



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ
ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μελέτη και Εφαρμογή Νέων Τεχνικών Κατάτμησης
Αρχαιομετρικών Εικόνων.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΡΟΥΣΣΗΣ Β. ΑΛΕΞΑΝΔΡΟΣ

Επιβλέπων: Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής Εθνικού Μετσόβιου Πολυτεχνείου Αθήνας



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ

ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΕΠΙΚΟΙΝΩΝΙΩΝ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ
ΣΥΣΤΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ

Μελέτη και Εφαρμογή Νέων Τεχνικών Κατάτμησης
Αρχαιομετρικών Εικόνων.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΡΟΥΣΣΗΣ Β. ΑΛΕΞΑΝΔΡΟΣ

Επιβλέπων: Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής Εθνικού Μετσόβιου Πολυτεχνείου Αθήνας

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή .
Αθήνα, Ιούλιος 2016

Κωνσταντίνος Παπαοδυσσεύς
Καθηγητής Ε.Μ.Π.

Βασίλειος Λούμος
Καθηγητής Ε.Μ.Π.

Ηλίας Κουκούτσης
Επικ.Καθηγητής Ε.Μ.Π

ΡΟΥΣΣΗΣ ΑΛΕΞΑΝΔΡΟΣ

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

**Copyright Ρούσσης Αλέξανδρος 2016 Copyright © Κωνσταντίνος
Παπαδουσσέυς 2016**

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 εκτύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό της 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.

Περίληψη

Σε αυτή τη διπλωματική γίνεται μια οικεία ενδελεχής μελέτη, ανάλυση και υλοποίηση μερικών σημαντικών αλγορίθμων ενός συνόλου σημαντικών αλγορίθμων κατάτμησης εικόνας .Οι αλγόριθμοι αυτοί εφαρμόζονται κυρίως σε εικόνες γραμμάτων από παπύρους αλλά και σε εικόνες βιοιατρικού περιεχομένου. Ειδικότερα αναλύονται και εφαρμόζονται οι αλγόριθμοι Watershed και Active Contours. Για να ελεγχθεί η απόδοση αυτών των αλγορίθμων , προηγείται καταρχήν η δημιουργία μιας στοιχειώδους βάσεως δεδομένων , η οποία περιλαμβάνει εικόνες κατάτμησης από τον χρήστη. Οι εικόνες αυτές ισοδυνάμως αποτελούν το πρότυπο με το οποίο θα συγκριθούν τα εκάστοτε αποτελέσματα της εφαρμοσθείσας μεθόδου αυτόματης κατάτμησης . Επίσης στη παρούσα διπλωματική γίνεται βιογραφική μελέτη αναφορικά με τις δυνατότητες εφαρμογής αυτών των αλγορίθμων κατάτμησης εικόνας στην αρχαιολογία και βιοιατρική .

Λέξεις Κλειδιά

Αλγόριθμος Watershed , Active Contours ,κατάτμηση, ομαδοποίηση, Βιοιατρική , πάπυρος

Abstract

In this thesis we make a familiar and in depth study, analysis and implementation of some very important segmentation algorithms. These algorithms may be applied in pictures from papyrus with illustrated letters but also in pictures with biomedical content. Particularly, Watershed and Active Contours algorithms are being analyzed and materialized. The precedent step of checking the algorithms' efficiency is to create a database which includes segmentation pictures from the user. These pictures constitute the pattern compared with the results from the automatic segmentation algorithms. Furthermore, in this thesis we make a biographical study pertaining to the segmentation algorithms application in archeology and biomedicine.

Key Words

Watershed Algorithm, Active Contours, segmentation, clustering, biomedicine, papyrus

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά όλους όσους συνέβαλαν στην επιτυχή ολοκλήρωση της παρούσας διπλωματικής και ιδιαίτερα:

Τον κ. Κωνσταντίνο Παπαοδυσσέα, καθηγητή του Τομέα Επικοινωνιών Ηλεκτρονικής και Συστημάτων Πληροφορικής του Εθνικού Μετσόβιου Πολυτεχνείου, για την ανάθεση της διπλωματικής εργασίας, το άριστο κλίμα συνεργασίας και το συνεχές ενδιαφέρον του σε όλη τη διάρκεια της συγγραφής της.

ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο 1

Εισαγωγή και σκοπός της παρούσας διπλωματικής

Κεφάλαιο 2

- 2.1 Η σπουδαιότητα της αυτόματης αναγνώρισης γραφέα κειμένων
- 2.2 Η ειδική περίπτωση της αυτόματης ταυτοποίησης γραφών αρχαίων κειμένων
 - 2.2.1 Η σπουδαιότητα
 - 2.2.2 Οι δυσκολίες

Κεφάλαιο 3

- 3.1 Μέθοδοι Clustering
 - 3.1.1 Τεχνικές Διαμέρισης (Partional Techniques)
 - 3.1.2 Ιεραρχικές Τεχνικές (Hierarchical Techniques)
 - 3.1.2.1 Γενική περιγραφή
 - 3.1.2.2 Μέθοδοι
 - 3.1.2.3 Δυσκολίες ιεραρχικών τεχνικών

Κεφάλαιο 4

- 4.1 Ομαδοποίηση Δεδομένων Με Δυναμικό Μοντέλο (Chameleon Data Clustering Algorithm)
 - 4.1.1 Γενική περιγραφή μεθόδου
 - 4.1.2 Λειτουργία του Chameleon

Κεφάλαιο 5

Τοπογραφική απόσταση και γραμμή watershed

- 5.1 Εισαγωγή
- 5.2 Ορισμός του watershed στον συνεχή χώρο.
 - 5.2.1 Η τοπογραφική μεταβολή
 - 5.2.2 Εφαρμογή σε δυαδικές εικόνες
 - 5.2.3 Ερμηνεία της συνάρτησης τοπογραφικής απόστασης
- 5.3 Τοπογραφική απόσταση και λεκάνες απορροής στον διακριτό χώρο.
 - 5.3.1 Η τοπογραφική συνάρτηση απόστασης
 - 5.3.2 Ορισμός λεκάνης απορροής

- 5.4 Η γραμμή watershed και οι αλγόριθμοι ελαχίστης διαδρομής
 - 5.4.1 Γεμίζοντας από τα τοπικά ελάχιστα, ή από τους markers
 - 5.4.2 Τροποποίηση της ομοιοπίας της εικόνας κλήσης

Κεφάλαιο 6

- 6.1 ACTIVE CONTOURS
 - 6.1.1 Parametric Active Contours
 - 6.1.1.1 Classic Active Contours (Snakes)
 - 6.1.1.2 Balloons 1989
 - 6.1.1.3 Gradient Vector Flow 1998
 - 6.1.2 Geometric Active Contours
 - 6.1.2.1 Geodesic Active Contours
 - 6.2.2.2 Βασικό Πρόβλημα
 - 6.2.2.3 Geodesic Active Contours

Κεφάλαιο 7

Ο αλγόριθμος κατάτμησης εικόνας που χρησιμοποιήθηκε

- 7.1 Ο αλγόριθμος που χρησιμοποιήσαμε
- 7.2 Εφαρμογή της μεθόδου

Κεφάλαιο 8

Συμπεράσματα

Κεφάλαιο 1

Εισαγωγή και σκοπός της παρούσας διπλωματικής

Στόχος της παρούσας διπλωματικής εργασίας είναι να εξετάσει τις εφαρμογές της πληροφορικής και των υπολογιστικών συστημάτων στον πολιτισμό και την αρχαιολογία. Συγκεκριμένα, στο παρόν έργο χρησιμοποιούνται πρωτότυπες τεχνικές επεξεργασίας εικόνας και προηγμένες τεχνικές μαθηματικών για την αντιμετώπιση σημαντικών αρχαιολογικών προβλημάτων, αλλά και τη μελέτη σχετικών ευρημάτων. Τα συστήματα και οι τεχνικές αυτές αποτελούνται, κυρίως, από αλγόριθμους ψηφιακής επεξεργασίας εικόνας, καθώς επίσης και από εφαρμογή γεωμετρικών μεθόδων και στατιστικής επεξεργασίας δεδομένων.

Ειδικότερα, ασχολούμαστε με τη μελέτη αρχαίων επιγραφών και την ειδική επεξεργασία κάθε γράμματος αυτών ξεχωριστά. Δοκιμάζουμε μία ανεπτυγμένη μέθοδο αυτόματης κατάτμησης εικόνας στη φωτογραφία κάθε γράμματος κάποιων επιγραφών με σκοπό την εξαγωγή των γραμμάτων από το υπόβαθρο και των περιγραμμάτων τους αντίστοιχα. Η κατάτμηση εικόνων είναι ένα από τα πιο σημαντικά και πιο δύσκολα προβλήματα στον τομέα της εφαρμογής υπολογιστικών συστημάτων σε εικόνες. Σε γενικές γραμμές, είναι η διαδικασία της απομόνωσης αντικειμένων στην εικόνα από το παρασκήνιο (υπόβαθρο), δηλαδή, ο κατακερματισμός της εικόνας σε ασυνεχείς περιοχές, κάθε μία από τις οποίες είναι διασυνδεδεμένη και ομογενής με βάση κάποιο κοινό χαρακτηριστικό-σημείο αναφοράς, όπως η τιμή του γκρι, η υφή, το χρώμα, η κίνηση, το βάθος. Καθώς ο όρος «ομογενής» είναι αρκετά ασαφής και επιτρέπει πολλές ερμηνείες της σημασίας του, ο σκοπός της κατάτμησης εικόνας φαίνεται ότι εξαρτάται τόσο από το κοινό χαρακτηριστικό ως προς το οποίο θέλουμε να εφαρμόσουμε τον κατακερματισμό αλλά και από την ίδια την εφαρμογή. Γενικά, θεωρείται ως ένα από τα πιο δύσκολα προβλήματα της όρασης υπολογιστών. Ακόμα και ένας άνθρωπος μπορεί να αντιμετωπίσει κάποιες σχετικές δυσκολίες στην κατάτμηση μίας εικόνας, και επίσης, μπορεί να προκύψει μία ποικιλία διαφορετικών κατατμήσεων από διαφορετικούς παρατηρητές.

Αξίζει να αναφέρουμε την αξία αυτής της εφαρμογής στις αρχαίες επιγραφές, καθώς ένας από τους βασικότερους τομείς της αρχαιολογίας είναι και η επιγραφολογία, δηλαδή ο κλάδος που ασχολείται ακριβώς με τις αρχαίες επιγραφές. Ο συγκεκριμένος τομέας έχει μεγάλο πεδίο εφαρμογής σε παγκόσμιο επίπεδο και είναι καθοριστικής σημασίας για την ορθή καταγραφή της ιστορίας. Αυτό ισχύει φυσικά σε παγκόσμιο επίπεδο και όχι μόνο για τις ελληνικές επιγραφές. Οι πληροφορίες που εξάγονται από τη μελέτη αυτών των επιγραφών είναι πολύτιμες, όχι μόνο για την αρχαιολογία καθ' αυτή, αλλά και για την ιστορία, καθώς και για άλλες ανθρωπιστικές επιστήμες που έχουν πεδίο έρευνας στην αρχαία εποχή.

Από την άλλη πλευρά, η ψηφιακή επεξεργασία εικόνας συνίσταται στη χρήση αλγορίθμων υπολογιστών για την πραγματοποίηση ψηφιακής εικόνας. Ως υποκατηγορία ή τομέας της ψηφιακής επεξεργασίας σήματος, η ψηφιακή επεξεργασία εικόνας έχει πολλά πλεονεκτήματα σε σχέση με την αναλογική επεξεργασία. Κατ' αρχάς, επιτρέπει ένα πολύ μεγαλύτερο φάσμα αλγορίθμων, οι οποίοι εφαρμόζονται στα δεδομένα εισόδου και βοηθούν στην αποφυγή προβλημάτων όπως η συσσώρευση θορύβου και η παραμόρφωση του σήματος κατά την επεξεργασία, και ως εκ τούτου, μπορούν να προσφέρουν πολύ πιο εξελιγμένες επιδόσεις σε απλές εργασίες, αλλά και εφαρμογές μεθόδων που θα ήταν αδύνατες με τον αναλογικό τρόπο. Ειδικότερα, η ψηφιακή επεξεργασία εικόνας βρίσκει πρακτικά εφαρμογή στην :

- Κατηγοριοποίηση
- Εξαγωγή χαρακτηριστικών
- Αναγνώριση προτύπων
- Προβολή – Σχεδίαση
- Multi-level ανάλυση σήματος

χρησιμοποιώντας τεχνικές όπως η ψηφιοποίηση, η ανάλυση κύριων συνιστωσών (principal components analysis), νευρωνικά δίκτυα και αρκετούς μετασχηματισμούς (Fourier, Laplace, κ.ά.).

Κατά συνέπεια, διαπιστώνουμε ότι η εισαγωγή και η χρήση των υπολογιστικών συστημάτων σε εργασίες και μελέτες που μέχρι πρότινος γινόταν, όπως λέγεται, «με το χέρι» και βάση του ενστίκτου και της εν γένη εμπειρίας των επιγραφολόγων, έχει πολύ μεγάλη αξία και παρέχει ιδιαίτερα μεγάλη βοήθεια στον τομέα αυτόν την έρευνας της αρχαιολογίας. Οι τρόποι ανάλυσης χωρίς τη χρήση πληροφοριακών συστημάτων, προσφέρουν αποτελέσματα, όμως κανείς εξ' αυτών δεν δίνει με συνέπεια αντικειμενικά και αδιαμφισβήτητα συμπεράσματα. Με άλλα λόγια, η προσπάθειά μας έγκειται στην εισαγωγή και εφαρμογή της ανθρώπινης εμπειρίας και κριτικής σκέψης σε ένα υπολογιστικό-πληροφοριακό σύστημα προς διευκόλυνση της εργασίας του ίδιου του ανθρώπου αλλά και εξαγωγή ταχύτερων και αντικειμενικότερων συμπερασμάτων.

Πιο συγκεκριμένα, μέσω της εξαγωγής των γραμμάτων και των περιγραμμάτων αυτών, πέραν του ότι μπορεί να αναγνωριστεί εύκολα το γράμμα της επιγραφής το οποίο μελετάται, σε πρώτο επίπεδο, τα αποτελέσματα μπορούν να χρησιμοποιηθούν και για την εξαγωγή ποικίλων άλλων συμπερασμάτων και περαιτέρω μελέτης. Για παράδειγμα, με εφαρμογή

μαθηματικοποιημένων κριτηρίων στο υπολογιστικό σύστημα, είναι δυνατή η επεξεργασία των γραμμάτων των επιγραφών ώστε να προκύψουν συμπεράσματα για την χρονολογική περίοδο στην οποία ανήκει η επιγραφή, αλλά και τον χαρακτήρα της. Τα μαθηματικοποιημένα αυτά κριτήρια, σε αντίθεση με όσα εφαρμόζονται μέχρι σήμερα, δίνουν ποσοτικά αποτελέσματα, τα οποία με τη σειρά τους επιτρέπουν το διαχωρισμό και την κατάταξη των επιγραφών, ως και την ταυτοποίηση του χαρακτήρα.

Κατόπιν όλων αυτών των αναφορών, διαπιστώνουμε ότι η κατάκτηση των εικόνων των φωτογραφιών των γραμμάτων επιγραφών με την οποία θα ασχοληθούμε στην παρούσα εργασία αποτελεί το πρώτο μόνο βήμα για τη συμβολή σε πολλές άλλες εφαρμογές και εξαγωγή πολλών αξιολογικών και ενδιαφέροντων αντικειμενικών συμπερασμάτων. Φυσικά, η εφαρμογή της μεθόδου μας είναι μία επίπονη και χρονοβόρα διαδικασία

Αν μία τέτοια διαδικασία γινόταν χωρίς τη χρήση υπολογιστικών πληροφοριακών συστημάτων, τότε, σε πολλές περιπτώσεις μπορεί να είναι ίσως και αδύνατη, καθώς υπόκειται στον ανθρώπινο παράγοντα, αλλά και λόγω της παρουσίας ποικίλων δυσκολιών, όπως η φθορά.

Αντίθετα, το δικό μας σύστημα έχει ως στόχο να μας δώσει όσο το δυνατόν καλύτερες και αντιπροσωπευτικές εικόνες των γραμμάτων που εμφανίζονται στις ψηφιακές φωτογραφίες των επιγραφών, καθώς έχει προσαρμοστεί όσο αυτό είναι δυνατό, στις ειδικές απαιτήσεις της δικής μας εφαρμογής.

Κεφάλαιο 2

2.1 Η σπουδαιότητα της αυτόματης αναγνώρισης γραφέα κειμένων

Η αναγνώριση χειρόγραφου μελετάται εδώ και αρκετές δεκαετίες από την ακαδημαϊκή κοινότητα με ιδιαίτερη προσήλωση και ζωηρότητα. Η βιομετρική αναγνώριση χειρόγραφου, ή αλλιώς η αναγνώριση γραφέα, δεν πρέπει να συγχέεται με την οπτική αναγνώριση χαρακτήρων. Στην πρώτη περίπτωση ο σκοπός είναι να γίνει αναγνώριση και ταυτοποίηση του γραφέα ενός δοσμένου χειρόγραφου κειμένου, ανεξάρτητα από το περιεχόμενο, ενώ στην δεύτερη ο σκοπός είναι η αναγνώριση του κειμένου και ουσιαστική ή ψηφιοποίηση των χαρακτήρων, ανεξάρτητα του γραφέα. Κάθε γραφέας μπορεί να χαρακτηριστεί από το δικό του τρόπο γραφής, από την αναπαραγωγή των λεπτομερών και του τρόπου που γράφει ασυναίσθητα. Για το λόγο αυτό, σε ορισμένες περιπτώσεις εμπειρογνωμοσύνης και ταυτοποίησης, τα δείγματα γραφής έχουν την ίδια αξία με εκείνη των δακτυλικών αποτυπωμάτων. Το πρόβλημα της αναγνώρισης γραφέα τίθεται συχνά στο Δικαστήριο των Ευρωπαϊκών Κοινοτήτων, όπου πρέπει κανείς να καταλήξει σε ένα συμπέρασμα σχετικά με τη γνησιότητα ενός εγγράφου (π.χ. διαθήκη). Προκύπτει, επίσης, σε τράπεζες για την επαλήθευση της υπογραφής, ή σε κάποια ιδρύματα που αναλύουν κείμενα παλαιότερων συγγραφέων, και ενδιαφέρονται για τη φύση των κειμένων αυτών, όπως για παράδειγμα τον προσδιορισμό των διαφόρων συγγραφέων που συμμετείχαν στην εκπόνηση ενός χειρόγραφου ή ποιος έκανε διορθώσεις πάνω στο κείμενο.

Η Αυτόματη Αναγνώριση Γραφέα μπορεί να χωριστεί σε δύο βασικές κατηγορίες. Την στατική και την δυναμική. Στην στατική, γνωστή και ως “off-line” το χειρόγραφο κείμενο αρχικά ψηφιοποιείται μέσω ενός scanner ή κάμερας και το σύστημα που θέτουμε σε εφαρμογή αναγνωρίζει το κείμενο και ταυτοποιεί τον γραφέα μέσω των μετρήσεων που παίρνει από τους χαρακτήρες και των συγκρίσεων που κάνει. Στην δυναμική κατηγορία, γνωστή και ως “on-line”, γίνεται χρήση κάποιου ηλεκτρονικού tablet ή γραφίδας με λειτουργία PDA και το κείμενο συλλέγεται και επεξεργάζεται σε πραγματικό χρόνο.

Η Αυτόματη Αναγνώριση Γραφέα μέσω ενός ηλεκτρονικού υπολογιστικού συστήματος δίνει πολλές λύσεις στις μέρες μας και σε καμία περίπτωση δεν έρχεται να αντικαταστήσει τη δουλειά ενός γραφολόγου, αλλά αντιθέτως είναι ένα χρήσιμο εργαλείο στα χέρια του που ελαχιστοποιεί τόσο την απόσταση όσο και το χρόνο. Τα παραδείγματα που δόθηκαν παραπάνω, η αναγνώριση μιας διαθήκης ή μιας υπογραφής, μπορούν να λυθούν πιο γρήγορα ή ακόμα να επιβεβαιωθούν με την βοήθεια ενός ηλεκτρονικού υπολογιστικού συστήματος. Επιπλέον κάποιες τραπεζικές ή ταχυδρομικές συναλλαγές που απαιτούν το γνήσιο της υπογραφής, μπορούν να γίνουν ευκολότερα με ένα δυναμικό σύστημα που αναφέραμε παραπάνω.

Τέλος υπάρχει τεράστιος όγκος τόσο αρχαίων, όσο και σύγχρονων κειμένων που δεν έχουν προλάβει να μελετήσουν και να ταυτοποιήσουν οι γραφολόγοι. Ένα

αξιόπιστο σύστημα αυτόματης αναγνώρισης γραφέα θα ήταν ένα σημαντικό εργαλείο τόσο στην ταυτοποίηση του συγγραφέα και της γνησιότητας του κειμένου, αλλά παράλληλα και στην χρονολόγηση του χειρόγραφου.

2.2

Η σπουδαιότητα της ταυτοποίησης γραφέα αρχαίων κειμένων και χειρόγραφων

2.2.1 Η σπουδαιότητα

Η έρευνα γύρω από την ταυτοποίηση γραφέα αρχαίων κειμένων αποτελεί ένα πολύ σημαντικό κεφάλαιο στην επιστήμη της ιστορίας και της αρχαιολογίας. Ιστορικές πηγές όπως διατηρημένα γραπτά έγγραφα σε λίθινες επιγραφές ή χειρόγραφα, προσφέρουν στους ιστορικούς ένα από τα πιο σημαντικά εργαλεία της εργασίας τους. Η μελέτη όμως μιας πηγής δεν μπορεί να γίνει σε κενό χρόνο. Είναι απολύτως απαραίτητο για τον ιστορικό επιστήμονα να γνωρίζει την ημερομηνία δημιουργίας της πηγής που μελετά καθώς αν αυτή δεν ερμηνευτεί μέσα στο ιστορικό πλαίσιο που γράφτηκε μπορεί είτε να οδηγήσει σε λάθος είτε σε ελλιπείς πληροφορίες. Όπως άλλωστε έγραψε πρόσφατα και ο μεγάλος ιστορικός Christian Habicht “σωστή χρήση των επιγραφών μπορεί να γίνει μόνο αν αυτές μπορούν να προσδιοριστούν χρονικά”. Ωστόσο, οι γραφείς δεν υπέγραφαν ποτέ τα κείμενα τους και δεν έβαζαν ημερομηνία.

Επίσης, το μεσαίωνα και μέχρι την εφεύρεση της τυπογραφίας, η λογοτεχνία του αρχαίου κόσμου διατηρούνταν και μεταφέρονταν μέσω παπύρων και περγαμηνών. Η πρακτική αυτή παράγει παρόμοια προβλήματα με τις αρχαίες επιγραφές. Παραδείγματος χάριν, το έπος Ιλιάδα του Ομήρου διασώθηκε μέσα στους αιώνες σε μορφή χειρόγραφων τόμων. Αυτοί οι τόμοι, γράφτηκαν στην Κωνσταντινούπολη αλλά σε διάφορες χρονικές περιόδους χρειάστηκε να αλλάξουν τοποθεσία και να ταξιδέψουν σε βιβλιοθήκες της Ευρώπης. Κατά τη διάρκεια όλων αυτών των χρόνων, στο έργο προστέθηκαν διάφορες σημειώσεις και σχόλια στα περιθώρια του κειμένου καθώς και μεταξύ των γραμμών.

Η σημαντική αυτή έρευνα, που επιχειρούμε να μελετήσουμε και να αναπτύξουμε εδώ, δηλαδή του χρονικού προσδιορισμού αρχαίων επιγραφών μέσω της ταυτοποίησης γραφέα μέχρι και μερικά χρόνια πριν θα ήταν εκ των πραγμάτων ανέφικτη. Αυτό γιατί τα απαιτούμενα τεχνολογικά μέσα δεν είχαν αναπτυχθεί. Εργαλεία όπως η ψηφιακή φωτογραφία, φτηνές και μεγάλης χωρητικότητας αποθηκευτικές συσκευές καθώς επίσης δίκτυα υψηλής ταχύτητας είναι επιτεύγματα που αναπτύχθηκαν πρόσφατα και χωρίς αυτά η ψηφιακή σύγκριση κειμένων ήταν απαγορευτική. Αντίθετα, στις μέρες μας η έρευνα αυτή, εκμεταλλευόμενη τα επιτεύγματα αυτά μπορεί να αναπτυχθεί και να παράξει πολύ σημαντικά

αποτελέσματα όσον αφορά την ποσοτική και στατιστική ανάλυση σχετικά με τους διαφορετικούς γραφείς αρχαίων επιγραφών και τη φύση των κειμένων-πηγών.

Η μέθοδος που αναπτύσσεται εδώ, επιτυγχάνει έμμεσα το χρονικό προσδιορισμό των αρχαίων επιγραφών, εκμεταλλευόμενη την επαγωγική ιδιότητα. Συγκεκριμένα, αναφορικά με τις αρχαίες επιγραφές μπορούμε να βασιστούμε σε μια πολύ σημαντική ιδιότητα της περιόδου για τη μελέτη μας. Συγκεκριμένα οι γραφείς που σμίλευαν τις πλάκες στην αρχαιότητα ήταν επάγγελμα. Επομένως ο αριθμός των διαφορετικών “χεριών” είναι ιδιαίτερα περιορισμένος, τόσο ώστε να μπορούμε να υποθέσουμε ότι οι δεκάδες χιλιάδες πλάκες αποδίδονται σε μερικές δεκάδες “χέρια”. Αν λοιπόν είμαστε βέβαιοι για τη χρονολογία μιας πλάκας και μπορέσουμε να αποδώσουμε με σιγουριά μια άλλη πλάκα στον ίδιο χαρακτή, τότε μπορούμε να προσδιορίσουμε χρονικά κατά προσέγγιση τη δεύτερη πλάκα σύμφωνα με τη χρονική περίοδο επαγγελματικής δράσης του χαρακτή. Αξίζει να σημειωθεί εδώ, ότι η επαγγελματική περίοδος ενός χαρακτή κάλυπτε περίπου είκοσι με είκοσι πέντε χρόνια. Όπως γίνεται εύκολα αντιληπτό, η πιθανότητα να επιτύχει η μέθοδος αυξάνει όσο αυξάνει ο αριθμός των διαφορετικών επιγραφών που αποδίδονται σε αυτό το χέρι. Αυτό συμβαίνει γιατί αυξάνει και η πιθανότητα κάποια από αυτές τις επιγραφές να μπορέσει να προσδιοριστεί χρονικά.

Τα παραπάνω, ισχύουν και στην περίπτωση χειρόγραφων. Όπως αναφέρθηκε ανωτέρω, το έργο του Ομήρου ταξίδεψε στην Ευρώπη και στους αιώνες και υπέστη πολλές προσθήκες σχολίων και σημειώσεων στο περιθώριο των πολυτελών αρχικών εκδόσεων. Συνεπώς, αν μπορέσουμε να προσδιορίσουμε με αυτόματα μέσα τα διαφορετικά “χέρια” που έγραψαν κείμενο, μπορούμε να αποκτήσουμε μια χρονολογική ταξινόμηση καθώς και επιπρόσθετες πολύ σημαντικές πληροφορίες. Σύμφωνα με τα παραπάνω η αξία της ανάπτυξης ενός πακέτου μεθόδων και ενός σχετικού συστήματος πληροφορίας που πραγματοποιεί σωστή ταυτοποίηση γραφέα είναι απόλυτα απαραίτητη. Την αναγκαιότητα αυτή φιλοδοξεί να καλύψει η παρούσα μελέτη χρησιμοποιώντας μεθόδους από τα πεδία των Μαθηματικών, Αναγνώρισης προτύπου, Επεξεργασίας Εικόνας καθώς και Επιστήμης Υπολογιστών εν γένει.

2.2.2 Οι δυσκολίες

Όπως σε κάθε καινοτόμα μεθοδολογία, έτσι και εδώ οι δυσκολίες δεν έλειψαν. Ο βαθμός δυσκολίας μάλιστα στη δικιά μας περίπτωση ήταν αυξημένες καθώς, η υποφαινόμενη μέθοδος αυτόματης ταυτοποίησης γραφέα παρουσιάζεται πρώτη φορά στην επιστημονική κοινότητα. Αυτό έχει ως αποτέλεσμα να μην υπάρχει σχετική βιβλιογραφία, να μην υπάρχουν κείμενα που μπορούν να χρησιμοποιηθούν ως πρότυπα και να μην υπάρχει μια βάση δεδομένων για συγκριτική και υποβοηθητική χρήση. Επίσης, εφόσον η σκοπιά από την οποία εξετάζουμε το αντικείμενο είναι αυτή το μηχανικού υπολογιστών και του μαθηματικού, είναι φυσιολογικό να απουσιάζει η ιστορική γνώση για τον αριθμό των διαφορετικών χειριών που είχαν γράψει τα σχετικά χειρόγραφα, μια επικουρική πληροφορία που θα ήταν πολύτιμη.

Εκτός από τις παραπάνω δυσκολίες, αντιμετωπίστηκαν και προβλήματα δομικής υφής. Για παράδειγμα, πολλές επιγραφές και βυζαντινοί κώδικες όπως είναι φυσιολογικό έχουν υποστεί σοβαρή φθορά στο πέρασμα των χρόνων. Επιπρόσθετα, μια ιδιομορφία που έχει να κάνει με τον ανθρώπινο γραφικό χαρακτήρα πρόσθεσε δυσκολίες στην εργασία μας. Πολλές φορές, δύο αλφαβητικά σύμβολα που έχουν γραφεί από τον ίδιο γραφέα παρουσιάζουν μικρότερη ομοιότητα απ' ό,τι δύο αλφαβητικά σύμβολα διαφορετικών γραφέων. Αυτό έχει ως αποτέλεσμα να εισάγεται ένας παράγοντας αβεβαιότητας στη μεθοδολογία ο οποίος αν και εξαλείφεται από τη στατιστική, χρήζει ιδιαίτερης προσοχής.

Ένα άλλο σημείο που έπρεπε να αντιμετωπιστεί, είναι ότι για πολλούς γραφείς υπάρχει ουσιώδης μεταβλητότητα στον τρόπο με τον οποίο ένας γραφέας σχηματίζει ένα αλφαβητικό σύμβολο σε ένα κείμενο. Για παράδειγμα, στο ίδιο έγγραφο και για τον ίδιο γραφέα ένα πλήθος υλοποιήσεων του “ω” μπορεί να είναι κλειστό, ενώ ένα άλλο πλήθος μπορεί να είναι ανοικτό όπως φαίνεται στην εικόνα 2.1. Η ίδια παρατήρηση ισχύει και για το γράμμα “ε” και άλλα συμβολογράμματα. Ένα χαρακτηριστικό παράδειγμα είναι το αλφαβητικό σύμβολο κάππα, το οποίο σε άλλες υλοποιήσεις μοιάζει με το λατινικό “u” ενώ σε άλλες μοιάζει με το κλασικό “κ”. Σε πολλές περιπτώσεις η μορφή του συμβολογράμματος που επιλέγεται από το γραφέα έχει να κάνει με τη θέση που εμφανίζεται αυτό μέσα σε μια λέξη (μέσα στη λέξη ή στο τέλος), ποιο γράμμα προηγείται ή ακολουθεί καθώς επίσης την κούραση του γραφέα τη στιγμή της υλοποίησης. Ένα άλλο χαρακτηριστικό παράδειγμα είναι το κεφαλαίο άλφα στις αρχαίες ελληνικές επιγραφές. Πολλοί γραφείς δε συνδέουν τη μεσαία οριζόντια γραμμή του “Α” με το αριστερό πόδι, άλλοι με το δεξί πόδι και άλλες φορές δεν τη συνδέουν με κανένα.



Εικόνα 2.1

Τέλος, ακόμα και αν η μορφή του συμβολογράμματος είναι σταθερή, πολύ σημαντικό ρόλο παίζει ο θόρυβος του περιγράμματος. Το φαινόμενο του θορύβου, που εξαρτάται σε μεγάλο βαθμό από την κούραση, τη διάθεση και την ηλικία του γραφέα, τα μέσα που αυτός χρησιμοποιεί όπως την ποιότητα του μαρμάρου ή του υποθέματος γραφής, τη συγκεκριμένη μορφή του οργάνου γραφής (μολύβι, σμίλη κλπ) και από τη φθορά του κειμένου, εμφανίζεται πολύ συχνότερα στις αρχαίες επιγραφές .

Κεφάλαιο 3

3.1 Μέθοδοι Clustering

Τα τελευταία χρόνια πολύ μεγάλη πρόοδο γνωρίζει ο κλάδος της επιστήμης των υπολογιστών που ασχολείται με την εξόρυξη δεδομένων data mining. Με την εξέλιξη αυτή έχουμε τον καταμερισμό ενός προβλήματος σε επιμέρους ζητήματα. Ένα από αυτά και πολύ σημαντικό μάλιστα είναι η ομαδοποίηση δεδομένων γνωστή ως data clustering. Αυτό αφορά την κατάταξη δεδομένων σε ομάδες εντός των οποίων υπάρχει ομοιότητα μεταξύ των στοιχείων τους, οι ομάδες αυτές όμως ωφελούν να είναι σαφώς διαχωρισμένες μεταξύ τους. Το clustering των δεδομένων είναι θεμελιώδους σημασίας για ποικίλα θέματα, όπως για παράδειγμα στην αναζήτηση δεδομένων σε βάσεις, αλλά και στην ορθή κατάταξη σε κατηγορίες στοιχείων για στατιστική ανάλυση και περαιτέρω επεξεργασία.

Κατά καιρούς έχουν προταθεί πολλές διαφορετικές μέθοδοι clustering λόγω της μεγάλης χρησιμότητας και εφαρμοσιμότητας που έχει η ομαδοποίηση δεδομένων. Η κάθε μια από τις μεθόδους αυτές παρουσιάζει πλεονεκτήματα ή μειονεκτήματα ανάλογα με την περίπτωση που καλείται να αντιμετωπίσει. Όλες αυτές οι μέθοδοι μπορούν να χωριστούν σε δύο κυρίαρχες τεχνικές clustering :

- Τις Τεχνικές Διαμέρισης ή Διαχωριστικές Τεχνικές (Partitional Techniques)
- Τις Ιεραρχικές Τεχνικές (Hierarchical Techniques)

3.1.1 Τεχνικές Διαμέρισης (Partional Techniques)

Στην αρχή στις τεχνικές διαμέρισης ακολουθήθηκε προσέγγιση με βάση το κέντρο βάρους της περιοχής (centroid). Μετά εμφανίστηκαν διαφορετικές προσεγγίσεις με βάση κάποια σημεία αντιπροσώπους (medoids). Οι τεχνικές με βάση το κέντρο βάρους της περιοχής, γνωστές και ως centroid-based ουσιαστικά αποδίδουν σημεία σε μία υποπεριοχή, όταν η απόστασή τους π.χ. Ευκλείδεια, τετραγωνική κτλ. από το κέντρο βάρους της υποπεριοχής αυτής ελαχιστοποιείται.

Μια τέτοια μέθοδος αρκετά διαδεδομένη είναι η K-means. Αυτή διαλέγει αρχικά τυχαία K περιοχές και στην συνέχεια αποδίδει σε κάθε μία από αυτές τα σημεία της αρχικής περιοχής με βάση το παραπάνω κριτήριο. Στον αλγόριθμο ISODATA επικρατεί αντίστοιχη λογική όπως και στις διάφορες παραλλαγές του.

Στις τεχνικές medoid-based ορίζονται σημεία αντιπρόσωποι (medoids), αντί για το κέντρο βάρους (centroid), έτσι ώστε η μέση διαφορά τους από τα υπόλοιπα σημεία του cluster να είναι ελάχιστη. Πρέπει να πούμε εδώ ότι με τον όρο διαφορά εννοούμε

τον βαθμό ομοιότητας μεταξύ των δεδομένων, που μπορεί να είναι είτε η Ευκλείδεια απόσταση είτε άλλου είδους αποστάσεις, π.χ. manhattan, minkowski.

Σε σχέση με την κατάτμηση που θα δώσει ο K-means οι αλγόριθμοι που βασίζονται σε medoids αντί των centroids (PAM, Clarans), δίνουν πιο αξιόπιστα αποτελέσματα, όμως και αυτοί δεν αποφεύγουν την αποτυχία σε clusters που διαφέρουν σημαντικά ως προς το μέγεθος, αλλά και σε αυτούς που έχουν κυρτά σχήματα. Επηρεάζονται λιγότερο από τον θόρυβο, όμως λειτουργούν καλύτερα, όπως και οι τεχνικές centroidbased, όταν οι υποπεριοχές έχουν κυκλικά σχήματα και παραπλήσια μεγέθη.

3.1.2 Ιεραρχικές Τεχνικές (Hierarchical Techniques)

3.1.2.1 Γενική περιγραφή

Οι ιεραρχικές τεχνικές ξεκινούν θεωρώντας όλα τα σημεία του συνόλου ως ξεχωριστές υποπεριοχές το καθένα και στη συνέχεια συγχωνεύουν τις δύο περιοχές που βρίσκονται κοντινότερα, σε μία. Η διαφορά αυτή των ιεραρχικών μεθόδων σε σχέση με τις διαχωριστικές τεχνικές είναι θεμελιώδης, αφού στις partitional μεθόδους οι τελικές περιοχές στις οποίες διαχωρίζονταν η συνολική ορίζονταν από την αρχή και αποδίδονταν σε κάθε μία τα σημεία. Λειτουργούν, δηλαδή συσσωρευτικά, αφού η διαδικασία της συγχώνευσης επαναλαμβάνεται κάθε φορά μέχρι να φτάσουμε στον επιθυμητό αριθμό περιοχών ή μέχρις ότου ικανοποιηθεί ένα κατάλληλο κατώφλι απόστασης μεταξύ των δεδομένων.

Τώρα θα αναφέρουμε τις γενικές αρχές που ισχύουν στις ιεραρχικές τεχνικές όταν έχουμε μία αρχική περιοχή και τα σημεία της .

Αρχικά ορίζουμε ένα μέτρο ομοιότητας μεταξύ αυτών των σημείων (π.χ. απόσταση) και κατασκευάζουμε έναν πίνακα ομοιότητας $n \times n$ όλων των σημείων μεταξύ τους. Έπειτα από τον πίνακα ομοιότητας, σε κάθε βήμα παράγεται ένας νέος τέτοιος πίνακας, ο οποίος συνήθως είναι αραιός πίνακας. Αυτό επιτυγχάνεται με την εισαγωγή του μηδενός στα στοιχεία του πίνακα που αντιστοιχούν σε clusters που είναι πολύ ανόμοιοι. Για το λόγο αυτό αναπαρίστανται από αραιούς γράφους. Οι κόμβοι του γράφου είναι τα δεδομένα και οι σταθμισμένες ακμές του γράφου είναι οι ομοιότητες των στοιχείων μεταξύ τους.

Οι μέθοδοι που έχουν παρουσιαστεί και οι οποίες βασίζονται σε agglomerative hierarchical clustering είναι πολλές, ενώ το κυρίαρχο στοιχείο το οποίο τις διαφοροποιεί μεταξύ τους είναι ο τρόπος με τον οποίο υπολογίζεται η ομοιότητα μεταξύ των υπάρχοντων clusters και των συγχωνευμένων, σε κάθε επανάληψη.

3.1.2.2 Μέθοδοι

Ανάλογα με το κριτήριο που χρησιμοποιείται κάθε φορά έχουμε τις παρακάτω παραλλαγές των ιεραρχικών τεχνικών:

- Ορίζεται σε κάθε cluster ένας αντιπρόσωπος (π.χ. κέντρο βάρους). Στην αρχή κάθε σημείο είναι και ο αντιπρόσωπος του εαυτού του. Μετράτε στη συνέχεια η ομοιότητα μεταξύ των αντιπροσώπων κάθε cluster και με βάση αυτό το κριτήριο και κάποιο συγκεκριμένο κατώφλι αποφασίζεται αν θα συγχωνευτούν δύο clusters.
- single link method: Κάθε cluster αντιπροσωπεύεται από όλα του τα στοιχεία και η ομοιότητα μεταξύ δύο clusters καθορίζεται από την ομοιότητα του πιο κοντινού ζεύγους στοιχείων που ανήκουν σε διαφορετικά clusters.
- CURE: Η μέθοδος αυτή είναι συνδυασμός των δύο προηγούμενων μεθόδων. Ορίζονται αντιπρόσωποι, αλλά περισσότεροι από ένας για κάθε cluster. Η ομοιότητα μετράτε μεταξύ των ζευγών των αντιπροσώπων των cluster και το κριτήριο συνένωσης δύο περιοχών εξαρτάται από τον βαθμό ομοιότητας των ζευγών αυτών που ανήκουν σε διαφορετικά clusters. Σε κάθε συγχώνευση clusters, δυναμικά ορίζονται νέοι αντιπρόσωποι του νέου cluster που προέκυψε. Επειδή ο αλγόριθμος αυτός αποδίδει πολλούς αντιπροσώπους σε κάθε περιοχή, αλλά και επειδή τους συρρικνώνει προς το κέντρο βάρους, αντιμετωπίζει καταστάσεις διαφορετικών μεγεθών και σχημάτων clusters με μεγαλύτερη επιτυχία. Ο αριθμός των αντιπροσώπων εξαρτάται από το μέγεθος, το σχήμα και την παρουσία θορύβου σε κάθε cluster.
- group-average method: Η μέθοδος αυτή υποθέτει διασύνδεση όλων των δεδομένων των clusters και υπολογίζει όλες τις αποστάσεις μεταξύ τους κανονικοποιώντας τις τιμές με nxm , όπου n είναι ο αριθμός των δεδομένων στην μία περιοχή και m ο αντίστοιχος στην δεύτερη, για να αντιμετωπίσει το πρόβλημα που προκύπτει όταν τα clusters έχουν διαφορετικό μέγεθος.
- ROCK: Ο αλγόριθμος αυτός μοιάζει με αυτόν που ακολουθείται στην μέθοδο group-average, κανονικοποιεί όμως με βάση κατώφλι ορισμένο από τον χρήστη, αντί του nxm .

3.1.2.3 Δυσκολίες ιεραρχικών τεχνικών

Όπως καταλαβαίνουμε οι ιεραρχικές τεχνικές παρουσιάζουν καλύτερα αποτελέσματα σε σχέση με τις διαχωριστικές αλλά παρόλα αυτά εμφανίζουν και αυτές ορισμένα μειονεκτήματα κυρίως διότι βασίζονται σε στατικά μοντέλα και δεν λαμβάνουν υπ' όψιν τις ιδιαιτερότητες κάθε cluster.

- Όταν χρησιμοποιούνται ως αντιπρόσωποι τα κέντρα βάρους των περιοχών,

η περίπτωση παρουσιάζει παρόμοιες αστοχίες με τις διαχωριστικές μεθόδους Kmeans

και K-medoids σε καταστάσεις διαφορετικών σχημάτων ή μεγεθών.

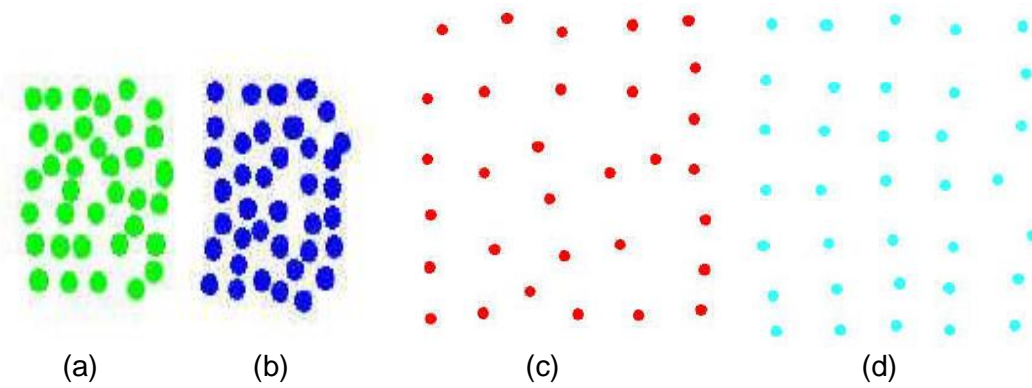
- Η single link method μπορεί να αντιμετωπίσει την παραπάνω δυσκολία αποτυγχάνει όμως όταν υπάρχει θόρυβος, αφού λαμβάνει υπόψιν της όλα τα ζεύγη δεδομένων.

- Ο CURE δεν δυσκολεύεται όταν καλείται να αποφασίσει για την συγχώνευση clusters ποικίλων μεγεθών και σχημάτων.

Όμως το γεγονός ότι λαμβάνει υπόψη του μόνο την ελάχιστη απόσταση μεταξύ των αντιπροσώπων και όχι τη συνολική διασύνδεση μεταξύ των δύο clusters οδηγεί σε λανθασμένες συγχωνεύσεις σε ειδικές περιπτώσεις.

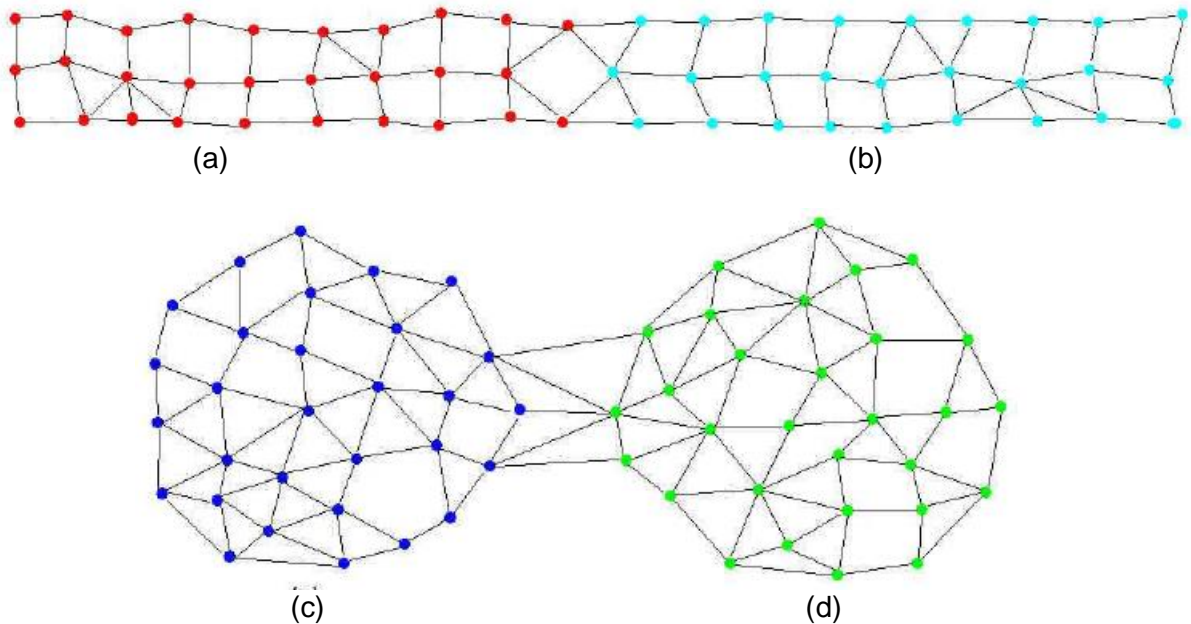
- Ο ROCK τώρα λαμβάνει υπόψη του την συνολική διασύνδεση των clusters και αδιαφορεί για την τιμή των βαρών των ακμών δια μέσου των clusters, οπότε και αυτός ο αλγόριθμος θα δίνει μη ικανοποιητικές συγχωνεύσεις σε συμπληρωματικές περιπτώσεις από αυτές που αστοχεί ο CURE.

Στις παρακάτω εικόνες 1 και 2 φαίνονται ειδικές περιπτώσεις όπου οι αλγόριθμοι CURE και ROCK παρουσιάζουν πιθανές αστοχίες κατά την απόφαση για συγχώνευση δύο περιοχών .



Εικόνα 1

Οι Cure και single line, θα συγχωνεύσουν τις περιοχές (α), (β) αντί των (c), (d), γιατί υπάρχουν μικρότερες αποστάσεις μεταξύ των αντιπροσώπων των περιοχών αυτών. Όμως οι αποστάσεις των αντιπροσώπων των περιοχών (c), (d) είναι αντίστοιχης τάξης με τις αποστάσεις των σημείων εντός των περιοχών επομένως η συγχώνευση των συγκεκριμένων περιοχών είναι προτιμότερη.

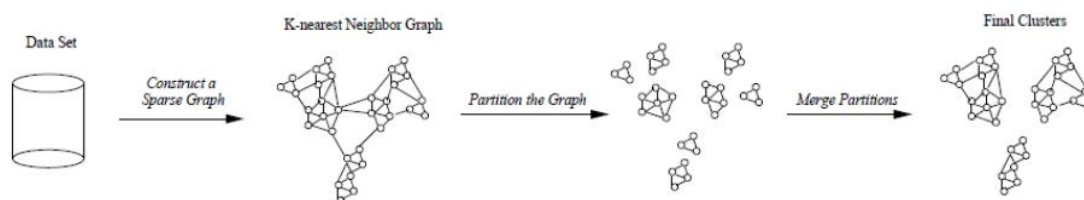


Ο ROCK θα ενώσει τα (c) και (d), καθώς λαμβάνει υπόψιν την διασύνδεση των περιοχών, αντί των (a) και (b) που έχουν μικρότερη απόσταση μεταξύ τους.

Κεφάλαιο 4

4.1 Ομαδοποίηση Δεδομένων Με Δυναμικό Μοντέλο (Chameleon Data Clustering Algorithm)

Ο αλγόριθμος chameleon, παρουσιάστηκε από τους Karypis, Han και Kumar εφαρμόζει δυναμικό μοντέλο αντί του στατικού που χρησιμοποιείται από τους αλγόριθμους που παρουσιάστηκαν μέχρι τώρα. Η βασική του ιδέα είναι ο συνδυασμός των ιδιοτήτων μιας περιοχής με την σχέση που συνδέει δυο περιοχές μεταξύ τους. Τέλος σε κάθε επανάληψη, το κατώφλι απόφασης θα εξαρτάται από τα δυναμικά χαρακτηριστικά των περιοχών. Στην εικόνα 1 παρουσιάζεται η γενική ιδέα της λειτουργίας του αλγόριθμου.



4.1.1 Γενική περιγραφή μεθόδου

Ο chameleon προκειμένου να μπορέσουν να συνδυαστούν σωστά η σχέση των δεδομένων μεταξύ τους και οι σχέσεις μεταξύ των διαφορετικών υποπεριοχών ορίζει την διασυνδεσιμότητα (inter-connectivity) και την εγγύτητα (closeness), οι οποίες καθορίζουν και την τελική ένωση των υποπεριοχών. Τώρα θα μιλήσουμε για τα βήματα που ακολουθούνται για την εκτέλεση του αλγορίθμου chameleon. Αρχικά κατασκευάζουμε τον γράφο που συσχετίζει τα σημεία με βάση το μέτρο ομοιότητας που έχουμε ορίσει μεταξύ τους. Έπειτα θεωρούμε μια περιοχή (κόμβο του γράφου) και για κάθε σημείο της περιοχής παίρνουμε τα K – κοντινότερα σε όλη την εικόνα τα οποία ενώνουμε με το αρχικό με ευθείες (edges) αποδίδοντας ταυτόχρονα σε κάθε edge ένα βάρος (ο βαθμός ομοιότητάς τους). Τα απομακρυσμένα σημεία δεν συνδέονται καθόλου στον γράφο. Στη συνέχεια χωρίζουμε τις ομαδοποιημένες υποπεριοχές του αρχικού cluster και τέλος συγχωνεύουμε αυτά τα sub-clusters σύμφωνα με τα κριτήρια ομοιότητας που ορίζει ο chameleon, δηλαδή λαμβάνοντας υπόψιν την interconnectivity και την closeness

Ένωση δύο περιοχών (sub-clusters) C_i, C_j

Το κριτήριο της συγχώνευσης δύο sub-clusters C_i, C_j είναι να συνδέονται ταυτόχρονα με υψηλή διασυνδεσιμότητα (inter-connectivity) και εγγύτητα (Closeness), ο ορισμός των οποίων δίνεται παρακάτω:

Σχετική Inter-connectivity: Για να υπολογιστεί η σχετική διασυνδεσιμότητα (relative inter-connectivity), ορίζονται πρώτα η απόλυτη inter-connectivity εντός ενός cluster καθώς επίσης και η απόλυτη inter-connectivity μεταξύ δύο cluster.

- Απόλυτη inter-connectivity μίας περιοχής: Ο αριθμός των βαρών των edges εντός μιας περιοχής C_i χρησιμοποιείται για τον υπολογισμό της εσωτερικής διασύνδεσης της C_i . Διχοτομούμε την περιοχή C_i και υπολογίζουμε το άθροισμα των βαρών των edges που συναντά αυτή η τομή. Η interconnectivity του cluster C_i είναι: EC_{C_i}

- Απόλυτη inter-connectivity μεταξύ δύο περιοχών: Παρόμοια με την απόλυτη inter-connectivity μίας περιοχής, υπολογίζουμε την απόλυτη interconnectivity μεταξύ δύο περιοχών C_i και C_j θεωρώντας την περιοχή που συμπεριλαμβάνει τις C_i και C_j και αθροίζοντας τα ζυγισμένα βάρη των edges που συναντά η τομή που χωρίζει το ευρύτερο cluster στις C_i και C_j .

Συμβολίζεται $EC_{\{C_i, C_j\}}$

- Τέλος κανονικοποιούμε την απόλυτη inter-connectivity μεταξύ των δύο περιοχών, με τις απόλυτες inter-connectivity της κάθε περιοχής, για να λάβουμε την τελική σχετική inter-connectivity των δύο cluster.

$$RI_{(C_i, C_j)} = \frac{|EC_{\{C_i, C_j\}}|}{\frac{|EC_{C_i}| + |EC_{C_j}|}{2}}$$

- Σχετική Closeness: Όπως και κατά τον υπολογισμό της σχετικής διασυνδεσιμότητας ορίσαμε τα απόλυτα μεγέθη, αντίστοιχα ορίζουμε την απόλυτη εγγύτητα μίας περιοχής και την απόλυτη εγγύτητα μεταξύ δύο περιοχών.

- Εσωτερική closeness ενός cluster: Εντός μίας περιοχής υπολογίζουμε τη μέση τιμή των βαρών των συνδέσεων μεταξύ των σημείων που αποτελούν τον κόμβο. Παρόμοια με τον υπολογισμό της inter-connectivity, λαμβάνονται

υπόψιν οι συνδέσεις πάνω στην ευθεία που χωρίζει το cluster σε δύο ίσα μέρη. Για τις περιοχές C_i και C_j αντίστοιχα είναι:

$$\bar{SEC}(C_i) \quad \bar{SEC}(C_j)$$

- Απόλυτη closeness μεταξύ δύο περιοχών: Ανάμεσα σε δύο clusters C_i και C_j μετράτε η μέση τιμή της ομοιότητας των στοιχείων του C_i που συνδέονται με αυτά του C_j και συμβολίζεται με:

$$\bar{SEC}(C_i, C_j)$$

- Τελικά η σχετική εγγύτητα μεταξύ των cluster C_i και C_j θα δίνεται από την σχέση:

$$RC(C_i, C_j) = \frac{\bar{SEC}(C_i, C_j)}{\frac{|C_i|}{|C_i| + |C_j|} \bar{SEC}(C_i) + \frac{|C_j|}{|C_i| + |C_j|} \bar{SEC}(C_j)}$$

- Όπου με $|C_i|$, $|C_j|$ συμβολίζονται οι αριθμοί των σημείων κάθε περιοχής και χρησιμοποιούνται για την κανονικοποίηση της σχετικής closeness.

4.1.2 Λειτουργία του Chameleon

Η λειτουργία του αλγορίθμου χωρίζεται στη Φάση 1η : Καθορισμός των αρχικών υποπεριοχών και στη Φάση 2η : Συγχώνευση των υποπεριοχών με χρήση του δυναμικού μοντέλου. Τώρα θα μιλήσουμε ποιο αναλυτικά για τις φάσεις αυτές.

Φάση 1η: Καθορισμός των αρχικών υποπεριοχών

Εδώ χωρίζεται η προς μελέτη περιοχή σε ένα αρχικό αριθμό υποπεριοχών, ώστε μετά να μπορεί να εφαρμοστεί το δυναμικό μοντέλο . Αν έχουμε δηλαδή ένα cluster C_i τον κατατέμνουμε σε δύο νέες υποπεριοχές $C_{i,A}$, $C_{i,B}$, για τις οποίες απαιτείται: να έχουν ικανό αριθμό στοιχείων , να είναι μικρότερες από τις μικρότερες περιοχές που αναμένεται να ληφθούν στο τέλος , να διαχωρίζονται έτσι ώστε η ομοιότητα μεταξύ των $C_{i,A}$, $C_{i,B}$ να είναι ελάχιστη και τέλος η κάθε μία από τις $C_{i,A}$, $C_{i,B}$ να έχει τουλάχιστον 25% των δεδομένων της αρχικής περιοχής . Για την ικανοποίηση αυτών των απαιτήσεων εφαρμόζονται αλγόριθμοι πολυεπίπεδου διαχωρισμού του γράφου των K γειτονικότερων σημείων της περιοχής. Η διαδικασία διαχωρισμού του μεγαλύτερου sub-cluster σε δύο ίσα μέρη συνεχίζεται και στους νέους sub-cluster μέχρις ότου πάψουν να πληρούνται οι προϋποθέσεις.

Φάση 2η: Συγχώνευση των υποπεριοχών με χρήση του δυναμικού μοντέλου

Σκοπός της φάσης αυτής είναι η συγχώνευση των περισσότερο όμοιων υποπεριοχών που έχουν προκύψει από την Φάση 1, εφαρμόζοντας βέβαια τις έννοιες που ορίστηκαν προηγουμένως της σχετικής διασυνδεσιμότητας και της σχετικής εγγύτητας.

Υπάρχουν δύο διαφορετικά Σχήματα που χρησιμοποιούνται:

1. Σχήμα (α): Κατώφλια για τις RI και RC
2. Σχήμα (β): Συνάρτηση των RI και RC

Σχήμα (α)

Για κάθε cluster C_i αναζητεί τον C_j που δίνει τιμές για την interconnectivity και την closeness μεγαλύτερες από τα προκαθορισμένα κατώφλια TRI και TRC .

Ελέγχονται,

δηλαδή οι clusters για τους οποίους ταυτόχρονα ικανοποιούνται οι συνθήκες:

$$RI(C_i, C_j) \geq TRI, RC(C_i, C_j) \geq TRC$$

Αφού γίνουν όλοι οι δυνατοί συνδυασμοί η διαδικασία επαναλαμβάνεται από την αρχή, μέχρι του σημείου όπου δεν ικανοποιούνται πλέον τα κατώφλια.

Σχήμα (β)

Ορίζεται συνάρτηση που συνδυάζει την inter-connectivity με την closeness και ζητείται η μεγιστοποίησή της. Το πιο άμεσο μοντέλο που μπορεί κανείς να υποθέσει για μια τέτοια συνάρτηση των RI, RC είναι το γινόμενό τους, αφού θέλουμε ταυτόχρονα την μεγιστοποίηση και των δύο κριτηρίων.

$$RI(C, C) * RC(C, C)^a$$

Η παράμετρος a καθορίζει σε ποιο από τα δύο κριτήρια θα δοθεί μεγαλύτερο βάρος. Προφανώς όταν $a > 1$ ο ρόλος της σχετικής closeness γίνεται σημαντικότερος, ενώ αντίθετα όταν $a < 1$ ο παράγοντας της inter-connectivity είναι εκείνος που θα επηρεάσει περισσότερο το τελικό κριτήριο.

Κεφάλαιο 5

Τοπογραφική απόσταση και γραμμή watershed

5.1 Εισαγωγή

Η γραμμή watershed είναι το κύριο εργαλείο που εξάχθηκε στο πλαίσιο της μαθηματικής μορφολογίας για κατάτμηση εικόνων. Ο αλγόριθμος watershed χρησιμοποιείται κυρίως σε εικόνες κλίσης (gradient images) και εντοπίζει τις λεκάνες απορροής (catchment basins). Μια καλή ερμηνεία του αλγόριθμου δίνεται από τον Beucher ο οποίος θεώρησε την εικόνα κλίσης σαν ένα τοπογραφικό ανάγλυφο. Οι πηγές τοποθετούνται στα τοπικά ελάχιστα και το ανάγλυφο πλημμυρίζεται με ομοιόμορφο τρόπο και ταχύτητα. Κατά την στιγμή όπου δύο διαφορετικές κοιλάδες αρχίζουν να ενώνονται, ένα φράγμα ανεγείρεται για να προφυλάξει την συνένωση των κοιλάδων. Η συνένωση όλων των φραγμάτων αποτελεί την γραμμή του watershed. Συχνά αυτά τα ελάχιστα είναι πολλά σε αριθμό και τα περισσότερα από αυτά δεν σχετίζονται με την επιθυμητή κατάτμηση, το οποίο οδηγεί σε υπερκατάτμηση (oversegmentation). Η λύση που προτάθηκε το 1982, βοηθά στην αντιμετώπιση αυτού του προβλήματος, αμέσως μόλις βρεθεί η πρώτη εκτίμηση. Αυτή η προσέγγιση καλείται δείκτες (markers), διότι περιέχει έναν δείκτη για κάθε αντικείμενο που θα ανιχνευτεί, συμπεριλαμβάνοντας το φόντο. Η τμηματοποίηση αποτελείται από δυο βήματα: α)στο πρώτο, ένα πλήθος από markers ανιχνεύεται για κάθε αντικείμενο και για το φόντο, β)στο δεύτερο βήμα κατασκευάζεται ο watershed της εικόνας κλίσης γεμίζοντας το ανάγλυφο από ένα πλήθος πηγών τοποθετημένων στους markers. Μετά το τέλος του γεμίσματος, είναι διαθέσιμα κομμάτια της εικόνας που σε κάθε ένα υπάρχει μόνο ένας marker.

5.2 Ορισμός του watershed στον συνεχή χώρο.

Ο ορισμός του watershed στον συνεχή χώρο πρέπει να γίνει με προσοχή, γιατί υλοποιείται σε εικόνες κλίσης οι οποίες συχνά περιέχουν θόρυβο και δεν είναι ομαλές. Για τον λόγο αυτό, πρέπει να ορισθεί ώστε να ισχύει ακόμα και αν η συνάρτηση δεν είναι αρκετά ομαλή.

5.2.1 Η τοπογραφική μεταβολή

Θεωρούμε συνάρτηση f από το R^n στο R . Έστω ότι $\text{supp}(f)$ είναι το support της συνάρτησης, T διάστημα στο R και γ μια συνεχή συνάρτηση από το T στο $\text{supp}(f)$, (T, γ) μια διαδρομή που περιέχεται στο support της f . Για λόγους συντομίας γράφουμε $\gamma_i = \gamma(t_i)$. Επιπλέον ορίζεται το στοιχειώδες erosion ϵ_i σαν το erosion από δίσκο ακτίνας ϵ_i , την γεωδαισιακή απόσταση ανάμεσα στα σημεία γ_i και γ_{i-1} .

Ορισμός 1

Η τοπογραφική μεταβολή της συνάρτησης f κατά μήκος της πολυγωνικής γραμμής η ορίζεται ως:

$$TV_\zeta = \sum_i [f(\gamma_i) - \epsilon_i f(\gamma_i)]$$

Ορισμός 2

Τοπογραφική μεταβολή της συνάρτησης f στη διαδρομή (T, γ) , ονομάζεται ο θετικός αριθμός (πεπερασμένος ή άπειρος) που ορίζεται από:

$$TV_\gamma = \sup TV_\zeta \text{ (για κάθε πεπερασμένο } \zeta \subset \gamma \text{)}$$

Αν $TV_\gamma < \infty$, τότε η f λέγεται ότι είναι πεπερασμένης τοπογραφικής μεταβολής στο T . Με αυτόν τον τρόπο, η τοπογραφική απόσταση (topographical distance) δυο σημείων p και q , ορίζεται θεωρώντας το σύνολο $\Gamma(p, q)$ το οποίο ανήκει στο support της f :

$$TD(p, q) = \inf_{\gamma \in \Gamma(p, q)} TV_\gamma$$

Ορισμός 3

Λεκάνη απορροής (catchment basin) $CB(m_i)$ ενός τοπικού ελαχίστου m_i , καλείται το σύνολο των σημείων $x \in \text{supp}(f)$, τα οποία είναι πιο κοντά στο m_i από κάθε άλλο τοπικό ελάχιστο για την τοπογραφική απόσταση:

$$\forall j \in I, j \neq i \Rightarrow TD(x, m_i) < TD(x, m_j)$$

Ορισμός 4

Η γραμμή watershed μίας συνάρτησης f , είναι το σύνολο των σημείων του $\text{supp}(f)$ τα οποία δεν ανήκουν σε καμία λεκάνη απορροής

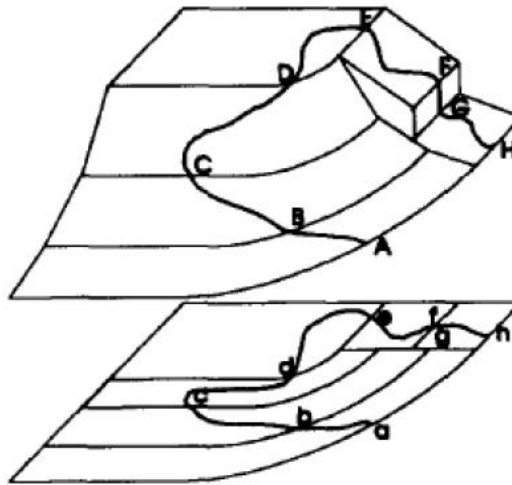
$$Wsh(f) = \text{supp}(f) \cap [\cup_i (CB(m_i))]^c$$

5.2.2 Εφαρμογή σε δυαδικές εικόνες

Μια απαίτηση για τον ορισμό της γραμμής watershed είναι η εφαρμογή σε δυαδικές εικόνες.

5.2.3 Ερμηνεία της συνάρτησης τοπογραφικής απόστασης

Εδώ θα δοθεί ένα παράδειγμα για την κατανόηση της τοπογραφικής απόστασης σε απεικονίσεις του γκρι. Στην εικόνα 1 απεικονίζεται μια συνάρτηση τόνων του γκρι με σταθερή κλίση σε κάθε επίπεδο της εικόνας. Παρακάτω υπολογίζεται η τοπογραφική απόσταση στο μονοπάτι (A, B, C, D, E, F, G, H) του οποίου η προβολή είναι (a, b, c, d, e, f, g, h) . Το κομμάτι της διαδρομής από το a στο b , ανήκει σε μια ζώνη σταθερής κλίσης α . Η τοπογραφική απόσταση για το κομμάτι (A, B) είναι ίση με την γεωδαισιακή απόσταση της αντίστοιχης προβολής (a, b) πολλαπλασιασμένη με το βάρος α . Αν οι κλίσεις των διαδοχικών κομματιών ab, bc, cd, de, ef, fg και gh είναι αντίστοιχα $\alpha, \beta, \gamma, \delta, \epsilon, \eta$ και θ , τότε θ τοπογραφική απόσταση κατά μήκος της διαδρομής (AH) είναι το άθροισμα $\alpha * d(a, b) + \beta * d(b, c) + \gamma * d(c, d) + \delta * d(d, e) + \epsilon * d(e, f) + \eta * d(f, g) + \theta * d(g, h)$. Από το D στο E , θ τοπογραφική απόσταση κα είναι μηδέν, εφόσον η κλίση δ είναι μηδέν, ενώ από το F στο G η κλίση είναι άπειρη γιατί θ απόσταση από το f στο g ισούται με μηδέν. Σε αυτήν την περίπτωση $\delta * d(f, g) = f(f) - f(g)$.



Εικόνα 1: Συσχέτιση της τοπογραφικής απόστασης σε μια διαδρομή, με την γεωδαισιακή απόσταση στην προβολή της

5.3 Τοπογραφική απόσταση και λεκάνες απορροής στον διακριτό χώρο.

Στον διακριτό χώρο, πάντα υπάρχουν τα ίδια προβλήματα: απειροελάχιστα δομικά στοιχεία δεν υπάρχουν στο ψηφιακό πλέγμα και οι γειτονικές σχέσεις είναι αδύναμες. Παρ' όλα αυτά είναι δυνατόν να γίνουν καλές προσεγγίσεις της γραμμής watershed.

5.3.1 Ερμηνεία της συνάρτησης τοπογραφικής απόστασης

Θεωρούμε μια συνάρτηση τόνων του γκρι f από το Z^n στο Z και το support της $\text{supp}(f)$. Έστω G το υποκείμενο πλέγμα, το οποίο μπορεί να είναι οποιοδήποτε τύπου. Συσχετίζεται στο G ένας γειτονικός γράφος U , όπου είναι ένα υποσύνολο $Z^n \times Z^n$ το οποίο ορίζεται από $(a,b) \in U$ αν και μόνο αν τα a και b είναι γειτονικά.

Ονομάζεται $N_U(p)$ τα γειτονικά εικονοστοιχεία ενός εικονοστοιχείου p , ως προς το U και το $\text{supp}(f)$: $N_U(p) = \{p' \in Z^n, (p,p') \in U \cap \text{supp}(f)\}$. Το υποσύνολο $B(p)$ όλων των εικονοστοιχείων του $N_U(p)$ που βρίσκονται σε απόσταση 1 από το p καλείται μοναδιαίος δίσκος μεγέθους 1. Το erosion της συνάρτησης f από αυτόν τον δίσκο είναι το στοιχειώδες erosion εf .

Ορισμός 5:

Ένα σύνολο π πλήθους n μεταξύ δυο εικονοστοιχείων p και q στο πλέγμα G , είναι μια πλειάδα n -εικονοστοιχείων (p_1, p_2, \dots, p_n) τέτοια ώστε $p_1=p, p_n=q$ και $\forall i [1, n-1], (p_i, p_{i+1}) \in G$.

Το μήκος της διαδρομής π , ορίζεται από $l(\pi) = \sum_i \text{dist}(p_i, p_{i+1})$

Η κλίση μεταξύ δυο εικονοστοιχείων p και p' για $f(p') < f(p)$ ορίζεται από

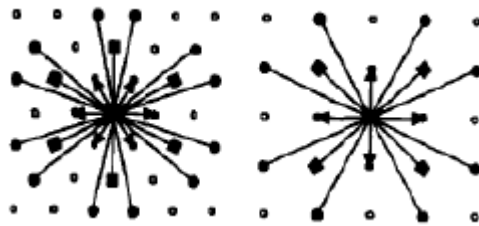
$$\text{slope}(p,p') = \frac{f(p) - f(p')}{\text{dist}(p,p')} .$$

Ορισμός 6:

Το σύνολο των κατώτερων γειτόνων του p (lower neighbours), για τα οποία $\text{slope}(p,p')$ είναι μέγιστη συμβολίζεται ως $\Gamma(p)$. Η τιμή αυτής της μέγιστης κλίσης ονομάζεται κλίση της συνάρτησης στο σημείο p :

$$LS(p) = \max \left[\frac{f(p) - f(p')}{\text{dist}(p,p')} \right] \text{ για } p' \in N \cup (p) \text{ και } f(p') < f(p)$$

Οι περισσότερο χρησιμοποιούμενοι γειτονικοί γράφοι, είναι αυτοί των 4-συνδέσεων στο τετραγωνικό πλέγμα και των 6-συνδέσεων στο εξαγωνικό. Μερικές φορές μπορεί να χρησιμοποιηθούν πιο σύνθετοι, όπως π.χ. 8 ή 16 γειτόνων στο τετραγωνικό πλέγμα, 12 ή 24 στο εξαγωνικό. Αυτές οι γειτονίες ονομάζονται chamfer και μερικά παραδείγματα παρουσιάζονται στην εικόνα 2. Χρησιμοποιούνται για την κατασκευή των συναρτήσεων απόστασης, όπου οι αποστάσεις από κάθε γειτονικό σημείο προς το κεντρικό υπολογίζονται με ακέραιους αρμούς. Ενδεικτικά, σε τετραγωνική διασύνδεση με 16 διασυνδέσεις οι αποστάσεις είναι 5, 7 και 11.



Εικόνα 2 :Βασικές γειτονικές σχέσεις για το τετραγωνικό και το εξαγωνικό πλέγμα

Ορισμός 7:

Η αντιστοίχιση (mapping) $\Gamma(p)$ επιτρέπει τον ορισμό ενός κατευθυνόμενου γράφου V , ως υπογράφο του γειτονικού γράφου U :

$$(p,p') \in V \Leftrightarrow p' \in \Gamma(p)$$

Ορισμός 8:

Ορίζεται το κόστος για την μετακίνηση στην τοπογραφική επιφάνεια από την θέση

$f(p_{i-1})$ στην γειτονική θέση $f(p_i)$:

$$f(p_{i-1}) > f(p_i) \Rightarrow \text{cost}(p_{i-1}, p_i) = LS(p_{i-1}) * \text{dist}(p_{i-1}, p_i),$$

$$f(p_{i-1}) < f(p_i) \Rightarrow \text{cost}(p_{i-1}, p_i) = LS(p_i) * \text{dist}(p_{i-1}, p_i),$$

$$f(p_{i-1}) = f(p_i) \Rightarrow \text{cost}(p_{i-1}, p_i) = \frac{LS_{p_{i-1}} + LS_{p_i}}{2} \times \text{dist}(p_{i-1}, p_i).$$

Ορισμός 9:

Έστω f μια συνάρτηση γκρι τόνου και π η διαδρομή $(p_1=p, p_2, \dots, p_n=q)$ μεταξύ δύο εικονοστοιχείων p και q μέσα στο $\text{supp}(f)$. Τότε η π - τοπογραφική απόσταση μεταξύ των p και q στην f κατά μήκος της διαδρομής π είναι η σταθμισμένη απόσταση (weighted distance) που ορίζεται από:

$$T_f^\pi(p, q) = \sum_{i>1} \text{cost}(p_{i-1}, p_i)$$

Ορισμός 10:

Η τοπογραφική απόσταση μεταξύ δύο εικονοστοιχείων p και q ορίζεται ως η ελάχιστη π -τοπογραφική απόσταση μεταξύ των δύο εικονοστοιχείων p και q μεταξύ όλων των διαδρομών π ανάμεσα στα p και q μέσα στο $\text{supp}(f)$:

$$T_r(p, q) = \inf(T_f^\pi(p, q))$$

Ορισμός 11:

Ένα εικονοστοιχείο p ανήκει στο upstream ενός εικονοστοιχείου q , εάν υπάρχει διαδρομή π απότομης κλίσης μεταξύ p και q :

$$\pi = (p_1=p, p_2, \dots, p_n=q) \text{ και } \forall i, i \leq n-1 \Gamma(p_i).$$

5.3.2 Ορισμός λεκάνης απορροής

Οι ορισμοί της λεκάνης απορροής και της γραμμής watershed παραμένουν οι ίδιοι στο διακριτό χώρο, αντικαθιστώντας την συνεχή απόσταση TD με την διακριτή T_f .

5.4 Η γραμμή watershed και οι αλγόριθμοι ελαχίστης διαδρομής

Στους κόμβους που ανήκουν σε τοπικά ελάχιστα αντιστοιχίζεται κόστος ίσο με το ύψος τους. Η κατασκευή της λεκάνης απορροής σε μια συνάρτηση του γκρι, ανάγεται στο πρόβλημα εύρεσης της διαδρομής με ελάχιστο κόστος μεταξύ κάθε εικονοστοιχείου και ενός τοπικού ελαχίστου. Όλα τα εικονοστοιχεία κατά μήκος μίας διαδρομής ελαχίστου κόστους, έχουν την ίδια 'ετικέτα' (label) με το τοπικό ελάχιστο στην αρχή της διαδρομής.

5.4.1 Γεμίζοντας από τα τοπικά ελάχιστα, ή από τους markers

Στις εικόνες κλίσης υπάρχει πληθώρα τοπικών ελαχίστων, αλλά ελάχιστα από αυτά είναι χρήσιμα για την κατάτμηση της, γεγονός που πολλές φορές οδηγεί σε υπερκατάτμηση (oversegmentation). Όπως αναφέρθηκε, η λύση σε αυτό το πρόβλημα είναι η τοποθέτηση ενός συνόλου από markers σε κάθε αντικείμενο και στο φόντο, από τους οποίους το ανάγλυφο της εικόνας πλημμυρίζεται. Οι

παρακάτω αλγόριθμοι περιγράφουν τον τρόπο κατασκευής της γραμμής του watershed που σχετίζεται με το σύνολο των τοπικών ελαχίστων. Στο πρόβλημα πως θα εισάγουμε τους markers, υπάρχει μια κλασική λύση από τους Beucher και Meyer η οποία ονομάζεται τροποποίηση ομοτοπίας (homotopy modification).

5.4.2 Τροποποίηση της ομοτοπίας της εικόνας κλίσης

Ο αλγόριθμος για αυτήν την διαδικασία είναι απλός και βασίζεται σε ανακατασκευή του γκρι. Η κλασική λύση αποτελείται από την αντικατάσταση της αρχικής κλίσης g , από μια καινούργια συνάρτηση g' , η οποία έχει μόνο τα ελάχιστα τα οποία αντιστοιχούν σε markers και έχουν την τιμή 0. Μετά από αυτόν τον μετασχηματισμό, όλοι οι κλασικοί αλγόριθμοι για την κατασκευή του watershed μπορούν να εφαρμοστούν.

Κεφάλαιο 6

6.1 ACTIVE CONTOURS

Θα αναπαραστήσουμε με τη βοήθεια μιας παραμέτρου p μια καμπύλη C στο επίπεδο ως εξής $C(p)=\{x(p),y(p)\}$ όπου το p ανήκει στο διάστημα $[0,1]$.

Κάθε τιμή της παραμέτρου p δίνει τις συντεταγμένες ενός σημείου στο επίπεδο και κατά συνέπεια ένα σημείο της καμπύλης C . Ορίζουμε σε κάθε σημείο της καμπύλης:

α) \vec{T} το εφαπτομενικό διάνυσμα σε αυτό το σημείο της καμπύλης β) \vec{N} το κάθετο διάνυσμα σε αυτό το σημείο

Αυτά τα δύο διανύσματα σχηματίζουν ένα ορθοκανονικό σύστημα συντεταγμένων διότι είναι κάθετα μεταξύ τους και τους αποδίδουμε συνήθως μοναδιαίο μέτρο

γ) $\kappa(p)$ η καμπυλότητα της καμπύλης η οποία ορίζεται ως η μεταβολή της γωνίας $\theta(p)$ του εφαπτομενικού διανύσματος T κατά μήκος της καμπύλης και είναι

$$\kappa(p) = \frac{d\theta}{ds}$$

Τα παραπάνω χαρακτηριστικά της καμπύλης συνδέονται με τις εξής σχέσεις

$$\vec{T}(p) = \frac{\vec{C}_p}{\|\vec{C}_p\|} - \frac{\vec{C}_p}{\sqrt{x^2 p - y^2 p}}$$

$$\begin{aligned} \vec{T}_p(p) &= \kappa \vec{N} \\ \vec{N}_p(p) &= -\kappa \vec{T} \end{aligned}$$

Υποθέτουμε πως μια καμπύλη C εξελίσσεται στο χρόνο t . Ονομάζουμε την αρχική καμπύλη $C_0=C(p,t=0)$ και την καμπύλη σε μια χρονική στιγμή $t>0$ ως $C=C(p,t)$. Η εξίσωση της μορφής $C=C(p,t)$ μας δείχνει την εξίσωση της καμπύλης στο χρόνο. Θεωρώ ότι κάθε σημείο της καμπύλης έχει ταχύτητα \vec{V} η οποία αποτελείται από δύο συντεταγμένες: μία V_t παράλληλη στο \vec{T} και μία V_n παράλληλη

στο \vec{N} οπότε θα είναι $V=(V_t, V_n)$ και άρα $\frac{\partial \vec{C}}{\partial t} = \vec{V}_t \vec{T} + \vec{V}_n \vec{N}$. Τα \vec{V}_t και \vec{V}_n

εξαρτώνται από την καμπυλότητα κ . Το \vec{V}_t είναι εφαπτομενική συνιστώσα και

το \vec{V}_n κάθετη. Η εξίσωση κίνησης γίνεται $\frac{\partial \vec{C}}{\partial t} = \vec{V}(\kappa) \vec{N}$. Η ποσότητα

$|\vec{C}_p| = \sqrt{x^2 p - y^2 p}$ δείχνει πόσο έντονος είναι ο τοπικός κυματισμός της

καμπύλης. Ορίζουμε το μέγεθος της ολικής ταλάντωσης μέσα από τη σχέση

$$Var(t) = \int_C^T |k(p,t)| |C_p| dp.$$

Η ποσότητα $\frac{dVar(t)}{dt}$ μας δείχνει πως εξελίσσεται η καμπύλη (εάν θα ομαλοποιηθεί ή εάν θα παραμορφωθεί) . Ισχύει ότι $\left. \frac{dV}{dk} \right|_{k=0} \leq 0 \Rightarrow \frac{dVar}{dt} \leq 0$ και ότι $\left. \frac{dV}{dk} \right|_{k=0} \geq 0 \Rightarrow \frac{dVar}{dt} \geq 0$ και εάν γνωρίζω τη $V(k)$ μπορώ να προβλέψω την εξέλιξη της καμπύλης.

Όπως αναφέραμε και παραπάνω σε μια τυπική περίπτωση, ένα αντικείμενο διαχωρίζεται από το υπόλοιπο μέρος της εικόνας εξαιτίας πχ μιας απότομης αλλαγής στη φωτεινότητα ή στο χρώμα. Σε κάθε περίπτωση, υπάρχει κάποια συνοριακή γραμμή ή καμπύλη η οποία αποτελεί το σύνορο του αντικειμένου. Μπορεί κανείς να ξεκινήσει από κάποια αρχική καμπύλη στο επίπεδο και στη συνέχεια να επιβάλλει σε αυτή να ακολουθήσει μια εξίσωση κίνησης, η οποία θα την αναγκάσει να καταλήξει στην καμπύλη που αποτελεί το σύνορο του αντικειμένου. Οι καμπύλες αυτές ονομάζονται Ενεργές Καμπύλες (active contours), διότι προσαρμόζονται σταδιακά στην εκάστοτε εικόνα. Για να γίνει αυτό πρέπει να θεσπιστούν κατάλληλα κριτήρια, τα οποία θα απορρέουν από την προς κατάτμηση εικόνα και θα επιδρούν με τέτοιο τρόπο στην αρχική καμπύλη ώστε αυτή θα κινείται προς τα σύνορα των αντικειμένων. Γενικά τα κριτήρια αυτά προκύπτουν όταν με κατάλληλη επεξεργασία της αρχικής εικόνας προκύπτει μια δύναμη F η οποία «έλκει» την αρχική καμπύλη προς τις ακμές των αντικειμένων. Το πώς θα γίνει αυτή η επεξεργασία ώστε να προκύψει μια αποτελεσματική δύναμη F είναι ένα θέμα που επιδέχεται πολλές λύσεις, άλλες λιγότερο και άλλες περισσότερο αποτελεσματικές. Στη συνέχεια θα εξετάσουμε μερικές τέτοιες προτάσεις και ιδέες.

6.1.1 Parametric Active Contours

6.1.1.1 Classic Active Contours (Snakes)

Μια πρώτη εμφάνιση των active contours έχουμε το 1987 από τους Kass, Witkin

και Terzopoulos. Ο νόμος κίνησης $\gamma \frac{\partial \vec{c}}{\partial t} = \vec{F}$ χαρακτήριζε τη δυναμική της καμπύλης.

Ισχύει γενικά ότι η F αποτελείται από δύο διανύσματα. Το διάνυσμα της δύναμης που απορρέει ανά πάσα στιγμή από την ίδια την καμπύλη και δεν έχει σχέση με την εικόνα της οποίας κάποιο σύνορο μέσα σε αυτή πρέπει να ανιχνευθεί και το διάνυσμα της δύναμης που απορρέει από την εικόνα και την έλκει προς τις ακμές

κάποιου αντικειμένου. Το πρώτο το συμβολίζουμε \vec{F}_{int} και το δεύτερο \vec{F}_{ext} .

Έτσι έχω ότι $\vec{F} = \vec{F}_{int} + \vec{F}_{ext}$

Το διάνυσμα \vec{F}_{int} αποτελείται από δύο άλλα διανύσματα δυνάμεων. Το ένα συμβολίζεται ως $\vec{F}_{elastic}$ και το άλλο ως \vec{F}_{rigid} και είναι $\vec{F}_{int} = \vec{F}_{elastic} + \vec{F}_{rigid}$.

Το $\vec{F}_{elastic}$ καθορίζεται από την ελαστικότητα της καμπύλης .Είναι δηλαδή μεγαλύτερη

όσο περισσότερο απέχουν διαδοχικά σημεία της καμπύλης και τείνει συνεχώς να μειώσει το μήκος της καμπύλης. Επίσης θεωρώντας ένα βάρος α ρυθμίζεται πόσο

έντονη θα είναι η επίδραση της στην εξέλιξη της καμπύλης : $\vec{F}_{elastic} = \alpha \vec{C}_{pp}$.

Η \vec{F}_{rigid} καθορίζεται από την ακαμψία της καμπύλης. Είναι μεγαλύτερη όσο περισσότερες «ρυτίδες» έχει η καμπύλη και τείνει συνεχώς να ομαλοποιήσει την καμπύλη και να εξαλείψει τις ανωμαλίες πάνω σε αυτή. Θεωρώντας ένα βάρος β είναι: $\vec{F}_{rigid} = \beta \vec{C}_{pp}$.

Η \vec{F}_{ext} (εξωτερική δύναμη) γράφεται στη μορφή $\vec{F}_{ext} = -grad U_{ext} = -\nabla U_{ext}$ και είναι αυτή που καθορίζει την κίνηση της καμπύλης προς τις ακμές της εικόνας. Πάμε τώρα να δούμε τη συμβαίνει με το U_{ext} μέσω του οποίου θα κατανοήσουμε καλύτερα την F_{ext} . Θεωρώ μια δισδιάστατη εικόνα $I(x,y)$. Το δυναμικό αυτό λοιπόν μπορεί να γραφεί ως εξής:

$$U_{ext} = -|\nabla I(x,y)|^2$$

$$U_{ext} = -|\nabla G_{\sigma}(x,y) * I(x,y)|^2$$

$$U_{ext} = -\frac{1}{1+|\nabla G_{\sigma}(x,y)*I(x,y)|^q} \text{ όπου } q=1 \text{ ή } q=2$$

$$U_{ext} = 1 - \frac{1}{2\pi\sqrt{\sigma}} \frac{|\nabla[G_{\sigma}(x,y)*I(x,y)]|^2}{2\sigma^2}$$

Οι παραπάνω συναρτήσεις έχουν την ιδιότητα να αναγκάζουν την καμπύλη να κινηθεί προς τις ακμές που είναι τα σημεία του ελαχίστου δυναμικού δηλαδή έχουν την ιδιότητα να ελαχιστοποιούνται στα σημεία που υπάρχουν ακμές.

Στους παραπάνω τύπους παρατηρούμε ότι εμφανίζεται η συνέλιξη $G_{\sigma}(x,y) * I(x,y)$. Αυτή βοηθά στη μείωση του θορύβου αλλά ταυτόχρονα και στη μείωση της ακρίβειας των ακμών. Τέλος άλλη μια συνάρτηση που υπάρχει για την U_{ext} θεωρώντας μια φθίνουσα θετική συνάρτηση $f(r)$ έχει τη μορφή $U_{ext} = \gamma f(\nabla I(c))$.

Συνοψίζοντας τα όσα έχουμε πει η εξίσωση της καμπύλης γράφεται στη μορφή $\frac{\partial \vec{c}}{\partial t} = \alpha \vec{C}_{pp} + \beta \vec{C}_{pp} - \nabla U_{ext}$.

Ο σκοπός είναι να βρεθεί η καμπύλη C η οποία θα οδηγεί στην κατάλληλη ισορροπία των δυνάμεων που επιδρούν στη καμπύλη. Η ισορροπία αυτή επιτυγχάνεται όταν η καμπύλη «φθάσει» στις ακμές της εικόνας. Τότε θα ισχύει: $\frac{\partial \vec{c}}{\partial t} = 0$ και $F_{int} + F_{ext} = 0$ και $\alpha \vec{C}_{pp} + \beta \vec{C}_{pp} - \nabla U_{ext} = 0$.

Η εξίσωση $\frac{\partial \vec{c}}{\partial t} = \alpha \vec{C}_{pp} + \beta \vec{C}_{pp} - \nabla U_{ext}$ μπορεί να λυθεί αριθμητικά διακριτοποιώντας την, και εκτελώντας μια επαναληπτική μέθοδο. Όταν η λύση σε αυτή την εξίσωση σταθεροποιηθεί, τότε θα ικανοποιείται και η εξίσωση Euler.

6.1.1.2 Balloons 1989

Από αυτό το σημείο ξεκινάνε οι διάφορες βελτιώσεις που μπορούν να προστεθούν στο αρχικό μοντέλο που αναφέρθηκε. Οι βελτιώσεις αυτές έγκεινται στην κατάλληλη επιλογή της μορφής των εξωτερικών δυνάμεων που θα αναγκάσουν την αρχική καμπύλη να πλησιάσει τις ακμές της εικόνας.

Ένα βασικό πρόβλημα που παρουσιάζει η κλασική μέθοδος των snakes είναι η αρχικοποίηση της καμπύλης, δηλαδή ποια θα είναι η αρχική καμπύλη η οποία θα εξελιχθεί. Υπάρχει το ενδεχόμενο αν επιλεγεί μια αρχική καμπύλη μακριά από τις ακμές της εικόνας, τότε να μην την έλκει αρκετά ώστε να κινηθεί προς αυτές. Επίσης, μια κακή αρχική καμπύλη μπορεί να έχει σαν αποτέλεσμα η καμπύλη να συρρικνωθεί σε ένα σημείο, αν δεν ασκηθούν κατάλληλες δυνάμεις. Ακόμα, εξαιτίας της διακριτοποίησης για την αριθμητική λύση του προβλήματος, μπορεί να παρουσιαστούν μεμονωμένα σημεία ως ακμές και η καμπύλη να «ξεγελαστεί» και να μείνει σε αυτά αγνοώντας τις πραγματικές (ισχυρές) ακμές της εικόνας. Για να λυθεί αυτό το πρόβλημα προτάθηκε [5] στην εξωτερική δύναμη να ασκηθεί επιπλέον μια δύναμη η οποία να ωθεί πάντα την καμπύλη προς τα έξω, σαν ένα μπαλόνι που σταδιακά φουσκώνει (από το γεγονός αυτό ονομάστηκαν τα snakes αυτά balloons). Δηλαδή η εξωτερική δύναμη να έχει τη μορφή: $\vec{F}_{ext} = k_1 \vec{n}(p) - k \frac{\nabla U_{ext}}{\|\nabla U_{ext}\|}$.

Στον τύπο αυτό $n(p)$ είναι το κάθετο διάνυσμα (με μοναδιαίο μέτρο) σε κάθε σημείο της καμπύλης.

6.1.1.3 Gradient Vector Flow 1998

Τα balloons λύνουν μεν τα προβλήματα που αναφέρονται στην προηγούμενη παράγραφο, όμως δεν μπορούν να αντιμετωπίσουν προβλήματα που προκύπτουν από τη γεωμετρία των ακμών της εικόνας, όπως είναι οι κοιλότητες.

Παραπάνω είπαμε ότι η \vec{F}_{ext} (εξωτερική δύναμη) γράφεται στη μορφή $\vec{F}_{ext} = -grad U_{ext} = -\nabla U_{ext}$. Τώρα Οι Xu και Prince θεώρησαν σαν εξωτερική δύναμη ένα γενικότερο πεδίο που έχει τη μορφή $\vec{F}_{ext} = \vec{v}(u(x, y), u(x, y))$. Αυτό το διάνυσμα ονομάστηκε Gradient Vector Flow. Με τη βοήθεια της ενέργειας θα είναι :

$$E = \iint [\mu(u_x^2 + u_y^2 + u_x^2 + u_y^2) + |\nabla f|^2 |v - \nabla f|^2] dx dy$$

Όταν το $|\nabla f|$ είναι μικρό ο δεύτερος όρος του αθροίσματος είναι μικρός και η ενέργεια κυριαρχείται από το άθροισμα των μερικών παραγώγων του πεδίου, οπότε η ελαχιστοποίηση έχει σαν αποτέλεσμα ένα αργά μεταβαλλόμενο πεδίο. Αντίθετα όταν το $|\nabla f|$ είναι μεγάλο, κυριαρχεί ο δεύτερος όρος του αθροίσματος, και η ενέργεια ελαχιστοποιείται όταν $v = |\nabla f|$.

6.1.2 Geometric Active Contours

Μέχρι το σημείο αυτό μελετήσαμε τα active contours θεωρώντας παραμετρική μορφή των καμπυλών και είπαμε ότι η εξίσωση της καμπύλης γράφεται στη μορφή

$\frac{\partial \vec{c}}{\partial t} = \alpha \vec{C}_{pp} + \beta \vec{C}_{pp} - \nabla U_{ext}$. Η μορφή αυτή δεν είναι μοναδική για την ίδια καμπύλη.

Το επιθυμητό είναι η εξέλιξη μιας καμπύλης να περιγράφεται από μια εξίσωση η οποία θα περιλαμβάνει μόνο τα γεωμετρικά της χαρακτηριστικά. Τέτοιες καμπύλες είναι οι Geometric Active Contours .Το 1993 προτάθηκε ως διαφορετική εξίσωση για

την εξέλιξη των καμπυλών η εξής : $\frac{\partial \vec{c}}{\partial t} = (fk)\vec{N}$. Η εξίσωση αυτή δίνει τα σημεία της καμπύλης καθώς πλησιάζει προς τις ακμές και εξαρτάται μόνο από τα γεωμετρικά χαρακτηριστικά και όχι από κάποια παραμετροποίηση. Ενώ στην παραμετρική περιγραφή απαιτείται να ασκούνται στην καμπύλη δυνάμεις για να διατηρηθεί η συνοχή της στη γεωμετρική περιγραφή το ρόλο αυτό ελέγχει η καμπυλότητα. Εάν προσθέσουμε έναν σταθερό όρο ταχύτητας (που το βλέπουμε στα Baloon snakes) δίνουμε μια σταθερή συνεχή ώθηση στην κίνηση της καμπύλης και μπορούμε έτσι να πάρουμε μια βελτιωμένη εκδοχή της παραπάνω εξίσωσης η οποία έχει τη μορφή : $\frac{\partial \vec{c}}{\partial t} = f(k + Vo)\vec{N}$. Αυτές οι εξισώσεις λειτουργούν αρκετά καλά σε εικόνες που έχουν δυνατό contrast.

Σε περιπτώσεις όμως που οι ακμές δεν είναι τόσο ευδιάκριτες ή υπάρχουν μικρά κενά , παρουσιάζεται το φαινόμενο της διαρροής κατά το οποίο η καμπύλη διαρρέει προς το εξωτερικό του συνόρου. Η λύση έρχεται σε αυτή τη περίπτωση με τα geodesic active contours .

6.1.2.1 Geodesic Active Contours

Θεωρώ καμπύλη την $C(p)=\{x(p),y(p)\}$ όπου το p ανήκει στο διάστημα $[0,1]$. Το μήκος της είναι το $L = \int_0^1 \|\vec{C}_p\| dp$.Εάν θεωρήσω k την καμπυλότητα της καμπύλης και N το μοναδιαίο κάθετο προς τα μέσα διάνυσμα σε αυτή σε κάθε σημείο της τότε η εξίσωση είναι εκείνη που εάν το μήκος της καμπύλης την ακολουθεί τότε ελαχιστοποιείται όσο το δυνατόν ταχύτερα. Η εξίσωση $\frac{\partial \vec{c}}{\partial t} = k\vec{N}$ είναι γεωμετρική και δεν εξαρτάται από την παραμετροποίηση της καμπύλης αλλά μόνο από την καμπυλότητα και το κάθετο διάνυσμα. Ο Grayson απέδειξε ότι μια καμπύλη που ακολουθεί αυτή την εξίσωση κίνησης(κινείται δηλαδή υπό την επίδραση της καμπυλότητας)θα συρρικνωθεί σε ένα σημείο χωρίς να δημιουργήσει κάποια ασυνέχεια. Για να χρησιμοποιήσουμε τη γεωμετρική περιγραφή των καμπυλών για την ανίχνευση ακμών, θα πρέπει να

«οδηγήσουμε» μια αρχική καμπύλη στην καμπύλη που περιγράφει τις ακμές της εικόνας, δηλαδή να ελαχιστοποιήσουμε την «απόσταση» ανάμεσα στις δύο αυτές καμπύλες.

Αντί να προσπαθήσουμε να ελαχιστοποιήσουμε το απλό μήκος της Καμπύλης $L = \int_0^1 \|\vec{C}_p\| dp$ μπορούμε να ελαχιστοποιήσουμε το γενικευμένο μήκος της καμπύλης

$L = \int_0^1 \|\vec{C}_p\| dp$ που εξαρτάται από τη συνάρτηση βάρους f .

Μπαίνει δηλαδή ο περιορισμός της συνάρτησης f στο πώς θα ελαχιστοποιηθεί το μήκος της καμπύλης. Τώρα αντί του απλού στοιχειώδους μήκους της καμπύλης dl για το οποίο ισχύει ότι : $dl = dx^2 + dy^2 = \|\vec{C}_p\| dp = \sqrt{x_p^2 + y_p^2} dp$ ορίζουμε το γενικευμένο στοιχειώδες μήκος dl_f για το οποίο ισχύει : $dl_f = \|\vec{C}_p\| f dp$ το οποίο συμπεριλαμβάνει στη συνάρτηση f τη γεωμετρική πληροφορία για τις ακμές της εικόνας.

Στις Geodesic Active Contours έχω την εξίσωση $\frac{\partial \vec{c}}{\partial t} = f(k + Vo) - [\nabla f \vec{N}] \vec{N}$. Η ποσότητα ∇f έλκει την καμπύλη προς τις ακμές της εικόνας αν αυτή διαφύγει προσωρινά στο εξωτερικό του συνόρου. Αντί για την εξίσωση αυτή μπορεί να συναντήσουμε την εξίσωση $\frac{\partial \vec{v}}{\partial t} = \mu \nabla^2 \vec{v} - f |\nabla f|^2 (\vec{v} - \nabla f)$ που είναι γενικευμένη (κοντά στις ακμές της εικόνας ισχύει ότι $\vec{v} = \nabla f$

6.2.2.2 Βασικό Πρόβλημα

Ένα βασικό πρόβλημα που εμφανίζουν τόσο η παραμετρική όσο και η γεωμετρική περιγραφή των active contours είναι το πρόβλημα της τοπολογίας. Αυτό σημαίνει ότι ξεκινώντας από κάποια αρχική καμπύλη C_0 , η τελική καμπύλη θα έχει μορφή παρόμοια με αυτή της C_0 . Για παράδειγμα δεν είναι δυνατό μία κλειστή καμπύλη να χωριστεί σε δύο κλειστές καμπύλες με φυσικό τρόπο διότι εμφανίζονται ασυνέχειες και «σπασίματα» στην εξέλιξή της. Έτσι σε μία εικόνα δεν είναι δυνατό να ανιχνευθούν ταυτόχρονα περισσότερα από ένα αντικείμενα. Το πρόβλημα αυτό έλυσαν το 1998 οι Osher και Sethian.

Η βασική ιδέα είναι η εξής. Μια οποιαδήποτε καμπύλη στο επίπεδο, μπορώ να την ορίσω σαν την τομή με ένα επίπεδο μιας τρισδιάστατης επιφάνειας. Δεδομένης μιας καμπύλης C_0 στο επίπεδο, μπορεί να κατασκευαστεί μια επιφάνεια

$\Phi: z = \Phi(x, y)$ σε 3 διαστάσεις τέτοια ώστε η τομή της με το επίπεδο $z=0$ (δηλαδή η $\Phi=0$) να είναι η καμπύλη C_0 . Αντί τώρα να αφήσουμε τη καμπύλη

C_0 να εξελιχθεί σύμφωνα με το νόμο $\frac{\partial \vec{c}}{\partial t} = \vec{F}(\vec{c})$ μπορούμε ισοδύναμα να εξελίξουμε την επιφάνεια Φ σύμφωνα με κάποιο αντίστοιχα ισοδύναμο νόμο

$\frac{\partial \vec{\Phi}}{\partial t} = \vec{G}(\vec{\Phi})$ είναι κατασκευασμένη με τέτοιο τρόπο ώστε σε κάθε φάση της

εξέλιξης η τομή $\Phi=0$ να δίνει την αντίστοιχη καμπύλη C , όπως αυτή θα εξελισσόταν σύμφωνα με την πρώτη εξίσωση.

Με αυτόν τον τρόπο μπορούμε να παρακολουθούμε μόνο την εξέλιξη της τρισδιάστατης επιφάνειας και ανά πάσα στιγμή να βρίσκουμε την επιθυμητή δισδιάστατη καμπύλη παίρνοντας απλώς $\Phi=0$ από τη λύση. Δηλαδή εκφράζουμε την καμπύλη C σαν τη zero level set function της επιφάνειας Φ .

Χρησιμοποιώντας την τεχνική αυτή οι τοπολογικές αλλαγές στην δισδιάστατη καμπύλη χειρίζονται αυτόματα, διότι αν και οι αλλαγές αυτές μπορεί να είναι απότομες και ασυνεχείς σε δύο διαστάσεις, σε 3 διαστάσεις οι αλλαγές εμφανίζουν ομαλή συμπεριφορά.

6.2.2.3 Geodesic Active Contours

Η θεωρία των level sets εμφανίστηκε χρονικά λίγο αφού διατυπώθηκαν τα πρώτα snakes και πριν τις Geometric active contours. Ο συνδυασμός των level sets και των geodesic active contours δημιούργησε ένα πολύ ισχυρό εργαλείο για την κατάτμηση των εικόνων. Η βασική αρχή βέβαια είναι η ίδια: ξεκινώντας από κάποια αρχική καμπύλη, την παραμορφώνουμε κατάλληλα ώστε να καταλήξει προς τις ακμές μιας εικόνας. Θα δούμε τώρα πώς συνδυάζονται αυτές οι δύο τεχνικές.

Χρησιμοποιώντας τις εξισώσεις που δείξαμε παραπάνω στα geometric active contours μπορούμε να γράψουμε τις αντίστοιχες μορφές τους για τα level sets. Για παράδειγμα η εξίσωση $\frac{\partial \vec{c}}{\partial t} = f(k + Vo)\vec{N}$ αντιστοιχεί στην

$\frac{\partial \vec{c}}{\partial t} = f(k + Vo)\|\nabla\Phi\|$ με την καμπυλότητα να δίνεται από τον τύπο

$$\kappa = -\text{div} \frac{\nabla\Phi}{|\nabla\Phi|} .$$

Η εξίσωση $\frac{\partial \vec{c}}{\partial t} = f(k + Vo) - [\nabla f \vec{N}] \vec{N}$ έχει στη level set έκδοση τη μορφή

$$\frac{\partial \Phi}{\partial t} = f(k + Vo)|\nabla\Phi| - \nabla f \nabla\Phi .$$

Κεφάλαιο 7

7.1 Ο αλγόριθμος κατάτμησης εικόνας που χρησιμοποιήθηκε

Για την εκτέλεση του αλγορίθμου σε πραγματικά δεδομένα, χρησιμοποιήθηκε ένας αλγόριθμος παρόμοιος με τον chameleon, βασισμένος στην ίδια ακριβώς φιλοσοφία. Στον αλγόριθμο αυτό έχουμε τα εξής:

➤ Η πρώτη κατάτμηση της αρχικής περιοχής γίνεται με κριτήριο την κοντινότερη απόσταση και η παράμετρος του ελάχιστου εμβαδού χρησιμοποιείται όπως στον chameleon για τον χωρισμό κάθε περιοχής C_i σε C_{iA} και C_{iB} .

➤ Υπολογίζουμε όλες τις αποστάσεις (χρωματικές) του κάθε pixel από όλα τα

υπόλοιπα και ορίζουμε μια σχέση μεταξύ τους (similarity). Το μέγιστο αυτού του βαθμού ομοιότητας μεταξύ δύο σημείων που ανήκουν στον χώρο V ορίζεται ως ο βαθμός εσωτερικής ομοιότητας του χώρου $Int(V)$ αντί της απόλυτης interconnectivity που χρησιμοποιούσε ο chameleon.

➤ Η διαφορά μεταξύ δύο περιοχών C_i και C_j ($dif(C_i, C_j)$) μπορεί να οριστεί ως

εξής: θεωρούμε όλες τις ακμές που ενώνουν τα C_i και C_j και όλα τα αντίστοιχα βάρη αυτών $w(k)$. Η διαφορά $dif(C_i, C_j)$

- (α) ορίζεται ως το άθροισμα των βαρών $w(k)$
- (β) θεωρείται ότι είναι το ελάχιστο των $w(k)$
- Το κριτήριο που αποφασίζει για την συγχώνευση δύο περιοχών είναι

όπως στον αλγόριθμο chameleon, με την διαφορά ότι οι δύο περιοχές συγχωνεύονται αν $dif(C_i, C_j) > \max(Int(C_i), Int(C_j))$, δηλαδή αν η διαφορά μεταξύ δύο περιοχών – κόμβων είναι μεγαλύτερη από το μέγιστο της εσωτερικής απόστασης των δύο περιοχών.

7.2 Εφαρμογή της μεθόδου

Ο αλγόριθμος για την ομαδοποίηση δεδομένων chameleon βρίσκει εφαρμογή και στο πρόβλημα κατάτμησης εικόνας, όπως και όλοι οι αντίστοιχοι αλγόριθμοι clustering. Σε αυτό το δύσκολο πρόβλημα εφαρμόσαμε και τον αλγόριθμο που περιγράφηκε ανωτέρω.

Συγκεκριμένα, ορίσαμε καταρχήν το μέτρο της απόστασης-ομοιότητας μεταξύ δύο περιοχών με τους εξής δύο τρόπους:

1. Εάν δεν μας ενδιαφέρει η χωρική απόσταση μεταξύ δύο περιοχών, αλλά μόνο

η χρωματική, γεγονός που ισχύει στις συντριπτικά περισσότερες περιπτώσεις εφαρμογών που απαιτούν κατάτμηση εικόνας, τότε η απόσταση μεταξύ δύο περιοχών C_i, C_j την $D(C_i, C_j) = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i + B_j)^2}$, όπου R_i, G_i, B_i είναι το χρωματικό περιεχόμενο της τυχαίας περιοχής C_i .

2. Εάν θέλουμε να λάβουμε υπόψη μας και την χωρική απόσταση μεταξύ των δύο περιοχών, μπορούμε στην παραπάνω απόσταση να προσθέσουμε και μια ποσότητα που εκφράζει αυτήν ακριβώς την απόσταση, όπως π.χ. την απόσταση των κέντρων βάρους των περιοχών ή την μικρότερη απόσταση μεταξύ αυτών.

Για τους λόγους της μεγάλης εφαρμοσιμότητας που προαναφέρθηκαν επιλέξαμε τον 1ο τρόπο.

Χρησιμοποιήσαμε σαν είσοδο στο πρόγραμμα δύο παραμέτρους και συγκεκριμένα την απόσταση όπως ορίστηκε ανωτέρω, αλλά και το ελάχιστο επιτρεπτό πλήθος στοιχείων που μπορεί να έχει η κάθε περιοχή C_i . Ο αλγόριθμος εφαρμόστηκε σε εικόνες (grayscale) αρχαίων ελληνικών επιγραφών, των οποίων η κατάτμηση είναι αρκετά δύσκολη.

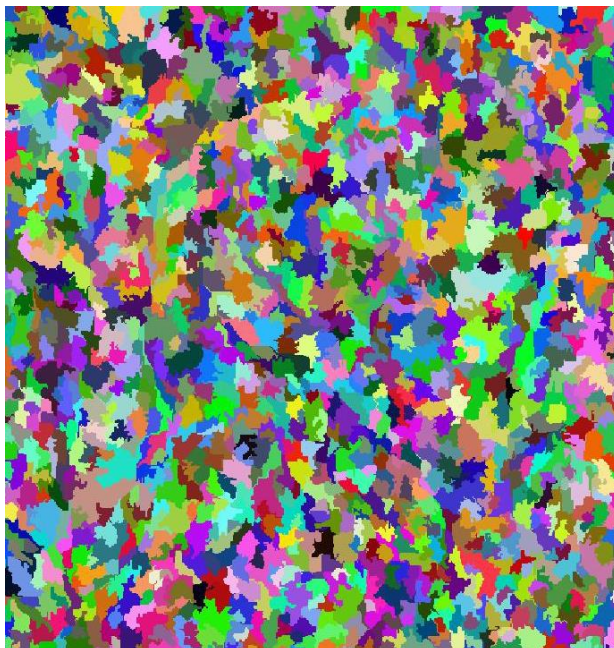
Για την πρώτη παράμετρο, το ελάχιστο επιτρεπτό εμβαδό δηλαδή, δόθηκαν οι τιμές 50 και 100 με τα αποτελέσματα της αρχικής κατάτμησης να παρουσιάζονται στις παρακάτω εικόνες.

Στη συνέχεια για κάθε μία από τις τιμές αυτές του ελάχιστου εμβαδού αποδόθηκαν τιμές για το κατώφλι απόστασης, που ήταν 5 και 10, για την περίπτωση του εμβαδού 50, ενώ για την αντίστοιχη του εμβαδού 100 προτιμήθηκαν οι τιμές 10 και 20 για την απόσταση, ώστε να υπάρχει μεγαλύτερη ανοχή, καθώς είχαμε μεγαλύτερες περιοχές. Τα αποτελέσματα των δοκιμών αυτών φαίνονται στις παρακάτω εικόνες.

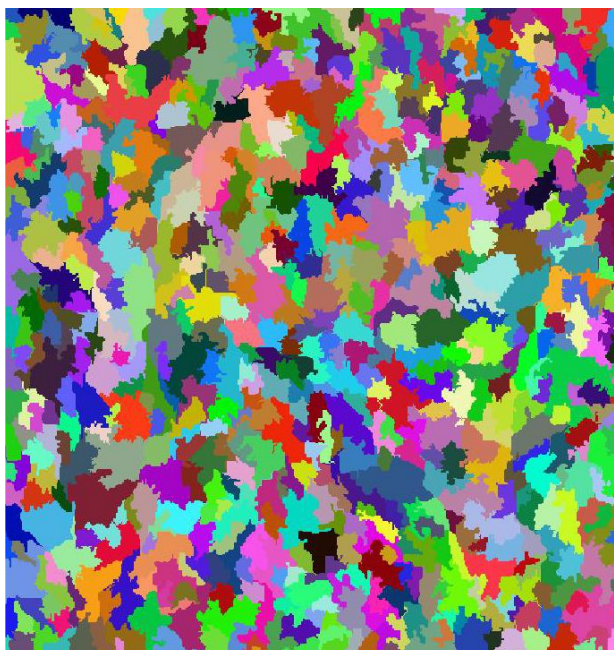


Εικόνα 7.2.α : Αρχική εικόνα την οποία θέλουμε να κατατμήσουμε σε όμοιες περιοχές

Φάση 1 : Καθορισμός αρχικών περιοχών

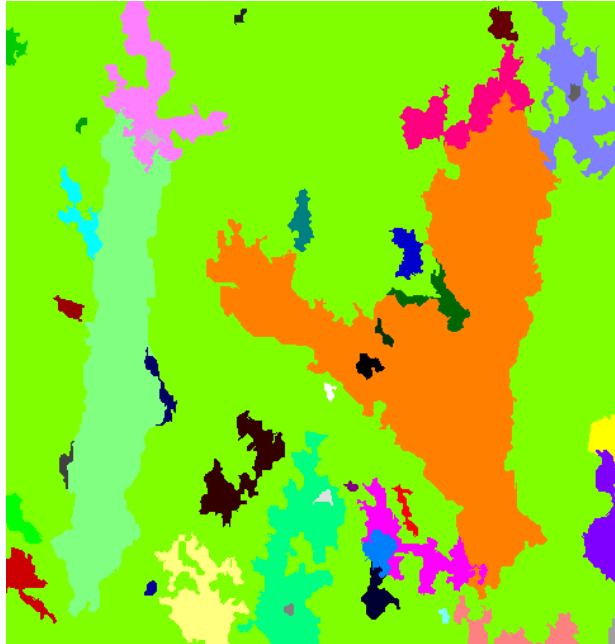


Εικόνα 7.2.β : Κατατμημένη εικόνα της αρχικής με ελάχιστο εμβαδόν 50 pixels

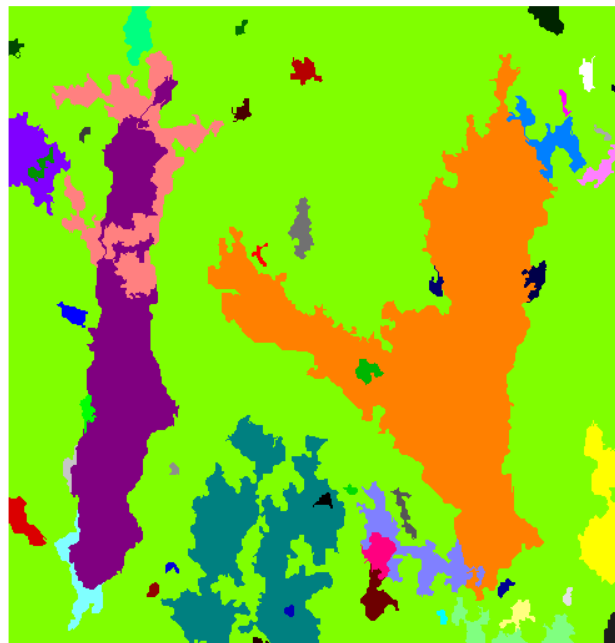


Εικόνα 7.2.γ : Κατατμημένη εικόνα της αρχικής με ελάχιστο εμβαδόν 100 pixels

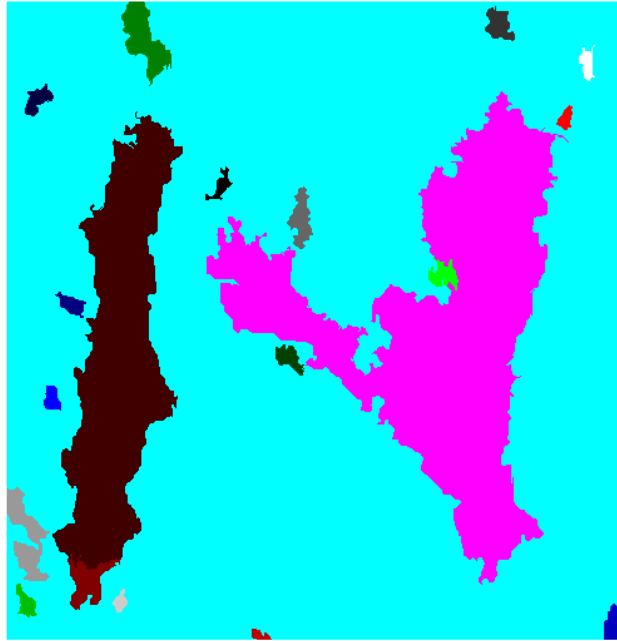
Φάση 2 : Συγχώνευση των περιοχών



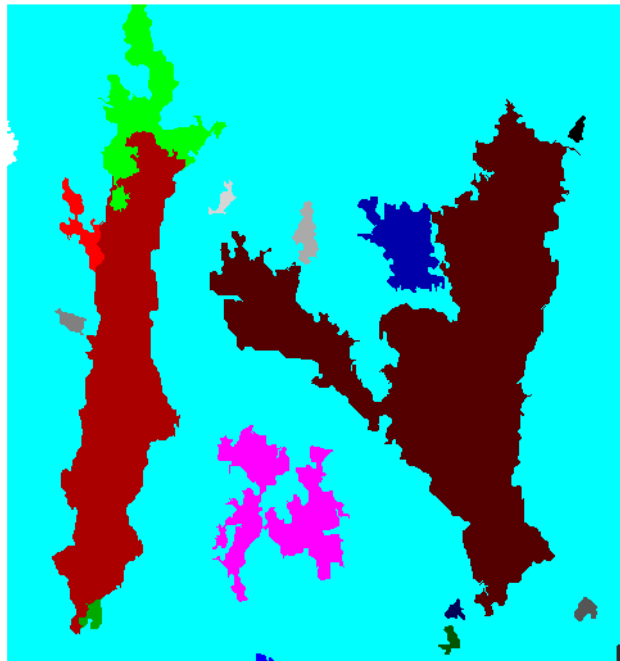
Εικόνα 7.2.δ : Συγχώνευση των περιοχών που προέκυψαν από την 1η Φάση με ελάχιστο εμβαδόν 50 και κατώφλι ομοιότητας 5.



Εικόνα 7.2.ε : Συγχώνευση των περιοχών που προέκυψαν από την 1η Φάση με ελάχιστο εμβαδόν 50 και κατώφλι ομοιότητας 10.



Εικόνα 7.2.ζ : Συγχώνευση των περιοχών που προέκυψαν από την 1η Φάση με ελάχιστο εμβαδόν 100 και κατώφλι ομοιότητας 10.



Εικόνα 7.2.η : Συγχώνευση των περιοχών που προέκυψαν από την 1η Φάση με ελάχιστο εμβαδόν 100 και κατώφλι ομοιότητας 20.

Κεφάλαιο 8

Συμπεράσματα

Παρότι η όλη διαδικασία είμαι αυστηρή , ο αλγόριθμος μας στη κατάτμηση εικόνων φαίνεται να λειτουργεί ικανοποιητικά ειδικά στην περίπτωση της εξαγωγής των γραμμάτων των επιγραφών από την αρχική φωτογραφία. Η κατάτμηση της εικόνας πρέπει να γίνεται με όσο το δυνατό μεγαλύτερη ακρίβεια, ενώ παράλληλα πρέπει να είναι όσο το δυνατό πιο αυτοματοποιημένη περιορίζοντας την ανθρώπινη παρέμβαση στο ελάχιστο.

Παρατηρήθηκε ότι για τις περισσότερες εικόνες οι βέλτιστες παράμετροι ήταν 100 pixels για το ελάχιστο επιτρεπτό εμβαδό κατά την πρώτη φάση αρχικού διαχωρισμού της εικόνας σε υποπεριοχές και 10 μονάδες χρωματική απόσταση για το κατώφλι σε ότι αφορά το κριτήριο συγχώνευσης της δεύτερης φάσης.

Βιβλιογραφία:

- ❖ Sethian, J.A., Level Set methods, Cambridge University Press 1996
- ❖ Sethian, J.A., Curvature and the evolution of fronts, *Comm. In Math. Phys.*, 101, pp. 487-499, 1985
- ❖ M.Kass, A.Witkin, D.Terzopoulos, Snakes: Active Contour Models, *International Journal of Computer Vision*, 321-331, 1988
- ❖ V.Caselles, R.Kimmel, G.Sapiro, Geodesic Active Contours, *International Journal of Computer Vision* 22(1), 61-79, 1997
- ❖ Cohen, L.D., On active contour models and balloons, In *computer vision, graphics and image processing: Image understanding*, 53(2):211-218, 1991
- ❖ L.D. Cohen, I. Cohen, Finite elements methods for active contour models and balloons for 2D and 3D images, *IEEE Transactions on pattern analysis and machine intelligence*, PAMI-15, 1993
- ❖ B.Leroy, I.Herlin, and L.D.Cohen, Multi resolution algorithms for active contour models, in *12 Int. Conf. Analysis and optimization of systems*
- ❖ C.Xu, J.L.Prince, Snakes, Shapes and Gradient Vector Flow, *IEEE transactions on image processing*, Vol. 7, No. 3, 1998
- ❖ R.Courant, D.Hilbert, *Methods of Mathematical Physics*, vol. 1, New York: Interscience, 1953
- ❖ C.Xu, J.L.Prince, Generalized gradient vector flow external forces for active contours, *Signal Processing* 71, (1998) 131-139
- ❖ V.Caselles, F.Catte, F.Dibos, A geometric model for active contours, *Numerische Mathematik*, 66:1-31, 1993
- ❖ R.Malladi, J.A.Sethian, B.C.Vemuri, Shape modeling with front propagation: A level set approach. *IEEE T. Patt. Anal. Mach. Intell.* 17(2):158-175, 1995
- ❖ A.Yezzi, S.Kichenassamy, A.Kumar, P.Olver, A.Tannenbaum A geometric snake model for segmentation of medical imagery. *IEEE T.Med. Imag.*, 16:199-209, 1997
- ❖ N.Paragios, O.Mellina-Gottardo, V.Ramesh, Gradient vector flow fast geodesic active contours, 2001
- ❖ S. Osher, J.A.Sethian, Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations, *Journal of Computational Physics*, 79, pp.12-49, 1998
- ❖ J.A.Sethian, Hypersurfaces moving with curvature dependent speed: hamilton- jacobi equations, conservation laws and numerical algorithms, *Journal of differential geometry*, 31, pp. 131-161, 1989

