(\partial_{ij}) &= L(\partial_{ij}) \cdot \sum_g(\partial_{ik}) \end{aligned} \quad (A'.11)$$

$$(A'.12)$$

Ομοίως,

$$L(\partial_{jk}) \cdot \sum_g(\partial_{ij}) \leq L(\partial_{ij}) \cdot \sum_g(\partial_{jk}) \quad (A'.13)$$

Συνδυάζοντας τις εξισώσεις (A'.11) και (A'.13) έχουμε:

$$\begin{aligned} (L(\partial_{ik}) + L(\partial_{jk})) \cdot \sum_g(\partial_{ij}) &\leq L(\partial_{ij}) \cdot (\sum_g(\partial_{ik}) + \sum_g(\partial_{jk})) \Rightarrow \\ \frac{\sum_g(\partial_{ij})}{L(\partial_{ij})} &\leq \frac{(\sum_g(\partial_{ik}) + \sum_g(\partial_{jk}))}{L(\partial_{ik}) + L(\partial_{jk})} \Rightarrow \\ \delta_g(R_i, R_j) &\leq \delta_g(R_i \cup R_j, R_k) \end{aligned} \quad (A'.14)$$

□





# Βιβλιογραφία

- [1] P. Arbelaez. Boundary extraction in natural images using ultrametric contour maps. In *Proceedings of POCV*, 2006.
- [2] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. From contours to regions: An empirical evaluation. In *Proceedings of Computer Vision and Pattern Recognition*, pages 2294–2301, 2009.
- [3] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 898–916, 2011.
- [4] Y. Avrithis and K. Rapantzikos. The medial feature detector: Stable regions from image boundaries (long version). Unpublished, 2010.
- [5] Y. Avrithis and K. Rapantzikos. The medial feature detector: Stable regions from image boundaries. In *Proceedings of International Conference on Computer Vision*, 2011.
- [6] S. Belongie, C. Carson, H. Greenspan, and J. Malik. Color- and texture-based image segmentation using em and its application to content-based image retrieval. In *Proceedings of IEEE International Conference on Computer Vision*, pages 675–692, 1998.
- [7] S. Beucher and F. Meyer. *The Morphological Approach to Segmentation : The Watershed Transformation*. In *Mathematical Morphology in Image Processing*, chapter 12, pages 433–481. Ed. E. R. Dougherty, 1993.
- [8] H. Blum. A transformation for extracting new descriptors of shape. In *Models for the Perception of Speech and Visual Form*. MIT Press, Cambridge, 1967.
- [9] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 679–698, 1986.
- [10] Fan R K Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.
- [11] Dorin Comaniciu and Peter Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

- [12] T. Cour, F. Bénézit, and J. Shi. Spectral segmentation with multiscale graph decomposition. In *Proceedings of Computer Vision and Pattern Recognition*, pages 1124–1131, 2005.
- [13] P.F. Felzenszwalb and D.P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
- [14] V. Ferrari, T. Tuytelaars, and L. Van Gool. Simultaneous object recognition and segmentation by image exploration. In *Proceedings of the European Conference on Computer Vision*, 2004.
- [15] M. Fussenegger, A. Opelt, A. Pinz, and P. Auer. Object recognition using segmentation for feature detection. In *In ICPR*, pages 41–44, 2004.
- [16] R. C. Gonzalez and R. E. Woods. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2006.
- [17] S. Kim. An  $O(N)$  Level Set Method for Eikonal Equations. *SIAM journal on scientific computing*, 22(6):2178–2193, 2001.
- [18] I. Kovacs and B. Julesz. A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences, USA*, 90:7495–7497, 1993.
- [19] V. Lamme. The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, 15:1605–1615, 1995.
- [20] T. S. Lee, D. Mumford, R. Romero, and V. Lamme. The role of the primary visual cortex in higher level vision. *Vision Research*, 38:2429–2454, 1998.
- [21] B. Leibe, A. Leonardis, and B. Schiele. Robust object detection with interleaved categorization and segmentation. *International Journal of Computer Vision*, 77:259–289, 2008.
- [22] Hwei-Jen Lin, Shu-Yi Wang, Shwu-Huey Yen, and Yang-Ta Kao. Face detection based on skin color segmentation and neural network. In *Proceedings of Neural Networks and Brain, 2005.*, volume 2, pages 1144 –1149, 2005.
- [23] L. Linardakis and N. Chrisochoides. A static medial axis domain decomposition for 2d geometries. *ACM Transactions on Mathematical Software*, 34(1):1–19, 2005.
- [24] P. Maragos and M.A. Butt. Curve evolution, differential morphology, and distance. *Fundamenta Informaticae*, 41:91–129, 2000.
- [25] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt and Co., Inc. New York, NY, USA, 1982.

- [26] D. R. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *ICCV*, pages 416–425, 2001.
- [27] David R. Martin, Charless Fowlkes, and Jitendra Malik. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 530–549, 2004.
- [28] B. M. Mehtre and B. Chatterjee. Segmentation of fingerprint images - a composite method. *Pattern Recognition*, 22(4):381–385, 1989.
- [29] H. Mobahi, S. Rao, A. Yang, S. Sastry, and Y. Ma. Segmentation of natural images by texture and boundary compression. *International Journal of Computer Vision*, pages 1–13, 2011.
- [30] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42:577–685, 1989.
- [31] R.L. Ogniewicz and O. Kübler. Hierarchic voronoi skeletons. *Pattern Recognition*, 28(3):343–359, 1995.
- [32] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9:62–66, 1979.
- [33] D. L. Pham, C. Xu, and J. L. Prince. A survey of current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315–338, 2000.
- [34] A. Rosenfeld and J. L. Pfaltz. Distance functions on digital pictures. *Pattern Recognition*, 1:33–61, 1968.
- [35] J.A. Sethian. Fast marching methods. *SIAM journal on scientific computing*, 41(2):199–235, 1999.
- [36] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:888–905, 1997.
- [37] George Stockman and Linda G. Shapiro. *Computer Vision*. Prentice Hall, 1st edition, 2001.
- [38] L. Vincent. Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Transactions on Image Processing*, 2: 176–201, 1993.
- [39] Chee S. W. A block-based map segmentation for image compressions. *IEEE Transactions on Circuits Syst. Video Techn.*, 8:592–601, 1998.

- [40] L. Wang, J. Shi, G. Song, and I. Shen. Object detection combining recognition and segmentation. In *Proceedings of the Asian conference on Computer vision*, pages 189–199, 2007.
- [41] Z. Wu and R. Leahy. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 15:1101–1113, 1993.
- [42] L. Yatziv, A. Bartesaghi, and G. Sapiro.  $O(N)$  Implementation of the fast marching algorithm. *Journal of Computational Physics*, 212(2):393–399, 2006.
- [43] Y. Yi, D. Qu, and F. Xu. Face detection method based on skin color segmentation and eyes verification. In *Proceedings of Artificial Intelligence and Computational Intelligence*, volume 3, pages 495–501, 2009.