



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ & ΡΟΜΠΟΤΙΚΗΣ

Παρακολούθηση Χειρονομιών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μιχαήλ Γ. Ράπτη

Επιβλέπων: Πέτρος Α. Μαραγκός
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2006



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΗΜΑΤΩΝ, ΕΛΕΓΧΟΥ &
ΡΟΜΠΟΤΙΚΗΣ

Παρακολούθηση Χειρονομιών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Μιχαήλ Γ. Ράπτη

Επιβλέπων: Πέτρος Α. Μαραγκός
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 17^η Ιουλίου 2006.

.....
Πέτρος Μαραγκός
Καθηγητής Ε.Μ.Π

.....
Νικόλαος Μαράτος
Καθηγητής Ε.Μ.Π

.....
Κων/νος Τζαφέστας
Λέκτορας Ε.Μ.Π

Αθήνα, Ιούλιος 2006

.....

Μιχαήλ Γ. Ράπτη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογίστων

© (2006) Εθνικό Μετσόβιο Πολυτεχνείο. 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.

Περίληψη

Η παρούσα διπλωματική εργασία επικεντρώνεται στην αυτόματη παρακολούθηση των χειρονομιών του χεριού σε μια ακολουθία εικόνων, με τελικό σκοπό την επίτευξη επικοινωνίας μεταξύ ανθρώπου και υπολογιστή μέσω χειρονομιών. Μετά από μια σύντομη ανασκόπηση των υπαρχόντων εφαρμογών για παρακολούθηση του χεριού, εξετάζεται αναλυτικά το μοντέλο των ενεργών καμπύλων. Συγκεκριμένα, περιγράφεται το μαθηματικό υπόβαθρο του, η διαδικασία διακριτοποίησης του και ορίζεται το πεδίο εξωτερικών δυνάμεων Gradient Vector Flow. Αυτό συγκρίνεται με το κλασσικό μοντέλο φιδιών, ενώ περιγράφεται και μεθοδολογία εύρεσης και απαλοιφής των μη “σημαντικών” ακμών από την εικόνα. Το προκαταρτικό μέρος της εργασίας συμπληρώνεται με την παρουσίαση του τρόπου δημιουργίας στατιστικού μοντέλου σχήματος, την κατασκευή του συνόλου εκπαίδευσης και την στατιστική ανάλυση των δεδομένων. Οι μεθοδολογίες αυτές συνδυάζονται με σκοπό την δημιουργία εφαρμογής προσανατολισμένης στο προσδιορισμό και την παρακολούθηση του περιγράμματος του χεριού. Το διάνυσμα θέσης της ενεργής καμπύλης προβάλλεται στο χώρο σχήματος, ο οποίος έχει προκύψει από την στατιστική ανάλυση, με αποτέλεσμα να περιορίζεται η καμπύλη σε σχήματα παραπλήσια με αυτά που εμφανίζονται στο σύνολο εκπαίδευσης. Με τον τρόπο αυτό συνδυάζουμε δυο ευρέως διαδεδομένα μοντέλα στην 'Οραση Υπολογιστών και τα πειραματικά αποτελέσματα αποδεικνύουν ότι ο συνδυασμός αυτός ισχυροποιεί σημαντικά τη διαδικασία παρακολούθησης.

Λέξεις–Κλειδιά

όραση υπολογιστών, επεξεργασία εικόνας, παρακολούθηση χειρονομιών, ενεργές καμπύλες, φίδια, στατιστικά μοντέλα σχήματος, χώρος σχήματος, πρωτεύουσες συνιστώσες, πεδίο δυνάμεων, περίγραμμα χεριού, ανοιχτό φίδι, ορόσημα.

Abstract

This thesis focuses on a number of issues in developing techniques and algorithms applied in automating the visual tracking of hand in a video sequence. Firstly, some of the existing applications for hand tracking are briefly considered. Afterwards, we present a detailed analysis of the active contour model, commonly called “snakes”, emphasizing in its mathematical background and its discrete model. Then, an external force for active contours called gradient vector flow (GVF), which solves some of the problems of the initial snake model is examined. Methods for edge detection are also analyzed combined with a proposal for simplification of the edge map. Additionally, we describe the method for building models, by learning patterns of variability from a training set of correctly annotated images. These statistical models of shape are combined with the active contour model in order to create an application orientated towards hand tracking. The hand boundaries in a new image are found by placing an initial snake and allowing deforming only according to the examined shape variations. In this way we combine two basic techniques used in Computer Vision, succeeding to make tracking more accurate and robust.

Keywords

computer vision, image processing, hand tracking, active contours, snakes, statistical models of shape, shape space, principal components, gradient vector flow, hand boundaries, open active curve, landmarks.

Ευχαριστίες

Κατ' αρχήν θα ήθελα να ευχαριστήσω το Καθηγητή του Ε.Μ.Π. Πέτρο Μαραγκό, του οποίου οι διαλέξεις στα πλαίσια των προπτυχιακών μαθημάτων 'Οραση Υπολογιστών και Αναγνώριση Προτύπων έπαιξαν σημαντικό ρόλο για την επιλογή μου να εργαστώ περαιτέρω στην ερευνητική περιοχή της 'Ορασης Υπολογιστών. Επιπλέον, θα ήθελα να τον ευχαριστήσω ιδιαιτέρα για τις κατευθύνσεις που μου έδωσε για την εκπόνηση της παρούσας διπλωματικής εργασίας.

Θα ήθελα επίσης να ευχαριστήσω τους συμφοιτητές μου Κώστα Πελεχρίνη, Χρήστο Μπρεργελέ και Παναγιώτη Παπαδημητρίου για τις πολύτιμες συμβουλές τους κατά την συγγραφή της διπλωματικής.

Ακόμα θέλω να ευχαριστήσω τους υποψηφίους διδάκτορες του εργαστηρίου 'Ορασης Υπολογιστών για τις χρήσιμες συμβουλές τους στα διαφορά θέματα που με απασχόλησαν κατά την διάρκεια της έρευνας για την παρούσα εργασία.

Τέλος, θα ήθελα να αφιερώσω την εργασία αυτή στην οικογένεια μου, που τόσο χρόνια μου έχει προσφέρει με οποιοδήποτε τρόπο βοήθεια και υποστήριξη.

Περιεχόμενα

1 Εισαγώγη	19
1.1 'Οραση Υπολογιστών και Σχετικά Επιστημονικά Πεδία	19
1.2 Κίνητρο της Διπλωματικής Εργασίας	22
1.2.1 Περιορισμοί Υπαρχόντων Συστημάτων	22
1.2.2 Περιεχόμενο παρούσας εργασίας	24
2 Παρόμοιες Εργασίες	25
2.1 Παρακολούθηση χεριού	25
2.1.1 Δυσδιάστατη παρακολούθηση χεριού	26
2.1.2 Παρακολούθηση χεριού με βάση μοντέλο	28
2.1.3 Αναγνώριση με βάση την εκτίμηση της Παρατήρησης . .	31
2.1.4 Εκτίμηση στάσης με μία παρατήρηση	34
3 Ενεργές Καμπύλες	35
3.1 Βασικές Έννοιες	36
3.2 Το δυναμικό μοντέλο των ενεργών καμπύλων	39
3.3 Διαχριτοποίηση του Μοντέλου	43
3.4 Παράμετροι της Ενεργής Καμπύλης	49
3.5 Ενεργές Καμπύλες: Πλεονεκτήματα και Μειονεκτήματα . . .	51
3.5.1 Πλεονεκτήματα	51
3.5.2 Μειονεκτήματα	53
4 Πεδίο Εξωτερικών δυνάμεων	55
4.1 Κλασσικό Μοντέλο Ενεργών Καμπύλων	56
4.1.1 Εξωτερικές Δυνάμεις	58
4.2 Διανυσματικό Πεδίο Ροής Κλίσης	60
4.2.1 Διαχριτοποίηση των Εξισώσεων και Εφαρμογή	61
4.2.2 Τροποποίηση του μοντέλου Ενεργών Καμπυλών	63

4.3	Εξέλιξης της Καμπύλης σύμφωνα με το Πεδίο Ροής Κλίσης	63
4.4	Χάρτης ακμών	66
4.4.1	Απλοποίηση του Χάρτη Ακμών	67
5	Στατιστικό Μοντέλο Σχήματος	71
5.1	Εισαγωγή	71
5.2	Στατιστική Ανάλυση Σχήματος	73
5.2.1	Τοποθέτηση των Ορόσημων	74
5.2.2	Ευθυγράμμιση Σχήματος	75
5.2.3	Ευθυγράμμιση Δυο Σχημάτων	76
5.2.4	Γενικευμένη Ανάλυση Προκρούστη	77
5.2.5	Στατιστική Ανάλυση των Ευθυγραμμισμένων Σχημάτων	78
5.2.6	Επιλογή των Πρωτευουσών Συνιστώσαν	82
6	Συνδυασμός Μοντέλων Σχήματος & Ενεργών Καμπυλών	85
6.1	Εισαγωγή	85
6.2	Προβολή στο Επιτρεπόμενο Χώρο Σχήματος	86
6.2.1	Εναλλακτικός Αλγόριθμος Προβολής Ενεργής Καμπύλης	88
6.3	Αποτέλεσμα Χρήσης Στατιστικού Μοντέλου	90
7	Αλγόριθμος και Πειραματική Διαδικασία	95
7.1	Αλγόριθμος	95
7.1.1	Στάδιο Εκπαίδευσης	95
7.1.2	Στάδιο Εφαρμογής	97
7.2	Πειραματική Διαδικασία	98
7.2.1	Πείραμα 1°	98
7.2.2	Πείραμα 2°	101
7.2.3	Πείραμα 3°	103
8	Συμπεράσματα και Κατευθύνσεις για Μελλοντική Έρευνα	107
8.1	Συμπεράσματα	107
8.1.1	Πλεονεκτήματα	107
8.1.2	Μειονεκτήματα και Περιορισμοί	108
8.2	Συνδρομή Εργασίας	109
8.3	Κατευθύνσεις για Μελλοντική Έρευνά	109
A	Δημιουργία Ανοιχτής Ενεργής Καμπύλης	111
A.1	Εισαγωγή Ασυνέχειας Θέσης	111

B Ευθυγράμμιση Δύο Σχημάτων

113

Κατάλογος σχημάτων

3.1	Μια εικόνα ως 3-Δ δυναμική επιφάνεια	37
3.1.1	37
3.1.2	37
3.2	Φιλτράρισμα εικόνας για τη παραγωγή κατάλληλης επιφάνειας .	38
3.2.1	38
3.2.2	38
4.1	Παράδειγμα κλασσικού μοντέλου ενεργής καμπύλης	57
4.1.1	$I(x, y)$	57
4.1.2	$- \nabla I(x, y) $	57
4.1.3	Λεπτομεριακή αναπαράσταση του Αντίχειρα	57
4.2	Παράδειγμα εξέλιξης του κλασσικού μοντέλου ενεργής καμπύλης	58
4.2.1	Αρχική Θέση	58
4.2.2	Εξέλιξη Καμπύλης	58
4.2.3	Τελική Θέση	58
4.3	Χρήση Gaussian φίλτρου για την μεταβολή του πεδίου δυνάμεων	59
4.3.1	Πεδίο δυνάμεων	59
4.3.2	Λεπτομέρεια Συνόρων	59
4.3.3	Τελική Θέση Καμπύλης	59
4.4	Εικόνα του πεδίου GVF	64
4.4.1	$f(x, y)$	64
4.4.2	GVF	64
4.4.3	Λεπτομεριακή αναπαράσταση του Αντίχειρα	64
4.5	Παράδειγμα εξέλιξης του GVF φιδιού	65
4.5.1	Αρχική Θέση	65
4.5.2	Εξέλιξη Καμπύλης	65
4.5.3	Τελική Θέση	65
4.6	Εικόνα μάσκας 8-συνεκτικότητας	68

4.7	Παράδειγμα Απλοποίησης της εικόνας ακμών	69
4.7.1	Αρχική Εικόνα Ακμών	69
4.7.2	Απλοποιημένη Εικόνα Ακμών	69
5.1	Παράδειγμα ορισμού Σχήματος	74
5.2	Ευθυγράμμιση του Συνόλου Εκπαίδευσης	79
5.2.1	Σχήματα Χειριών	79
5.2.2	Ευθυγραμμισμένα Σχήματα	79
5.3	Μέσο Σχήμα και παρουσίαση της Διακύμανσης κάθε σημείου .	81
5.3.1	Μέσο Σχήμα	81
5.3.2	Διακύμανση κάθε σημείου	81
5.4	Παράδειγμα παραμορφώσεων του μοντέλου σχήματος	83
5.4.1	$b_1 = +2\sqrt{\lambda_1}$	83
5.4.2	$b = 0$	83
5.4.3	$b_1 = -1\sqrt{\lambda_1}$	83
5.4.4	$b_2 = +2\sqrt{\lambda_2}$	83
5.4.5	$b = 0$	83
5.4.6	$b_2 = -1\sqrt{\lambda_2}$	83
6.1	Παράδειγμα Χρήσης Στατιστικού Μοντέλου	90
6.1.1	Φίδι χωρίς περιορισμούς	90
6.1.2	Χρήση Στατιστικού Μοντέλου	90
6.2	Απεικόνιση αποτελέσματος χρήσης Στατιστικού Μοντέλου .	91
6.2.1	Εικόνα Ενεργής Καμπύλης	91
6.2.2	Εικόνα Περιορισμένης Ενεργής Καμπύλης	91
6.3	Δεύτερο Παράδειγμα Χρήσης Περιορισμών	92
6.3.1	Φίδι χωρίς Περιορισμούς	92
6.3.2	Χρήση Στατιστικού Μοντέλου	92
6.3.3	Εικόνα Ενεργής Καμπύλης	92
6.3.4	Εικόνα Περιορισμένης Ενεργής Καμπύλης	92
7.1	Διάγραμμα Ροής Σταδίου Εκπαίδευσης	96
7.2	Διάγραμμα Ροής Σταδίου Εφαρμογής	97
7.3	Απεικόνιση Αποτελεσμάτων 1 ^{ου} πειράματος	99
7.3.1	99
7.3.2	99
7.3.3	99
7.3.4	99

7.3.5	99
7.3.6	99
7.4 Απεικόνιση Αδυναμίας Παρακολούθησης	100
7.4.1	100
7.4.2	100
7.4.3	100
7.4.4	100
7.5 Απεικόνιση Αποτελεσμάτων 2 ^{ου} πειράματος	101
7.5.1	101
7.5.2	101
7.5.3	101
7.5.4	101
7.6 Απεικόνιση Αποτελεσμάτων 2 ^{ου} πειράματος	102
7.6.1	102
7.6.2	102
7.6.3	102
7.6.4	102
7.6.5	102
7.6.6	102
7.6.7	102
7.6.8	102
7.7 Απεικόνιση Αποτελεσμάτων 3 ^{ου} πειράματος	103
7.7.1	103
7.7.2	103
7.7.3	103
7.7.4	103
7.8 Απεικόνιση Αποτελεσμάτων 3 ^{ου} πειράματος	104
7.8.1	104
7.8.2	104
7.8.3	104
7.8.4	104
7.8.5	104
7.8.6	104
7.8.7	104
7.8.8	104

Κεφάλαιο 1

Εισαγώγη

1.1 Όραση Υπολογιστών και Σχετικά Επιστημονικά Πεδία

Η οπτική αντίληψη διαδραματίζει σπουδαίο ρόλο στην ζωή των ανθρώπων αλλά και των περισσοτέρων ζώντων οργανισμών, επιτρέποντας την αναπαράσταση του περιβάλλοντος τους με τέτοιο τρόπο, ώστε να είναι δυνατή η πραγματοποίηση σύνθετων εργασιών όπως η μετακίνηση, η παρακολούθηση, ο χειρισμός και αναγνώριση αντικειμένων. Η εμφάνιση τέτοιων δυνατοτήτων σε μηχανικά συστήματα συνάδει με την ικανότητα αίσθησης, επεξεργασίας, επικοινωνίας, ελέγχου και κίνησης. Ενώ στα βιολογικά συστήματα οι ικανότητες αυτές είναι πλήρως συνδυασμένες στα μηχανικά είναι σχεδόν ανεξάρτητες. Τις τελευταίες η μεγάλη τεχνολογική πρόοδος μας έχει δώσει την δυνατότητα να γνωρίζουμε πως να φτιάξουμε αισθητήρες, πως να φτιάξουμε υπολογιστές, πως να φτιάξουμε κινούμενα robot. Παρόλα αυτά δεν είμαστε ακόμα σε θέση να δημιουργήσουμε ένα μηχανικό σύστημα που να αντιλαμβάνεται τον ορατό κόσμο με ανθρώπινη ευφυΐα, αυτός είναι ο απώτερος στόχος της επιστημονικής περιοχής που ονομάζεται 'Όραση Υπολογιστών.

Η 'Όραση Υπολογιστών [15] άρχισε ως επιστημονικό πεδίο την δεκαετία του 1960 από προσπάθειες ερευνητών της Τεχνητής Νοημοσύνης να δημιουργήσουν υπολογιστές που βλέπουν. Δημιουργήθηκε και μετέπειτα ωρίμασε κατά τις δυο πρώτες δεκαετίες από την συμβολή τριών κυρίως επιστημονικών πεδίων:

- Τεχνητή Νοημοσύνη
- Επεξεργασία Σημάτων

• Αναγνώριση Προτύπων

Η Όραση Υπολογιστών μπορεί να θεωρηθεί ως παρακλάδι της Τεχνητής Νοημοσύνης με τα δεδομένα εικόνας να παίζουν το ρόλο δεδομένων εισόδου για τον έλεγχο της συμπεριφοράς του συστήματος. Μερικές από τις μεθόδους οι οποίες χρησιμοποιούνται στην Όραση Υπολογιστών είναι βασισμένες στις μεθόδους εκμάθησης που αναπτυχθήκαν στο τομέα της Τεχνητής Νοημοσύνης.

Καθώς η κάμερα μπορεί να θεωρηθεί ως ένας αισθητήρας φωτός, υπάρχουν διάφορες μέθοδοι στην Όραση Υπολογιστών που αναλύουν φυσικά φαινόμενα που σχετίζονται με το φως κάνοντας χρήση των εικόνων των φαινομένων αυτών. Για παράδειγμα, είναι δυνατό να εξάγουμε πληροφορίες για την κίνηση μέσα σε υγρά αλλά και για την κίνηση χυμάτων με κατάλληλη ανάλυση των εικόνων που έχουν προέλθει από τα φαινόμενα αυτά. Ακόμα ένας κλάδος της Όρασης Υπολογιστών ασχολείται λεπτομερώς με την σχέση μεταξύ της προσπίτουσας ακτινοβολίας στην επιφάνεια των αντικειμένων, που προέρχεται από την φωτεινή ακτινοβολία της πηγής, και της εκπεμπόμενης φωτεινής ακτινοβολίας από την επιφάνεια των αντικειμένων, λαμβάνοντας υπόψη την σχετική γεωμετρία μεταξύ πηγής, επιφανείας και κάμερας καθώς και τις οπτικές ιδιότητες των επιφανειών. Γίνεται, λοιπόν, φανερό, ότι η Όραση υπολογιστών μπορεί να θεωρηθεί ως μια προέκταση της Φυσικής, και πιο συγκεκριμένα της οπτικής φυσικής.

Ένα τρίτο επιστημονικό πεδίο το οποίο διαδραματίζει σημαντικό ρόλο είναι η νευροβιολογία, πιο συγκεκριμένα η μελέτη των βιολογικών συστημάτων όρασης. Κατά την διάρκεια του τελευταίου αιώνα, έχουν γίνει εκτεταμένες έρευνες για το τρόπο λειτουργίας των ματιών, του νευρικού συστήματος και των εγκεφαλικών δομών που είναι υπεύθυνα για την λήψη και επεξεργασία του οπτικού σήματος τόσο στους ανθρώπους όσο και σε διάφορα ζώα. Η έρευνες αυτές έχουν οδηγήσει σε μια λεπτομερή γνώση του πως τα βιολογικά συστήματα όρασης λειτουργούν, με σκοπό να επιλύσουν συγκεκριμένα προβλήματα. Με βάση τα αποτελέσματα αυτά γίνεται προσπάθεια δημιουργίας τεχνητών συστημάτων, τα οποία θα μιμούνται το τρόπο επεξεργασίας και τη συμπεριφορά των βιολογικών συστημάτων, σε διάφορα επίπεδα πολυπλοκότητας.

Ακόμα, όπως αναφέρθηκε και παραπάνω, η όραση υπολογιστών είναι στενά συνδεδεμένη με τη επεξεργασία σημάτων. Πολλές μέθοδοι για επεξεργασία σημάτων μιας μεταβλητής μπορούν να επεκταθούν για επεξεργασία σημάτων δυο ή περισσοτέρων μεταβλητών στην όραση υπολογιστών. Παρόλα αυτά, λόγω της συγκεκριμένης φύσης των εικόνων έχουν αναπτυχθεί αρκετές μέθοδοι, οι οποίες δεν έχουν άμεση αντιστοιχία με μεθόδους της επεξεργασίας σημάτων.

Ο διακριτός χαρακτήρας των μεθόδων αυτών, που πηγάζει από το γεγονός ότι είναι μη γραμμικές και την πολυδιάστατη μορφή των σημάτων, ορίζει ένα υποπεδίο της επεξεργασίας σημάτων ως τμήμα της όρασης υπολογιστών.

Η πλειοψηφία των ερευνητικών προσπαθειών, στο τομέα της όρασης υπολογιστών, μπορούν ακόμα να μελετηθούν και από μια εντελώς μαθηματική όψη. Πολλοί από τους τρόπους εξαγωγής πληροφορίας από εικόνες βασίζονται στη μαθηματική ανάλυση των δεδομένων. Πιο αναλυτικά υπάρχουν επιρροές από κλάδους όπως:

- **Στατιστική :** Στατιστικές μέθοδοι έχουν εισαχθεί τα τελευταία χρόνια στην 'Όραση Υπολογιστών τόσο για την ανάλυση όσο και για την σύνθεση εικόνων. Χαρακτηριστικά αναφέρονται οι ακολουθιακές Bayesian μέθοδοι, όπως Kalman filters και particle filters, που βελτιώνουν σημαντικά τη συμπεριφορά αλγορίθμων παρακολούθησης αντικειμένων ακόμα και όταν το περιβάλλον κίνησης είναι ιδιαιτέρα πολύπλοκο. Ανάλυση σε πρωτεύοντες άξονες μας επιτρέπει να δημιουργήσουμε ενεργά σχήματα με βάση ένα σύνολο εκπαίδευσης. Πολλά στατιστικά μοντέλα χειρίζονται την εικόνα ως ένα σύνολο σωματιδίων, όπου το καθένα αλληλεπιδρά με τα γειτονικά του, δημιουργώντας ένα ισχυρό μοντέλο στο οποίο μπορούν να εφαρμοστούν στατιστικές ιδέες όπως τα Markov Random Fields.
- **Μαθηματική Μορφολογία :** Το επιστημονικό πεδίο της μαθηματικής μορφολογίας συνεισφέρει στην επεξεργασία εικόνας, με ένα ευρύ πλήθος τελεστών βασισμένων στη θεωρία συνόλων. Οι βασικοί μορφολογικοί τελεστές είναι μη-γραμμικά φίλτρα τύπου sup/inf/rank. Η μαθηματική μορφολογία εξετάζει τη γεωμετρική δομή της εικόνας με τη βοήθεια δομικών στοιχείων, με μεταβαλλόμενο μέγεθος και σχήμα, και είναι στενά συνδεδεμένη με μεθόδους επιπεδοσυνόλων. Έχει χρησιμοποιηθεί σε εφαρμογές όπως η εύρεση του σκελετού ενός σχήματος, ο εντοπισμός γεωμετρικών χαρακτηριστικών σε πολλαπλές κλίμακες, η κατάτμηση εικόνων καθώς και η αναγνώριση αντικειμένων.
- **Μαθηματική Ανάλυση:** Οι Μερικές Διαφορικές Εξισώσεις (PDEs), η Συναρτησιακή Ανάλυση και Λογισμός των Μεταβολών έχουν χρησιμοποιηθεί επιτυχώς στην μοντελοποίηση φυσικών φαινόμενων. Τα εργαλεία αυτά βρίσκονται ευρεία εφαρμογή σε προβλήματα της 'Όρασης Υπολογιστών, όπως οι ενεργές καμπύλες, τα γεωδαιτικά ενεργά περιγράμματα, η ανάλυση εικόνων σε πολλαπλές κλίμακες και η κατάτμηση εικόνων.

1.2 Κίνητρο της Διπλωματικής Εργασίας

Το κύριο θέμα της διπλωματικής εργασίας είναι η δυσδιάστατη παρακολούθηση της όψης ενός χεριού σε μια ταινία. Η παρακολούθηση αντικειμένων σε διαδοχικές εικόνες είναι ένα από τα δομικά προβλήματα στην 'Οραση Υπολογιστών και ειδικά η παρακολούθηση αρθρωτής κίνησης υπάρχει ως τομέας έρευνας από το 1980 [17]. Το συγκεκριμένο πρόβλημα της παρακολούθησης χεριού εξετάζεται πάνω από μια δεκαετία και συνεχίζει να προσελκύει το ερευνητικό ενδιαφέρον, καθώς πιθανές εφαρμογές έρχονται στο προσκήνιο, όπως η προοπτική υπολογιστών ή ρομπότ, που είναι ικανά να αντιλαμβάνονται τις κινήσεις του ανθρώπινου χεριού, έχει εμπνεύσει ερευνητές αλλά και ανθρώπους της βιομηχανίας διασκέδασης. Στο κέντρο του ενδιαφέροντος είναι η ανάπτυξη human computer interfaces (HCI), μέσω των οποίων ο υπολογιστής θα μπορεί να αντιληφθεί κινήσεις χεριού και χειρονομίες και θα ανταποκρίνεται με συγκεκριμένες ενέργειες σε αυτές. Μερικά παραδείγματα είναι, η καθοδήγηση ενός εικονικού αντικειμένου και ο έλεγχος ενός εικονικού ή ακόμα και μηχανικού χεριού. Η δυνατότητα, ο υπολογιστής, να καταλαβαίνει με ακρίβεια την θέση και την χειρονομία του χεριού θα μπορούσε να μεταφραστεί σε δυνατότητα επιλογής ενός συγκεκριμένου αντικειμένου, είτε σε μια οθόνη, είτε στο πραγματικό περιβάλλον, και μια κατάλληλη ενέργεια θα μπορούσε να αντιστοιχηθεί με την επιλογή αυτή. 'Ενα ψηφιακό τραπέζι θα ήταν μια χρήσιμη εφαρμογή σε ένα μουσείο για εκπαιδευτικούς σκοπούς ή σε άλλους δημόσιους χώρους. Ακόμα η είσοδος δεδομένων χωρίς χειριστήριο είναι μια ιδιαιτέρα ενδιαφέρουσα εφαρμογή για παιχνίδι υπολογιστών, όπου ήδη υπάρχει στη αγορά διαθέσιμη μια σχετικά απλοποιημένη εφαρμογή (Eye Toy, Sony Playstation 2).

1.2.1 Περιορισμοί Υπαρχόντων Συστημάτων

Οι περισσότερες προσεγγίσεις του θέματος της παρακολούθησης αντικειμένων και ιδιαιτέρα των χεριών χρησιμοποιούν ένα αριθμό από υποθέσεις, οι οποίες απλοποιούν το πρόβλημα της ανίχνευσης και παρακολούθησης του χεριού. Αυτές εισαγάγουν τους παρακάτω περιορισμούς:

- Ομοιόμορφο ή στατικό φόντο :** Η χρήση ενός απλοποιημένου φόντου (π.χ. μαύρο ή μπλε) επιτρέπει με ακρίβεια τη κατάτυπη της εικόνας και το διαχωρισμό του χεριού από την υπόλοιπη εικόνα[2][29]. Αν η κάμερα είναι στατική, το φόντο μπορεί να διαχωριστεί, υποθέτοντας ότι

η εμφάνιση του χεριού είναι ιδιαίτερα διαφορετική από αυτή του φόντου [22]. Στις περιπτώσεις εφαρμογών παρακολούθησης με βάση τις ακμές, η απουσία ακμών του φόντου κάνουν πιο εύρωστη την εύρεση των ακμών του αντικειμένου[27]. Όμως, η χρήση ομοιόμορφου φόντου περιορίζει ιδιαίτερα το σύστημα, και η τεχνική της απαλοιφής του φόντου είναι μη εφικτή όταν στην εικόνα υπάρχουν σκιές και ο φωτισμός της εικόνας αλλάζει κατά την διάρκεια της κίνησης.

- **Κατάτμηση με βάση το χρώμα του δέρματος :** Ο σκοπός της κατάτμησης, με βάση το χρώμα του δέρματος, είναι η απόκτηση του περιγράμματος του χεριού, το οποίο στην συνέχεια μπορεί να χρησιμοποιηθεί για γίνει ταυτοποίηση του σχήματος. Έχει δειχθεί ότι σε πολλές περιπτώσεις το χρώμα του δέρματος έχει χαρακτηριστική κατανομή, η οποία μπορεί να διαχωριστεί από τις χρωματικές κατανομές του φόντου. Παρόλα αυτά, μια δυσκολία της μεθόδου είναι ότι δεν διαχωρίζονται μόνο τα χέρια αλλά και άλλες περιοχές του σώματος καθώς και πιθανά άλλα αντικείμενα που έχουν παραπλήσια χρωματική κατανομή.
- **Αρχικοποίηση της εφαρμογής:** Σε αρκετές εφαρμογές υποθέτουμε ότι το μοντέλο είναι σωστά ευθυγραμμισμένο στην πρώτη από την ακολουθία των εικόνων [27]. Το πρόβλημα μπορεί να διατυπωθεί ως πρόβλημα βελτιστοποίησης, στο οποίο οι παραμετροί του μοντέλου μεταβάλλονται σε κάθε εικόνα, ώστε να ελαχιστοποιούν μια συνάρτηση σφάλματος. Όμως για εμπορική χρήση, οι εφαρμογές είναι επιθυμητό να αρχικοποιούνται με αυτόματη διαδικασία.
- **Ταχύτητα κίνησης Χεριού:** Γνωρίζοντας ότι η κίνηση του χεριού είναι αργή, η προσέγγιση ακολουθιακής βελτιστοποίησης των παραμέτρων του μοντέλου χεριού, αποδεικνύεται αποδοτική. Η λύση του προηγούμενου καρέ χρησιμοποιείται ως αρχικοποίηση στο τρέχον.[27] Παρόλα αυτά λόγω ενός αριθμού από τοπικά ελάχιστα, η προσέγγιση αυτή δε λειτουργεί σωστά όταν η κίνηση του χεριού είναι γρήγορη, ή όταν υπάρχει επικάλυψη του χεριού από κάποιο άλλο αντικείμενο.

Οι υποθέσεις αυτές ισχύουν σε συγκεκριμένες περιπτώσεις, με αποτέλεσμα την μη αποδοτική συμπεριφορά των μεθόδων στη γενική περίπτωση. Για παράδειγμα, όταν δίνεται μια εικόνα ως είσοδος, δεν υπάρχει καμία γνώση για το φόντο της σκηνής. Όμοια δεν υπάρχει εγγύηση ότι δεν υπάρχει κανένα

άλλο αντικείμενο με χρώμα παρόμοιο του δέρματος στην εικόνα. Αντιλαμβανόμαστε λοιπόν, ότι το πρόβλημα παραχολούθησης χεριού είναι ιδιαίτερα πολύπλοκο και διάφορες ερευνητικές ομάδες έχουν ασχοληθεί εκτενώς με την επίλυση του προβλήματος αυτού.

1.2.2 Περιεχόμενο παρούσας εργασίας

Στην εργασία αυτή ο τρόπος προσέγγισης του προβλήματος της παραχολούθησης χεριού βασίζεται στις ενεργές καμπύλες (Active Contours)[18]. Η ιδέα των ενεργών καμπύλων εισήχθη αρχικά για την εύρεση των ορίων ενός αντικειμένου σε μια εικόνα. Στην συνέχεια, όμως, διάφορες μέθοδοι χρησιμοποιήσαν τις ενεργές καμπύλες για την παραχολούθηση άκαμπτων αντικειμένων [35] [25][26]. Αυτές οι μέθοδοι εισήγαγαν περιορισμούς σχήματος στο τρόπο εξέλιξης της καμπύλης αλλά και πιθανοτικά μοντέλα για την αύξηση της ευρωστίας του συστήματος, όπως το Kalman filter. Παρόλα αυτά, οι συγκεκριμένες μέθοδοι αδυνατούσαν να περιγράψουν χωρίς πρωθύστερη γνώση το πολύπλοκο σχήμα της παλάμης, το οποίο είναι ιδιαίτερα ευμετάβλητο και οι ακμές αλληλοεπηρεάζονται έντονα.

Με σκοπό να βελτιώσουμε την ικανότητα των ενεργών καμπύλων να βρίσκουν και να ακολουθούν τα όρια της παλάμης εισαγάγαμε ιδέες που προκύπτουν από τα μοντέλα κατανομής σημείων (Point Distribution Models), τα οποία είναι η "καρδιά" των ενεργών μοντέλων σχήματος (Active Shape Models) [12] [11]. Το μοντέλο που δημιουργούμε κωδικοποιεί πληροφορία για το δυσδιάστατο σχήμα του αντικειμένου το οποίο επιθυμούμε να ακολουθήσουμε. Πιο συγκεκριμένα, δημιουργούμε ένα σύνολο εκπαίδευσης που περιέχει διάφορες εκφάνσεις του σχήματος ενός χεριού, το οποίο επεξεργαζόμαστε στατιστικά. Με το τρόπο αυτό προκύπτει ο χώρος επιτρεπτού σχήματος μέσα στο οποίο όταν κινείται η ενεργή καμπύλη λαμβάνει μορφές σχημάτων παραπλήσιων εκείνων του συνόλου εκπαίδευσης.

Ο συνδυασμός των δυο αναφερθέντων ιδεών επιτυγχάνεται προβάλλοντας την ενεργή καμπύλη μετά από κάθε εξέλιξή της πάνω στην εικόνα στο επιτρεπτό χώρο [16]. Αυτό έχει ως συνέπεια η προβάλλουσα καμπύλη να διαθέτει τα χαρακτηριστικά του σχήματος που επιθυμούμε και ταυτόχρονα τείνει προς τα όρια του επιθυμητού αντικειμένου.

Κεφάλαιο 2

Παρόμοιες Εργασίες

Στο κεφάλαιο αυτό παρατίθεται μια σύντομη αναφορά στις μεθόδους και τις εφαρμογές που έχουν αναπτυχθεί με σκοπό την παρακολούθηση χεριού. Η ανασκόπηση αυτή βασίζεται σε αντίστοιχο κεφαλαιό που δημοσιεύεται στην διδακτορική εργασία του R. Stenger [30]. Στις παρακάτω ενότητες θα αναφερθούμε στις ιδέες: “παρακολούθηση με βάση μοντέλο” (model-based tracking), “αναγνώριση με βάση την παρατήρηση” (view-based recognition) και “εκτίμηση της στάσης από μια παρατήρηση” (single-view pose estimation).

2.1 Παρακολούθηση χεριού

Η έρευνα που γίνεται στον τομέα της παρακολούθησης χεριού μπορεί να χωρισθεί σε δυο κύρια ρεύματα: στην προσέγγιση, που στηρίζεται σε μοντέλα και στην προσέγγιση, που στηρίζεται στην εκτίμηση της όψης. Οι προσεγγίσεις που χρησιμοποιούν μοντέλα κάνουν χρήση ενός αρθρωτού τρισδιάστατου μοντέλου για το χέρι. Το τρισδιάστατο αυτό μοντέλο προβάλλεται σε μια εικόνα και στη συνέχεια υπολογίζουμε μια συνάρτηση σφάλματος, αξιολογώντας έτσι την ποιότητα του ταιριάσματος. Οι παράμετροι του μοντέλου στη συνέχεια προσαρμόζονται, ώστε να ελαχιστοποιήσουμε το κόστος. Συνήθως υποθέτουμε ότι η διαμόρφωση του μοντέλου στο προηγούμενο καρέ είναι γνωστή και μόνο μια μικρή ανανέωση των παραμέτρων χρειάζεται. Σαν αποτέλεσμα αυτής της υπόθεσης η αρχικοποίηση του μοντέλου υπήρξε ένα μεγάλο εμπόδιο. Στις περισσότερες περιπτώσεις το μοντέλο ευθυγραμμίζεται χειρονακτικά στο πρώτο καρέ. Στην παράγραφο 2.1.2 θα κάνουμε μια ανασκόπηση για τις μεθόδους που χρησιμοποιούνται για παρακολούθηση χεριού και στηρίζονται

σε μοντέλα.

Στις προσεγγίσεις που στηρίζονται στην εκτίμηση όψης το πρόβλημα διατυπώνεται ως ένα πρόβλημα κατάταξης. 'Ενα σύνολο από στοιχεία του χεριού μαρκάρονται σε μια συγκεκριμένη στάση του χεριού και χρησιμοποιούνται για την εκμάθηση ενός κατηγοριοποιητή (classifier). Τέτοιες τεχνικές έχουν χρησιμοποιηθεί για την αναγνώριση χειρονομιών, σε προβλήματα όμως στα οποία έχουμε λίγες στάσεις του χεριού προς μάθηση. Στην παράγραφο 2.1.3 θα παρουσιάσουμε μια γενική εποπτεία των μεθόδων εκτίμηση για την αναγνώριση χειρονομιών.

Πρόσφατα, τα σύνορα που χώριζαν τις δυο παραπάνω προσεγγίσεις έχουν γίνει λιγότερο ευδιάκριτα. Σε αρκετά άρθρα [2][32] τρισδιάστατα μοντέλα έχουν χρησιμοποιηθεί για να δημιουργήσουν ένα διακριτό σύνολο από δυσδιάστατες εμφανίσεις, οι οποίες στην συνέχεια ερευνάται εάν ταιριάζουν στην εικόνα. Αυτές οι δυσδιάστατες απεικονίσεις χρησιμοποιούνται για να δημιουργήθει μια αυθαίρετη ποσότητα από δεδομένα εκμάθησης. Επιπλέον δημιουργείται αυτόματα από το μοντέλο προβολής μια αντιστοιχία μεταξύ τρισδιάστατων πινάκων χειρονομιών και δυσδιάστατων απεικονίσεων. Η αντίστροφη χαρτογράφηση μπορεί να βρεθεί πολύ εύκολα με αναγνώριση βασισμένη σε εκτίμηση της πόζας, δηλαδή εκτιμώμενη πόζα δίνεται από την δυσδιάστατη απεικόνιση που ταιριάζει καλύτερα. Αυτές οι μέθοδοι έχουν την δυνατότητα να λύνουν πολλά προβλήματα που είναι έμφυτα στις μεθόδους παρακολούθησης, που στηρίζονται σε μοντέλα.

'Ένα πλήθος από συστήματα παρακολούθησης δουλεύουν με βάση τη δυσδιάστατη παρακολούθηση χεριού. Στην επόμενη ενότητα παραθέτουμε μερικές από αυτές τις μεθόδους.

2.1.1 Δυσδιάστατη παρακολούθηση χεριού

Η παρακολούθηση του χεριού σε δυο διαστάσεις έχει ένα μεγάλο αριθμό από εφαρμογές, όπως για παράδειγμα σε απλές διεπαφές υπολογιστή με ανθρώπους. Στην περίπτωσή μας αναφερόμαστε σε δυσδιάστατη παρακολούθηση χεριού με μεθόδους που χρησιμοποιούν μία μόνο κάμερα για να παρακολουθούν το χέρι - πιθανόν με κάποιες αλλαγές κλίμακας - και δεν ανακτούν τρισδιάστατη πληροφορία για την περιστροφή εκτός επιπέδου εικόνας ή για την διαμόρφωση των δαχτύλων. Αυτό είναι ένα πολύ πιο απλό πρόβλημα από την πλήρη τρισδιάστατη παρακολούθηση για αρκετούς λόγους. Πρώτα από όλα οι διαστάσεις του χώρου παρακολούθησης είναι πολύ μικρότερες. 'Έχουμε

μόνο 4 βαθμούς ελευθερίας κίνησης, οι οποίες είναι μετατόπιση, κλιμάκωση και περιστροφή στο επίπεδο της εικόνας, ενώ στην τρισδιάστατη περίπτωση έχουμε 6 βαθμούς ελευθερίας για άκαμπτη κίνηση. Ένας άλλος παράγοντας, που καθιστά το πρόβλημά στις δυο διαστάσεις πιο εύκολο, είναι ότι οι στάσεις του χεριού μπορούν να αναγνωριστούν πιο εύκολα όταν το χέρι είναι τοποθετημένο παράλληλα μπροστά από την κάμερα, γιατί η ορατή περιοχή είναι σχετικά μεγαλύτερη και η αλληλοκάλυψη μεταξύ διαφορετικών δάχτυλων μικρότερη. Ένα παράδειγμα εφαρμογής διεπαφής μεταξύ υπολογιστή και ανθρώπου είναι το σύστημα ζωγραφικής που βασίζεται στην όραση, που παρουσιάστηκε από τους MacCormick και Isard [23]. Το σχήμα του χεριού μοντελοποιήθηκε χρησιμοποιώντας μια σπλήνα 2ης τάξης (B-spline) και ένα μικρό φίλτρο χρησιμοποιήθηκε για να παρακολουθήσει το περίγραμμα του χεριού. Για μεγαλύτερη αποδοτικότητα ο χώρος των καταστάσεων διατρέθηκε, με την απόζευξη της κίνησης του δείκτη και του αντίχειρα. Για να περιοριστεί η περιοχή αναζήτησης γίνεται εκτίμηση του φόντου και έτσι 7 βαθμοί ελευθερίας παρακολουθούνται σε πραγματικό χρόνο. Η εφαρμογή αρχικοποιείται με αναζήτηση στις περιοχές που είχαν κατηγοριοποιηθεί σαν χρώμα δέρματος. Ένας περιορισμός ωστόσο της συγκεκριμένης μεθόδου είναι το γεγονός ότι το περίγραμμα από μόνο του δεν είναι ένα αξιόπιστο χαρακτηριστικό κατά την διάρκεια γρήγορων κινήσεων του χεριού.

Οι Laptev και Lindeberg [19] και Bretzner [7] παρουσίασαν ένα σύστημα, το οποίο θα μπορούσε να παρακολουθήσει ένα ανοιχτό χέρι και να αναγνωρίσει την όψη αυτού ανάμεσα σε 5 διαφορετικές καταστάσεις του χεριού. Χρησιμοποιήσαν ειραρχική ανάλυση για να βρίσκουν την ένταση ή το χρώμα των χαρακτηριστικών κηλίδων σε διαφορετικές κλίμακες, που ανταποκρίνονται στην παλάμη και στα δάχτυλα. Κατασκευάζονται στη συνέχεια πίνακες πιθανοτήτων για αυτά τα χαρακτηριστικά και χρησιμοποιούνται για να αποτιμήσουν πέντε καταστάσεις του χεριού που αναφέρονται σε συγκεκριμένα λυγίσματα και εκτάσεις συγκεκριμένων δακτύλων. Η παρακολούθηση γίνεται επιτυχής χρησιμοποιώντας particle filter και το σύστημα επεξεργάζεται 10 καρέ το δευτερόλεπτο. Το όλο σύστημα χρησιμοποιήθηκε σε μία εφαρμογή απομακρυσμένου ελέγχου τηλεόρασης.

Οι Von Hardenberg και Berard έδειξαν την χρήση ενός απλού αλλά αποδοτικού ακροδέκτη-ανιχνευτή σε ορισμένες HCI εφαρμογές. Το σύστημα εκτιμά προσαρμοστικά τον φόντο και ψάχνει στο προσκήνιο για σχήματα που μοιάζουνε με δάχτυλα σε μία κλίμακα. Αυτή η μέθοδος δουλεύει μόνο για εφαρμογές με ακίνητη κάμερα και με καθαρή κατάτμηση του προσκήνιου, προϋ-

ποθέσεις που είναι δύσκολο να ισχύουν σε μια γενική περίπτωση. Παρόλα αυτά δεν απαιτεί αρχικοποίηση και επιτρέπει την παρακολούθηση γρήγορων κινήσεων του χεριού.

Συνοφίζοντας μπορούμε να πούμε ότι ένας μεγάλος αριθμός δυσδιάστατων συστημάτων έχουν αποδειχτεί ότι δουλεύουν σε εφαρμογές HCI σε πραγματικό χρόνο. Παρόλα αυτά είναι δύσκολο να επεκταθούν σε τρεις διαστάσεις χωρίς την χρησιμοποίηση ενός γεωμετρικού μοντέλου για το χέρι.

2.1.2 Παρακολούθηση χεριού με βάση μοντέλο

Ένα σύστημα παρακολούθησης χεριού βασιζόμενο σε μοντέλο γενικά αποτελείται από ένα αριθμό σταδίων επεξεργασίας, όπως είναι το στάδιο επεξεργασίας της εικόνας και η εξαγωγή των χαρακτηριστικών της, η διαδικασία τατιράσματος με την απεικόνιση που έχει προκύψει από το μοντέλο, τον υπολογισμό της συνάρτησης σφάλματος και σύμφωνα με αυτή η βελτιστοποίηση των παραμέτρων του μοντέλου. Το γεωμετρικό μοντέλο του χεριού συνήθως κατασκευάζεται από τον προγραμματιστή, αλλά είναι δυνατό να αποκτηθεί και από μεθόδους ανακατασκευής. Τα μοντέλα που χρησιμοποιούνται σε τέτοιες εφαρμογές βασίζονται σε απλά γεωμετρικά σχήματα, όπως είναι κύλινδροι και κώνοι.

Η κινηματική δομή των μοντέλων αυτών είναι βασισμένη στους μηχανισμούς του βιολογικού χεριού. Κάθε δάκτυλο μοντελοποιείται σαν μια κινηματική αλυσίδα με 4 βαθμούς ελευθερίας τοποθετημένη πάνω στη παλάμη, και συνήθως ο αντίχειρας μοντελοποιείται με παρόμοιο τρόπο αλλά με τέσσερις ή πέντε βαθμούς ελευθερίας. Μαζί με την άκαμπτη κίνηση του χεριού που προέρχεται από την κίνηση όλου του σώματος, πρέπει να εκτιμηθούν 27 βαθμοί ελευθερίας. Αντιλαμβανόμαστε, ότι η διάσταση του προβλήματος είναι εξαιρετικά μεγάλη δημιουργώντας αντίστοιχα προβλήματα, τα χαρακτηριστικά σημεία τα οποία θα μπορούσαν να ακολουθηθούν με μεγάλη αξιοπιστία είναι ελάχιστα, αλληλοεπικαλύψεις συμβαίνουν συχνά. Παρόλα αυτά ο ανατομικός σχεδιασμός του ανθρώπου θέτει συγκεκριμένους περιορισμούς στην κίνηση των χεριών. Αυτοί οι περιορισμοί αξιοποιούνται με σκοπό την μείωση της διάστασης του προβλήματος. Ένας περιορισμός κίνησης είναι η περιορισμένη γωνία περιστροφής του αγκώνα, που ορίζει το εύρος της κίνησης. Ένας άλλος σημαντικός περιορισμός είναι η συσχέτιση της κίνησης των αρθρώσεων. Οι αρθρώσεις του ίδιου δακτύλου, αλλά και μερικών δακτύλων μεταξύ τους, δεν είναι ανεξάρτητες. Αυτοί οι περιορισμοί μπορούν να ενεργοποιηθούν στο σύ-

στημα με διάφορους τρόπους, όπως η συσχέτιση ορισμένων παραμέτρων του μοντέλου ή με τη μείωση του υποχώρου του προβλήματος, π.χ με χρήση PCA.

Διάφορα δυναμικά μοντέλα έχουν χρησιμοποιηθεί για να προσεγγίσουν την κίνηση του χεριού. Συνήθης τεχνικές, όπως το διευρυμένο Kalman φίλτρο (extended Kalman filter) με σταθερή ταχύτητα, αποτελούν μέρος των δυναμικών μοντέλων. Ένα πιο σύνθετο μοντέλο είναι το “ιδιοδυναμικό” μοντέλο που πρότειναν οι Zhu et al., στο οποίο η κίνηση κάθε δακτύλου είναι μοντελοποιημένη ως ένα δισδιάστατο δυναμικό μοντέλο σε μειωμένης διάστασης χώρο ιδιοδιανυσμάτων. Για δισδιάστατες κινήσεις ο Isard και Blake [5] δημιούργησαν μοντέλα διαφορετικών κινήσεων τα οποία χαρακτηρίζουν μορφές κίνησης. Η κίνηση όμως του χεριού μπορεί να είναι γρήγορη σε σχέση με την κίνηση της εικόνας (ταχύτητα λήψης) και να προκύπτουν αρκετά προβλήματα λόγω έλλειψης ομαλότητας. Έρευνες έχουν γίνει που έχουν αναδείξει ότι το λύγισμα ενός δακτύλου μπορεί να επιτευχθεί σε 0.1 δευτερόλεπτα, αντιστοιχώντας μόνο σε τρία καρέ σε μια ακολουθία από εικόνες. Το δεδομένο αυτό αυξάνει κατακόρυφα τη δυσκολία μοντελοποίησης της δυναμικής κίνησης του χεριού. Μια προσέγγιση έχει προτείνει την αναγνώριση ορισμένων χαρακτηριστικών μορφών του χεριού στην ακολουθία των εικόνων και στην συνέχεα την παρεμβολή της κίνησης του χεριού σύμφωνα με ένα μοντέλο μεταξύ των χαρακτηριστικών καρέ που έχουν βρεθεί.

Ένα από τα πιο σημαντικά σταδία των εφαρμογών βασισμένων σε μοντέλα, είναι το στάδιο του μοντέλου παρατήρησης, πιο συγκεκριμένα η επιλογή των χαρακτηριστικών τα οποία θα χρησιμοποιηθούν για να ανανεώσουν το μοντέλο. Πιθανά χαρακτηριστικά είναι οι ακμές εντάσεως της εικόνας, οι χρωματικές ακμές ή οι σιλουέτες που προκύπτουν από την κατάτμηση σύμφωνα με το χρώμα. Η αναγνώριση καθορισμένων σημείων έχει επιχειρηθεί, όπως για παράδειγμα η αναγνώριση των ακροδακτύλων με έρευνα στο σχήμα της σιλουέτας.

Ένα από τα πρώτα συστήματα για παρακολούθηση χεριού χωρίς κάποιο εξωτερικό δείκτη ήταν αυτό που αναπτύχθηκε από τους Cipolla και Hollinghurst, οι οποίοι χρησιμοποίησαν B-spline φίδια για τη παρακολούθηση ενός χεριού από δυο μη βαθμονομημένες όψεις. Από κάθε όψη εκτιμούταν ο αφινικός μετασχηματισμός του φιδιού και υπό την δράση ενός δυναμικού βρισκόταν η τελική θέση του φιδιού. Το σύστημα απαιτούσε απλοποιημένο φόντο, αλλά μπορούσε να λειτουργήσει σε πραγματικό χρόνο.

Η πρώτη εργασία για τρισδιάστατη αρθρωτή κίνηση χεριού πραγματοποιήθηκε από τους Regh και Kanade [28] με το λεγόμενο DigitEyes σύστημα

παρακολούθησης. Στο σύστημα χρησιμοποιήθηκε ένα τρισδιάστατο μοντέλο χεριού, το οποίο κατασκευάστηκε από χυλίνδρους και υποστηριζόταν από ένα κινηματικό μοντέλο με 27 βαθμούς ελευθερίας. Η παρακολούθηση πραγματοποιείτο με την ελαχιστοποίηση μιας συνάρτησης σφάλματος, η οποία μετρούσε πόσο καλά το μοντέλο ευθυγραμμιζόταν στην εικόνα. Δυο συναρτήσεις σφάλματος προτάθηκαν, η μια βασιζόταν στις ακμές και η άλλη σε ένα πρότυπο εγγραφής. Στην πρώτη μέθοδο, οι άξονες των χυλίνδρων του μοντέλου προβάλλονταν στην εικόνα και βρίσκονταν οι τοπικές ακμές. Ο όρος σφάλματος ήταν το άθροισμα των τετραγώνων των αποστάσεων μεταξύ των σημείων των προβολών και των χαρακτηριστικών σημείων της εικόνας. Η ελαχιστοποίηση επιτυγχάνεται με την χρήση του αλγόριθμου Gauss-Newton. Η δεύτερη μέθοδος χρησιμοποιείτο με σκοπό την αντιμετώπιση των αλληλοεπικαλύψεων. Με την χρήση δυο καμερών και εξειδικευμένου υπολογιστικού συστήματος, το σύστημα κατάφερνε να επεξεργαστεί 10 καρέ το δευτερόλεπτο. Όμως, ήταν αναγκαία η ύπαρξη ουδέτερου φόντου για την αξιόπιστη εύρεση των χαρακτηριστικών της κάθε εικόνας.

Οι Shimada et al. χρησιμοποίησαν μια προσέγγιση βασισμένη στη σιλουέτα για την παρακολούθηση με μια κάμερα. Οι θέσεις των δακτύλων αποκόπτονταν από την σιλουέτα και γινόταν προσπάθεια ταιριάσματος αυτών με το προβολή ενός τρισδιάστατου μοντέλου. Ένα σύνολο από πιθανά διανύσματα πόζας παράγονταν και κάθε ένα από αυτά ανανεωνόταν με την χρήση διευρυμένου kalman φίλτρου. Η βέλτιστη οδός βρισκόταν για το σύνολο της ακολουθίας εικόνων. Η δυσκολία αυτής της μεθόδου είναι η αναγνώριση του σχήματος των δακτύλων όταν υπάρχει κάποιο είδος αλληλοεπικάλυψης.

Ο Stenger [31] πρότεινε την δημιουργία ενός τρισδιάστατου μοντέλου χεριού, το οποίο κωδικοποιεί πληροφορίες για το σχήμα, την κίνηση το χρώμα που ενδέχεται το χέρι να έχει σε μια εικόνα. Μεταβλητές κατάστασης συνδέονται με το μοντέλο αυτό, οι οποίες εκτιμούνται από φίλτρα που συνδέονται προσωρινές πληροφορίες με παρατηρήσεις από την εικόνα. Οι παρατηρήσεις αυτές είναι οι ακμές των συνόρων των αντικειμένων που περιέχονται στην εικόνα και το χρώμα του δέρματος. Στην εργασία του Stenger et al. προτείνονται δυο τρόποι επίλυσης του προβλήματος εκτίμησης των παραμέτρων. Ο πρώτος τρόπος βασίζεται σε ένα αναδρομικό Bayesian φίλτρο, μια προέκταση του Kalman φίλτρου, το unscented Kalman filter. Αυτή η προσέγγιση κάνει χρήση ενός δυναμικού μοντέλου για την πρόβλεψη της κίνησης του χεριού σε κάθε καρέ και συνδυάζει την πρόβλεψη με τη παρατήρηση για να καταλήξει στην εκτίμηση της πόζας-στάσης του χεριού. Αυτή η προσέγγιση λειτουργεί

σωστά εάν γίνουν κάποιες υποθέσεις, οι οποίες είναι ιδιαίτερα περιοριστικές για το γενικό πλαίσιο του προβλήματος. Η δεύτερη προσέγγιση συνδυάζει ιδέες από τεχνικές εύρεσης και Bayesian φιλτραρίσματος. Η προσέγγιση των τεχνικών εύρεσης επιλύει το πρόβλημα αρχικοποίησης, όταν δίνεται μια εικόνα χωρίς προγενέστερη γνώση για τη θέση του χεριού. Αυτό επιτυγχάνεται με μέθοδο ταιριάσματος των πρότυπων που παράγονται από το τρισδιάστατο μοντέλο με τα αντικείμενα της εικόνας. Έχοντας όμως γνώση της αρχικής εικόνας και κάνοντας χρήση της πληροφορίας που προκύπτει γίνεται χρήση ενός λειτουργεί, αλλά μόνο σε σχετικά αργές παραμορφώσεις του χεριού και συνήθως σε περιορισμένα περιβάλλοντα, όπως είναι τα ουδέτερα φόντα ή ο συγκεκριμένος φωτισμός. Το πρόβλημα είναι μη καλά ορισμένο με είσοδο μόνο μιας όψης του χεριού, γιατί μερικές διατάξεις του χεριού είναι εξαιρετικά δύσκολο να τις παρατηρήσουμε. Για να αντιμετωπιστεί αυτό το πρόβλημα πολλές από τις μεθόδους χρησιμοποιούν πολλαπλές όψεις του χεριού. Ένα επιπλέον πρόβλημα είναι η αρχικοποίηση του 3-Δ μοντέλου. Όταν αρχίζει η διαδικασία παρακολούθησης ή όταν χάνεται το αντικείμενο παρακολούθησης είναι απαραίτητο να υπάρχει ένας μηχανισμός αρχικοποίησης. Γενικά, το πρόβλημα αυτό μπορεί να αντιμετωπιστεί ως πρόβλημα εύρεσης χεριού, το οποίο επιλύεται με μεθόδους βασιζόμενες στην παρατήρηση (view-based).

2.1.3 Αναγνώριση με βάση την εκτίμηση της Παρατήρησης

Οι συγκεκριμένες μέθοδοι χρησιμοποιούνται παραδοσιακά για αναγνώριση χειρονομιών. Το πρόβλημα αυτό πολύ συχνά αναφέρεται σαν πρόβλημα αναγνώρισης προτύπων, το οποίο μπορεί να διασπαστεί σε διάφορα άλλα υποπροβλήματα. Αυτές οι μέθοδοι πολλές φορές αναφέρονται ως μέθοδοι "από κάτω προς τα πάνω" γιατί χαρακτηριστικά χαμηλού επιπέδου χρησιμοποιούνται για να εξαγάγουμε συμπεράσματα για χαρακτηριστικά υψηλοτέρου επιπέδου. Τα βασικά προβλήματα που συνοδεύουν τις συγκεκριμένες μεθόδους είναι, το πώς θα κατατυμήσουμε το χέρι από ένα γενικό φόντο, αλλά και ποια χαρακτη-

ριστικά θα εξαγάγουμε από την κατατμημένη περιοχή. Η κατηγοριοποίηση γίνεται συνήθως χρησιμοποιώντας τον πιο κοντινό γειτονικό κατηγοριοποιητή ή χρησιμοποιώντας άλλες καθιερωμένες μεθόδους κατηγοριοποίησης. Γενικά ο κατηγοριοποιητής (classifier) μαθαίνει από ένα σύνολο από παραδείγματα εκπαίδευσης και τοποθετεί την εικόνα εισόδου σε μία από τις δυνατές κατηγορίες. Στις επόμενες παραγράφους θα παρουσιάσουμε μόνο μερικά επιλεγμένα άρθρα για την αναγνώριση χειρονομίας.

Οι Cui και Weng [13] παρουσίασαν ένα σύστημα για αναγνώριση νευμάτων του χεριού, το οποίο εμπεριείχε και την κατάτμηση και την αναγνώριση. Κατά την φάση εκπαίδευσης, πραγματοποιείται χαρτογράφηση από τοπικά υπο-“παράθυρα” του χεριού σε μια μάσκα κατάτμησης. Κατά την φάση της αναγνώρισης, αναγνωρίσθηκαν περιοχές με κίνηση και τοπικά υπο-“παράθυρα” διαφορετικού μεγέθους χρησιμοποιήθηκαν για να γίνουν υποθέσεις για τον αριθμό των τμημάτων. Κάθε μία από τις υποθέσεις αυτές επαληθεύεται με τη βοήθεια ενός κατηγοριοποιητή στην περιοχή κατάτμησης. Έγινε επιλογή των χαρακτηριστικών με τη βοήθεια διακριτής ανάλυσης και χρησιμοποιήθηκε ένα δένδρο απόφασης για την κατηγοριοποίηση. Το σύστημα αναγνώριζε 28 διαφορετικά νεύματα με αποτέλεσμα αναγνώρισης της τάξης του 93%. Όμως η φάση κατάτμησης η οποία διήρκησε περίπου 1 λεπτό ανά καρέ δημιουργεί προβλήματα επίδοσης. Σε αντιδιαστολή η φάση κατηγοριοποίησης διήρκησε περίπου 0.1 δευτερόλεπτα. Άλλος ένας περιορισμός αυτής της προσέγγισης είναι η υπόθεση ότι το χέρι είναι το μόνο κινούμενο αντικείμενο στο όλο σκηνικό.

Οι Triesch και Von der Malsburg δημιούργησαν μια διεπαφή χειρονομιών για τον έλεγχο ενός ρομπότ. Χρησιμοποιήθηκε μέθοδος ταιριάσματος ελαστικών δυσδιάστατων γράφων ώστε να συγχριθούν 12 διαφορετικές χειρονομίες σε μια εικόνα. Για να μπορέσει το σύστημα να λειτουργήσει σε ένα ακατάστατο φόντο χρησιμοποιήθηκαν φίλτρα Gabor και έγχρωμες cues. Το συγκεκριμένο σύστημα επιτυγχάνει επίσης ποσοστό αναγνώρισης 93% σε απλά φόντα και 86% σε ακατάστατα φόντα.

Οι Wu και Huang [151] εισήγαγαν μία μέθοδο που στηρίζεται στον αλγόριθμο EM (Expectation-maximization) και συνδυάζει δεδομένα εκμάθησης, τα οποία έχουν χαρακτηριστεί με κάποια ετικέτα αλλά και δεδομένα εκμάθησης τα οποία δεν έχουν χαρακτηριστεί. Χρησιμοποιώντας αυτή την προσέγγιση μπόρεσαν να κατηγοριοποιήσουν 14 διαφορετικές στάσεις του χεριού σε διαφορετικές όψεις. Δύο τύποι από διανύσματα χαρακτηριστικών συγχρίθηκαν: το ένα διάνυσμα αποτελεί συνδυασμό υφής και περιγραφών σχήματος και το άλλο διάνυσμα προέρχεται από την ανάλυση πρωτευουσών συνιστώσων (Prin-

cipal Components Analysis) της εικόνας του χεριού. Τα χαρακτηριστικά που παράγθηκαν από την PCA εμφάνισαν καλύτερες επιδόσεις με ένα ποσοστό σωστής κατηγοριοποίησης της τάξης του 92%.

Οι Lockton και Fitzgibbon [21] δημιούργησαν ένα σύστημα αναγνώρισης χειρονομιών πραγματικού χρόνου για 46 διαφορετικές στάσεις του χεριού, συμπεριλαμβάνοντας γράμματα από την Αμερικάνική νοηματική γλώσσα. Η μέθοδος τους στηρίζεται στην κατάτμηση με βάση το χρώμα του δέρματος και χρησιμοποιεί έναν boosting αλγόριθμο για γρήγορη κατηγοριοποίηση. Ο χρήστης του συστήματος πρέπει να φοράει ένα βραχιόλι έτσι ώστε το σχήμα του χεριού να μπορεί να χαρτογραφηθεί ποιο εύκολα σε κανονικό καρέ. Το σύστημα αυτό παρουσιάζει το υψηλότερο ποσοστό ορθής αναγνώρισης της τάξης του 99.87%. Ωστόσο το όλο σύστημα εμφανίζει έναν περιορισμό, απαιτείται να είναι υπό έλεγχο οι συνθήκες φωτισμού.

Ο Tomasi et al. χρησιμοποίησαν μια μέθοδο κατηγοριοποίησης μαζί με παρεμβολή παραμέτρων για να παρακολουθήσουν την κίνηση του χεριού. Δεδομένα της έντασης της εικόνας χρησιμοποιήθηκαν για την εκπαίδευση ενός ιεραρχικού κατηγοριοποιητή του συντομότερου γείτονα (24 στάσεις σε 15 σκοπίες), ταξινομώντας κάθε ένα καρέ σαν μια από τις 360 σκοπιές. Στη συνέχεια χρησιμοποιήθηκε ένα τρισδιάστατο μοντέλο για το χέρι για την οπτικοποίηση της κίνησης. Οι παράμετροι των 24 στάσεων καθορίστηκαν πρόχειρα με το χέρι και οι παράμετροι του μοντέλου παρεμβλήθηκαν για τη μετάβαση μεταξύ 2 στάσεων. Το σύστημα ναι μεν μπορούσε να ανταπεξέλθει σε γρήγορες κινήσεις του χεριού, αλλά είχε τους περιορισμούς για καθαρό χρώμα ώστε να επιτυγχάνεται κατάτμηση του δέρματος, καθώς και ελεγχόμενες συνθήκες φωτισμού, ακριβώς όπως και στην περίπτωση των Lockton και Fitzgibbon.

Πρέπει επίσης να σημειώσουμε ότι πολλά άρθρα στην αναγνώριση χειρονομίας, όπως στο σύστημα αναγνώρισης νοηματικής γλώσσας των Staner et al., κάνουν εκτεταμένη χρήση προσωρινής πληροφορίας. Η Αμερικανική νοηματική γλώσσα έχει ένα λεξιλόγιο με περίπου 6000 χειρονομιών, κάθε μία από τις οποίες παριστάνει μια λέξη. Αυτό διαφέρει από την ορθογραφία με τα δάχτυλα η οποία χρησιμοποιείται για υπαγόρευση αγνώστων λέξεων ή κύριων ουσιαστικών. Η ανίχνευση χεριού στη μέθοδο αυτή δεν προσπαθεί να ανακτήσει την λεπτομέρεια, αλλά ανακτά μόνο μια χοντρική περιγραφή του σχήματος του χεριού, της θέσης του καθώς και του προσανατολισμού του. Μέθοδοι από την αναγνώριση φωνής, όπως τα κρυφά Μαρκοβιανά μοντέλα, μπορούν να χρησιμοποιηθούν για τέτοιες εργασίες. Ο Starner et al. πέτυχαν ένα ποσοστό αναγνώρισης 98% χρησιμοποιώντας λεξιλόγιο 40 λέξεων.

Συνοψίζοντας, μπορούμε να πούμε ότι οι μέθοδοι που στηρίζονται στην παρατήρηση μπορούν να είναι αποδοτικές στον διαχωρισμό μεταξύ ενός συγκεκριμένου αριθμού στάσεων του χεριού, κάτι το οποίο είναι ικανοποιητικό για ένα πλήθος εφαρμογών αναγνώρισης χειρονομιών. Ένα από τα προβλήματα αυτών των μεθόδων είναι η φάση κατάτμησης. Αυτή η φάση πραγματοποιείται κυρίως με κατάτμηση που βασίζεται στο χρώμα του δέρματος, κάτι το οποίο προϋποθέτει από τον χρήστη να φορά ένα βραχιόλι ώστε να διαχωρίζεται η παλάμη από το υπόλοιπο χέρι. Μια άλλη δυνατότητα είναι να δοκιμάζονται διάφορες πιθανές κατατμήσεις, κάτι το οποίο όμως θα αυξήσει τον χρόνο αναγνώρισης.

2.1.4 Εκτίμηση στάσης με μία παρατήρηση

Τελευταία έχουν προταθεί τεχνικές εκτίμησης της στάσης του χεριού με τη βοήθεια μιας μόνο παρατήρησης. Σε αυτές τις τεχνικές ένα τρισδιάστατο μοντέλο χρησιμοποιείται για να πάρουμε δυσδιάστατες εμφανίσεις, οι οποίες ταιριάζουν στην εικόνα και η αντιστοίχηση μεταξύ των τριών διαστάσεων και των δυο γίνεται με προβολή του μοντέλου στην εικόνα. Οι απεικονίσεις δυο διαστάσεων αποθηκεύονται και δημιουργούν μια βάση δεδομένων.

Οι τεχνικές αυτές στηρίζονται στην παρατήρηση του περιγράμματος. Μπορούν να αναζητηθούν πολλά άρθρα στην βιβλιογραφία σχετικά με αντίστοιχα θέματα όπως στο [2]. Οι τεχνικές της εκτίμησης με μία παρατήρηση εμφανίζουν μεγάλο ενδιαφέρον, γιατί μπορούν να λύσουν το πρόβλημα της αρχικοποίησης, που παρουσιάζονται σε εφαρμογές που στηρίζονται σε μοντέλα. Στην πράξη αν ήταν αρκετά αποτελεσματικές οι τεχνικές αυτές θα μπορούσαν να χρησιμοποιηθούν για την εκτίμηση της κίνησης σε κάθε καρέ ξεχωριστά.

Και στην περίπτωση αυτή όμως το μεγάλο πρόβλημα εμφανίζεται στην φάση κατάτμησης. Οι μέθοδοι εκτίμησης στάσης από μια παρατήρηση απαιτούν είτε καθαρό φόντο είτε σαφή κατάτμηση του χεριού.

Πρέπει να επισημάνουμε ότι πολλές μέθοδοι για την παρακολούθηση χεριού πηγάζουν από εφαρμογές και προσπάθειες παρακολούθησης του συνολικού σώματος του ανθρώπου και τα δυο αυτά προβλήματα αποτελούν υπο-προβλήματα της παρακολούθησης μη άκαμπτων αρθρωτών κινήσεων.

Κεφάλαιο 3

Ενεργές Καμπύλες

Το μοντέλο των ενεργών καμπύλων (Active Contours) ή το μοντέλο “φίδια” (snakes), όπως χαρακτηριστικά αποκαλούνται, γεννήθηκε από τις μεθόδους κανονικοποίησης, που εφαρμόζονται στην Όραση Υπολογιστών, στις οποίες η όραση προσεγγίζεται ως ένα αντίστροφο πρόβλημα οπτικής [34], για το λόγω αυτό είναι γενικά μη καλά ορισμένο. Εντός του πλαισίου κανονικοποίησης, το πρόβλημα είναι πλέον περιορισμένο, θέτοντας περιορισμούς ομαλότητας για το αντικείμενο, στο οποίο τα δεδομένα θέλουμε να ταιριάξουν. Καθώς ενδιαφερόμαστε κυρίως για την εύρεση και δημιουργία του περιγράμματος ενός αντικειμένου σε εικόνες, στη δικιά μας περίπτωση το αντικείμενο είναι το χέρι, περιοριζόμαστε στο πρόβλημα εύρεσης ενός δυσδιάστατου αντικειμένου. Οι ενεργές καμπύλες προτάθηκαν από τους Kass et al.[18] ως πιθανό εργαλείο για την επίλυση του προβλήματος κατάτμησης. Αποτελούνται από μια καμπύλη ενεργειακής ελαχιστοποίησης, η οποία καθοδηγείται από εξωτερικές δυνάμεις και επηρεάζεται από δυνάμεις δυναμικών πεδίων, οι οποίες την οδηγούν προς τα επιθυμητά στοιχειά της εικόνας. Επιπλέον, εσωτερικοί περιορισμοί μπορούν να χρησιμοποιηθούν για να ενισχύσουν την ομαλότητα της καμπύλης ή να εισάγουν ασυνέχεια σε ένα σημείο της καμπύλης. Η μέθοδος των ”φιδιών” σχεδιάστηκε αρχικά για να εφαρμοστεί σε γραφικές παραστάσεις συναρτήσεων με σκοπό την εύρεση των τοπικών ελαχίστων των συναρτήσεων αυτών. Μια εικόνα, όμως, μπορεί να θεωρηθεί ως μια συνάρτηση δύο μεταβλητών και με μια κατάλληλη επεξεργασία, σημαντικά χαρακτηριστικά της εικόνας, όπως οι ακμές και τα περιγράμματα, μπορούν να αντιστοιχούν σε τοπικά ελάχιστα της γραφικής παράστασης της συγκεκριμένης συνάρτησης.

Η προσέγγιση των συνόρων του αντικειμένου επιτυγχάνεται με την ελαχιστοποίηση ενός ενεργειακού συναρτησιακού της καμπύλης. Αν η θέση της

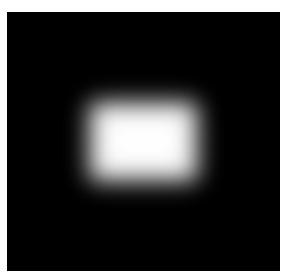
καμπύλης είναι αρκετά κοντά σε περίγραμμα της εικόνας, έτσι ώστε οι δυνάμεις του δυναμικού πεδίου να σχετίζονται άμεσα με τα συγκεκριμένα περιγράμματα, το "φίδι" θα κινηθεί και τελικά θα φτάσει το τοπικό ενεργειακό ελάχιστο που αντιστοιχεί στο περίγραμμα. Για το λόγω αυτό, το "φίδι" εξαρτάται από εξωτερικούς μηχανισμούς για την τοποθέτηση του σχετικά κοντά στα όρια κάποιου αντικειμένου. Παρόλο, λοιπόν, που αυτή η προσέγγιση δεν επιλύει απόλυτα το πρόβλημα της κατάτμησης της εικόνας, ο δυναμικός χαρακτήρας της ενεργής καμπύλης αποδεικνύεται ικανός για την παρακολούθηση παραμορφώσιμων αντικειμένων.

Το πρόβλημα επιλογής της αρχικής θέσης της ενεργής καμπύλης γίνεται σχεδόν αμελητέο, όταν η μέθοδος παρακολούθησης γίνεται σε διαδοχικές εικόνες. Για παράδειγμα, όταν θέλουμε να ακολουθήσουμε την κίνηση ενός παραμορφώσιμου αντικειμένου, σε τέτοιες περιπτώσεις, μόνο στο πρώτο καρέ της κίνησης αντιμετωπίζουμε πραγματικό πρόβλημα αρχικοποίησης, το οποίο αντιμετωπίζεται συνήθως με την βοήθεια του χρήστη της εφαρμογής. Για τα επόμενα καρέ, η "βέλτιστη" θέση του "φιδιού" που έχει βρεθεί στο αμέσως προηγούμενο καρέ μπορεί να χρησιμοποιηθεί ως αρκετά καλή αρχικοποίηση για την ανάλυση του τρέχοντος καρέ. Αυτή η τεχνική θα αποδειχθεί ιδιαιτέρα ικανοποιητική αν κανείς υποθέσει ότι μεγάλες μεταβολές δεν συμβαίνουν από τη μια εικόνα στην άλλη. Μεγάλη μεταβολή εννοούμε τη μεταβολή για την οποία οι δυνάμεις του δυναμικού πεδίου, που αντιστοιχούν στην μεταβολή του περιγράμματος, δεν θα επηρεάζουν τη ενεργή καμπύλη. Για το λόγω αυτό θεωρούμε ότι το χέρι δεν θα κινείται ιδιαιτέρα γρήγορα, ώστε η αρχικοποίηση της ενεργής καμπύλης να βρίσκεται στο εσωτερικό του δυναμικού πεδίου που δημιουργείται από το δεδομένο αντικείμενο που ακολουθούμε σε κάθε καινούργια εικόνα.

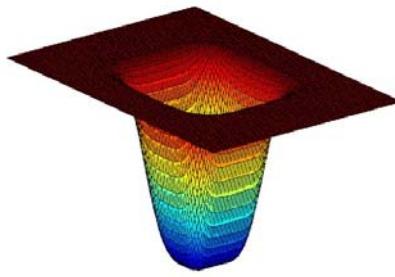
Στο κεφάλαιο αυτό αρχικά θα αναφερθούμε στα βασικά στοιχεία των ενεργών καμπύλων.

3.1 Βασικές Έννοιες

'Ενα "φίδι" είναι ένα μοντέλο καμπύλης που είναι ικανή να παραμορφωθεί. Θα μπορούσαμε να την φανταστούμε σαν ένα σώμα το οποίο αποτελείται από κοινάτια ελαστικών υλικών, το καθένα με χαρακτηριστικές ιδιότητες. 'Έτσι, η ελαστική καμπύλη του μοντέλου μπορούμε να θεωρήσουμε απλοϊκά, ότι δημιουργείται από συνδυασμό σχοινιού και ξύλινης βέργας. Το πρώτο κάνει το "φίδι" να αντιστέκεται στην επιμήκυνση και το δεύτερο να αντιστέκεται στο



3.1.1

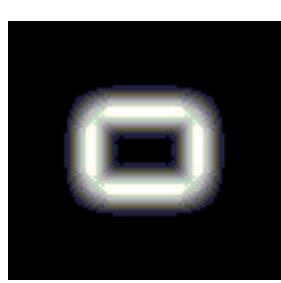


3.1.2

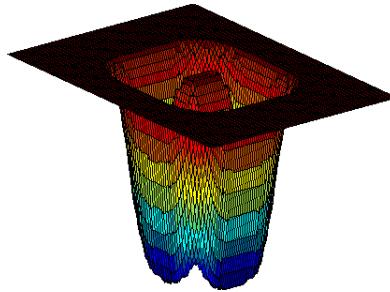
Εικόνα 3.1: Εικόνα ως 3-Δ επιφάνεια. Στην (3.1.1) παρουσιάζεται η εικάνα ενός θολωμένου ορθογωνιού και στην (3.1.2) η αναπαράσταση του ως 3-Δ επιφάνεια. Αυτή η αναπαράσταση αντιστοιχεί στη γραφική παράσταση της $-I(x, y)$.

λύγισμα. Αυτή η παραμορφώσιμη καμπύλη ενεργοποιείται με τη εφαρμογή της στη γραφική παράσταση μιας συνάρτησης δύο μεταβλητών (όπως η γραφική παράσταση μιας εικόνας $I(x, y)$ στο χώρο \mathbb{R}^3). Μια τέτοια γραφική παράσταση μπορεί να θεωρηθεί ως μια τρισδιάστατη επιφάνεια, H . Το φίδι ακουμπάει πάνω στην H υπό τη επίδραση κάποιας μορφής βαρυτικής δύναμης g . Με αλλά λόγια, ένα βάρος ανατίθεται στο φίδι ώστε να ακολουθεί την κλίση της επιφανείας H .

Ανάλογα με τον στόχο, τον οποίο έχει μια μέθοδος, η φύση των επιφανειών στις οποίες εφαρμόζονται οι ενεργές καμπύλες είναι διαφορετική. Μια τυπική περίπτωση είναι τα δεδομένα της εικόνας να κατασκευάζουν την επιφάνεια H . Η εικόνα $I(x, y)$ μπορεί να αναπαρασταθεί ως μια 3-Δ επιφάνεια με συντεταγμένες (x, y) στο καρτεσιανό επιπεδοχώρο. Η επιφάνεια $H(x, y)$ μπορεί να αντιστοιχεί στην ένταση της εικόνας, $H = \pm I(x, y)$, ή στις τιμές της αντίθεσης της εικόνας, $H = \nabla I(x, y)$. Στην επιφάνεια που έχει, λοιπόν, προκύψει επιθυμούμε να εφαρμόσουμε την ενεργή καμπύλη και να της επιτρέπουμε να παραμορφώνεται σύμφωνα με την τοπολογία της επιφανείας. Στην εικόνα 2.1 παρουσιάζεται πως μια εικόνα μπορεί να παρασταθεί ως μια 3-Δ επιφάνεια. Το “φίδι” τοποθετείται στην επιφάνεια και κάτω από την επίδραση της βαρύτητας κατευθύνεται προς κοιλάδες της επιφάνειας μέχρι να επιτευχθεί κατάσταση ισορροπίας. Αυτή η συμπεριφορά προκαλείται με την συσχέτιση της καμπύλης με ένα ενεργειακό συναρτησιακό, το οποίο εξαρτάται από το ύφος της επιφανείας στην περιοχή που βρίσκεται το φίδι (παρόμοια με τη δυναμική ενέργεια). Το φίδι, λοιπόν, θα προσπαθήσει να ελαχιστοποιήσει την ενέργεια



3.2.1



3.2.2

Εικόνα 3.2: Παράδειγμα επιφάνειας, η οποία προκύπτει από φιλτράρισμα της αρχικής εικόνας 3.1, ώστε να έχει τα χαρακτηριστικά που επιθυμεί η μέθοδος των ενεργών καμπύλων. Στην εικόνα (3.2.1) έχουμε τη κλίση του θολωμένου ορθογωνίου. Στη (3.2.2) παρουσιάζεται η αναπαράσταση της ως τρισδιάστατη επιφάνεια. Η αναπαράσταση αυτή αντιστοιχεί στη γραφική παράσταση της συνάρτησης $-|\nabla I(x, y)|$. Το “φίδι” στη περίπτωση αυτή θα κατέληγε στο ελάχιστο της γραφικής παράστασης και θα ταυτίζόταν με το περίγραμμα του αρχικού ορθογωνίου.

του με την εύρεση ενός τοπικού ελάχιστου στην γειτονική του περιοχή.

Σημαντικό βήμα στην μέθοδο το ενεργών καμπύλων είναι η δημιουργία μιας επιφάνειας, όπου οι βαθιές κοιλάδες να αντιστοιχούν σε σημαντικά χαρακτηριστικά της συνάρτησης που αναπαριστά η επιφάνεια. Σημαντικά και ιδιαιτέρα ενδιαφέροντα στοιχεία μιας εικόνας είναι οι ακμές και τα περιγράμματα, περιοχές δηλαδή υψηλής αντίθεσης ή περιοχές ομογενείς που διαθέτουν κάποιο κοινό χαρακτηριστικό. Μια κατάλληλη επιφάνεια θα έπρεπε να έχει κοιλάδες στις θέσεις που αντιστοιχούν σε ακμές, περιγράμματα ή σύνορα ομοιομόρφων περιοχών. Γι' αυτό το λόγω, γενικά, η αρχική εικόνα, πρέπει πρώτα να επεξεργαστεί με κατάλληλα φιλτραρίσματα, ώστε να έχει η εικόνα τα παραπάνω κατάλληλα χαρακτηριστικά. Ένα παράδειγμα καταλλήλου φιλτραρίσματος δίνεται στην εικόνα 3.2.

Ένα κρίσιμο σημείο στην μέθοδο των ενεργών καμπύλων είναι το πώς αποκρίνεται σε ένα πιθανό εγκλωβισμό του φιδιού σε ένα τοπικό ελάχιστο ή σε μια κοιλάδα που οφείλεται σε θόρυβο ή κάποιο μη εξεταζόμενο χαρακτηριστικό της εικόνας. Το αρχικό μοντέλο των ενεργών καμπύλων [18] είναι εφοδιασμένο με έμμεσο τρόπο επίλυσης τέτοιων προβλημάτων. Επιπρόσθετες εξωτερικές δυνάμεις θα μετακινήσουν τη καμπύλη από ανεπιθύμητα σημεία. Συμπληρωματικές εξωτερικές δυνάμεις επιτρέπουν ακόμα στη καμπύλη να έλκεται από

συγκεκριμένα σημαντικά χαρακτηριστικά της εικόνας.

Ανακεφαλαιώνοντας, το μοντέλο που έχουμε είναι μια ελαστική καμπύλη τοποθετημένη σε μια τρισδιάστατη επιφάνεια και στην οποία μπορούμε να εφαρμόσουμε εξωτερικές δυνάμεις. Επιπλέον, η καμπύλη έχει “μάζα” και επηρεάζεται από ένα βαρυτικό πεδίο. Η επιφάνεια H , στην οποία εφάπτεται η καμπύλη μπορεί να θεωρηθεί ότι ορίζει την συνάρτηση της βαρυτικής ενέργειας. Έτσι, λοιπόν, η θέση, δηλαδή το συγκεκριμένο ύψος, στην οποία βρίσκεται το φίδι επηρεάζει άμεσα την δυναμική ενέργεια του φιδιού.

Παρατηρούμε ότι το μοντέλο των ενεργών καμπύλων μπορεί να παρομοιαστεί με φυσικά συστήματα. Όπως και σε αυτά, μέρος της δυναμικής ενέργειας του φιδιού μετατρέπεται σε κινητική ενέργεια καθώς κινείται με σκοπό να μειώσει την δυναμική του ενέργεια που οφείλεται στο βαρυτικό πεδίο. Στην συνέχεια η κινητική ενέργεια αποσβένει, με αποτέλεσμα τελικά το φίδι να καταλήγει σε κατάσταση ισορροπίας, η οποία αντιστοιχεί σε μικρότερη δυναμική ενέργεια και άρα μικρότερο ύψος. Η αναλογία αυτή οδήγησε στη περιγραφή αρχικά των ενεργών καμπύλων με μερικές διαφορικές εξισώσεις (PDEs), όπως αυτές προκύπτουν από την Lagrangian μοντελοποίηση της κίνησης [33]. Παρακάτω παρατίθεται αναλυτικά το δυναμικό μοντέλο της κίνησης των καμπύλων στο συνεχή χώρο και στην συνέχεια παρουσιάζεται η διαχριτοποίηση του μοντέλου, η οποία χρησιμοποιείται στις εφαρμογές.

3.2 Το δυναμικό μοντέλο των ενεργών καμπύλων σε συνεχές πεδίο ορισμού

Θεωρούμε την παραμορφώσιμη καμπύλη $v(s, t)$, οπού s χωρική παράμετρος και t χρονική παράμετρος, ορισμένες σε ανοιχτά σύνολα Ω και T αντίστοιχα. Η παραμορφώσιμη αυτή καμπύλη είναι συνάρτηση δυο μεταβλητών, των χωρικών συντεταγμένων x και y :

$$v(s, t) = (x(s, t), y(s, t)) : s \in \Omega, t \in T, \quad (3.1)$$

Η συνάρτηση δυναμικής ενέργειας του “φιδιού”, $E_{\text{snake}}(v)$, ορίζεται ως εξής [18]:

$$E_{\text{snake}}(v) = \frac{1}{2} \int_{\Omega} [E_{\text{int}}(v) + E_{\text{ext}}(v) + E_{\text{field}}(v)] ds, \quad (3.2)$$

Η E_{int} παριστάνει την εσωτερική δυναμική ενέργεια της καμπύλης. Είναι συνάρτηση, των δυνάμεων που τείνουν να λυγίσουν την καμπύλη αλλά και

των δυνάμεων που τείνουν να την τεντώσουν. Η ενέργεια E_{ext} οφείλεται στις εξωτερικές δυνάμεις και η E_{field} στο βαρυτικό πεδίο.

Η εσωτερική δυναμική ενέργεια, E_{int} , ορίζεται ως εξής:

$$E_{int}(v(s)) = a(s)|v_s|^2 + b(s)|v_{ss}|^2, \quad (3.3)$$

όπου $v_s \equiv \frac{\partial v}{\partial s}$ και $v_{ss} \equiv \frac{\partial^2 v}{\partial s^2}$. Ο πρώτος όρος $a(s)|v_s|^2$ κάνει το φίδι να αντιστέκεται στο τέντωμα, όπως ένα σχοινί, ενώ ο δεύτερος όρος $b(s)|v_{ss}|^2$ το κάνει να αντιστέκεται στο λύγισμα, όπως μια βέργα. Η παράμετρος $a(s)$ ρυθμίζει την τάση μεταξύ των σημείων του φιδιού, ενώ η παράμετρος $b(s)$ ρυθμίζει την ακαμψία. Με την ανάθεση της μηδενικής τιμής σε μια από τις δύο παραμέτρους σε κάποιο σημείο εισάγονται στο μοντέλο δύο ειδών ασυνέχειες, ασυνέχεια της συνάρτησης της καμπύλης, ή ασυνέχεια της πρώτης παραγωγού της συνάρτησης της καμπύλης [20]. Αν θέσουμε $b(s_0) = 0$ αφαιρούμε κάθε περιορισμό ακαμψίας στο σημείο s_0 , έτσι στο σημείο αυτό εισάγεται δυνατότητα ύπαρξης ασυνέχειας της εφαπτομένης της καμπύλης. Επιπλέον, εάν θέσουμε $a(s_0) = 0$ και $b(s_0) = 0$ κανένας εσωτερικός περιορισμός δεν υφίσταται στο s_0 και ως αποτέλεσμα προκύπτει ασυνέχεια στο συγκεκριμένο σημείο της καμπύλης.

Η εξωτερική δυναμική ενέργεια E_{ext} δημιουργείται με τέτοιο τρόπο, ώστε να παίρνει μικρότερες τιμές στα σημεία της εικόνας, που αντιστοιχούν σε χαρακτηριστικά, όπως τα σύνορα των αντικειμένων. Δοσμένης, λοιπόν, της εικόνας $I(x, y)$ σε κλίμακες του γκρι, συνήθως οι εξωτερικές δυνάμεις σχεδιάζονται ώστε να οδηγούν τις ενεργές καμπύλες προς τις ακμές της εικόνας. Μερικές χαρακτηριστικές ενέργειες είναι οι εξής:

$$\begin{aligned} E_{ext}^1 &= -|\nabla I(x, y)|^2 \\ E_{ext}^2 &= -|\nabla(G_\sigma(x, y) * I(x, y))|^2 \end{aligned}$$

οπού $G_\sigma(x, y)$ είναι η δυσδιάστατη γκαουσιανή (Gaussian) συνάρτηση με τυπική απόκλιση σ και ∇ είναι ο τελεστής κλίσης. Αν η εικόνα είναι ασπρόμαυρη, τότε κατάλληλες εξωτερικές ενέργειες είναι:

$$\begin{aligned} E_{ext}^3 &= I(x, y) \\ E_{ext}^4 &= G_\sigma(x, y) * I(x, y) \end{aligned}$$

Τα πεδία δυνάμεων που προκύπτουν δεν είναι ιδιαίτερα αποτελεσματικά σε συγκεκριμένες καταστάσεις, για το λόγο αυτό οι Chenyang Xu και Jerry L.

Prince πρότειναν το πεδίο δυνάμεων “Gradient Vector Flow”, το οποίο θα αναλύσουμε σε επόμενο κεφάλαιο.

Η δυναμική ενέργεια του βαρυτικού πεδίου, E_{field} , παράγεται όπως και στα φυσικά συστήματα από την εξίσωση:

$$E_{field} = \mu G z(v(s, t)), \quad (3.4)$$

Οπού μ είναι η πυκνότητα μάζας του “φιδιού”, G είναι το μέτρο της επιτάχυνσης της βαρύτητας και $z(v(s, t))$ είναι το ύψος του φιδιού, σε ένα σημείο s τη χρονική στιγμή t , πάνω στην επιφάνεια H .

Έχοντας, λοιπόν, την δυναμική ενέργεια του φιδιού, E_{snake} (εξίσωση 3.2), για μια συγκεκριμένη αρχική θέση μπορεί μια διαδικασία ελαχιστοποίησης να εφαρμοστεί με σκοπό η καμπύλη να καταλήξει σε μια πιο ευσταθή ενεργειακή κατάσταση. Αυτό επιτυγχάνεται μέσω μετατροπής της δυναμικής ενέργειας σε κινητική ενέργεια, η οποία στην συνέχεια μειώνεται λόγω αποσβέσεων[33].

Η διαδικασία ελαχιστοποίησης μπορεί εύκολα να αναλυθεί ως ένα πρόβλημα κλασσικής μηχανικής. Σύμφωνα με την Αρχή της Ελάχιστης Ενέργειας, η κίνηση της παραμόρφωσης καμπύλης κάτω από την επίδραση των διάφορων συντηρητικών δυνάμεων σε ένα χρονικό διάστημα T , περιγράφεται από τις συναρτήσεις $v(s, t)$ για τις οποίες το επικαμπύλιο ολοκλήρωμα I_L :

$$I_L = \int_T L(v) dt, \quad (3.5)$$

ελαχιστοποιείται. Η λαγκρασιάνη (Lagrangian) $L(u)$ ισούται με $L(v) = T(v) - U(v)$, όπου $T(u)$ εκφράζει την κινητική ενέργεια της καμπύλης, v , και $U(v)$ εκφράζει την δυναμική ενέργεια. Η κινητική ενέργεια ορίζεται ως εξής:

$$T(v) = \frac{1}{2} \int_{\Omega} \mu |v_t|^2 ds, \quad (3.6)$$

όπου $v_t \equiv \frac{\partial v}{\partial t}$. Η δυναμική ενέργεια $U(v)$ ταυτίζεται με την ενέργεια του φιδιού $E_{snake}(v)$. Ο συνδυασμός, λοιπόν, των εξισώσεων (3.2), (3.5) και (3.6) δίδει :

$$I_L = \frac{1}{2} \int_T \int_{\Omega} [\mu |v_t|^2 - E_{int}(v(s, t)) - E_{ext}(v(s, t)) - E_{field}(v(s, t))] ds dt. \quad (3.7)$$

Από την μαθηματική ανάλυση είναι γνωστό ότι ένα ακρότατο της I_L πρέπει να ικανοποιεί τις εξισώσεις Euler–Lagrange. Από τις εξισώσεις (3.3), (3.4) και (3.6) καταλήγουμε ότι η λαγκρασιανή είναι συνάρτηση των :

$$L = L(s, t, x(s, t), y(s, t), x_t, y_t, x_s, y_s, x_{ss}, y_{ss}). \quad (3.8)$$

Οι εξισώσεις Euler-Lagrange που προκύπτουν είναι οι εξής:

$$\begin{aligned} \frac{\partial}{\partial x}(L) - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial x_t}(L)\right) - \frac{\partial}{\partial s}\left(\frac{\partial}{\partial x_s}(L)\right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial}{\partial x_{ss}}(L)\right) &= 0, \\ \frac{\partial}{\partial y}(L) - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial y_t}(L)\right) - \frac{\partial}{\partial s}\left(\frac{\partial}{\partial y_s}(L)\right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial}{\partial y_{ss}}(L)\right) &= 0, \end{aligned} \quad (3.9)$$

Στο σύστημα, που ως τώρα έχουμε περιγράψει, οι δυνάμεις που ενεργούν είναι συντηρητικές, αφού προέρχονται από δυναμικά πεδία. Σε τέτοια συστήματα η δυναμική ενέργεια μετατρέπεται εξ' ολοκλήρου σε κινητική ενέργεια, με αποτέλεσμα η ταχύτητα να αυξάνει διαρκώς. Όμως, για να φτάσει η καμπύλη μας σε μια κατάσταση ισορροπίας ($v_t = 0$) που αντιστοιχεί σε μικρότερη δυναμική ενέργεια θα πρέπει η κινητική ενέργεια να αποσβαίνει. Για την περιγραφή ενός τέτοιου συστήματος αρκεί να εισαχθούν στις εξισώσεις κίνησης του συστήματος (εξίσωση 3.9) δυνάμεις τριβής, με μέτρο ανάλογο της ταχύτητας των σημείων της καμπύλης, οι δυνάμεις αυτές είναι μη-συντηρητικές. Οι δυνάμεις τριβής μπορούν να παραχθούν από την εξής συνάρτηση:

$$D(v_t) = \frac{1}{2} \int_{\Omega} \gamma |v_t|^2 ds, \quad (3.10)$$

όπου γ θεωρείται ο συντελεστής τριβής. Έτσι, οι δυνάμεις τριβής που ενεργούν στην ενεργή καμπύλη είναι:

$$\begin{aligned} D_f(x_t) &= -\frac{\partial}{\partial x_t}(D(v_t)) = -\gamma x_t \\ D_f(y_t) &= -\frac{\partial}{\partial y_t}(D(v_t)) = -\gamma y_t \end{aligned} \quad (3.11)$$

Εισάγοντας τις δυνάμεις αυτές στις εξισώσεις Euler-Lagrange της κίνησης, έχουμε:

$$\begin{aligned} \frac{\partial}{\partial x}(L) - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial x_t}(L)\right) - \frac{\partial}{\partial s}\left(\frac{\partial}{\partial x_s}(L)\right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial}{\partial x_{ss}}(L)\right) &= -\gamma x_t, \\ \frac{\partial}{\partial y}(L) - \frac{\partial}{\partial t}\left(\frac{\partial}{\partial y_t}(L)\right) - \frac{\partial}{\partial s}\left(\frac{\partial}{\partial y_s}(L)\right) + \frac{\partial^2}{\partial s^2}\left(\frac{\partial}{\partial y_{ss}}(L)\right) &= \gamma y_t. \end{aligned} \quad (3.12)$$

Η λαγκρασιανή L μπορεί με βάση τις εξισώσεις 3.3,3.4 είναι :

$$L = T - U = \frac{1}{2} \int_{\Omega} [u|v_t|^2 - a(s)|v_s|^2 - b(s)|v_{ss}|^2 - E_{ext} - E_{field}] ds \quad (3.13)$$

Αντικαθιστώντας την λαγκρασιανή στην εξίσωση (3.12) λαμβάνουμε:

$$\mu x_{tt} + \gamma x_t - \frac{\partial}{\partial s}(a(s)x_s) - \frac{\partial^2}{\partial s^2}(b(s)x_{ss}) = -\frac{1}{2}(E_{ext_x}(v) + E_{field_x}(v)),$$

$$\mu y_{tt} + \gamma y_t - \frac{\partial}{\partial s}(a(s)y_s) - \frac{\partial^2}{\partial s^2}(b(s)y_{ss}) = -\frac{1}{2}(E_{ext_y}(v) + E_{field_y}(v)). \quad (3.14)$$

όπου $x_{tt} \equiv \frac{\partial^2 x}{\partial t^2}$, $y_{tt} \equiv \frac{\partial^2 y}{\partial t^2}$, $E_{ext_x} \equiv \frac{\partial}{\partial x}(E_{ext})$, $E_{ext_y} \equiv \frac{\partial}{\partial y}(E_{ext})$, $E_{field_x} \equiv \frac{\partial}{\partial x}(E_{field})$ και $E_{field_y} \equiv \frac{\partial}{\partial y}(E_{field})$. Για την επίλυση των παραπάνω μερικών διαφορικών εξισώσεων είναι αναγκαίος ο καθορισμός αρχικών συνθηκών για $t = 0$ και στα άκρα του διαστήματος Ω .

Οι εξισώσεις (3.14) περιγράφουν το δυναμικό μοντέλο των ενεργών καμπύλων. Συνοφίζοντας, ο όρος mu_{tt} στο πρώτο μέλος των εξισώσεων μοντελοποιεί την αδράνεια, ενώ ο όρος γv_t περιγράφει την απόσβεση, ο πρώτος είναι ανάλογος της επιτάχυνσης της καμπύλης και ο δεύτερος ανάλογος της ταχύτητας. Οι επόμενοι δυο όροι του πρώτου μέλους περιγράφουν την τάση $(a(s)v_s)$ και την ακαμψία $(b(s)v_{ss})$. Ο όρος $a(s)v_s$ είναι ανάλογος του διαστήματος μεταξύ των σημείων του φιδιού (“snaxels”), ενώ ο όρος $b(s)v_{ss}$ είναι ανάλογος της καμπυλότητας του φιδιού. Στο δεύτερο μέλος των εξισώσεων ο πρώτος όρος αναπαριστάνει την επίδραση των εξωτερικών περιορισμών ως συνάρτηση της πρώτης χωρικής παραγωγού της ενέργειας E_{ext} . Ο δεύτερος όρος μοντελοποιεί την επίδραση της κλίσης του επιφάνεια H , που έχει προκύψει από την εικόνα. Έτσι, το πρώτο μέλος των εξισώσεων αναπαριστάνει τις εσωτερικές αλληλεπιδράσεις μεταξύ των σημείων της καμπύλης, ενώ το δεύτερο μέλος τις εξωτερικές δυνάμεις που είναι ανεξάρτητες από την κατάσταση του φιδιού.

Για το λόγο του ότι οι εξισώσεις του μοντέλου που προέκυψαν δεν μπορούν να χρησιμοποιηθούν αυτούσιες σε εφαρμογές, αφού περιγράφονται σε συνεχές πεδίο ορισμού, στην συνέχεια παρουσιάζεται η διακριτοποίηση του μοντέλου, η οποία θα χρησιμοποιηθεί στην εφαρμογή μας.

3.3 Διακριτοποίηση του Μοντέλου

Η λειτουργία των ηλεκτρονικών υπολογιστών βασίζεται στην ψηφιακή τεχνολογία, αυτό έχει ως αποτέλεσμα για την επίλυση ενός προβλήματος με τη

βοήθεια ηλεκτρονικού υπολογιστή να είναι απαραίτητη η διακριτοποίηση του.

Η διακριτοποίηση της περιοχής E του χώρου \mathbb{R}^3 , μέσα στην οποία ορίζεται η επιφάνεια H , γίνεται με τον ορισμό τριών διακριτών συνόλων: \overline{X} και \overline{Y} για τις χωρικές συντεταγμένες, και Z για το ύψος της επιφάνειας. Με το τρόπο αυτό οι μεταβλητές $(x, y, z) \in E$ αντιστοιχούν σε ένα σύνολο διακριτών τιμών $(\bar{x}, \bar{y}, \bar{z}) \in \overline{E}$. Το σύνολο \overline{X} και \overline{Y} αποτελείται από ακέραιους αριθμούς, οι οποίοι παίρνουν τιμές από 0 μέχρι μια μέγιστη τιμή $M_{\overline{X}}$ και $M_{\overline{Y}}$ αντίστοιχα: $\{\bar{x} \in \overline{X} : \overline{X} = \{0, \dots, M_{\overline{X}}\}\}$ και $\{\bar{y} \in \overline{Y} : \overline{Y} = \{0, \dots, M_{\overline{Y}}\}\}$. Το τρίτο σύνολο αποτελείται από πραγματικούς αριθμούς μεταξύ ενός δεδομένου ελάχιστου \overline{Z}_{min} και ενός μέγιστου, \overline{Z}_{max} , έτσι $\{\bar{z} \in \overline{Z} : \overline{Z} = \{\overline{Z}_{min}, \dots, \overline{Z}_{max}\}\}$. Ένα παράδειγμα ενός τέτοιου διακριτού συνόλου \overline{E} είναι η φημιακή φωτογραφία, στην οποία προκύπτει το πεδίο I .

Όπως η περιοχή E διακριτοποιήθηκε, έτσι και οι εξωτερικές δυνάμεις πρέπει να διακριτοποιηθούν χωρικά. Αυτό επιτυγχάνεται χρησιμοποιώντας τις διακριτές χωρικές συντεταγμένες $((\bar{x}, \bar{y}) : \bar{x} \in \overline{X} \text{ και } \bar{y} \in \overline{Y})$ κατά το υπολογισμό των εξωτερικών δυνάμεων.

Το φίδι πρέπει, όμως, να διακριτοποιηθεί και ως προς το χώρο και το χρόνο. Η χωρική διακριτοποίηση γίνεται με το χωρισμό της καμπύλης σε πεπερασμένο αριθμό από κόμβους, προκύπτουν λοιπόν $M_{\bar{s}}$ διακριτά σημεία του φιδιού, τα οποία στην βιβλιογραφία αναφέρονται ως “snaxels”. Οδηγούμαστε, λοιπόν, σε ένα διακριτό σύνολο $\overline{\Omega}(\{\bar{s} \in \overline{\Omega} : \overline{\Omega} = \{0, \dots, M_{\bar{s}} - 1\}\})$. Η χρονική διακριτοποίηση επιτυγχάνεται θεωρώντας ένα διακριτό χρονικό διάστημα \overline{T} , με αρχή το 0. Αποτελείται από χρονικές στιγμές \bar{t} , οι οποίες διαφέρουν κατά $\Delta \bar{t}$, προκύπτει, λοιπόν, $\{\bar{t} \in \overline{T} : \overline{T} = \{0, \Delta \bar{t}, 2\Delta \bar{t}, \dots\}\}$. Έχοντας, διακριτοποίηση και τις δύο παραμέτρους s και t που περιγράφουν το φίδι, η εξίσωση 3.1 παίρνει την μορφή:

$$\bar{v}(\bar{s}, \bar{t}) = (\bar{x}(\bar{s}, \bar{t}), \bar{y}(\bar{s}, \bar{t})). \quad (3.15)$$

Το επόμενο βήμα είναι η διακριτοποίηση των εξισώσεων της κίνησης του φιδιού (εξίσωση 3.14). Όπως και προηγουμένως διακριτοποίηση χρειάζεται και στις χωρικές μεταβλητές αλλά και στις χρονικές. Η μέθοδος διακριτοποίησης θα προσεγγίσει την πρώτη και δεύτερη παραγωγό στις εξισώσεις (3.14).

Αρχικά, διακριτοποιούμε της χρονικές παραγωγούς που υπάρχουν στους δύο πρώτους όρους των εξισώσεων. Καθώς δεν υπάρχει πρωθύστερη γνώση για την μελλοντική θέση του φιδιού, ο υπολογισμός γίνεται την χρονική στιγμή $t = \bar{t} - \Delta \bar{t}$, οπού η θέση είναι γνωστή. Προκύπτουν, λοιπόν, οι εξής εξισώσεις:

$$\mu v_{tt}(s, \bar{t} - \Delta \bar{t}) \approx \frac{\mu}{(\Delta \bar{t})^2} [\bar{v}(\bar{s}, \bar{t}) - 2\bar{v}(\bar{s}, \bar{t} - \Delta \bar{t}) + \bar{v}(\bar{s}, \bar{t} - 2\Delta \bar{t})],$$

$$\gamma v_t(s, \bar{t} - \Delta \bar{t}) \approx \frac{\gamma}{2(\Delta \bar{t})} [\bar{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{s}, \bar{t} - 2\Delta \bar{t})], \quad (3.16)$$

Οι αρχικές θέσεις που αντιστοιχούν στις χρονικές στιγμές $\bar{t} = -\Delta \bar{t}$ και $\bar{t} = -2\Delta \bar{t}$ θεωρούνται γνωστές. Για απλοποίηση αντικαθιστούμε το χρονικό βήμα $\bar{t} = -\Delta \bar{t}$ με 1 [33]. Έτσι καταλήγουμε στην εξίσωση:

$$\mu v_{tt} + \gamma v_t \approx \bar{v}(\bar{s}, \bar{t})[\mu + \frac{\gamma}{2}] + \bar{v}(\bar{s}, \bar{t} - 1)[-2\mu] + \bar{v}(\bar{s}, \bar{t} - 2)[\mu - \frac{\gamma}{2}], \quad (3.17)$$

Από την εξίσωση αυτή καταλήγουμε στο ότι η διακριτοποίηση των χρονικών παραγώγων έχει ως αποτέλεσμα η επίδραση της αδράνειας και των αποσβέσεων να εξαρτάται από την τρέχουσα θέση του φιδιού αλλά και των δυο ακριβώς προηγούμενων θέσεων του.

Στη συνέχεια, διακριτοποιούμε τις παραγώγους ως προς τις χωρικές συντεταγμένες, s , της ενεργής καμπύλης στους δυο τελευταίους όρους του πρώτου μέλους των εξισώσεων 3.14, την χρονική στιγμή $t = \bar{t}$. Υποθέτοντας ότι έχουμε γνώση για τη θέση των γειτονικών σημείων γύρω από κάθε σημείο του φιδιού που εξετάζουμε, μετατρέπουμε τις παραγωγούς σε εξίσωση διαφορών. Στην ανάλυση που ακολουθείται θεωρούμε ότι η καμπύλη είναι κλειστή, δηλαδή το πρώτο σημείο του φιδιού είναι γειτονικό με το τελευταίο. Στο παράρτημα A αναλύεται λεπτομερειακά ο τρόπος εισαγωγής ασυνέχειας που επιτρέπει την εξέλιξη μιας ανοιχτής ενεργής καμπύλης, όπως αυτή που χρησιμοποιούμε στην εφαρμογή “παρακολούθησης χεριού”. Η διακριτοποιημένη εξίσωση που λαμβάνουμε είναι η εξής:

$$\begin{aligned} & -\frac{\partial}{\partial s}(a(s)v_s) - \frac{\partial^2}{\partial s^2}(b(s)v_{ss}) \approx \\ & \frac{1}{(\Delta \bar{s})^2} \{a(\bar{s})[\bar{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{s} - \Delta \bar{s}, \bar{t})]\} + \\ & \frac{1}{(\Delta \bar{s})^2} \{a(\bar{s} + \Delta \bar{s})[\bar{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{s} + \Delta \bar{s}, \bar{t})]\} + \\ & \frac{1}{(\Delta \bar{s})^4} \{b(\bar{s} - \Delta \bar{s})[\bar{v}(\bar{s} - 2\Delta \bar{s}, \bar{t}) - \bar{v}(\bar{s} - \Delta \bar{s}, \bar{t}) + \bar{v}(\bar{s}, \bar{t})]\} - \\ & \frac{2}{(\Delta \bar{s})^4} \{b(\bar{s})[\bar{v}(\bar{s} - \Delta \bar{s}, \bar{t}) - \bar{v}(\bar{s}, \bar{t}) + \bar{v}(\bar{s} + \Delta \bar{s}, \bar{t})]\} + \\ & \frac{1}{(\Delta \bar{s})^4} \{b(\bar{s} + \Delta \bar{s})[\bar{v}(\bar{s}, \bar{t}) - \bar{v}(\bar{s} + \Delta \bar{s}, \bar{t}) + \bar{v}(\bar{s} + 2\Delta \bar{s}, \bar{t})]\}, \end{aligned} \quad (3.18)$$

στην περίπτωση κλειστής καμπύλης προκύπτει η ανάγκη ορισμού των εξής συνοριακών συνθηκών: $\bar{v}(-2\Delta \bar{s}, \bar{t}) = \bar{v}(M_{\bar{s}} - 2\Delta \bar{s}, \bar{t})$, $\bar{v}(-\Delta \bar{s}, \bar{t}) = \bar{v}(M_{\bar{s}} - \Delta \bar{s}, \bar{t})$,

$\bar{v}(0, \bar{t}) = \bar{v}(M_{\bar{s}}, \bar{t})$, $\bar{v}(\Delta \bar{s}, \bar{t}) = \bar{v}(M_{\bar{s}} + \Delta \bar{s}, \bar{t})$, $b(-\Delta \bar{s}) = b(M_{\bar{s}} - \Delta \bar{s})$, $b(M_{\bar{s}}) = b(0)$ και $a(M_{\bar{s}}) = a(0)$. Όπως και στην περίπτωση των χρονικών παραγώγων, ορίζουμε το $\Delta \bar{s} = 1$, για να απλοποιήσουμε την εξίσωση (3.18). Για να επιτύχουμε καλή προσέγγιση των παραγώγων θα πρέπει η απόσταση μεταξύ των διακριτών σημείων του φιδιού, $\Delta \bar{s}$, να παραμένει μικρή και σχετικά σταθερή. Με το τρόπο αυτό τα αριθμητικά λάθη που είναι ανάλογα του $(\Delta \bar{s})^2$ κατά τον υπολογισμό του $\frac{\partial}{\partial s}(a(s)y_s)$ και ανάλογα του $(\Delta \bar{s})^4$ κατά τον υπολογισμό του $\frac{\partial^2}{\partial s^2}(b(s)y_{ss})$ διατηρούνται σχετικά μικρά. Με την αντικατάσταση του $\Delta \bar{s} = 1$ η εξίσωση (3.18) γίνεται:

$$\begin{aligned} -\frac{\partial}{\partial s}(a(s)v_s) - \frac{\partial^2}{\partial s^2}(b(s)v_{ss}) &\approx \\ \bar{v}(\bar{s}-2, \bar{t})b(\bar{s}-1) + & \\ \bar{v}(\bar{s}-1, \bar{t})[-2b(\bar{s}-1) - 2b(\bar{s}) - a(\bar{s})] + & \\ \bar{v}(\bar{s}, \bar{t})[b(\bar{s}-1) + 4b(\bar{s}) + b(\bar{s}+1) + a(\bar{s}) + a(\bar{s}+1)] + & \\ v(\bar{s}+1, \bar{t})[-2b(\bar{s}) - 2b(\bar{s}+1) - a(\bar{s}+1)] + & \\ \bar{v}(\bar{s}+2, \bar{t})b(\bar{s}+1), & \end{aligned} \quad (3.19)$$

Οι αντίστοιχες συνοριακές συνθήκες που προκύπτουν είναι: $\bar{v}(-2, \bar{t}) = \bar{v}(M_{\bar{s}} - 2, \bar{t})$, $\bar{v}(-1, \bar{t}) = \bar{v}(M_{\bar{s}} - 1, \bar{t})$, $\bar{v}(0, \bar{t}) = \bar{v}(M_{\bar{s}}, \bar{t})$, $\bar{v}(1, \bar{t}) = \bar{v}(M_{\bar{s}} + 1, \bar{t})$, $b(-1) = b(M_{\bar{s}} - 1)$, $b(M_{\bar{s}}) = b(0)$ και $a(M_{\bar{s}}) = a(0)$. Λαμβάνοντας υπόψη τις παραπάνω οριακές συνθήκες μπορούμε να εκφράσουμε την εξίσωση (3.19) ως ένα σύστημα γραμμικών εξισώσεων για όλα τα \bar{s} σε μορφή πίνακα:

$$-\frac{\partial}{\partial s}(a(s)v_s) - \frac{\partial^2}{\partial s^2}(b(s)v_{ss}) \approx \mathbf{K}\bar{V}(\bar{s}, \bar{t}), \quad (3.20)$$

όπου \mathbf{K} είναι ο πίνακας “ακαμψίας”, και περιέχει πληροφορίες για τις σχέσεις ελαστικότητας που έχουν μεταξύ τους όλα τα διακριτά σημεία του φιδιού (snaxels), $\bar{s} = 0, \dots, M_{\bar{s}} - 1$:

$$\mathbf{K} = \begin{pmatrix} c_0 & \beta_0 & \alpha_0 & & & \alpha_{M_{\bar{s}}-2} & \beta_{M_{\bar{s}}-1} \\ \beta_0 & c_1 & \beta_1 & \alpha_1 & & & \alpha_{M_{\bar{s}}-1} \\ \alpha_0 & \beta_1 & c_2 & \beta_2 & \alpha_2 & & \\ \alpha_1 & \beta_2 & c_3 & \beta_3 & \alpha_3 & & \\ & & & & \ddots & & \\ & & & \alpha_{M_{\bar{s}}-5} & \beta_{M_{\bar{s}}-4} & c_{M_{\bar{s}}-3} & \beta_{M_{\bar{s}}-3} & \alpha_{M_{\bar{s}}-3} \\ & & & & \alpha_{M_{\bar{s}}-4} & \beta_{M_{\bar{s}}-3} & c_{M_{\bar{s}}-2} & \beta_{M_{\bar{s}}-1} \\ \alpha_{M_{\bar{s}}-2} & & & & \alpha_{M_{\bar{s}}-3} & \beta_{M_{\bar{s}}-2} & c_{M_{\bar{s}}-1} & \\ \beta_{M_{\bar{s}}-1} & \alpha_{M_{\bar{s}}-1} & & & & & & \end{pmatrix} \quad (3.21)$$

Ο πίνακας αυτός είναι συμμετρικός και “pentadiagonal” κάτι που αποδεικνύεται ιδιαιτέρα χρήσιμο κάτι την επίλυση των διακριτών εξισώσεων κίνησης της ενεργής καμπύλης. Οι παράμετροι του πίνακα σχετίζονται με τις παραμέτρους των εξισώσεων κίνησης ως εξής:

$$\begin{aligned}\alpha_{\bar{s}} &= b(\bar{s} + 1), \\ \beta_{\bar{s}} &= -2b(\bar{s}) - 2b(\bar{s} + 1) - a(\bar{s} + 1), \\ c_{\bar{s}} &= b(\bar{s} - 1) + 4b(\bar{s}) + b(\bar{s} + 1) + a(\bar{s}) + a(\bar{s} + 1).\end{aligned}\quad (3.22)$$

Τέλος, το διάνυσμα \bar{V} αποτελείται από τις θέσεις των snaxels:

$$\bar{V}(\bar{s}, \bar{t}) = (v(0, \bar{t}) \bar{v}(1, \bar{t}) \dots \bar{v}(M_{\bar{s}-1}, \bar{t}))^T \quad (3.23)$$

Από το παραπάνω σύστημα εξισώσεων παρατηρούμε ότι η διακριτοποίηση επηρεάζει το τρόπο αλληλεπίδρασης των σημείων της καμπύλης. Κάθε διακριτό σημείο \bar{s} είναι άμεσα συνδεδεμένο με τέσσερα γειτονικά σημεία του φιδιού ($\bar{s} - 2, \bar{s} - 1, \bar{s} + 1$ και $\bar{s} + 2$). Τα βάρη που δίνονται σε κάθε σύνδεση μέσω των a και b είναι συμμετρικά κατανεμημένα γύρω από κάθε σημείο. Η συμμετρία αυτή επιτρέπει την χρήση γρήγορων μεθόδων για την επίλυση του συστήματος εξισώσεων. Ακόμα, λόγω της τοπικής φύσης των αλληλεπιδράσεων είναι ιδιαιτέρα εύκολη η εισαγωγή ασυνεχειών στη ενεργή καμπύλη, σπάζοντας μετρικούς από τους γειτονικούς συνδέσμους. Δυο είδη ασυνέχειας μπορούν να εισαχθούν στο μοντέλο των ενεργών καμπύλων [20]: είτε ασυνέχεια της συνάρτησης εφαπτομένης της καμπύλης, δηλαδή ασυνέχεια της πρώτης παραγώγου της συνάρτησης της καμπύλης, είτε ασυνέχεια θέσης, δηλαδή μη αλληλεπίδραση μεταξύ δύο διαδοχικών σημείων της καμπύλης, με αποτέλεσμα το σπάσιμο της καμπύλης μεταξύ των δύο αυτών σημείων. Ασυνέχεια της συνάρτησης εφαπτομένης της καμπύλης επιτυγχάνεται με τη μη εισαγωγή ακαμψίας ($b(\bar{s})$) στους συνδέσμους ενός δεδομένου snaxel. Ασυνέχεια θέσης επιτυγχάνεται με την μη εισαγωγή ακαμψίας αλλά και τάσης ($b = a = 0$) στους συνδέσμους μεταξύ δύο δοσμένων snaxels. Η ασυνέχεια της συνάρτησης εφαπτομένης επιτρέπει στο φίδι να δημιουργεί γωνίες, ενώ η ασυνέχεια θέσης επιτρέπει τη δημιουργία ανοιχτής καμπύλης. Οι ασυνέχειες αυτές πραγματοποιούνται με την μεταβολή κάποιων συγκεκριμένων στοιχείων του πίνακα “ακαμψίας” K . Λεπτομερειακή ανάλυση της μεταβολής αυτής γίνεται στο Παράρτημα A και αυτή η μορφή πίνακα K τελικά χρησιμοποιείται στην εφαρμογή μας.

Αυτό που απομένει είναι η διακριτοποίηση του δευτέρου μέλους της εξισώσης κίνησης (3.14). Είναι αναγκαία η διακριτοποίηση των παραγώγων των

εξωτερικών δυνάμεων και των δυνάμεων του δυναμικού πεδίου. Επειδή, όμως, όπως ήδη έχουμε αναφέρει δεν έχουμε πρωθύστερη γνώση για τη μελλοντική θέση της καμπύλης, μπορούμε μόνο να προσεγγίσουμε τις παραγώγους με βάση την τελευταία γνωστή θέση του φιδιού, τη χρονική στιγμή $\bar{t} - 1$. Το δεύτερο μέλος της εξίσωσης του κίνησης γίνεται:

$$\begin{aligned} E_{ext_x}(v(s, t)) + E_{field_x}(v(s, t)) &\approx E_{ext_x}(\bar{v}(\bar{s}, \bar{t} - 1)) + E_{field_x}(\bar{v}(\bar{s}, \bar{t} - 1)), \\ E_{ext_y}(v(s, t)) + E_{field_y}(v(s, t)) &\approx E_{ext_y}(\bar{v}(\bar{s}, \bar{t} - 1)) + E_{field_y}(\bar{v}(\bar{s}, \bar{t} - 1)). \end{aligned} \quad (3.24)$$

Καταλήγουμε, λοιπόν, στη διακριτή μορφή των εξισώσεων κίνησης της ενεργής καμπύλης. Το αποτέλεσμα είναι δυο σύνολα $M_{\bar{s}}$ γραμμικών εξισώσεων:

$$\begin{aligned} A\bar{X}(\bar{s}, \bar{t}) &= B_{\bar{x}}(\bar{s}, \bar{t} - 1, \bar{t} - 2), \\ A\bar{Y}(\bar{s}, \bar{t}) &= B_{\bar{y}}(\bar{s}, \bar{t} - 1, \bar{t} - 2), \end{aligned} \quad (3.25)$$

όπου στο πρώτο μέλος οι όροι ισούνται με :

$$\begin{aligned} A &= \left[\left(\frac{\gamma}{2} + \mu \right) \mathbf{I} \right] + \mathbf{K}, \\ \bar{X}(\bar{s}, \bar{t}) &= (\bar{x}(0, \bar{t}), \bar{x}(1, \bar{t}), \dots, \bar{x}(M_{\bar{s}-1}, \bar{t}))^T, \\ \bar{Y}(\bar{s}, \bar{t}) &= (\bar{y}(0, \bar{t}), \bar{y}(1, \bar{t}), \dots, \bar{y}(M_{\bar{s}-1}, \bar{t}))^T, \end{aligned}$$

Παρατηρούμε ότι ο πίνακας A έχει τα ίδια μη διαγώνια στοιχεία με το K , ενώ η μόνη διαφορά στη κύρια διαγώνιο είναι ο όρος $[\frac{\gamma}{2} + \mu]$, ο οποίος προστίθεται. Για αυτό, ο πίνακας A έχει τις ίδιες ιδιότητες με το πίνακα K , είναι δηλαδή συμμετρικός και “pentadiagonal”. Επιπλέον, μπορεί να δειχθεί ότι ο A είναι θετικά ορισμένος. Στις εξισώσεις (3.25) τα στοιχεία του δευτέρου μέλους $B_{\bar{x}}$ και $B_{\bar{y}}$ ορίζονται ως εξής:

$$\begin{aligned} B_{\bar{x}}(\bar{s}, \bar{t} - 1, \bar{t} - 2) &= \\ &- \frac{1}{2} (E_{ext_x}(\bar{v}(\bar{s}, \bar{t} - 1)) + E_{field_x}(\bar{v}(\bar{s}, \bar{t} - 1))) + \\ &[2\mu]\bar{x}(\bar{s}, \bar{t} - 1) + [\frac{\gamma}{2} - \mu]\bar{x}(\bar{s}, \bar{t} - 2), \end{aligned}$$

$$\begin{aligned} B_{\bar{y}}(\bar{s}, \bar{t} - 1, \bar{t} - 2) &= \\ &- \frac{1}{2} (E_{ext_y}(\bar{v}(\bar{s}, \bar{t} - 1)) + E_{field_y}(\bar{v}(\bar{s}, \bar{t} - 1))) + \\ &[2\mu]\bar{y}(\bar{s}, \bar{t} - 1) + [\frac{\gamma}{2} - \mu]\bar{y}(\bar{s}, \bar{t} - 2), \end{aligned}$$

με $\bar{s} = 0, \dots, M_{\bar{s}} - 1$. Η λύση του συστήματος εξισώσεων 3.25 μας δίνει την νέα θέση του φιδιού τη χρονική στιγμή \bar{t} :

$$\bar{X} = A^{-1}B_{\bar{x}}, \bar{Y} = A^{-1}B_{\bar{y}} \quad (3.26)$$

3.4 Παράμετροι της Ενεργής Καμπύλης

Στο σημείο αυτό αναλύουμε την επίδραση των διαφόρων παραμέτρων που υπάρχουν στις διακριτές εξισώσεις κίνησης της ενεργής καμπύλης[18].

Παρατηρώντας τις διακριτές εξισώσεις κίνησης (εξισώσεις 3.25), αντιλαμβανόμαστε ότι οι παράμετροι μ και γ έχουν ως αποτέλεσμα την εισαγωγή μνήμης στο μοντέλο, με το να εκτιμούν την νέα θέση του φιδιού $\bar{V}(\bar{t})$ ανάλογα με τις δυο αμέσως προηγούμενες θέσεις του $\bar{V}(\bar{t}-1)$ και $\bar{V}(\bar{t}-2)$. Το μ και το γ είναι θετικοί ή μηδενικοί πραγματικοί αριθμοί: $\mu, \gamma \geq 0 : \mu, \gamma \in \mathbb{R}$. Με το να μεταβάλλουμε τις τιμές των παραμέτρων αυτών, μεταβάλλουμε τη βαρύτητα που έχουν οι προηγούμενες θέσεις της καμπύλης. Αν θέσουμε και τις δυο αυτές παραμέτρους ίσες με το μηδέν απαλείφουμε την ιδιότητα της μνήμης από το μοντέλο, το φίδι κινείται μόνο βάσει των εσωτερικών αλληλεπιδράσεων και των εξωτερικών δυνάμεων. Αξίζει να σημειωθεί ότι και το μ και το γ τα ορίζουμε στην εφαρμογή μας να είναι σταθερά για όλα τα διακριτά σημεία του φιδιού, όπως και στο πρωταρχικό μοντέλο του Kass et al., ώστε να απλοποιείται η επίλυση των εξισώσεων κίνησης. Όμως, σε ένα περισσότερο γενικό μοντέλο ενεργών καμπύλων θα μπορούσαμε και η πυκνότητα μάζας μ να μην είναι σταθερή αλλά και ο συντελεστής τριβής γ να μην είναι σταθερός, με το τρόπο αυτό θα μοντελοποιούσαμε μια καμπύλη η οποία θα αποτελείται από διαφορετικά υλικά και κάθε διακριτό σημείο της θα είχε τις δικές του ιδιότητες. Για παράδειγμα, ο συντελεστής τριβής θα μπορούσε να είναι συνάρτηση του s και να επιβραδύνει κάθε snaxel ανάλογα με την δικιά του ταχύτητα. Το βασικό μειονέκτημα είναι ότι δεν θα μπορούσε να γίνει εύκολα παραγοντοποίηση των εξισώσεων κίνησης για όλα τα s . Ακόμα, η μη δυνατότητα παραγοντοποίησης θα εισάγει και άλλες παραγώγους στις εξισώσεις κίνησης, πράγμα που δεν θα εξασφαλίσει την ευστάθεια του μοντέλου, αφού οι νέοι παράγωγοι θα προσεγγιστούν με εξισώσεις διαφορών, που είναι επιπλέον πηγή αστάθειας. Για τους λογούς αυτούς είναι προτιμότερο από την πλευρά πολυπλοκότητας του μοντέλου να διατηρούνται σταθεροί οι παράμετροι μ και γ .

Η επίδραση των μεταβολών στα βάρη της τάσης και της ακαμψίας, a και b , γίνεται κατανοητή λαμβάνοντας υπόψη την ελαχιστοποίηση του ενεργειακού

συναρτησιακού που έχει ανατεθεί στην καμπύλη. Διαισθητικά, λοιπόν, μπορούμε να θεωρήσουμε ότι σκοπός μας είναι να ελαχιστοποιήσουμε ξεχωριστά τα εξής επικαμπύλια ολοκληρώματα:

$$I_a = \int_{\Omega} a(s) |v_s|^2 ds, I_b = \int_{\Omega} b(s) |v_{ss}|^2 ds.$$

Αν $a(s) > 0$, τότε $I_a > 0$ και για να το ελαχιστοποιήσουμε θα πρέπει να μειώσουμε το $|v_s|^2$. Συμπεραίνουμε, λοιπόν, ότι θα πρέπει να μειώσουμε το $\Delta \bar{v}^2$, δηλαδή θα πρέπει να ελαττωθεί η μέση απόσταση των διακριτών σημείων της ενεργής καμπύλης, με αποτέλεσμα η καμπύλη να έχει την τάση να συστέλλεται. Αντίθετα, εάν $a(s) < 0$, τότε $I_a < 0$ και για να ελαχιστοποιηθεί θα πρέπει το $\Delta \bar{v}^2$ να αυξηθεί, επομένως η καμπύλη θα έχει την τάση να διαστέλλεται. Παρόμοια, εάν $b(s) > 0$ τότε $I_b > 0$ και για να το ελαχιστοποιήσουμε θα πρέπει το $|v_{ss}|^2$ να μειωθεί. Αυτό έχει ως αποτέλεσμα να χρειάζεται να μειωθεί η καμπυλότητα του φιδιού¹. Αν $b(s) < 0$ τότε $I_b < 0$ και για να το ελαχιστοποιηθεί το $|v_{ss}|^2$ να μειωθεί. Στην περίπτωση αυτή όμως η επίπτωση στην συμπεριφορά της καμπύλης δεν είναι ευδιάκριτη, αφού η αύξηση του αθροίσματος $x_{ss}^2 + y_{ss}^2$ δεν αντιστοιχεί σε παρόμοια συμπεριφορά της καμπυλότητας του φιδιού.

Φυσικά, στην επίδραση των α και β δεν πρέπει να αμελούμε την εισαγωγή ασυνεχειών. 'Οπως αναφέραμε και προηγουμένως, δύο ειδών ασυνεχειών ορίζονται στο μοντέλο: ασυνέχεια της συνάρτησης εφαπτομένης και ασυνέχεια της συνάρτησης της καμπύλης (ασυνέχεια θέσης). Το πρώτο είδος εμφανίζεται όταν $b(s) = 0$, ενώ το άλλο όταν και $b(s)$ και $a(s)$ είναι μηδέν. Το σημαντικό, όμως, είναι σε ποια σημεία της καμπύλης θα εισαγάγουμε τις ασυνέχειες αυτές. Η ασυνέχεια θέσης μπορεί να εισαχθεί αυθαίρετα μεταξύ οποιωνδήποτε σημείων, ανάλογα με την αρχικοποίηση της καμπύλης και τα χαρακτηριστικά του σχήματος που επιθυμούμε τα σύνορα του να προσεγγιστούν από την ενεργή καμπύλη. Από την άλλη πλευρά, η εισαγωγή ασυνεχειών στην πρώτη παράγωγο της καμπύλης δίνει τη δυνατότητα στο φίδι να ακολουθεί και να παραμορφώνεται κατάλληλα σε γωνίες και ακμές όταν αυτές εμφανίζονται στα αντικείμενα που επιθυμούμε να προσεγγίσουμε τα σύνορα τους. 'Όμως, χρειάζεται ένα ενδιάμεσο στάδιο το οποίο θα δίνει πληροφορίες για το που είναι απαραίτητο η εισαγωγή τέτοιων ασυνεχειών. 'Ενα τέτοιο ενδιάμεσο στάδιο θα μπορούσε να διερευνά την περεταίρω μείωση του ενεργειακού συναρτησιακού

¹Η καμπυλότητα της v δίδεται από το τύπο: $k = \frac{x_s y_{ss} - y_s x_{ss}}{v_s^3}$. Καθώς τα x_{ss} και y_{ss} τείνουν στο μηδέν και η καμπυλότητα k τείνει στο μηδέν

E_{snake} όταν εισάγονται ασυνέχειες εφαπτομένης. Για παράδειγμα, όταν το φίδι φτάνει σε κατάσταση ισορροπίας ($\bar{v}_t = 0$), θα μπορούσε να εξετάζει την καμπυλότητα του φιδιού ψάχνοντας σημεία, όπου το φίδι λυγίζει ιδιαίτερα. Σε αυτά τα σημεία θα μπορούσε να εισαχθεί ασυνέχεια της εφαπτομένης. Όταν το φίδι φτάσει στην νέα κατάσταση ισορροπίας, μπορούμε να διαπιστώσουμε την ορθότητα των νέων παραμέτρων του μοντέλου εάν η ενέργεια του είναι μικρότερη. Οι Blake και Zisserman [4] πρότειναν τον ορισμό μιας συνάρτησης σφάλματος, η οποία πληροφορεί την ενεργή καμπύλη, εάν η πιθανή μεταβολή των παραμέτρων, μειώνει την ενέργεια της καμπύλης.

Ακόμα, οι εξωτερικές δυνάμεις πρέπει να οριστούν κατάλληλα στις περιοχές της εικόνας, ώστε η ενεργή καμπύλη να έλκεται από τα επιθυμητά χαρακτηριστικά της εικόνας και η καμπύλη να μην εγκλωβίζεται σε τοπικά ελάχιστα που δημιουργούνται λόγω της παρουσίας θορύβου στην εικόνα ή άλλων ανεπιθύμητων χαρακτηριστικών. Στο επόμενο κεφάλαιο, θα αναλύσουμε το τρόπο ορισμού των εξωτερικών δυνάμεων. Ένα άλλο σημαντικό θέμα, είναι ο ορισμός της “δυναμικής επιφάνειας” H : η εικόνα I πρέπει να επεξεργαστεί κατάλληλα ώστε οι κοιλάδες και οι κορυφές να αντιστοιχούν σε σημαντικά χαρακτηριστικά της εικόνας, η επεξεργασία αυτή συνήθως αποτελείται από μεθόδους εύρεσης ακμών, ανάλυσης της εικόνας σε πολλαπλά επίπεδα και κατάτμησης της εικόνας.

3.5 Το μοντέλο Ενεργών Καμπύλων: Πλεονεκτήματα και Μειονεκτήματα

Το “φίδι” είναι μια ενεργή καμπύλη με ενσωματωμένο ένα ενεργειακό συναρτησιακό, του οποίου τα ελάχιστα περικλείουν το σύνολο των χαρακτηριστικών που έχουν προκύψει από επεξεργασία σε υψηλότερο επίπεδο [18]. Ένα τέτοιο μοντέλο έχει και πλεονεκτήματα αλλά και μειονεκτήματα κατά την εφαρμογή του σε προβλήματα αναφερόμενα σε συναρτήσεις του χώρου της εικόνας I .

3.5.1 Πλεονεκτήματα

Το μοντέλο των ενεργών καμπύλων, όπως ήδη έχουμε αναφέρει, βασίζεται σε μια διαδικασία ελαχιστοποίησης για την εξαγωγή των χαρακτηριστικών μιας εικόνας, διατηρώντας ταυτόχρονα σχετικές πληροφορίες που μπορούν να

παρουσιάζονται στην εικόνα. Αυτό είναι ριζικά διαφορετικό από άλλες τεχνικές που χρησιμοποιούνται για εύρεση ακμών και συνόρων αντικειμένων, οι περισσότερες των οποίων βασίζονται σε κατωφλιοποίηση, ώστε να οριστικοποιήσουν τις επιθυμητές ακμές. Οι τεχνικές αυτές, σαν αποτέλεσμα, δεν είναι δυναμικές, σε αντίθεση με το μοντέλο των ενεργών καμπύλων, που προσεγγίζει τα σύνορα των αντικειμένων χωρίς να είναι εξειδικευμένο για αυτά.

Άλλο ένα πλεονέκτημα του μοντέλου των ενεργών καμπύλων σε σχέση με άλλες μεθόδους εξαγωγής ακμών και περιγραμμάτων είναι η ιδιότητα της συνεχείας που έχει η συνάρτηση της καμπύλης. Η ιδιότητα αυτή δίνει την δυνατότητα ένα σημείο της καμπύλης να δίνει πληροφορίες για την θέση των γειτονικών του. Αυτό είναι ιδιαιτέρα χρήσιμο όταν η εικόνα είναι θορυβώδης ή είναι σκηνή στο φυσικό περιβάλλον. Σε τέτοιες περιπτώσεις πολλές φορές τα περιγράμματα δεν ορίζονται επακριβώς και απουσιάζουν κομμάτια, είτε λόγω χαμηλής αντίθεσης του αντικειμένου με το περιβάλλον είτε λόγω σημαντικής αλλοίωσης της εικόνας από το θόρυβο. Η εφαρμογή ενός φιδιού σε τέτοια προβλήματα συνήθως περιγράφει αποτελεσματικά το σύνορο του αντικειμένου γεφυρώνοντας τα κενά που φαινομενικά υπάρχουν.

Ακόμα η σαφής περιγραφή των εξωτερικών δυνάμεων και η δυνατότητα μεταβολής των εσωτερικών αλληλεπιδράσεων του μοντέλου με την αλλαγή των ελαστικών ιδιοτήτων ή με την εισαγωγή ασυνεχειών κάνει το μοντέλο των ενεργών καμπύλων ιδιαίτερα ευμετάβλητο και ικανό να συνεργαστεί με μεθόδους επεξεργασίας υψηλοτέρου επιπέδου. Φυσικά, η ύπαρξη των διαφόρων παραμέτρων οδηγεί και σε περιορισμούς στις εφαρμογές των “φιδιών”. Όμως, ο συνδυασμός των φιδιών με άλλες υπολογιστικές τεχνικές γίνεται αναγκαίος σε διάφορες πολύπλοκες εφαρμογές στην Όραση Υπολογιστών.

Λόγω του δυναμικού χαρακτήρα των ενεργών καμπύλων, ένα φίδι είναι κατάλληλο για την παρακολούθηση παραμορφώσεων και κινήσεων ευλύγιστων φυσικών σχημάτων. Το φίδι θα ακολουθεί το ίδιο τοπικό ελάχιστο, δηλαδή το ίδιο περίγραμμα, από χαρέ σε χαρέ σε μια ακολουθία εικόνων, υποθέτοντας ότι οι παραμορφώσεις δεν είναι “μεγάλες”. Η ίδια ιδέα έχει εφαρμοστεί σε μια προσπάθεια επίλυσης του προβλήματος της στερέοφης, όπου ως διαδοχικές εικόνες χρησιμοποιούνται τα δυο χαρέ που δίνονται από τις δυο διαφορετικές κάμερες και ένας επιπλέον όρος ενέργειας ενσωματώνεται για να ερμηνεύσει τις ανομοιότητες μεταξύ των δυο όψεων [18]. Η δυνατότητα παρακολούθησης παραμορφώσεων, βρίσκει ακόμα εφαρμογή σε ανάλυση σε πολλαπλές κλίμακες, όπου τα χαρέ προκύπτουν κατά σειρά από τις χονδροειδείς αναλύσεις μέχρι τις ακριβείς.

Η προσέγγιση προβλημάτων με μοντέλα ενεργών καμπύλων εφαρμόζεται σε διάφορες μεθόδους εξαγωγής περιγραμμάτων, παρακολούθησης περιγραμμάτων, όπως η εφαρμογή αυτής της διπλωματικής εργασίας, στερέοψης, εύρεσης του σχήματος σκελετού αντικειμένου, κατάτμησης εικόνας αλλά και σχεδιασμός κίνησης ρομπότ. Ακόμα, το μοντέλο των ενεργών καμπύλων γενικεύεται και σε περισσότερες διαστάσεις, όπως οι ενεργές επιφάνειες (active surface) [34].

3.5.2 Μειονεκτήματα

Ο τρόπος ορισμού του μοντέλου φυσικά εισάγει εκτός από πλεονεκτήματα και κάποιους περιορισμούς. Στην κατασκευή του μοντέλου έχουν υποτεθεί περιορισμοί ομαλότητας κατά την προσέγγιση της καμπύλης στις ακμές του αντικειμένου, κάτι που μπορεί να μην είναι συμβατό με την πραγματική φύση του αντικειμένου. Ακόμα, το μοντέλο απαιτεί η δυναμική επιφάνεια πάνω στην οποία θα κινηθεί να είναι αρκετά ομαλή, χωρίς για δυο λογούς : το φίδι να είναι ικανό να “βλέπει” την επιθυμητή λύση και οι υπολογισμοί που γίνονται κατά την εξέλιξη του φιδιού να είναι αριθμητικά ευσταθής και καλά ορισμένοι. Όμως και σε περιπτώσεις που υπάρχουν οι περιορισμοί ομαλότητας το φίδι μπορεί να παγιδευτεί σε ένα τοπικό ελάχιστο. Γίνεται, λοιπόν, σαφές ότι το μοντέλο των ενεργών καμπύλων δεν μπορεί να εφαρμοστεί απευθείας σε εικόνες με ανώμαλη υφή, οι οποίες δημιουργούν συνεχώς κορυφές και κοιλάδες, και είναι απαραίτητη η προ-επεξεργασία της εικόνας.

Ένα άλλο μειονέκτημα του μοντέλου είναι το γεγονός ότι χρειάζεται αυστηρή διαδικασία αρχικοποίησης. Η αρχική καμπύλη πρέπει να βρίσκεται αρκετά κοντά στην επιθυμητή λύση ώστε να επηρεάζεται από τα κατάλληλα πεδία δυνάμεων. Προκύπτουν, λοιπόν, περιορισμοί σε εφαρμογές παρακολούθησης αντικειμένων, οι οποίοι για να ξεπεραστούν απαιτείται η καμπύλη να είναι κατάλληλα τοποθετημένη στο αρχικό καρέ και στην συνέχεια η κίνηση του αντικειμένου να μην είναι ιδιαίτερα γρήγορη.

Το γεγονός ότι το μοντέλο των ενεργών καμπύλων βασίζεται σε επεξεργασία υψηλότερου επιπέδου της εικόνας ή σε προγενέστερη γνώση της μορφής των δομών που περιέχονται στην εικόνα, προκειμένου να προσδιοριστούν οι τιμές των παραμέτρων του και οι θέσεις των εξωτερικών δυνάμεων, περιορίζει το μοντέλο όταν το πρόβλημα είναι πολύπλοκο και δεν έχουμε περισσότερες γνώσεις για αυτό. Αυτή η κατάσταση είναι συχνή όταν αντιμετωπίζουμε το πρώτο καρέ από ένα σύνολο διαδοχικών εικόνων. Η εφαρμογή του μοντέλου

των ενεργών καμπύλων είναι περισσότερο συμβατή σε συγκεκριμένα προβλήματα όπου γνωρίζουμε την ακριβή φύση του αντικειμένου που θέλουμε να προσεγγίσουμε.

Κεφάλαιο 4

Πεδίο Εξωτερικών δυνάμεων

Στο προηγούμενο κεφάλαιο αναλύθηκε η μέθοδος των ενεργών καμπύλων, χωρίς όμως να περιγραφεί λεπτομερειακά ο τρόπος ορισμού των εξωτερικών δυνάμεων. Η θέση, η κατεύθυνση, το μέτρο των εξωτερικών δυνάμεων κρίνεται σε σημαντικό βαθμό την αποτελεσματικότητα της μεθόδου. Τα προβλήματα της αρχικοποίησης, της μη ακριβής εύρεσης των συνόρων του αντικειμένου, σε περιπτώσεις απότομων κοιλάδων μπορούν να επιλυθούν σε ικανοποιητικό βαθμό με τον ορισμό κατάλληλου πεδίου εξωτερικών δυνάμεων.

Το πρόβλημα της αρχικοποίησης προσπάθησαν να αντιμετωπίσουν διάφορες μέθοδοι, όπως η μέθοδος πολλαπλών-αναλύσεων (multiresolution)[3], οι δυνάμεις πίεσης και τα δυναμικά αποστάσεων (distance potentials)[9]. Όμως η πλειονότητα των μεθόδων αυτών επιλύουν ένα μέρος του προβλήματος και ταυτόχρονα δημιουργούν νέες δυσκολίες κατά την εφαρμογή. Για παράδειγμα, στη μέθοδο πολλαπλών αναλύσεων επιλύεται το πρόβλημα της αρχικής απόστασης της καμπύλης από το αντικείμενο, αλλά ο τρόπος με τον οποίο το φίδι θα κινηθεί από την μια ανάλυση στην άλλη παραμένει προβληματικός. Ακόμα στην μέθοδο των δυνάμεων πίεσης, πρέπει να οριστεί η κατεύθυνση πίεσης (προς το εξωτερικό ή το εσωτερικό της καμπύλης), το οποίο σημαίνει, ότι πρέπει να έχουμε γνώση για την σχέση της αρχικής θέσης του φιδιού και την θέση του αντικειμένου, του οποίου επιθυμούμε τα σύνορα να προσδιοριστούν.

Τα πλεονεκτήματα που προκύπτουν από την χρήση του GVF, είναι ότι η εξέλιξη της ενεργής καμπύλης δεν επηρεάζεται τόσο από την αρχικοποίηση της και έχει την ικανότητα να κινηθεί εντός κοιλοτήτων που δημιουργούνται από τα σύνορα, όπως για παράδειγμα μεταξύ των δακτύλων του χεριού. Σε περιπτώσεις που το αντικείμενο είναι απλοϊκού σχήματος, όπως απλά γεωμετρικά σχήματα,- τετράπλευρο, κύκλος, έλλειψη - η αρχικοποίηση μπορεί να γίνει

εντός / εκτός του σχήματος, ή ακόμα και τέμνοντας αυτό. Στην εφαρμογή μας, όμως, πρέπει να είμαστε ιδιαίτερα προσεκτικοί κατά την αρχικοποίηση γιατί το σχήμα της παλάμης και οι παραμορφώσεις που μπορεί να λάβει είναι αρκετά πολύπλοκη και παρά την βελτίωση της συμπεριφοράς του μοντέλου, με μη κατάλληλη αρχικοποίηση, η καμπύλη αδυνατεί να ταυτιστεί με το περιγραμμα. Επιπλέον, σε αντίθεση με τις δυνάμεις πίεσης, δεν είναι απαραίτητη η γνώση για το εάν πρέπει να συσταλεί ή να διασταλεί η καμπύλη, ώστε να ταυτιστεί με τις ακμές. Ακόμα, το φίδι μπορεί να τοποθετηθεί αρκετά μακριά από το αντικείμενο και χωρίς να χρειάζεται να θολώσει την εικόνα να κινηθεί κατάλληλα.

4.1 Κλασσικό Μοντέλο Ενεργών Καμπύλων

Σύμφωνα με το κλασσικό μοντέλο ενεργών καμπύλων που περιγράφεται στο Κεφάλαιο 2, το φίδι, το οποίο ελαχιστοποιεί το ενεργειακό συναρτησιακό, πρέπει να ικανοποιεί τη παρακάτω εξίσωση Euler.²

$$ax''(s) - bx'''(s) - \nabla E_{ext} = 0, \quad (4.1)$$

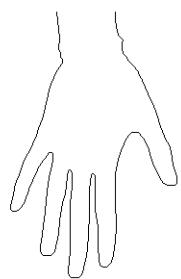
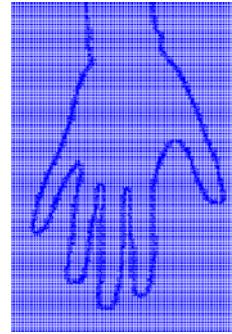
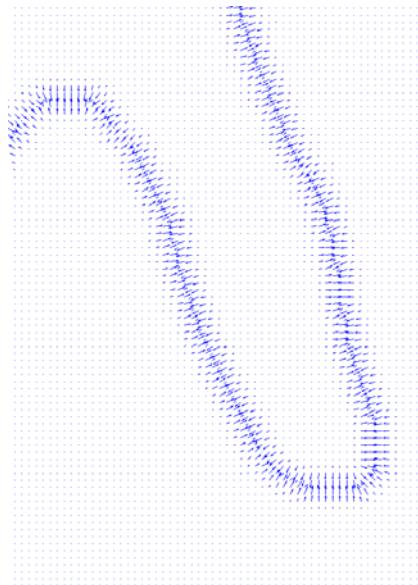
Η εξίσωση αυτή μπορεί να ερμηνευθεί ως μια εξίσωση ισορροπίας δυνάμεων:

$$F_{int} + F_{ext} = 0 \quad (4.2)$$

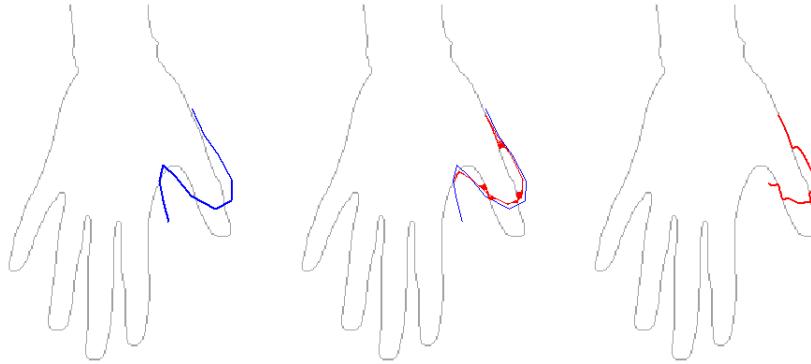
όπου $F_{int} = ax''(s) - bx'''(s)$ και $F_{ext} = -\nabla E_{ext}$. Όπως φαίνεται και στην εικόνα 4.1.2, το πεδίο των εξωτερικών δυνάμεων έχει κατεύθυνση προς τις ακμές των αντικειμένων. Το συγκεκριμένο πεδίο δυνάμεων δημιουργείται βρίσκοντας την κλίση της εικόνας 4.1.1, η οποία είναι η τοπολογία των ακμών της εικόνας που επιθυμούμε να επεξεργαστούμε, αφού όμως πρώτα έχουμε συνέλιξη της εικόνας αυτή με Gaussian φίλτρο. Στην συγκεκριμένη περίπτωση $\sigma = 1$ και η εικόνα δεν επηρεάζεται.

Ένα παράδειγμα της συμπεριφοράς του κλασσικού μοντέλου εμφανίζεται στην εικόνα 4.1. Η αρχική θέση της ανοιχτής καμπύλης 4.2.1 τέμνει τον αντίχειρα και είναι τοποθετημένη, όπως φαίνεται, σχετικά κοντά στα σύνορα του. Παρόλα αυτά, η σύγκλιση του φιδιού, η οποία φαίνεται στην εικόνα 4.2.3, δεν είναι ικανοποιητική. Ο λόγος είναι ότι οι εξωτερικές δυνάμεις έχουν διεύθυνση κάθετη προς τις ακμές με αποτέλεσμα η ενεργή καμπύλη να μην εξελίσσεται

² Στην εφαρμογή μας, όπως και στην πλειοψηφία των εφαρμογών των ενεργών καμπύλων οι παράμετροι $a(s)$ και $b(s)$ θεωρούνται σταθεροί

4.1.1 $I(x, y)$ 4.1.2 $-|\nabla I(x, y)|$ 4.1.3 Λεπτομεριακή αναπαράσταση
του Αντίχειρα

Εικόνα 4.1: Στην εικόνα (4.1.1) παρουσιάζεται το ακριβές σύνορο ενός χεριού, η επιφάνεια αυτή θα αποτελέσει την δυναμική επιφάνεια πάνω στην οποία θα κινηθεί το φίδι. Στην (4.1.2) παρουσιάζεται το πεδίο των εξωτερικών δυνάμεων και στην (4.1.3) φαίνεται με λεπτομέρεια ο τρόπος με τον οποίο τοποθετούνται οι δυνάμεις



4.2.1 Αρχική Θέση

4.2.2 Εξέλιξη Καμπύλης

4.2.3 Τελική Θέση

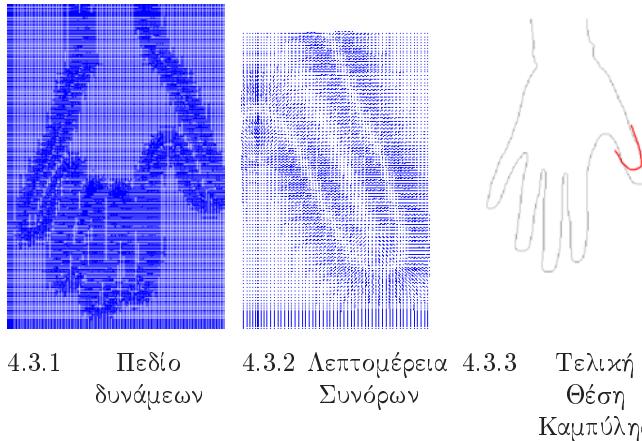
Εικόνα 4.2: Στις εικόνες αυτές παρουσιάζεται η εξέλιξη του “φιδιού” υπό την επίδραση των δυνάμεων $f_{ext} = -|\nabla I(x, y)|$, (4.1)

στην κοιλότητα του δακτύλου. Επιπλέον, ο χώρος γύρω από το αντικείμενο, στο οποίο εμφανίζονται οι εξωτερικές δυνάμεις, είναι περιορισμένος, για αυτό μέρος της καμπύλης δεν ταυτίζεται με το αντικείμενο. Η μεταβολή των παραμέτρων α και b δεν βελτιώνει αισθητά το αποτέλεσμα. Μόνο όταν η ελαστικότητα μεταξύ των σημείων της καμπύλης αυξάνει, τότε παρουσιάζεται το φαινόμενο, το σύνολο της καμπύλης να μετακινείται προς τον αντίχειρα, αλλά και μέρος της να παραμένει στο σύνορο του “δείκτη”.

Το πρόβλημα της ραγδαίας μείωσης του μέτρου των εξωτερικών δυνάμεων μακριά από το αντικείμενο μπορεί να αντιμετωπιστεί αυξάνοντας τη τυπική απόκλιση του Gaussian φίλτρου. Όμως, το φιλτράρισμα αυτό έχει ως άμεσο αποτέλεσμα να θολωθούν οι ακμές και κατά συνέπεια να μην είναι δυνατή η με ακρίβεια προσέγγιση των συνόρων του αντικειμένου. Ακόμα, ο τρόπος αυτός δεν επιλύει το πρόβλημα της εξέλιξης της καμπύλης στο εσωτερικό ενός κοιλώματος, όπως φαίνεται και στην εικόνα (4.3).

4.1.1 Εξωτερικές Δυνάμεις

Οι τελικές θέσεις των ενεργών καμπυλών στις εικόνες (4.2.3) και (4.3.3) ικανοποιούν τις εξισώσεις Euler που έχουν προκύψει για το μοντέλο, όμως αντιστοιχούν σε ένα τοπικό ελάχιστο του ενεργειακού συναρτησιακού και όχι στο γενικό ελάχιστο. Για την επίλυση του προβλήματος αυτού, πολλές μέθοδοι



Εικόνα 4.3: Στην εικόνα (4.3.1) παρουσιάζεται η μεταβολή του δυναμικού πεδίου των εξωτερικών δυνάμεων, όταν το $\sigma = 5\text{pixel}$ στο Gaussian φίλτρο, και στη (4.3.3) η νέα τελική μορφή της εξέλιξης της καμπύλης

έχουν εισάγει στο μοντέλο γενικευμένες μορφές εξωτερικών δυνάμεων και μέσω της εξίσωσης ισορροπίας δυνάμεων καταλήγουν στις επιθυμητές λύσεις.

$$F_{int} + F_{ext}^{(g)} = 0 \quad (4.3)$$

Η επιλογή της εξωτερικής δύναμης $F_{ext}^{(g)}$ διαδραματίζει σπουδαίο ρόλο στην συμπεριφορά του φιδιού. Οι εξωτερικές δυνάμεις μπορούν να χωριστούν σε δυο κατηγορίες: στατικές και δυναμικές. Στατικές χαρακτηρίζονται οι δυνάμεις οι οποίες υπολογίζονται άμεσα από τα δεδομένα της εικόνας και δεν αλλάζουν κατά την διάρκεια της εξέλιξης της καμπύλης. Οι δυνάμεις που περιγράφηκαν παραπάνω είναι στατικές δυνάμεις. Αντίθετα οι δυναμικές δυνάμεις αλλάζουν κατά την εξέλιξη της καμπύλης.

Διάφοροι τύποι δυναμικών εξωτερικών δυνάμεων έχουν αναπτυχθεί με σκοπό την βελτίωση του μοντέλου των ενεργών καμπυλών. Για παράδειγμα οι δυνάμεις που χρησιμοποιούνται στις ενεργές καμπύλες πολλαπλών-αναλύσεων αλλά και οι δυνάμεις πίεσης που προτάθηκαν από το Cohen [10]. Όμως, στην εφαρμογή μας επιλέξαμε τις στατικές εξωτερικές δυνάμεις που προτάθηκαν από τους Chenyang Xu and Jerry L. Prince [38].

Η βάση των δυνάμεων αυτών είναι το θεώρημα του Helmholtz, το οποίο δηλώνει ότι τα πιο γενικά στατικά διανυσματικά πεδία μπορούν να αποσυντεθούν σε δυο όρους, ένας μη περιστροφικός ($\nabla \times F = 0$) και ένας σωληνοειδής ($\nabla \cdot F = 0$). Το πεδίο δυνάμεων που δημιουργείται σύμφωνα με το κλασσικό

μοντέλο των ενεργών καμπυλών είναι ένα στατικό μη περιστροφικό πεδίο, αφού προκύπτει ως η κλίση ενός βαθμωτού δυναμικού πεδίου. Για το λόγο αυτό, ένα περισσότερο γενικό στατικό πεδίο θα πρέπει να επιτρέπει την δυνατότητα ύπαρξης και των δυο όρων, και του μη περιστροφικού και του σωληνοειδούς.

4.2 Διανυσματικό Πεδίο Ροής Κλίσης

Το διανυσματικό πεδίο ροής κλίσης (Gradient Vector Flow Field) ορίζεται ως ένα διανυσματικό πεδίο $\mathbf{v}(x, y) = (u(x, y), v(x, y))$ το οποίο ελαχιστοποιεί το εξής ενεργειακό συναρτησιακό:

$$\mathcal{E} = \int \int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |\mathbf{v} - \nabla f|^2 dx dy \quad (4.4)$$

Η μορφή αυτού του ενεργειακού συναρτησιακού είναι τέτοια ώστε να ομαλοποιεί τα αποτελέσματα στις περιοχές που δεν υπάρχουν δεδομένα, ενώ στις περιοχές των ακμών δημιουργεί μεταβαλλόμενο μέτρο και διεύθυνσης δυνάμεις. Πιο συγκεκριμένα, όταν η κλίση ∇f είναι μικρή, το άθροισμα των τετραγώνων των μερικών διαφορικών των διανυσμάτων του πεδίου υπερισχύει παράγοντας ένα αργά μεταβαλλόμενο πεδίο. Από την άλλη πλευρά, όταν η κλίση ∇f είναι μεγάλη, ο δεύτερος όρος επικρατεί στο ολοκλήρωμα και ελαχιστοποιείται όταν $\mathbf{v} = \nabla f$. Με το τρόπο αυτό παράγουμε το επιθυμητό αποτέλεσμα, δηλαδή να διατηρούμε το \mathbf{v} περίπου ίσο με τη κλίση της εικόνας των ακμών, όταν αυτή είναι μεγάλη, και να μεταβάλλεται αργά στις ομογενείς περιοχές. Η παράμετρος μ κανονικοποιεί το ενεργειακό συναρτησιακό ορίζοντας την σχέση μεταξύ πρώτου και δεύτερου όρου. Η παράμετρος αυτή πρέπει να λαμβάνει τιμές ανάλογα με το ποσοστό θορύβου που παρουσιάζεται στην εικόνα (όσο μεγαλύτερος είναι ο θόρυβος, τόσο πρέπει να αυξήσουμε το μ).

Είναι δυνατό, με την βοήθεια της θεωρίας του λογισμού μεταβολών, να δειχθεί ότι το πεδίο GVF μπορεί να βρεθεί επιλύοντας τις εξής εξισώσεις Euler:

$$\begin{aligned} \mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) &= 0 \\ \mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) &= 0 \end{aligned} \quad (4.5)$$

όπου ∇^2 είναι ο λαπλασιανός τελεστής. Αυτές οι εξισώσεις παρέχουν περισσότερες πληροφορίες για την διαισθητική κατανόηση της δημιουργίας του GVF πεδίου. Παρατηρούμε, λοιπόν, ότι σε ομογενής περιοχές, δηλαδή σε περιοχές όπου η εικόνα $I(x, y)$ είναι σταθερή, ο δεύτερος όρος και των δυο εξισώσεων

είναι μηδέν γιατί η κλίση της $f(x, y)$ είναι μηδέν. Για το λόγο αυτό, στις περιοχές αυτές, τα u και v προσδιορίζονται από μια εξίσωση Laplace, με συνοριακές συνθήκες τις τιμές του πεδίου στα σύνορα των περιοχών. Αυτό εξηγεί γιατί η μέθοδος αυτή επιτυγχάνει οι δυνάμεις να έχουν διεύθυνση προς το εσωτερικό των κοιλωμάτων που εμφανίζονται στα σύνορα του αντικειμένου.

4.2.1 Διακριτοποίηση των Εξισώσεων και Εφαρμογή

Οι εξισώσεις (4.5) μπορούν να επιλυθούν θεωρώντας τα u και v σαν συνάρτηση του χρόνου και λύνοντας τις εξισώσεις:

$$u_t(x, y, t) = \mu \nabla^2 u - (u(x, y, t) - f_x(x, y))(f_x(x, y)^2 + f_y(x, y)^2) \quad (4.6a)$$

$$v_t(x, y, t) = \mu \nabla^2 v - (v(x, y, t) - f_y(x, y))(f_x(x, y)^2 + f_y(x, y)^2) \quad (4.6b)$$

Η λύση σταθερής κατάστασης των παραπάνω εξισώσεων είναι η επιθυμητή λύση για τις εξισώσεις Euler (4.5). Η εξισώσεις αυτές είναι σχετικά εύκολο να επιλυθούν αφού δεν βρίσκονται σε πεπλεγμένη μορφή. Όπως και οι εξισώσεις κίνησης του μοντέλου των ενεργών καμπυλών έτσι και οι εξισώσεις αυτές πρέπει να διακριτοποιηθούν, ώστε να είναι δυνατή η υλοποίηση τους σε αλγόριθμο, σε κάποια γλώσσα υπολογιστή. Αντικαθιστούμε, λοιπόν, τις μεταβλητές x , y και t με τις i , j και n αντίστοιχα και ορίζουμε το διάστημα μεταξύ των pixel να είναι Δx και Δy και ο χρόνος μεταξύ των επαναλήψεων να είναι ίσος με Δt . Έτσι έχουμε:

$$\begin{aligned} u_t &= \frac{1}{\Delta t}(u_{i,j}^{n+1} - u_{i,j}^n) \\ v_t &= \frac{1}{\Delta t}(v_{i,j}^{n+1} - v_{i,j}^n) \\ \nabla^2 u &= \frac{1}{\Delta x \Delta y}(u_{i+1,j} + u_{i,j+1} + u_{i-1,j} + u_{i,j-1} - 4u_{i,j}) \\ \nabla^2 v &= \frac{1}{\Delta x \Delta y}(v_{i+1,j} + v_{i,j+1} + v_{i-1,j} + v_{i,j-1} - 4v_{i,j}) \end{aligned}$$

Οι εξισώσεις (4.6a) και (4.6b) γίνονται:

$$u_{i,j}^{n+1} = (1 - b_{i,j} \Delta t)u_{i,j}^n + r(u_{i+1,j}^n + u_{i,j+1}^n + u_{i-1,j}^n + u_{i,j-1}^n - 4u_{i,j}^n) + c_{i,j}^1 \Delta t \quad (4.7a)$$

$$v_{i,j}^{n+1} = (1 - b_{i,j} \Delta t)v_{i,j}^n + r(v_{i+1,j}^n + v_{i,j+1}^n + v_{i-1,j}^n + v_{i,j-1}^n - 4v_{i,j}^n) + c_{i,j}^2 \Delta t \quad (4.7b)$$

όπου

$$\begin{aligned} r &\equiv \frac{\mu \Delta t}{\Delta x \Delta y} \\ b(x, y) &\equiv f_x(x, y)^2 + f_y(x, y)^2 \\ c^1(x, y) &\equiv b(x, y) f_x(x, y) \\ c^2(x, y) &\equiv b(x, y) f_y(x, y) \end{aligned}$$

Η σύγκλιση της παραπάνω επαναληπτικής μεθόδου είναι εγγυημένη από την θεωρία Αριθμητικής Ανάλυσης. Γνωρίζοντας ότι τα b, c^1, c^2 είναι φραγμένα, η εξίσωση (4.6) είναι ευσταθής όταν το βήμα r παραμένει μικρότερο από $1/4$. Έχοντας τα Δx και Δy ορισμένα και το μ ορισμένο με βάση το τρόπο συσχέτισης των περιοχών της εικόνας, όπως αναφέρθηκε παραπάνω, προκύπτει περιορισμός για το χρονικό βήμα:

$$\Delta t \leq \frac{\Delta x \Delta y}{4\mu} \quad (4.8)$$

Διαισθητικά, αντιλαμβανόμαστε ότι η σύγκλιση επιτυγχάνεται ταχυτέρα σε χαμηλής ανάλυσης εικόνες, στις οποίες η απόσταση μεταξύ των pixel είναι μεγαλύτερες, με αποτέλεσμα τα Δx και Δy να είναι μεγαλύτερα. Ακόμα, όταν το μ είναι μεγάλο, ο ρυθμός σύγκλισης είναι μεγαλύτερος λόγω του ότι θα πρέπει το Δt να διατηρηθεί μικρό, ενώ το πεδίο GVF προβλέπεται να είναι ομαλό. Πιο συγκεκριμένα, ο υπολογισμός του πεδίου GVF σε κώδικα Matlab διαρκεί περίπου 42 δευτερόλεπτα για εικόνα $N = 640 \times 480$ pixel σε Centrino 2GHz. Ο υπολογισμός είναι ιδιαίτερα χρονοβόρος, διότι ο αλγόριθμος απαιτεί αρκετές επαναλήψεις μέχρι να συγκλίνει, πράγμα που είναι ασύμβατο με το τρόπο λειτουργίας της πλατφόρμας Matlab. Η διάρκεια υλοποίησης σε κώδικα Matlab είναι ιδιαίτερα μεγάλη, αν λέβουμε υπόψη και το γεγονός ότι στην εφαρμογή παρακολούθησης χεριού θα πρέπει να υπολογιστεί το πεδίο για όλα τα καρέ. Ο χρόνος, λοιπόν, υλοποίησης του συνολικού αλγορίθμου θα ήταν εξαιρετικά μεγάλος. Η υλοποίηση του αλγόριθμου κατασκευής του GVF σε C μειώνει σημαντικά το χρόνο, το πεδίο για την ίδια εικόνα παράγεται σε περίπου 2 δευτερόλεπτα. Η ταχύτητα του αλγόριθμου μπορεί να μειωθεί περαιτέρω είτε με την χρήση μεθόδου πολλαπλών πλεγμάτων (multigrid methods) είτε με την εύρεση του πεδίου στην περιοχή ενδιαφέροντος και όχι σε ολόκληρη την εικόνα.

4.2.2 Τροποποίηση του μοντέλου Ενεργών Καμπυλών

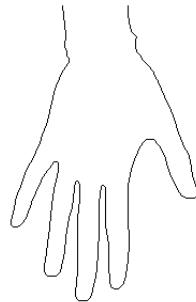
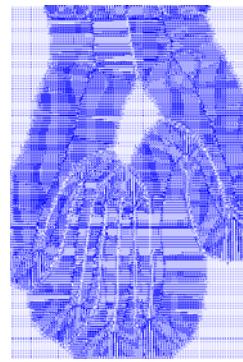
Έχοντας παράγει το πεδίο εξωτερικών δυνάμεων $F_{ext}^{(g)} = \mathbf{v}(x, y)$, η δυναμική εξίσωση κίνησης της ενεργής καμπύλης, που προκύπτει από την εξίσωση (4.1) ελαχιστοποίησης του ενεργειακού συναρτησιακού, τροποποιείται:

$$x_t(s, t) = ax''(s, t) - bx'''(s, t) + \mathbf{v} \quad (4.9)$$

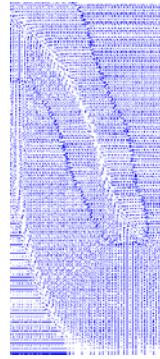
Η επίλυση της δυναμικής αυτής εξίσωσης, δίνει την νέα θέση του φιδιού, το οποίο ονομάζεται GVF snake [38]. Παρότι η τελική θέση του φιδιού θα ικανοποιεί την συνθήκη ισορροπίας δυνάμεων, η εξίσωση (4.9) δεν μπορεί να αναπαραστήσει τις εξισώσεις Euler για την ελαχιστοποίηση του ενεργειακού συναρτησιακού του μοντέλου. Η αιτία είναι ότι το $\mathbf{v}(x, y)$ δεν είναι γενικά μη περιστροφικό πεδίο. Παρόλα αυτά η συμπεριφορά των ενεργών καμπυλών βελτιώνεται αισθητά με την χρήση του πεδίου ροής κλίσης.

4.3 Εξέλιξης της Καμπύλης υπό την επίδραση του Πεδίου Ροής Κλίσης

Στην ενότητα αυτή παρουσιάζεται αρχικά η εικόνα του GVF πεδίου, όπως αυτό δημιουργείται με βάση την εικόνα (4.1.1), που χρησιμοποιήθηκε στο προηγούμενο παράδειγμα παραμόρφωσης του φιδιού. Παρατηρούμε στην εικόνα (4.4.2) ότι στις περιοχές γύρω από κάθε δάκτυλο, δημιουργείται ένα έντονο πεδίο, το οποίο εκτείνεται σε αρκετή απόσταση από το σύνορο κάθε δακτύλου. Παρά το πολύπλοκο της τοπολογίας ακμών της παλάμης, το πεδίο GVF διαχωρίζει τις περιοχές έλξης, είναι σαφώς ορισμένη η περιοχή της εικόνας που τα σημεία της έλκονται από κάποιο δάκτυλο, κάθε σημείο του φιδιού θα έλκεται από ένα συγκεκριμένο σύνορο. Επιπλέον, στην εικόνα (4.4.3) γίνεται φανερό ότι το πεδίο GVF ενεργεί σωστά και στις περιπτώσεις κοιλοτήτων του συνόρου του αντικειμένου. Από την αρχή του αντίχειρα υπάρχουν δυνάμεις που δεν κατευθύνονται κάθεται προς το σύνορο, όπως στα κλασσικά πεδία δυνάμεων, αλλά έχουν κατάλληλες κλίσης, ώστε να οδηγούν την ενεργή καμπύλη και προς το εσωτερικό της κοιλότητας. Στις άσπρες περιοχές της εικόνας (4.4.3) το πεδίο είναι εξαιρετικά αραιό και σε ορισμένα σημεία μηδενικό, λόγω της αλληλεξουδετέρωσης των δυνάμεων που δημιουργούν οι γειτονικές ακμές του αντικειμένου. Τα σημεία αυτά δημιουργούν το διαχωρισμό των περιοχών έλξης, που αναφέραμε προηγουμένως. Γίνεται αντιληπτό ότι το πεδίο δυνάμεων ροής κλίσης, επιλύει σε σημαντικό βαθμό τα πρόβλημα του κλασσικού

4.4.1 $f(x, y)$ 

4.4.2 GVF

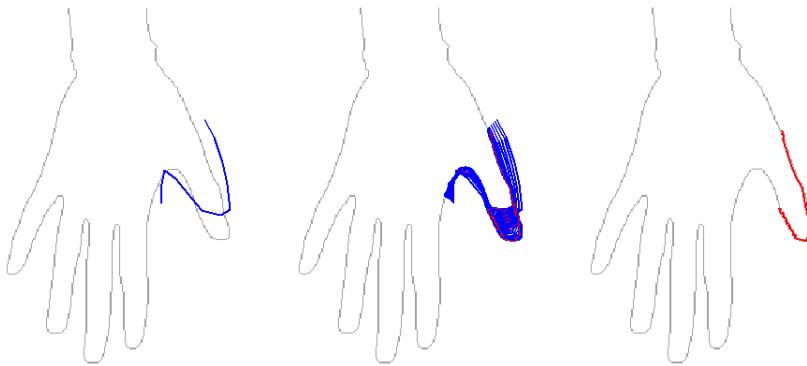


4.4.3 Λεπτομεριακή αναπαράσταση του Αντίχειρα

Εικόνα 4.4: Στην εικόνα (4.4.1) παρουσιάζεται το ακριβές σύνορο ενός χειριού. Στην (4.4.2) παρουσιάζεται το πεδίο GVF και στην (4.4.3) φαίνεται με λεπτομέρεια ο τρόπος με τον οποίο τοποθετούνται οι δυνάμεις

μοντέλου ενεργών καμπυλών: (α) απόσταση αρχικοποίηση της καμπύλης από το επιθυμητό σύνορο, (β) παραμόρφωση της καμπύλης στο εσωτερικό κοιλωμάτων, (γ) διαχωρισμός των περιοχών έλξης.

Το παράδειγμα εξέλιξης της ενεργής καμπύλης στην εικόνα (4.5) υπό την επίδραση του GVF, φανερώνει την αποτελεσματικότητα του πεδίου αυτού σε σχέση με τα αντίστοιχα παραδείγματα εξέλιξης της καμπύλης που παρουσιάστηκαν στις εικόνες (4.2) και (4.3). Πρέπει να τονίσουμε ότι και στις τρεις περιπτώσεις η αρχικοποίηση της καμπύλης ήταν όμοια, οι τιμές των παραμέτρων a και b ήταν ίδιες και ο αριθμός των επαναλήψεων της διαδικασίας παραμόρφωσης ήταν και αυτός ίδιος, πιο συγκεκριμένα $a = 0.5$ και $b = 0$. Δηλαδή, με $b = 0$ επιτρέπουμε στην καμπύλη να παρουσιάζει ασυνέχειες στην



4.5.1 Αρχική Θέση

4.5.2 Εξέλιξη Καμπύλης

4.5.3 Τελική Θέση

Εικόνα 4.5: Στις εικόνες αυτές παρουσιάζεται η εξέλιξη του “φιδιού” υπό την επίδραση του πεδίου *GVF*

συνάρτηση της εφαπτόμενης της. Ακόμα, ο συντελεστής τριβής ορίζεται ίσος με την μονάδα, ενώ το φίδι θεωρείται χωρίς μάζα. Η τελική θέση της ενεργής καμπύλης ταυτίζεται απόλυτα με το σύνορο του δακτύλου και όπως φαίνεται στην εικόνα (4.5.2) το φίδι μετακινείται άμεσα στην κατεύθυνση των ακμών και προς το εσωτερικό του κοιλώματος που δημιουργείται στον αντίχειρα. Αν η ελαστικότητα των σημείων του φιδιού ήταν αυξημένη, δηλαδή η τιμή της παραμέτρου *a* ήταν μικρότερη τότε τα ακραία σημεία του φιδιού θα παράμεναν κοντά στις αρχικές του θέσεις. Όμως, λόγω των αλληλεπιδράσεων μεταξύ των γειτονικών σημείων, τα σημεία που βρίσκονταν αρχικά στο εσωτερικό του δακτύλου έλκουν προς τα έξω και τα υπόλοιπα σημεία. Έτσι, ενώ αρχικά τα ακραία σημεία του φιδιού ταυτίζονται με το σύνορο παλάμης-δακτύλου στην συνέχεια οι αλληλεπιδράσεις μεταξύ των σημείων αναγκάζουν την μια άκρη του φιδιού να κινηθεί μακριά από την παλάμη. Όμως, καθ' όλη την διάρκεια της μετακίνησης αυτής το φίδι περιγράφει τοπικά ελάχιστα της εικόνας. Την κίνηση αυτή του φιδιού θα μπορούσαμε να την παρομοιάσουμε σαν την ροή ενός ποταμού στο εσωτερικό μιας κοιλάδας.

Ένα σημαντικό πρόβλημα, το οποίο προκύπτει με το πεδίο *GVF*, είναι ότι στις περιπτώσεις πολύπλοκων εικόνων, όπου παρουσιάζονται πολλές ακμές που οφείλονται σε θόρυβο ή σε αντικείμενα μειωμένου ενδιαφέροντος, προκύπτουν πολλές περιοχές με μηδενικό πεδίο, αλλά και πολλές φορές το πεδίο κατευ-

θύνεται προς “λανθασμένες” ακμές. Για παράδειγμα, σε εικόνες με έντονο φωτισμό είναι δυνατό το περίγραμμα του νυχιού στα δάκτυλα να είναι ορατό και να αποτελεί μέρος της τοπολογίας ακμών στην εικόνα. Αυτό το φαινόμενο κάνει αδύνατη την ταύτιση της ενεργής καμπύλης με το εξωτερικό σύνορο των δακτύλων. Για το λόγο αυτό στην επόμενη ενότητα αναλύουμε το τρόπο εξαγωγής της τοπολογίας ακμών από την εικόνα και την προσπάθεια μας να απλοποιήσουμε την τοπολογία αυτή, ώστε να προσανατολίζεται στην εφαρμογή παρακολούθησης χεριού.

4.4 Χάρτης ακμών

Η εύρεση των ακμών σε μια εικόνα είναι ένα πρόβλημα που έχει ερευνηθεί εκτενώς και έχουν αναπτυχθεί διάφορες μέθοδοι για την επίλυση του όπως: Roberts Cross, Prewitt, Sobel, Canny, Marr-Hildreth, zero-crossings. Η μέθοδος Canny [8] είναι η πιο συνήθης μέθοδος εύρεσης ακμών και είναι αυτή που χρησιμοποιούμε και στην εφαρμογή μας. Ο Canny αρχικά μελέτησε το σχεδιασμό ενός βέλτιστου φίλτρου ομαλοποίησης της εικόνας για την μετέπειτα εύρεση των ακμών, στην συνέχεια κατέληξε ότι αυτό το φίλτρο θα μπορούσε να προσεγγιστεί από τη πρώτη παραγωγό ενός Gaussian πυρήνα. Ο αλγόριθμος Canny χρησιμοποιεί 4 μάσκες για την εύρεση των οριζόντιων, των καθέτων και των διαγωνίων ακμών. Το μεγαλύτερο αποτέλεσμα που προκύπτει από την συνέλιξη των τεσσάρων αυτών μασκών με την εικόνα αποθηκεύεται, καθώς και η αντιστοιχία του με μια από τις μάσκες. Έτσι για ολόκληρη την εικόνα δημιουργείται ένας χάρτης με δεδομένα, την ένταση της κλίσης σε κάθε σημείο της εικόνας και την διεύθυνση της κλίσης αυτής. Όμως, δεν υπάρχει δεδομένη τιμή έντασης άνω της οποίας ένα σημείο είναι απαραίτητα σημείο ακμής, για αυτό το λόγο η μέθοδος χρησιμοποιεί κατωφλιοπόίηση με υστέρηση. Η υστέρηση με κατωφλιοπόίηση χρειάζεται δύο κατώφλια, ένα ανώτατο και ένα κατώτατο. Κάνοντας την υπόθεση ότι οι “σημαντικές” ακμές θα είναι συνεχής γραμμές, μας επιτρέπει να ακολουθούμε τις ακμές ακόμα και όταν αυτές μειώνονται σε ένταση και ταυτόχρονα να απορρίπτουμε μερικά σημεία της εικόνας με υψηλή ένταση κλίσης, που όμως δεν ανήκουν σε κάποια γραμμή. Έτσι, ο αλγόριθμος αρχικά εφαρμόζει το υψηλό κατώφλι, το οποίο προσδιορίζει τις ακμές που είναι σχεδόν σίγουρο ότι είναι πραγματικές. Έχοντας αυτές τις ακμές και κάνοντας χρήση τις πληροφορίες κατεύθυνσης, που είχαν αποθηκευτεί προηγουμένως, ακολουθούμε τις γραμμές των ακμών, οι οποίες διακόπτονται όταν δεν ικανοποιούνται το χαμηλό κατώφλι που έχει οριστεί.

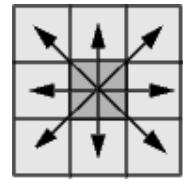
Ο αλγόριθμος Canny περιέχει έναν αριθμό από παραμέτρους, οι οποίοι επηρεάζουν την αποτελεσματικότητα του αλγόριθμου.

- Η τιμή της διακύμανσης του Gaussian φίλτρου: το φίλτρο ομαλοποίησης που χρησιμοποιείται στην πρώτη φάση επηρεάζει τα αποτελέσματα του αλγόριθμου. Μικρότερο φίλτρο δημιουργεί λιγότερη θόλωση και επιτρέπει τη εύρεση μικρών και λεπτών γραμμών. Αντίθετα, ένα μεγάλο φίλτρο προκαλεί θόλωση στην εικόνα, ομαλοποιώντας τις τιμές των pixels σε γειτονικές περιοχές, και έχει σαν αποτέλεσμα να βρίσκονται μόνο οι μεγάλες και ομαλές ακμές της εικόνας.
- Κατωφλιοποίηση: η χρήση δυο κατωφλίων επιτρέπει μεγαλύτερη προσαρμοστικότητα σε σχέση με την χρήση ενός μόνο κατωφλίου, αλλά δημιουργεί παρόμοια προβλήματα. Αν το ανώτατο κατώφλι είναι ιδιαίτερα υψηλό μπορεί να χαθεί πολύτιμη πληροφορία, αν όμως τεθεί χαμηλά μπορεί να ονομάσει πληροφορίες ως ακμές χωρίς αυτές να είναι. Είναι, λοιπόν, αδύνατο να οριστεί ένα κατώφλι το οποίο να λειτουργεί εξίσου καλά για όλες τις εικόνες.

4.4.1 Απλοποίηση του Χάρτη Ακμών

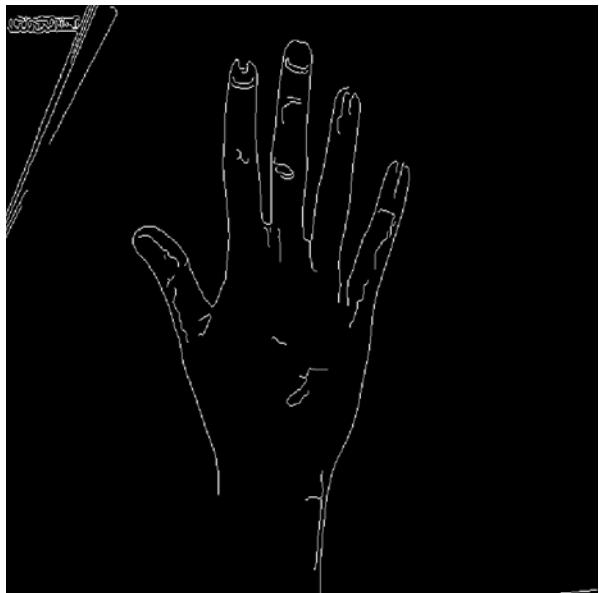
Το πεδίο των εξωτερικών δυνάμεων, GVF, όπως αναφέραμε, δημιουργείται με βάση το χάρτη ακμών της εικόνας $f(x, y)$. Γίνεται, λοιπόν, σαφής η σημαντικότητα του προσδιορισμού του κατάλληλου χάρτη ακμών, που θα παράγει ένα τέτοιο πεδίο εξωτερικών δυνάμεων, που θα καθιστά δυνατή την εξέλιξη της καμπύλης, ώστε να μπορέσει να ταυτιστεί με το σύνορο του χεριού. Μια πιθανή ακμή στο εσωτερικό ενός δάκτυλου μπορεί να οδηγήσει σε λανθασμένα συμπεράσματα για την θέση των δακτύλων και κατά συνέπεια της χειρονομίας που κάνει ο χρηστής της εφαρμογής. Για το λόγο αυτό αναπτύσσουμε έναν επιπλέον αλγόριθμο για την επεξεργασία του χάρτη ακμών.

Ο χάρτης ακμών που παράγεται από τον αλγόριθμο του Canny είναι μια δυαδική εικόνα. Αυτό μας επιτρέπει εύκολα να ορίσουμε τα γειτονικά pixel συνελίσσοντας το χάρτη ακμών με τις μάσκες συνεκτικότητας. Στο αλγόριθμο, που αναπτύξαμε, ορίζουμε ότι γειτονικά είναι τα pixel που είτε οι ακμές τους είτε οι γωνίες τους ακουμπούν, δηλαδή 8- συνεκτικότητα. Έχοντας βρει τα pixel του χάρτη ακμών που είναι γειτονικά, ομαδοποιούμε τα γειτονικά pixel δίνοντας τους ένα χαρακτηριστικό όνομα. Με τον τρόπο αυτό έχουμε την δυνατότητα να μετρήσουμε το “μήκος” κάθε ακμής, αφού όπως υποθέτει και ο



Εικόνα 4.6: Μορφή της μάσκας 8-συνεκτικότητας. Τα pixel γειτονεύουν εαν οι ακμές τους ή οι γωνίες του ακουμπούν.

Canny στον αλγόριθμο του οι ακμές δημιουργούν συνεχής γραμμές. Λαμβάνοντας υπόψη ότι το περίγραμμα του χεριού θα έχει ένα ιδιαίτερα μεγάλο μήκος, απλοποιούμε το χάρτη ακμών απαλείφοντας τις ακμές εκείνες που έχουν το μήκος τους, σε σχέση με το μήκος της μεγαλύτερης ακμής, μικρότερο από ένα ορισμένο κατώφλι. Το κατώφλι αυτό και πάλι όπως και στον αλγόριθμο ακμών ορίζεται από τη χρήση και διαφοροποιείται από εικόνα σε εικόνα, βέβαια επειδή οι εικόνες χεριού έχουν ιδιαίτερα χαρακτηριστικά και το κατώφλι έχει μορφή ποσοστού, το κατώφλι αυτό είναι εύρωστο και δεν χρειάστηκε να το αλλάξουμε στα διάφορα πειράματα που θα αναλύσουμε στην συνέχεια.



4.7.1 Αρχική Εικόνα Ακμών



4.7.2 Απλοποιημένη Εικόνα Ακμών

Εικόνα 4.7: Στην εικόνα (4.7.1) παρουσιάζεται ο χάρτης των ακμών όπως δημιουργήθηκε από τον αλγόριθμο Canny, ενώ στην δεύτερη (4.7.2) η απλοποιημένη εικόνα ακμών μετά την επεξεργασία

Κεφάλαιο 5

Στατιστικό Μοντέλο Σχήματος

5.1 Εισαγωγή

Ο τελικός στόχος της 'Ορασης Υπολογιστών είναι η κατανόηση των χαρακτηριστικών της εικόνας. Δηλαδή, η δυνατότητα όχι απλά να εξάγει τις δομές της εικόνας, αλλά να κατανοεί αυτόματα και τι αναπαριστάνουν οι δομές αυτές. Έτσι, ο στόχος της εργασίας αυτής είναι η παρακολούθηση ενός χεριού και η κατανόηση των χειρονομιών που αυτό κάνει. Για την επίτευξη αυτού του στόχου είναι σχεδόν βέβαιο ότι απαιτείται η χρήση κάποιου είδους μοντέλου, το οποίο θα αναπαριστάνει και θα ταυτοποιεί μια συγκεκριμένη δομή του φυσικού κόσμου. Περισσότερο από μια δεκαετία, μέθοδοι όρασης βασιζόμενες σε μοντέλα έχουν εφαρμοστεί επιτυχώς. Αρχικά για την ανακατασκευή απλοϊκών δομών, αλλά τελευταία και πολύπλοκων δομών όπως το πρόσωπο του ανθρώπου. Έχει αποδειχθεί ότι η ανάπτυξη εφαρμογών βασιζόμενων σε μοντέλα είναι εξαιρετικά δύσκολη, ειδικά για την εξαγωγή πληροφοριών από εικόνες που περιέχουν πολύπλοκες και μεταβαλλόμενες δομές, όπως κινούμενα πρόσωπα, κινούμενους ανθρώπους. Σε αυτές τις περιπτώσεις θα πρέπει να είναι εφικτό το μοντέλο να μπορεί να μεταβληθεί, να είναι ικανό να παραμορφωθεί σε "επιτρεπόμενα" σχήματα, που μπορεί το φυσικό αντικείμενο να λάβει. Αυτές τις ιδιότητες έχουν τα στατιστικά μοντέλα, τα οποία αναπτύχθηκαν ευρέως την τελευταία δεκαετία.

Οι μέθοδοι, οι οποίες χρησιμοποιούν μοντέλα, επιλύουν σε μεγάλο βαθμό τις δυσκολίες που προκύπτουν λόγω θορύβου, έλλειψης πληροφοριών και πολυπλοκότητας της δομής. Η πρωθύστερη γνώση του προβλήματος, μπορεί να χρησιμοποιηθεί για να επιλύσει την πιθανή σύγχυση, που προκαλείται από τα σύνθετα σχήματα, παρέχοντας ανοχή στο θόρυβο, στην έλλειψη δεδομένων

και παρέχοντας με έμμεσο τρόπο μερική αναγνώριση του σχήματος, το οποίο βρίσκεται στην εικόνα. Δυο είναι τα κύρια χαρακτηριστικά που είναι επιθυμητό τα μοντέλα να κατέχουν. Πρώτον, πρέπει να είναι γενικά, δηλαδή να είναι ικανά να παράγουν κάθε αληθοφανές παράδειγμα της κατηγόριας αντικειμένων που αναπαριστάνουν. Δεύτερον, και πιο κύριο, θα πρέπει να είναι συγκεκριμένα, δηλαδή θα πρέπει να παράγουν μόνο “νόμιμα” παραδείγματα, διότι μόνο τότε θα είναι ικανά να περιορίσουν το πρόβλημα σύμφωνα με την πρωθύστερη γνώση που έχουν. Αυτός είναι ο κύριος σκοπός εισαγωγής μοντέλου στην εφαρμογή της παρακολούθησης χεριού. Το χέρι, και κυρίως η παλάμη του χεριού, είναι ένα από τα πλέον παραμορφώσιμα μέρη του σώματος ενός ανθρώπου. Τα διαφορετικά σχήματα που το χέρι μπορεί να λάβει σε μια εικόνα είναι αναρίθμητα, μπορεί όλα τα δάκτυλα να είναι κλειστά, ανοιχτά ή σε άλλη θέση, να έχει κάποια κλίση ως προς την κάμερα, να φαίνεται από την εσωτερική του πλευρά ή από την εξωτερική του. Είναι, λοιπόν, εύλογο να προσπαθήσουμε να μοντελοποιήσουμε το σχήμα του χεριού, ώστε να παρέχουμε επιπλέον πληροφορία στην ενεργή καμπύλη, η οποία θα προσπαθεί να ταυτιστεί με το περίγραμμα του χεριού.

Μια μεγάλη ποικιλία από μεθόδους βασιζόμενες σε μοντέλα έχουν αναπτυχθεί. Χαρακτηριστικά θα αναφέρουμε μερικές από αυτές. Ο Yuille et al.[1] δημιούργησαν ένα μοντέλο για το ανθρώπινο μάτι χρησιμοποιώντας συνδυασμούς από παραμετροποιημένους κύκλους και τόξα. Η μέθοδος αυτή σε αρκετές περιπτώσεις ήταν αποτελεσματική, αλλά ήταν πολύπλοκη και έπρεπε να εξειδικεύεται σε κάθε εφαρμογή. Ο Turk και ο Pentland [37] χρησιμοποίησαν την ανάλυση σε πρωτεύουσες συνιστώσες για να περιγράψουν τα πρότυπα των εικόνων προσώπων, δημιουργώντας ένα σύνολο από συναρτήσεις βάσης, οι οποίες χαρακτηριστικά ονομάζονται ‘eigenfaces’. Παρόλο που οι επιτρεπόμενες μεταβολές προκύπτουν από σύνολο εκπαίδευσης, η αναπαράσταση δεν είναι εύρωστη στις αλλαγές του σχήματος καθώς και στις αλλαγές της όψης του προσώπου. Η μέθοδος αυτή, όμως, μπορεί να εφαρμοστεί στις εικόνες εύκολα με την βοήθεια συνέλιξης.

Οι Philp Torr et al [30] [31] ανέπτυξαν ένα τρισδιάστατο μοντέλο χεριού, το οποίο κατασκευάζεται από κωνικές επιφάνειες, κυλίνδρους και ελλειψοειδή, το οποίο χρησιμοποιείται για την δημιουργία πιθανών περιγραμμάτων του χεριού. Τα περιγράμματα που παράγονται συγχρίνονται με τα περιγράμματα ακμών που εμφανίζονται σε εικόνες. Έτσι, το πρόβλημα της παρακολούθησης χεριού μετατρέπεται σε πρόβλημα “εκτίμησης κατάστασης”, όπου οι παράμετροι του μοντέλου ορίζουν την εσωτερική κατάσταση, η οποία θα εκτιμηθεί με βάση τις

παρατηρήσεις που έχουν προκύψει από τις εικόνες.

Στην εργασία αυτή θα επικεντρωθούμε σε μια στατιστική προσέγγιση, όπως αυτή αναπτύχθηκε από τους Cootes et al. [12], στην οποία το μοντέλο φτιάχνεται ύστερα από την ανάλυση ενός συνόλου σημείων που περιγράφουν το σχήμα των αντικειμένων. Η στατιστική ανάλυση σε σχήματα είναι μια μέθοδος, της οποίας τις βάσεις έθεσε η εργασία του Sir D'Arcy Wentworth Thompson [36]. Η μέθοδος αυτή χρησιμοποιήθηκε στη μέθοδο Μοντέλων Ενεργών Σχημάτων (Active Shape Models) [11] και στην συνεχεία στη μέθοδο Μοντέλων Ενεργών Όψεων (Active Appearance Models) [24]. Τα πλεονεκτήματα της μεθόδου αυτής είναι τα εξής:

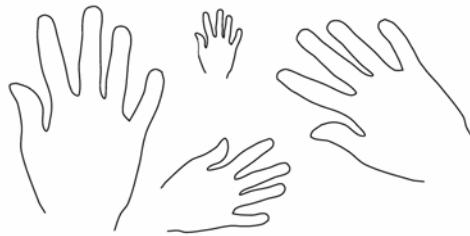
- Είναι ευρέως εφαρμόσιμη. Ο ίδιος αλγόριθμος μπορεί να εφαρμοστεί σε πολλά και διαφορετικά προβλήματα, απλώς διαφοροποιώντας το σύνολο εκπαίδευσης.
- Στο σύνολο εκπαίδευσης μπορούμε να ενσωματώσουμε πληροφορίες κατά τη διάρκεια της δημιουργίας του, τοποθετώντας περισσότερα σημεία στις περιοχές του αντικειμένου όπου το σχήμα είναι σύνθετο.
- Το μοντέλο περιέχει πληροφορίες για τις επιτρεπόμενες μεταβολές με συμπυκνωμένο τρόπο, αλλά ταυτόχρονα καλά ορισμένο, ώστε να μην επιτρέπει αυθαίρετες μεταβολές, διαφορετικές από αυτές που παρουσιάζονται στο σύνολο εκπαίδευσης.

5.2 Στατιστική Ανάλυση Σχήματος

Ο όρος σχήμα στην καθημερινή ομιλία χρησιμοποιείται συχνά με διάφορες σημασιολογικές έννοιες, στη εργασία όμως αποδεχόμαστε τον ορισμό που δόθηκε από το D.G Kendall[14] :

Ορισμός 1 : Σχήμα είναι η γεωμετρική πληροφορία, η οποία παραμένει, όταν η επίδραση της θέσης, της χλίμακας και της περιστροφής απαλειφθούν από ένα αντικείμενο.

Το σχήμα ενός αντικειμένου, λοιπόν, είναι αμετάβλητο σε Ευκλείδειους μετασχηματισμούς οιμοιότητας (αυτό σημαίνει ότι δύο αντικείμενα έχουν ακριβώς το ίδιο σχήμα εάν το ένα μπορούμε να το ταυτίσουμε με το άλλο, αφού πρώτα το μετατοπίσουμε, το μεγεθύνουμε ή το σμικρύνουμε και το περιστρέψουμε κατάλληλα). Η εικόνα (5.1) παρουσιάζει διαφορετικά αντικείμενα που έχουν



Εικόνα 5.1: Διαφορετικά χέρια με το ίδιο σχήμα

ακριβώς το ίδιο σχήμα. Ένας τρόπος για να περιγράψουμε το σχήμα είναι με την τοποθέτηση ενός πεπερασμένου αριθμού σημείων στο περίγραμμα του αντικειμένου. Σημεία τα οποία επισημαίνουν τη θέση σημαντικών χαρακτηριστικών του αντικειμένου ονομάζονται *Ορόσημα*, στη διεθνή βιβλιογραφία *landmarks*. Η μαθηματική αναπαράσταση ενός n -σημείων σχήματος σε k διαστάσεις μπορεί να γίνει ενσωματώνοντας κάθε διάσταση σε ένα $k \times n$ -διάνυσμα. Το διάνυσμα αναπαράστασης ενός επιπέδου σχήματος ($k = 2$), όπως είναι αυτά που θα ασχοληθούμε στην εργασία αυτή, έχει την εξής μορφή:

$$\mathbf{x} = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]^T \quad (5.1)$$

5.2.1 Τοποθέτηση των Ορόσημων

Η τοποθέτηση των σημείων-ορόσημων σε κάθε μια από τις εικόνες του συνόλου εκπαίδευσης είναι εξαιρετικά σημαντικό κομμάτι της διαδικασίας. Κάθε σημείο θα πρέπει να προσδιορίζει ένα συγκεκριμένο μέρος του αντικειμένου ή του συνόρου του. Στο χέρι, θα πρέπει σε όλες τις εικόνες τα ορόσημα - σημεία να προσδιορίζουν τα συγκεκριμένα σημεία του χεριού, για παράδειγμα εάν το πρώτο σημείο βρίσκεται στην κορυφή του αντίχειρα, θα πρέπει σε όλες τις εικόνες να προσδιορίζει αυτό το συγκεκριμένο μέρος. Η μέθοδος λειτουργεί μοντελοποιώντας το τρόπο με τον οποίο σημεία που προσδιορίζουν διαφορετικές θέσεις τείνουν να κινηθούν μαζί καθώς το σχήμα μεταβάλλεται. Αν η τοποθέτηση των “ορόσημων” δεν είναι ορθή, π.χ. ένα συγκεκριμένο σημείο - ορόσημο, εάν τοποθετείται σε διαφορετικές θέσεις πάνω στις εικόνες του συνόλου εκπαίδευσης, η μέθοδος θα αποτύχει να αντιληφθεί τις διακυμάνσεις του σχήματος.

Αντιλαμβανόμαστε, λοιπόν, το λόγο για τον οποίο η τοποθέτηση των σημείων-ορόσημων είναι σημαντική. Η διαδικασία αυτή, συνήθως απαιτεί κάποιον εξοι-

κειωμένο με την εφαρμογή, για την επιλογή των κατάλληλων θέσεων πάνω στο αντικείμενο αλλά και την ταυτόχρονη ορθή τοποθέτηση τους στις διαφορετικές εικόνες του συνόλου εκπαίδευσης. Παρότι η τοποθέτηση των σημείων έγινε με την βοήθεια της πλατφόρμας Matlab (πιο συγκεκριμένα με το Control Point Selection Tool) ήταν ιδιαίτερα χρονοβόρος. Πρέπει να τονιστεί όμως, ότι μόνο αυτό το μέρος της διαδικασίας γίνεται με το χέρι, η εύρεση του μέσου σχήματος και των βασικών αξόνων της διακύμανσης γίνεται αυτόματα.

5.2.2 Ευθυγράμμιση Σχήματος

Σύμφωνα με τον ορισμό της έννοιας σχήματος, για λάβουμε την πραγματική αναπαράσταση του σχήματος πρέπει να απαλειφτεί η θέση, η κλίμακα και η περιστροφή. Αυτό επιτυγχάνεται με τον ορισμό ενός συστήματος αναφοράς, στο οποίο όλα τα σχήματα θα είναι ευθυγράμμισμένα. Παρακάτω περιγράφεται η μέθοδος ευθυγράμμισης από την οποία προκύπτει το επιθυμητό σύστημα αναφοράς. Αυτή η διαδικασία είναι γνωστή ως Ανάλυση Προκρούστη³, και μεταφέρει το σύνολο των σχημάτων στο χώρο σχήματος.

Ορισμός 2 Ο Χώρος Σχήματος είναι το σύνολο όλων των πιθανών σχημάτων τα οποία μπορεί να λάβει το αντικείμενο που μελετάται. Ο χώρος σχήματος Σ_k^n είναι το συνολικό σχήμα των μη ταυτιζομένων σχηματισμών που αποτελούνται από n -σημεία στο χώρο \mathbb{R}^k υπό την επίδραση των Ευκλείδειων μετασχηματισμών ομοιότητας.

Η διάσταση του χώρου σχήματος δεν είναι kn , όπως κάποιος θα μπορούσε να ισχυριστεί, επειδή ο χώρος σχήματος δημιουργείται από m χορηγούς σημείων διανύσματα στο \mathbb{R}^k , λόγω της διαδικασίας ευθυγράμμισης, η οποία μειώνει τις διαστάσεις. Η μετατόπιση μειώνει k διαστάσεις, η κλιμάκωση μια διάσταση και η περιστροφή $\frac{1}{2}k(k - 1)$. Έτσι, η διάσταση του χώρου σχήματος είναι:

$$M = kn - k - 1 - \frac{k(k - 1)}{2} \quad (5.2)$$

³Ο Προκρούστη ήταν τρομερός ληστής και φονιάς, ηταν γνωστός και με άλλα ονόματα, όπως όλοι οι εγκληματίες άλλωστε. Τον ονόμαζαν και Δαμάστη ή Πολυπήμονα. Ο Προκρούστης υποδύθαν το ξενοδόχο. Είχε στήσει δυο κρεβάτια στο δρόμο που οδηγούσε για την Αθήνα και όποιον ταλαιπωρημένο οδοιπόρο έβλεπε να περνά τον προσκαλούσε να ξεκουραστεί. Τα κρεβάτια αυτά ήταν μακρύ και το άλλο κοντό. Ο “φιλόξενος” Προκρούστης πρόσφερε στους κοντούς το μακρύ κρεβάτι και στους ψηλούς το κοντό. Όπως ήταν επόμενο οι άνθρωποι δε ταίριαζαν σε κανένα κρεβάτι. Ο Προκρούστης τότε αποκαθιστούσε τις “ισορροπίες”. Ο Προκρούστης γνώρισε το θάνατο από τον Θησέα, επάνω σ' ένα από τα κρεβάτια του, όπως τα θύματα του. Ο όρος Ανάλυση Προκρούστη δόθηκε από τους Hurley και Cattell το 1962

Μια έννοια απόστασης είναι απαραίτητη για να οριστεί πλήρως ο μη-Ευκλείδειος χώρος σχήματος. Οι αποστάσεις που συνήθως χρησιμοποιούνται είναι η Hausdorff απόσταση και η απόσταση Προκρούστη. Η πρώτη συγχρίνει σχήματα με άνισο αριθμό σημείων, ενώ η απόσταση Προκρούστη απαιτεί τα σύνολα των σημείων να είναι αντίστοιχα. Στην συνεχεία, χρησιμοποιούμε τη απόσταση Προκρούστη.

5.2.3 Ευθυγράμμιση Δυο Σχημάτων

Πρώτα θα ασχοληθούμε με την ευθυγράμμιση ενός ζευγαριού σχημάτων [12]. Όπως αναφέραμε και παραπάνω θεωρούμε ότι το διάνυσμα \mathbf{x}_i περιγράφει τα n σημεία του i σχήματος :

$$\mathbf{x} = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]^T$$

Θεωρούμε ότι ο τελεστής $M_j[x_j]$ προκαλεί περιστροφή κατά θ_j και κλιμάκωση κατά s_j . Δοσμένων των δυο παρομοίων σχημάτων, x_i και x_j μπορούμε να βρούμε τα κατάλληλα θ_j , s_j και την κατάλληλη μεταπίση $(t_x, t_y)_j$ που κατευθύνουν το διάνυσμα x_j πάνω στο $M_j[x_j]$ με τέτοιο τρόπο ώστε να ελαχιστοποιείται η απόσταση Προκρούστη, δηλαδή να ελαχιστοποιεί το παρακάτω άθροισμα:

$$E_j = (x_j - M_j[x_j])^T \mathbf{W} (x_j - M_j[x_j]) \quad (5.3)$$

όπου

$$M_j \begin{pmatrix} x_{jk} \\ y_{ik} \end{pmatrix} = \begin{pmatrix} (s_j \cos \theta) x_{jk} - (s_j \sin \theta) y_{jk} + t_{jx} \\ (s_j \sin \theta) x_{jk} + (s_j \cos \theta) y_{jk} + t_{jy} \end{pmatrix} \quad (5.4)$$

και \mathbf{W} είναι ένας διαγώνιος πίνακας με βάρη για κάθε σημείο-ορόσημο.

Περισσότερες λεπτομέρειες δίνονται στο Παράρτημα B.

Τα βάρη επιλέγονται ώστε να δίνουν μεγαλύτερη βαρύτητα στα σημεία εκείνα τα οποία τείνουν να είναι πιο “σταθερά” μέσα στο σύνολο. Δηλαδή, αυτά τα σημεία, που κινούνται λιγότερο σε σχέση με τα άλλα σημεία στο σχήμα, επιθυμούμε να θεωρούνται πιο σημαντικά για την δομή του σχήματος. Ο πίνακας βαρών ορίζεται ως εξής: θεωρούμε R_{kl} να είναι η απόσταση μεταξύ των σημείων k και l στο σχήμα, ακόμα θεωρούμε ότι $V_{R_{kl}}$ είναι η διακύμανση της απόστασης αυτής στο σύνολο των σχημάτων. Και επιλέγουμε το βάρος, w_k , για το k σημείο να είναι ίσο με :

$$w_k = \left(\sum_{l=0}^{n-1} V_{R_{kl}} \right)^{-1} \quad (5.5)$$

Αν ένα σημείο έχει την τάση να κινείται αρκετά σε σχέση με τα αλλά σημεία που αποτελούν το σχήμα, το άθροισμα των διακυμάνσεων θα είναι μεγάλο και κατά συνέπεια το βάρος του σημείου θα είναι μικρό. Αν, από την άλλη πλευρά, ένα σημείο τείνει να παραμένει σταθερό σε μια συγκεκριμένη θέση σε σχέση με τα υπόλοιπα σημεία, το άθροισμα των διακυμάνσεων θα είναι μικρό και αντίστοιχα το βάρος του σημείου αυτό θα είναι μεγάλο. Αυτό θα έχει ως αποτέλεσμα το ταίριασμα των σημείων αυτών στα διάφορα σχήματα του συνόλου εκπαίδευσης να είναι σε πρώτη προτεραιότητα.

Σε μορφή αλγόριθμου η ευθυγράμμιση δυο σχημάτων θα περιείχε τα εξής βήματα:

1. Υπολογισμός του κέντρου βάρους κάθε σχήματος.
2. Κλιμάκωση κάθε σχήματος για να έχουν το ίδιο μέγεθος.
3. Ευθυγράμμιση με σκοπό τα κέντρα βαρών και των δυο σχημάτων να είναι στην ίδια θέση.
4. Ευθυγράμμιση του προσανατολισμού των δυο σχημάτων με περιστροφή.

5.2.4 Γενικευμένη Ανάλυση Προκρούστη

Για να ευθυγραμμίσουμε όλα τα σχήματα που ανήκουν στο σύνολο εκπαίδευσης πρέπει να ακολουθήσουμε τον εξής αλγόριθμο:

1. Περιστροφή, κλιμάκωση και μετατόπιση όλων των σχημάτων του συνόλου με σκοπό την ευθυγράμμιση τους με το πρώτο σχήμα.

Βρόχος:

2. Υπολογισμός του μέσου σχήματος των μετασχηματισμένων σχημάτων
3. *Eίτε*
 - (i) Προσαρμογή του μέσου σχήματος σε μια δεδομένη κλίμακα, προσανατολισμό και αρχική θέση.
Είτε
 - (ii) Περιστροφή, κλιμάκωση και μετατόπιση του μέσου σχήματος για να ευθυγραμμιστεί με το πρώτο σχήμα.

4. Περιστροφή, κλιμάκωση και μετατόπιση όλων των σχημάτων ξανά ώστε να ταιριάζουν στο προσαρμοσμένο μέσο σχήμα.

Επανάληψη μέχρι να επιτευχθεί σύγκλιση.

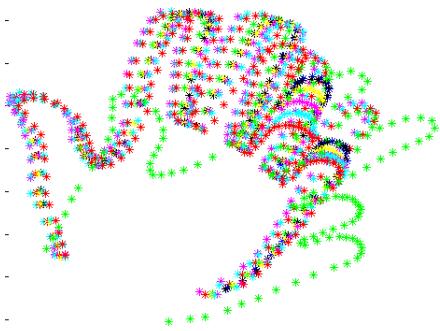
Το τρίτο στάδιο του αλγόριθμου στο εσωτερικό του βρόχου είναι απαραίτητο για την επανακανονικοποίηση του μέσου σχήματος. Χωρίς αυτό ο αλγόριθμος είναι μη-καλά ορισμένος – υπάρχουν $4(N_s - 1)$ περιορισμοί για $4N_s$ μεταβλητές (θ, s, t_x, t_y για κάθε σχήμα) – και δεν θα συγχλίνει, το μέσο σχήμα θα συρρικνώνεται και θα περιστρέφεται και οι επαναλήψεις θα τείνουν στο άπειρο. Περιορισμοί για την όψη και την κλίμακα του μέσου σχήματος επιτρέπουν στον αλγόριθμο να έχει μοναδική λύση. Για το λόγο αυτό, είτε κλιμακώνουμε, περιστρέφουμε και μετατοπίζουμε το μέσο σχήμα, ώστε να ταιριάζει στο πρώτο σχήμα του συνόλου, είτε αυθαίρετα επιλέγουμε περιορισμούς για το μέσο σχήμα, οι οποίοι μπορεί να έχουν ως σκοπό, είτε το κέντρο βάρους τους μέσου σχήματος να είναι σε κάποιο συγκεκριμένο σημείο, είτε να έχει κάποιο συγκεκριμένο προσανατολισμό, είτε η απόσταση μεταξύ δυο σημείων να είναι μοναδιαία στο μέσο σχήμα. Η επιλογή αυτή, εξαρτάται από το είδος της εφαρμογής στην οποία θέλουμε να χρησιμοποιήσουμε το μοντέλο. Στη εφαρμογή που αναπτύξαμε το μέσο σχήμα το ταιριάζαμε με το πρώτο σχήμα του συνόλου εκπαίδευσης.

Η συνθήκη σύγκλισης μπορεί να ελεγχθεί εξετάζοντας εάν μεταβάλλονται αισθητά οι παράμετροι των μετασχηματισμών κατά την ευθυγράμμιση του συνόλου των σημείων με το μέσο σχήμα που κάθε φορά προκύπτει. Έχει δειχθεί πειραματικά [6] ότι δυο (2) επαναλήψεις του βρόχου είναι επαρκής για τις περισσότερες περιπτώσεις.

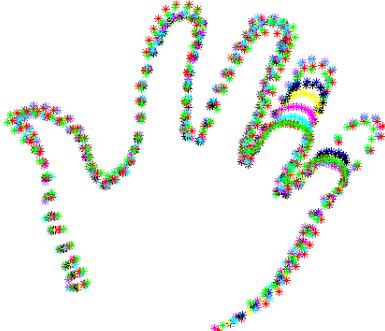
Παρακάτω παρουσιάζεται το αποτέλεσμα της ευθυγράμμισης του συνόλου εκπαίδευσης ενός από τα πειράματα που διεξήχθησαν για την εφαρμογή της παρακολούθησης ενός χεριού. Το σχήμα που αναπαριστάται είναι η παλάμη ενός χεριού, το οποίο παραμορφώνεται κατά την εκτέλεση μιας χειρονομίας. Για την δημιουργία του σχήματος χρησιμοποιήθηκαν 118 ορόσημα.

5.2.5 Στατιστική Ανάλυση των Ευθυγραμμισμένων Σχημάτων

Μόλις το σύνολο των ευθυγραμμισμένων σχημάτων είναι έτοιμο, το μέσο σχήμα και η διακύμανση μπορούν να προσδιοριστούν. Το μέσο σχήμα, \bar{x} ,



5.2.1 Σχήματα Χεριών



5.2.2 Ευθυγραμμισμένα Σχήματα

Εικόνα 5.2: Στην εικόνα (5.2.1) παρουσιάζεται το σύνολο εκπαίδευσης όπως αυτό προέκυψε έπειτα από την τοποθέτηση των ορόσημων στις διάφορες εικόνες. Στην (5.2.2) παρουσιάζονται τα ευθυγραμμισμένα σχήματα

υπολογίζεται ως εξής:

$$\bar{x} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i \quad (5.6)$$

Η τάση για διακύμανση, δηλαδή ο τρόπος με τον οποίο τα σημεία των σχημάτων τείνουν να κινηθούν μαζί, μπορεί να βρεθεί με την εφαρμογή Ανάλυσης Πρωτευουσών Συνιστωσών (Principal Components Analysis). Το μαθηματικό υπόβαθρο της ανάλυσης αυτής θα αναλυθεί παρακάτω.

Στο σύνολο των ευθυγραμμισμένων σχημάτων πάντοτε θα παρατηρούμε ένα βαθμό συσχέτισης μεταξύ των αντιστοίχων σημείων. Αν η συσχέτιση αυτή δεν παρουσιάζεται, είτε (α) δεν υπάρχει διακύμανση, είτε (β) τα σημεία είναι πλήρως τυχαία, και στις δύο περιπτώσεις σημαίνει ότι τα σημεία δεν είναι ορόσημα. Έτσι οδηγούμαστε στην ιδέα να δημιουργήσουμε μια αναπαράσταση του σχήματος, όπου θα λαμβανόταν υπόψη η συσχέτιση μεταξύ των σημείων. Αν η κίνηση μερικών σημείων ήταν αριθμητικά συσχετιζόμενη, αυτό θα μπορούσε να αξιοποιηθεί για την μείωση της διάσταση του χώρου σχήματος.

Στην περίπτωση μας ψάχνουμε έναν γραμμικό μετασχηματισμό των δεδομένων:

$$y = Mx \quad (5.7)$$

Θεωρούμε το μέσο σχήμα $\bar{x} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i$ και το πίνακα συνδιακύμανσης του

σχήματος:

$$\Sigma_x = \frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})(x_i - \bar{x})^T \quad (5.8)$$

Ο μέσος όρος των μεταβλητών y μπορεί να εκφραστεί ως :

$$\bar{y} = \frac{1}{N_s} \sum_{i=1}^{N_s} y_i = \frac{1}{N_s} \sum_{i=1}^{N_s} Mx_i = M\bar{x} \quad (5.9)$$

Και ως αποτέλεσμα η συνδιακύμανση των μεταβλητών y :

$$\begin{aligned} \Sigma_y &= \frac{1}{N_s} \sum_{i=1}^{N_s} (y_i - \bar{y})(y_i - \bar{y})^T \\ &= \frac{1}{N_s} \sum_{i=1}^{N_s} (Mx_i - M\bar{x})(Mx_i - M\bar{x})^T \\ &= \frac{1}{N_s} \sum_{i=1}^{N_s} M(x_i - \bar{x})(M(x_i - \bar{x}))^T \\ &= \frac{1}{N_s} \sum_{i=1}^{N_s} M(x_i - \bar{x})(x_i - \bar{x})^T M^T \\ &= \frac{1}{N_s} M \left(\sum_{i=1}^{N_s} (x_i - \bar{x})(x_i - \bar{x})^T \right) M^T \\ &= M\Sigma_x M^T \end{aligned} \quad (5.10)$$

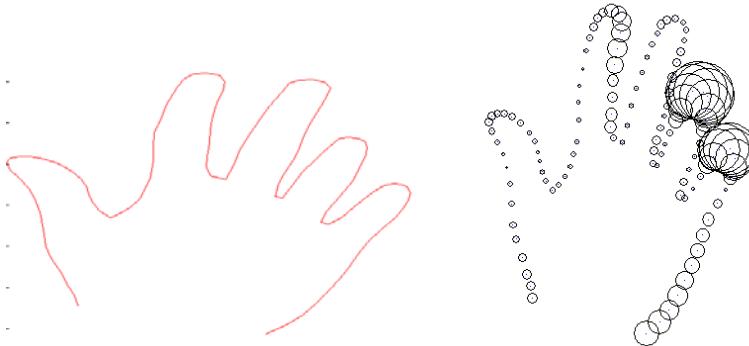
Αν περιοριστούμε σε ορθογώνιους μετασχηματισμούς ($M^{-1} = M^T$), τότε πολλαπλασιάζοντας από δεξιά την εξίσωση (5.10) με M^T έχουμε:

$$M^T \Sigma_y = \Sigma_x M^T \quad (5.11)$$

Αντικαθιστώντας το M^T με Φ παίρνουμε:

$$\Sigma_x \Phi = \Phi \Sigma_y \quad (5.12)$$

Από την παραπάνω σχέση συμπεραίνουμε ότι εάν ο πίνακας Φ επιλεχθεί ως τα ιδιοδιανύσματα του συμμετρικού πίνακα $\Sigma_x, \Phi = (\varphi_1 | \varphi_2 | \dots | \varphi_n)$ τότε η συνδιακύμανση των μετασχηματισμένων σχημάτων, Σ_y , μετατρέπεται στο διαγώνιο πίνακα των ιδιοτιμών, λ_i . Στην περίπτωση λοιπόν των συσχετιζόμενων σημείων η μικρότερη ιδιοτιμή θα ήταν κοντά στο μηδέν και θα αντιστοιχούσε



5.3.1 Μέσο Σχήμα

5.3.2 Διακύμανση κάθε σημείου

Εικόνα 5.3: Στην εικόνα (5.3.1) παρουσιάζεται το μέσο σχήμα όπως αυτό προέκυψε έπειτα από την επεξεργασία του συνόλου εκπαίδευσης. Στην (5.3.2) παρουσιάζονται σχηματικά η διακύμανση γύρω από κάθε σημείο.

σε ένα ιδιοδιάνυσμα, το οποίο θα μπορούσε να παραληφθεί από τον πίνακα Φ , μειώνοντας το μέγεθος του y .

Για να επανέλθουμε στις αρχικές μας μεταβλητές από το νέο σύνολο μεταβλητών y , αντιστρέφουμε την σχέση (5.7), γνωρίζοντας ότι το M είναι ορθογώνιο:

$$x = M^{-1}y = M^T y = \Phi y \quad (5.13)$$

Στην εφαρμογή μας εφαρμόζουμε PCA στις αποστάσεις από το μέσο σχήμα, $dx_i = x_i - \bar{x}$:

$$y = M(x - \bar{x}) \quad , \quad x = \bar{x} + \Phi x \quad (5.14)$$

Αν στο πίνακα Φ διατηρήσουμε, μόνο τα τι ιδιοδιανύσματα που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές, τότε μπορούμε να χρησιμοποιήσουμε το νέο πίνακα για να προσεγγίσουμε το σύνολο εκπαίδευσης, x χρησιμοποιώντας την σχέση:

$$x \approx \bar{x} + \Phi b \quad (5.15)$$

όπου b είναι ένα διάνυσμα διάστασης t , το οποίο δίνεται από την σχέση $b = \Phi^T(x - \bar{x})$. Το διάνυσμα b ορίζει το σύνολο των παραμέτρων του παραμορφώσιμου μοντέλου. Με την μεταβολή των στοιχείων του b μεταβάλλεται το σχήμα x , το οποίο προκύπτει από την εξίσωση (5.15). Η διακύμανση της ιστής παραμέτρου, b_i του συνόλου εκπαίδευσης είναι ίση με την ι-οστή ιδιοτιμή λ_i . Περιορίζοντας την παράμετρο b_i μεταξύ του $\pm 3\sqrt{\lambda_i}$ εξασφαλίζουμε ότι το

σχήμα, το οποίο παράγεται είναι παρόμοιο με εκείνα που ανήκουν στο σύνολο εκπαίδευσης.

5.2.6 Επιλογή των Πρωτευουσών Συνιστωσών

Ο αριθμός των ιδιοδιανυσμάτων t , που διατηρούνται στο μοντέλο μπορεί να επιλεχθεί με διάφορους τρόπους. Ένας από αυτούς είναι το t να επιλέγεται ώστε το μοντέλο να είναι ικανό να αναπαριστάνει ποσοστό (π.χ 97%) της συνολικής διακύμανσης των δεδομένων.

Όπως έχουμε ορίσει και παραπάνω, θεωρούμε λ_i τις ιδιοτιμές του πίνακα συνδιακύμανσης των δεδομένων εκπαίδευσης. Κάθε ιδιοτιμή εκφράζει την διακύμανση των δεδομένων γύρω από τη μέση τιμή στη κατεύθυνση που ορίζει το αντίστοιχο ιδιοδιανυσματα. Η συνολική διακύμανση του συνόλου εκπαίδευσης είναι το άθροισμα όλων των ιδιοτιμών, $V_T = \sum_{i=1}^N \lambda_i$. Μπορούμε, λοιπόν, να επιλέξουμε τις t μεγαλύτερες ιδιοτιμές οι οποίες ικανοποιούν την ανισότητα:

$$\sum_{i=1}^t \lambda_i \geq f_v V_T \quad (5.16)$$

όπου f_v εκφράζει το ποσοστό της συνολικής διακύμανσης που επιθυμούμε το μοντέλο να αναπαριστάνει.

Αν γνωρίζαμε, ότι ο θόρυβος των μετρήσεων μας, δηλαδή της διαδικασίας τοποθέτησης των ορόσημων, παρουσιάζει διακύμανση σ_n^2 , τότε θα μπορούσαμε να επιλέξουμε να διατηρούμε τα ιδιοδιανύσματα, των οποίων οι αντίστοιχες ιδιοτιμές είναι μεγαλύτερες από τη διακύμανση του θορύβου, $\lambda_t > \sigma_n^2$.

Μια εναλλακτική προσέγγιση είναι η επιλογή των πρωτευουσών συνιστωσών, ώστε το μοντέλο να είναι ικανό να αναπαραστήσει τα σχήματα του συνόλου εκπαίδευσης με ορισμένη ακρίβεια. Για παράδειγμα, θα μπορούσε να οριστεί ικανοποιητικό ένα μοντέλο, εάν μπορεί να προσεγγίσει όλα τα σημεία του συνόλου εκπαίδευσης με ακρίβεια ορισμένων pixel. Για να επιτευχθεί αυτό, το μοντέλο δημιουργείται κλιμακωτά, δηλαδή αυξάνουμε τον αριθμό των ιδιοδιανυσμάτων μέχρι να πληρούνται τα κριτήρια που έχουμε θέσει.

Στην εφαρμογή μας κάνουμε χρήση της πρώτης μεθόδου που περιγράφτηκε. Συνήθως στα πειράματα που εκτελέσαμε το ποσοστό της διακύμανσης που διατηρούσαμε ήταν ίσο με 97%.

Στην εικόνα (5.4) παρουσιάζεται ο τρόπος παραμόρφωσης του σχήματος αλλάζοντας τις τιμές του διανύσματος b . Παρατηρούμε, λοιπόν, ότι η αλλαγή

5.4.1 $b_1 = +2\sqrt{\lambda_1}$ 5.4.2 $b = 0$ 5.4.3 $b_1 = -1\sqrt{\lambda_1}$ 5.4.4 $b_2 = +2\sqrt{\lambda_2}$ 5.4.5 $b = 0$ 5.4.6 $b_2 = -1\sqrt{\lambda_2}$

Εικόνα 5.4: Παραμορφώσεις του μέσου σχήματος κάνοντας χρήση της πρώτης και δεύτερης πρωτεύουσας συνιστώσας

του βάρους της πρώτης συνιστώσας μεταβάλλει την θέση των δυο ακριανών δακτύλων: αυξάνοντας το b_1 , μεγαλώνουν τα δάκτυλα, ενώ όταν μικραίνει μικραίνουν και τα δάκτυλα. Ακόμα, μεταβάλλοντας το συντελεστή της δεύτερης πρωτεύουσας συνιστώσας b_2 , το σύνολο της παλάμης μετακινείται προς τα κάτω και σμικρύνεται. Βέβαια, κατά την ερμηνεία των εικόνων αυτών πρέπει να είμαστε ιδιαίτερα προσεκτικοί γιατί έχει προηγηθεί η διαδικασία ευθυγράμμισης, η οποία έχει κλιμακώσει τα σχήματα και ορισμένες λεπτομέρειες δεν είναι άμεσα ορατές με την μεταβολή ενός μόνο στοιχείου του διανύσματος b .

Κεφάλαιο 6

Συνδυασμός Μοντέλου Σχήματος και Μοντέλου Ενεργών Καμπυλών

6.1 Εισαγωγή

Ο προσδιορισμός των συνόρων του χεριού και στη συνέχεια η παρακολούθηση του σε μια ακολουθία εικόνων είναι, για πολλούς λόγους, ένα πολύπλοκο πρόβλημα, μερικοί από τους οποίους έχουν ήδη αναφερθεί, όπως η σύνθετη μορφή του σχήματος του χεριού και η ταυτόχρονη δυνατότητα του να παραμορφωθεί με ποικίλους τρόπους. Η μέθοδος των ενεργών καμπυλών, από μόνη της, θα ήταν ιδιαίτερα δύσκολο, να προσεγγίσει και να ταυτιστεί με το περίγραμμα της παλάμης. Είναι χαρακτηριστικό, ότι αρκετές εφαρμογές παρακολουθήσεις αντικειμένων έχουν πραγματοποιηθεί με βάση το μοντέλο των ενεργών καμπυλών, αλλά τα αντικείμενα αυτά, στην πλειονότητα των εφαρμογών αυτών, είναι σχετικά απλά σε μορφή, όπως το σχήμα ενός λευκοκύτταρου [25][26], ή το περίγραμμα του χεριού χωρίς να λαμβάνονται υπόψη οι κινήσεις των δάκτυλων [5].

Στην περίπτωση του σχήματος ενός λευκοκύτταρου, ο αλγόριθμος που προτείνεται [25] εισάγει στο μοντέλο των ενεργών καμπυλών περιορισμούς σχήματος και μεγέθους. Οι περιορισμοί αυτοί έχουν την μορφή ενεργειακού συναρτησιακού, το οποίο συνυπολογίζεται στο συνολικό ενεργειακό συναρτησιακό του φιδιού. Η ελαχιστοποίηση αυτού του ενεργειακού, οδηγεί το φίδι να έχει τη γεωμετρική μορφή κύκλου και να έχει εμβαδόν παραπλήσιο με το εμβαδόν ενός λευκοκύτταρου, το οποίο εισάγεται ως πρωθύστερη γνώση στο μοντέλο. Η διαφορά, όμως, της εφαρμογής εύρεσης ενός λευκοκύτταρου, με τη περίπτωση του προσδιορισμού των συνόρων της παλάμης, είναι ότι το σχήμα

του κύτταρου μπορεί να περιγραφεί με μαθηματικές εξισώσεις, πράγμα που είναι σχεδόν αδύνατο για το χέρι. Η δημιουργία ενός ενεργειακού συναρτησιακού, του οποίου η ελαχιστοποίηση θα έδινε χαρακτηριστικές μορφολογίες του χεριού, φαντάζει σχεδόν αδύνατη.

Από την άλλη πλευρά ο Blake και ο Isard [5] ασχολήθηκαν με περιπτώσεις μη παραμορφώσιμου χεριού, δηλαδή αντιμετωπίζουν το χέρι ως ένα αντικείμενο, του οποίου η μορφολογία δεν αλλάζει κατά την διάρκεια της εφαρμογής. Δημιουργούν έτσι, σύνολα εκπαίδευσης που περιγράφουν άκαμπτες κινήσεις, τον γραμμικό συνδυασμό των οποίων, η μέθοδος που προτείνουν, μπορεί να παρακολουθήσει. Βέβαια, η μέθοδος αυτή είναι πιθανοτική και έχει αξιόλογα δυναμικά χαρακτηριστικά, τα οποία βελτιώνουν ιδιαίτερα την ικανότητα παρακολούθησης της κίνησης. Το πιθανοτικό αυτό πλαίσιο θα μπορούσε να είναι μια προέκταση της εργασίας της διπλωματικής αυτής.

Με σκοπό, λοιπόν, να μειώσουμε τα προβλήματα που θα συναντούσε η ενεργή καμπύλη στην προσπάθεια της να προσδιορίσει το περίγραμμα του χεριού, συνδυάζουμε το μοντέλο της ενεργής καμπύλης με το στατιστικό μοντέλο σχήματος. Το στατιστικό μοντέλο, θα παρέχει τον επιτρεπόμενο Χώρο Σχήματος μετά την κατάλληλη στατιστική ανάλυση του συνόλου εκπαίδευσης, ενώ η ενεργή καμπύλη θα εξελίσσεται με βάση τον επιτρεπόμενο Χώρο Σχήματος. Πιο συγκεκριμένα, η ενεργή καμπύλη θα εξελίσσεται με σκοπό την ικανοποίηση της συνθήκης ισορροπίας δυνάμεων και σε τακτά χρονικά διαστήματα θα προβάλλεται στον επιτρεπόμενο Χώρο Σχήματος, ώστε να διασφαλίζεται ότι η μορφή του φιδιού παραμένει παραπλήσια με εκείνη που εμφανίζεται στο σύνολο εκπαίδευσης.

6.2 Προβολή στο Επιτρεπόμενο Χώρο Σχήματος

Στο Κεφάλαιο 3 αναλύθηκε ο τρόπος εξέλιξης της ενεργής καμπύλης υπό την επίδραση των διαφόρων δυναμικών και προέκυψε η εξίσωση (3.26), η επίλυση της οποίας δίνει τη νέα θέση του φιδιού. Στο σημείο αυτό θεωρούμε ότι οι συντεταγμένες των διακριτών σημείων του φιδιού δίνονται στη ίδια μορφή, όπως και τα διανύσματα σχήματος κατά την στατιστική ανάλυση, $\mathbf{x} = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]^T$. Ο επιτρεπόμενος χώρος σχήματος, δηλαδή ο υποχώρος που δημιουργείται από τις κύριες συνιστώσες του συνόλου εκπαίδευσης, δίνεται από την σχέση:

$$X = \bar{X} + \Phi b, \quad \text{με } |b_i| \leq 3\sqrt{\lambda_i}$$

όπου $\Phi = (\varphi_1 | \varphi_2 | \dots | \varphi_\tau)$.

Το διάνυσμα θέσης, των διακριτών σημείων του φιδιού, προβάλλεται στο χώρο σχήματος και η επιτρεπτή θέση της καμπύλης προκύπτει από τη προβολή του αρχικού διανύσματος. Η προβολή αυτή επιτυγχάνεται προσδιορίζοντας αρχικά το βαθμωτό διάνυσμα παραμέτρων b και στην συνέχεια με την ανακατασκευή του διανύσματος εντός του επιτρεπτού χώρου:

$$b = (\Phi^T \Phi)^{-1} \Phi^T (X - \bar{X}) \quad (6.1)$$

Το επιτρεπτό διάνυσμα καμπύλης που προκύπτει είναι:

$$X_{proj} = \bar{X} + \Phi b \quad (6.2)$$

Με το τρόπο αυτό περιορίζουμε την ενεργή καμπύλη σε απόσταση μικρότερη από $3\sqrt{\lambda_i}$, θεωρώντας ότι στις περισσότερες εφαρμογές, μια αληθιφανής μορφολογία χεριού, βρίσκεται εντός του διαστήματος που ορίζεται σε απόσταση τριών τυπικών αποκλίσεων από το μέσο σχήμα στη κατεύθυνση των πρωτευουσών συνιστωσών.

Αναλυτικά ο αλγόριθμος που χρησιμοποιείται κατά την προβολή της καμπύλης στο επιτρεπτό επίπεδο είναι ο εξής:

1. Ευθυγράμμιση του μέσου σχήματος του συνόλου εκπαίδευσης με την υπάρχουσα καμπύλη με σκοπό το προσδιορισμό των παραμέτρων κλιμάκωσης, περιστροφής και μετατόπισης.
2. Προσδιορισμός του βαθμωτού διανύσματος παραμέτρων
 $b = (\Phi^T \Phi)^{-1} \Phi^T (X - \bar{X})$
3. Περιορισμός των στοιχείων του b στα όρια τριών τυπικών αποκλίσεων :
 $|b_i| \leq 3\sqrt{\lambda_i}$
4. Παραγωγή της προβολής : $X_{proj} = \bar{X} + \Phi b$
5. Αντίστροφη διαδικασία κλιμάκωσης, περιστροφής και μετατόπισης της X_{proj} με βάση τις παραμέτρους που προέκυψαν στο Βήμα 1

Ακολουθώντας, όμως, τον αλγόριθμο αυτό, γίνεται φανερό ότι εμφανίζεται το πρόβλημα ότι η καμπύλη πρέπει να έχει “σταθερά” διακριτά σημεία (snaxels). Σταθερά, με την έννοια ότι το διάνυσμα θέσης πρέπει να έχει σταθερό μέγεθος, ώστε στις ίδιες θέσεις του να υπάρχουν οι συντεταγμένες σημείου

που περιγράφει συγκεκριμένα χαρακτηριστικά του περιγράμματος του αντικειμένου, το οποίο θέλουμε να προσεγγίσουμε, στην δικιά μας περίπτωση το χέρι. Για αυτό το λόγο, η εισαγωγή του αλγόριθμου αυτού συνεπάγεται την κατάργηση της ιδιότητας του μοντέλου ενεργών καμπυλών, να παρεμβάλλει νέα σημεία καθώς τα snaxels απομακρύνονται το ένα από το άλλο, με σκοπό την εξασφάλιση της συνέχειας της συνάρτησης της καμπύλης.

Το πρόβλημα αυτό επιλύεται με την τοποθέτηση μεγάλου αριθμού ορόσημων κατά την δημιουργία του συνόλου εκπαίδευσης. Σε ορισμένες όμως περιπτώσεις, όπως η αλλαγή της κλίμακας του χεριού στην εικόνα, είναι πιθανόν να προκύπτει μη ακριβής ταύτιση με το περίγραμμα λόγω έλλειψης ικανού αριθμού διακριτών σημείων. Στην εφαρμογή μας ακολουθήθηκε αυτός ο αλγόριθμος και στα πειράματα που διεξήχθησαν δεν προέκυψαν προβλήματα που να οφείλονται στο γεγονός αυτό. Ένας εναλλακτικός αλγόριθμος για τη μεθοδολογία προβολής παρατίθεται παρακάτω.

6.2.1 Εναλλακτικός Αλγόριθμος Προβολής Ενεργής Καμπύλης

Ο αλγόριθμος αυτός παρουσιάστηκε από τους G. Hamarneh και T. Gustavsson [16] με σκοπό την βελτίωση της αναπαράστασης της δεξιάς κοιλίας της καρδιάς από μια ενεργή καμπύλη. Σύμφωνα, λοιπόν, με τον αλγόριθμο αυτό κατά την διάρκεια του σταδίου εκπαίδευσης εφαρμόζουμε στα διανύσματα των σχημάτων, που έχουν δημιουργηθεί, Διακριτό Μετασχηματισμό Συνημίτονου. Η διαδικασία αυτή φέρνει τα διανύσματα στο χώρο συχνοτήτων, κατά συνέπεια τα ορόσημα που τοποθετούνται στις εικόνες εκπαίδευσης μετασχηματίζονται σε αντιστοίχου αριθμού συχνοτήτες.

Στο χώρο συχνοτήτων εκτελούμε Ανάλυση Πρωτευουσών Συνιστωσών (PCA) με ακριβώς την ίδια λογική που αναφέρθηκε στο Κεφαλαίο 5, δηλαδή, εύρεση των κυρίων συνιστωσών διακύμανσης των σημείων των σχημάτων και ταυτόχρονη μείωση της διάστασης του χώρου σχήματος. Με το τρόπο αυτό ταυτόχρονα ορίζεται και ο επιτρεπόμενος χώρος. Το φίδι εξελίσσεται σύμφωνα με το κλασσικό τρόπο των ενεργών καμπυλών, επιτρέποντας στον αλγόριθμο να παρεμβάλλει σημεία στα διαστήματα του φιδιού, όταν αυτό χρίνεται αναγκαίο. Όμως, ανά τακτά χρονικά διαστήματα, στο διάνυσμα θέσης της καμπύλης μεταβλητού μήκους εφαρμόζεται Διακριτό Μετασχηματισμός Συνημίτονου, διατηρώντας μόνο τους N πρώτους συντελεστές συχνοτήτων, όπου N ο αριθμός των συχνοτήτων που έχουν προκύψει από το μετασχηματισμό

των σχημάτων του συνόλου εκπαίδευσης. Οι συντελεστές αυτοί αναπαριστάστούν αντίστοιχα χωρικά χαρακτηριστικά της καμπύλης, έτσι επιτυγχάνεται η δυνατότητα προβολής σε έναν υποχώρο παρότι δεν υπήρχε ακριβής χωρική αντιστοιχία μεταξύ του διανύσματος θέσης της καμπύλης και των διανυσμάτων σχήματος του συνόλου εκπαίδευσης. Στην συνέχεια, το διάνυσμα των συντελεστών του μετασχηματισμού συνημίτονου προβάλλεται στον επιτρεπτό χώρο. Στο διάνυσμα προβολής εφαρμόζεται Αντίστροφος μετασχηματισμός Συνημίτονου και λαμβάνουμε το νέο περιορισμένο διάνυσμα θέσης της καμπύλης. Η εφαρμογή, λοιπόν, έχει τα εξής αλγορίθμικά βήματα:

1. Δημιουργία Συνόλου εκπαίδευσης και Εφαρμογή Διακριτού Μετασχηματισμού Συνημίτονου
2. Εφαρμογή PCA στους συντελεστές του μετασχηματισμού συνημίτονου και παραγωγή Επιτρεπόμενου Χώρου

Εφαρμογή :

- (a) Εξέλιξη Ενεργής Καμπύλης
- (b) Εφαρμογή Διακριτού Μετασχηματισμού Συνημίτονου
- (c) Προβολή των Συντελεστών στον Επιτρεπτό Χώρο
- (d) Αντίστροφος Διακριτός Μετασχηματισμός Συνημίτονου και Επαναπροσδιορισμός της θέσης της ενεργής καμπύλης

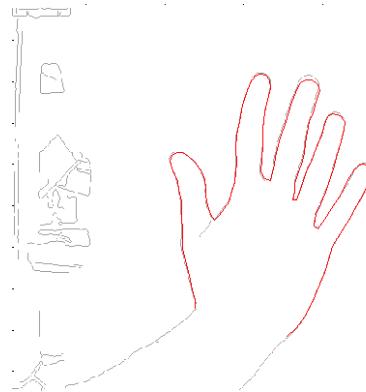
Ο λόγος για τον οποίο ο αλγόριθμος αυτός δεν χρησιμοποιήθηκε στην τελική εφαρμογή παρακολούθησης χεριού,- παρότι διαισθητικά δίνει περισσότερα πλεονεκτήματα στην εξέλιξη της ενεργής καμπύλης επειδή δεν της περιορίζει το αριθμό των διακριτών σημείων,- είναι ότι κατά την πειραματική διαδικασία ο προηγούμενος αλγόριθμος είχε πιο ικανοποιητικά αποτελέσματα. Στο χώρο συχνότητας δεν είναι δυνατός ο ορισμός ευθυγράμμισης σχήματος με αποτέλεσμα ο χώρος συχνοτήτων να αντικατοπτρίζει μόνο το ευθυγραμμισμένο σύνολο εκπαίδευσης. Δηλαδή, δεν είναι εφικτό να βρεθούν οι παράμετροι μετατόπισης μεταξύ των συχνοτήτων που αντιστοιχούν στο μέσο σχήμα και των συχνοτήτων που αντιστοιχούν σε μια παραμόρφωση του φιδιού. Η προβολή της θέσης του φιδιού, λοιπόν, μετατοπίζεται, κλιμακώνεται και περιστρέφεται γύρω από το μέσο σχήμα και δεν υπάρχει η δυνατότητα εύρεσης της αντίστροφης πορείας, ωστε η επιτρεπόμενη καμπύλη να τοποθετηθεί ξανά στο χώρο που είχε προσεγγίσει κάτω από την επίδραση των δυναμικών.

6.3 Αποτέλεσμα Χρήσης Στατιστικού Μοντέλου

Η χρήση του στατιστικού μοντέλου έχει ως αποτέλεσμα το περιορισμό της καμπύλης σε έναν υποχώρο που δημιουργείται βάσει της στατιστικής ανάλυσης του συνόλου εκπαίδευσης. Παρακάτω παρατίθενται παραδείγματα εύρεσης του περιγράμματος του χεριού με χρήση μόνο του μοντέλου των ενεργών καμπυλών αλλά και συνδυασμού αυτού με στατιστικό μοντέλο.



6.1.1 Φίδι χωρίς περιορισμούς



6.1.2 Χρήση Στατιστικού Μοντέλου

Εικόνα 6.1: Στην εικόνα (6.1.1) παρουσιάζεται η τελική θέση του φιδιού πάνω στην εικόνα ακμών, όπως αυτή προέκυψε σύμφωνα με το μοντέλο ενεργών καμπύλων. Ενώ στην (6.1.2) παρουσιάζεται η τελική θέση του φιδιού με την χρήση του συνδυασμένου μοντέλου ενεργών καμπύλων και στατιστικού μοντέλου

Γίνεται άμεσα αντιληπτό από τις εικόνες, ότι η προσέγγιση της καμπύλης με το συνδυασμό των δυο μοντέλων βελτιώνεται, περιορίζομενη μόνο σε παραλλαγές σχημάτων που υπάρχουν στο σύνολο εκπαίδευσης. Οι εικόνες που παρουσιάζονται προέρχονται από δυο διαφορετικά πείραμα. Οι εικόνες (6.1) και (6.2) προέρχονται από ένα πείραμα, στο οποίο το περίγραμμα του χεριού είναι με ακρίβεια ορισμένο. Επειδή η αρχικοποίηση του φιδιού είναι ιδιαίτερα κοντά στην επιθυμητή θέση και στη περίπτωση της χρήσης μόνο του μοντέλου των ενεργών καμπυλών, η προσέγγιση του περιγράμματος είναι αρκετά ικανοποιητική, όπως φαίνεται στην εικόνα (6.2.1). Παρόλα αυτά παρατηρούμε ότι στις άκρες των δακτύλων παρουσιάζονται ατέλειες, όπως και στις καμπυλώσεις μεταξύ των δακτύλων. Αντίθετα στην εικόνα (6.2.2) εμφανίζεται το αποτέλεσμα του συνδυασμού των δυο μοντέλων, το οποίο είναι βελτιωμένο. Πρέπει να λάβουμε ακόμα υπόψη ότι οι ατέλειες αυτές που εμφανίζονται θα



6.2.1 Εικόνα Ενεργής Καμπύλης



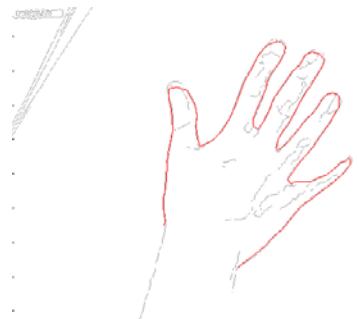
6.2.2 Εικόνα Περιορισμένης Ενεργής Καμπύλης

Εικόνα 6.2: Στις εικόνες αυτές απεικονίζεται η βελτίωση της προσέγγισης του χεριού η οποία προκύπτει με το περιορισμό της ενεργής καμπύλης στον επιτρεπόμενο χώρο σχήματος

λειτουργούν σχεδόν αθροιστικά κατά την εφαρμογή της παρακολούθησης του χεριού, αφού η τελική θέση του ενός καρέ θα αποτελεί την αρχή του άλλου και κατά συνέπεια μετά από έναν αριθμό καρέ το σχήμα του φιδιού να έχει χάσει τη μορφή του χεριού.

Στην επόμενη σειρά εικόνων (εικόνα: 6.3) απεικονίζεται μια άλλη εικόνα χεριού, της οποίας η φωτεινότητα και το φόντο δεν επιτρέπουν στη μέθοδο εύρεσης ακμών να προσδιορίσει με σαφήνεια το περίγραμμα του χεριού. Στη περιοχή του μικρού δακτύλου ένα κομμάτι του εξωτερικού περιγράμματος απουσιάζει. Αυτό έχει ως άμεση συνέπεια η τελική θέση της καμπύλης που εξελίσσεται χωρίς τους στατιστικούς περιορισμούς να προσεγγίζει μια εσωτερική ακμή που εμφανίζεται, όμοια προβλήματα εμφανίζονται και στον αντίχειρα, όπως γίνεται αντιληπτό στις εικόνες (6.3.1) και (6.3.3). Τα προβλήματα αυτά εξαλείφονται με την εισαγωγή του στατιστικού μοντέλου και το σύνολο του συνόρου του χεριού προσεγγίζεται με ακρίβεια, όπως φαίνεται στις εικόνες (6.3.2) και (6.3.4).

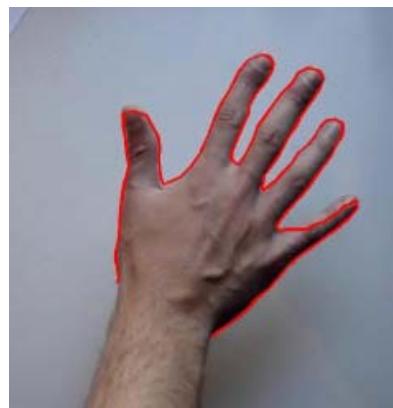
Συμπερασματικά, λοιπόν, η εισαγωγή του στατιστικού μοντέλου έχει ως αποτέλεσμα το κλασσικό μοντέλο ενεργών καμπυλών να περιορίζεται και η μορφή των σχημάτων που λαμβάνει να είναι παραπλήσια με εκείνα που υπάρχουν στο σύνολο εκπαίδευσης. Τα πλεονεκτήματα που προκύπτουν είναι αρ-



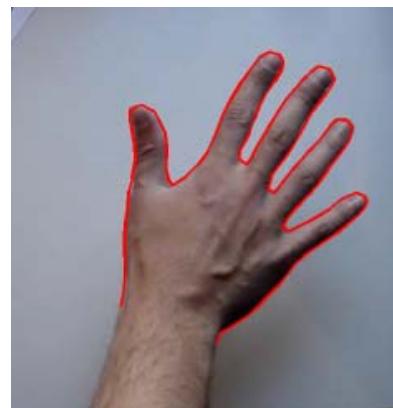
6.3.1 Φίδι χωρίς Περιορισμούς



6.3.2 Χρήση Στατιστικού Μοντέλου



6.3.3 Εικόνα Ενεργής Καμπύλης



6.3.4 Εικόνα Περιορισμένης Ενεργής Καμπύλης

Εικόνα 6.3: Στην εικόνα (6.3.1) παρουσιάζεται η θέση της καμπύλης στην εικόνα ακμών, η οποία δεν ορίζει με ακρίβεια το περίγραμμα του χεριού, στην συνέχεια εμφανίζεται η σαφώς βελτιωμένη “περιορισμένη καμπύλη” στην (6.3.2). Στις δύο επόμενες φαίνεται με σαφήνεια το πλεονέκτημα της προβολής στον επιτρεπτό υποχώρο

χετά και γίνονται αντιληπτά και στις παραπάνω εικόνες αλλά και στις εικόνες του επόμενου κεφαλαίου. Όμως, η εισαγωγή του μοντέλου αυτού προϋποθέτει την κατασκευή του συνόλου εκπαίδευσης, το οποίο όπως είπαμε και στο Κεφαλαίο 5 είναι χρονοβόρο και απαιτεί κάποιες ιδιαίτερες γνώσεις από τον χειριστή. Πρέπει να τονιστεί εδώ ότι στην δική μας εφαρμογή η διαδικασία τοποθέτησης των ορόσημων γίνεται ακόμα πιο σημαντική αφού τα διακριτά αυτά σημεία είναι εκείνα που θα παραμείνουν και κατά την εξέλιξη της καμπύλης και επομένως πρέπει να οριοθετούν όλα τα “κρίσιμα σημεία” του χεριού. Βέβαια, η διαδικασία κατασκευής του συνόλου εκπαίδευσης δεν είναι απαραίτητο να είναι εξατομικευμένη, αλλά είναι σημαντικό η χειρονομία του χεριού που επιθυμούμε να προσεγγίσουμε να υπάρχει στο σύνολο εκπαίδευσης, ώστε το μοντέλο να έχει την αντίστοιχη συνιστώσα παραμόρφωσης.

Κεφάλαιο 7

Αλγόριθμος και Πειραματική Διαδικασία

Στο κεφαλαίο αυτό θα παρουσιαστούν, τα βήματα του συνολικού αλγόριθμου, ένας αριθμός από πειράματα που διεξήχθησαν και θα αναλυθούν τα αποτελέσματα και οι διάφοροι περιορισμοί που προέκυψαν για την εφαρμογή που αναπτύχθηκε στα πλαίσια της εργασίας αυτής.

7.1 Αλγόριθμος

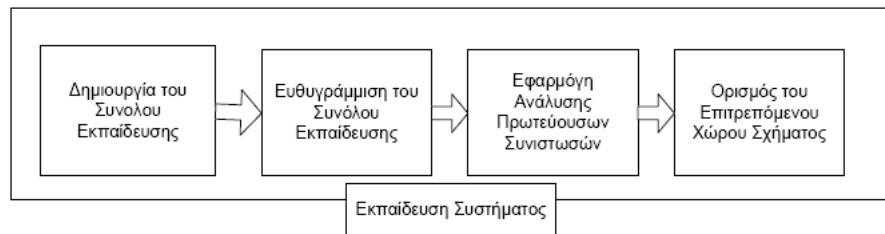
Ο αλγόριθμος της μεθοδολογίας παρακολούθησης χεριού που προτείνουμε μπορεί να χωριστεί σε δύο στάδια:

- Στάδιο Εκπαίδευσης
- Στάδιο Εφαρμογής

Κατά το στάδιο της εκπαίδευσης δημιουργείται το στατιστικό μοντέλο, όπως αυτό περιγράφεται στο Κεφάλαιο 5, ενώ κατά το στάδιο της εφαρμογής γίνεται η αρχικοποίηση της ενεργής καμπύλης, η εξέλιξη της πάνω στην δυναμική επιφάνεια της επεξεργασμένης εικόνας και η προβολή της καμπύλης στο επιτρεπτό χώρο σχήματος.

7.1.1 Στάδιο Εκπαίδευσης

Το πρώτο βήμα στο στάδιο εκπαίδευσης είναι η επιλογή των εικόνων βάσει των οποίων θα δημιουργηθεί το σύνολο εκπαίδευσης. Υπάρχουν δύο επιλογές για τη δημιουργία του συνόλου εκπαίδευσης, είτε να είναι εξειδικευμένο σε



Εικόνα 7.1: Το διάγραμμα ροής του Σταδίου Εκπαίδευσης

μια ακολουθία εικόνων, είτε να είναι γενικευμένο με διάφορες εικόνες χεριών. Όπως, είναι φυσικό η περίπτωση του εξειδικευμένου συνόλου εκπαίδευσης είναι περισσότερο αποτελεσματική, επειδή το στατιστικό μοντέλο είναι ειδικευμένο στο σχήμα του χεριού που επιθυμούμε να ακολουθήσουμε και οι κύριες συνιστώσες του περιέχουν μόνο τις παραμορφώσεις που θα εμφανιστούν στη δεδομένη ακολουθία εικόνων. Φυσικά, το μειονέκτημα που προκύπτει είναι ότι σε κάθε εφαρμογή πρέπει να δημιουργείται ένα ανεξάρτητο σύνολο εκπαίδευσης. Σε παρόμοιες εφαρμογές στατιστικού μοντέλου, όπως το Μοντέλο Ενεργών Σχημάτων [11], το γενικευμένο σύνολο εκπαίδευσης δημιουργείται από περισσότερες από 100 εικόνες που έχουν διαφορετικά σχήματα και παραμορφώσεις διάφορων χεριών. Λόγω, λοιπόν, και της χρονοβόρου διαδικασίας τοποθέτησης των ορόσημων προτιμήθηκε η δημιουργία εξειδικευμένων συνόλων εκπαίδευσης.

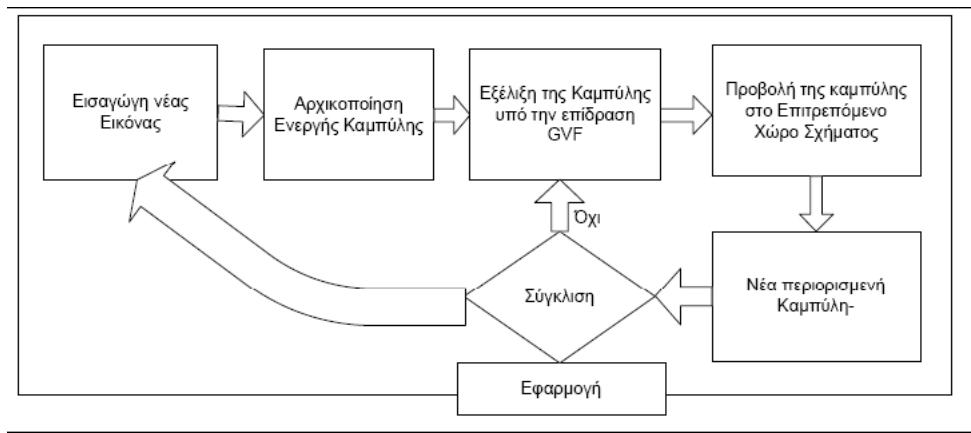
Σημαντικό, ακόμα, είναι οι εικόνες που επιλέγονται να απεικονίζουν τις κύριες παραμορφώσεις που επιθυμούμε να προσδιορίσουμε και να ακολουθήσουμε κατά την εφαρμογή. Πρέπει το σύνολο εκπαίδευσης να περιέχει δηλαδή τις εικόνες κλειδιά στην διαδικασία παραμόρφωσης.

Το δεύτερο βήμα είναι η τοποθέτηση των χαρακτηριστικών σημείων – ορόσημων στις εικόνες. Η απόσταση των ορόσημων είναι απαραίτητο να είναι τέτοια ώστε τα γειτονικά snaxels να προσδιορίζουν παραπλήσια μέρη του αντικειμένου. Όπως αναφέρθηκε στο προηγούμενο κεφαλαίο, επειδή δεν παρεμβάλλουμε νέα σημεία μεταξύ των ορόσημων, αυτό έχει σαν αποτέλεσμα να είναι αναγκαία η προσεκτική τοποθέτηση των ορόσημων. Ειδικότερα, στις άκρες των δακτύλων και στις περιοχές ανάμεσα στα δάχτυλα, πρέπει να τοποθετούνται τόσα σημεία όσα για να περιγραφεί με λεπτομέρεια το σχήμα της συγκεκριμένης περιοχής. Στα πειράματα που διεξήχθησαν ο αριθμός των

σημείων κυμαίνοταν από 102 μέχρι 118.

Μετά τη κατασκευή του συνόλου εκπαίδευσης, ακολουθεί η διαδικασία ευθυγράμμισης των σχημάτων και η στατιστική ανάλυση των μετασχηματισμένων σχημάτων, δηλαδή, ο προσδιορισμός του μέσου σχήματος και η Ανάλυση Πρωτευουσών Συνιστώσων. Στο σημείο αυτό ορίζουμε και το ποσοστό της διακύμανσης που επιθυμούμε το σύστημα να μοντελοποιεί. Κατά την πειραματική διαδικασία το ποσοστό που επιλέξαμε είναι το 97%, το οποίο πρακτικά σημαίνει την μείωση κατά μια διάσταση του πίνακα των ιδιοδιανυσμάτων, άρα και τον επιτρεπτό χώρο σχήματος.

7.1.2 Στάδιο Εφαρμογής



Εικόνα 7.2: Το διάγραμμα ροής του Σταδίου Εφαρμογής

Το πρώτο βήμα στο στάδιο της Εφαρμογής, είναι η εισαγωγή της αρχικής εικόνας από την ακολουθία των εικόνων. Η θέση, στην οποία θα αρχικοποιηθεί η ενεργή καμπύλη, είναι η βασική δυσκολία που συναντάει ο αλγόριθμος, εάν η καμπύλη βρίσκεται σε μεγάλη απόσταση από το χέρι το πεδίο δυνάμεων GVF αδυνατεί να κατευθύνει το φίδι στην επιθυμητή θέση. Για τον λόγο αυτό, η αρχικοποίηση στο πρώτο καρέ γίνεται χειρωνακτικά. Συνήθως, το πρώτο καρέ ήταν μέρος του συνόλου εκπαίδευσης και η αρχικοποίηση ταυτίζόταν με το διάνυσμα των ορόσημων που είχε προκύψει κατά την διάρκεια του σταδίου εκπαίδευσης. Για τα επόμενα καρέ η αρχικοποίηση της καμπύλης είναι η τελική θέση της στο προηγούμενο ακριβώς καρέ. Δυσκολίες παρουσιάζονται, όπως θα φανεί παρακάτω κατά την πειραματική διαδικασία, όταν η κίνηση του χεριού

είναι σχετικά γρήγορη.

Στο σημείο αυτό, γίνεται η επεξεργασία της εικόνας εισόδου, ώστε να εξαχθούν τα χαρακτηριστικά που επιθυμούμε η καμπύλη να προσδιορίσει. Στην εικόνα, λοιπόν, εφαρμόζεται αλγόριθμος εύρεσης ακμών και στην συνέχεια, με τη βοήθεια του αλγορίθμου που αναλύθηκε, απαλείφουμε τις ακμές μικρότερου μήκους.

Στο επόμενο βήμα η ενεργή καμπύλη εξελίσσεται υπό την επίδραση του πεδίου GVF σύμφωνα με το μοντέλο των ενεργών καμπύλων. Η επίλυση των εξισώσεων για την επίτευξη της συνθήκης ισορροπίας δυνάμεων γίνεται επαναληπτικά. Έτσι σε τακτά χρονικά διαστήματα προβάλλεται η καμπύλη στον επιτρεπτό χώρο σχήματος, η επιλογή των διαστημάτων αυτών κρίνει το βαθμό ελευθερίας που θέλουμε να έχει η καμπύλη από το στατιστικό μοντέλο. Προκύπτει, έτσι, μια νέα επιτρεπτή θέση της καμπύλης. Η νέα αυτή θέση συγχρίνεται με την προηγουμένη επιτρεπτή θέση της καμπύλης και εάν οι συντεταγμένες δύον των σημείων δεν έχουν αλλάξει σημαντικά, θεωρούμε ότι η συνθήκη ισορροπίας έχει επιτευχθεί και προχωρούμε στην ανάλυση του επόμενου καρέ.

7.2 Πειραματική Διαδικασία

Στις επόμενες παραγράφους θα παρουσιαστούν τα αποτελέσματα τριών ενδεικτικών πειραμάτων και θα σχολιαστούν τα αποτελέσματα τους.

7.2.1 Πείραμα 1^o

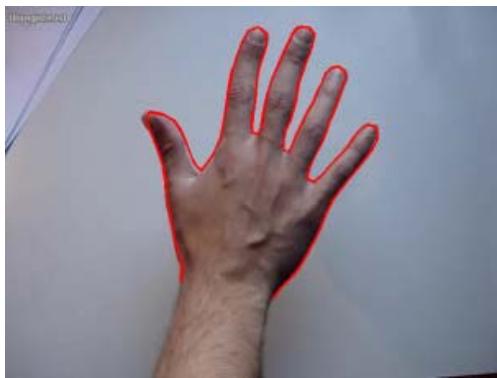
Η ακολουθία των εικόνων του πρώτου πειράματος απεικονίζει ένα χέρι το οποίο βρίσκεται σε ουδέτερο φόντο με μη ελεγχόμενο φωτισμό. Η φωτεινότητα της εικόνας οδηγεί, όπως θα φανεί, στη μη ακριβή εύρεση των ακμών του χεριού και άρα στην μη ορθή δημιουργία του πεδίου δυνάμεων.(Αναλυτικά έχει παρουσιαστεί το πρόβλημα αυτό στην Εικόνα (6.3)). Το σχήμα του χεριού στην ακολουθία των συγκεκριμένων εικόνων δεν παραμορφώνεται αισθητά, η κίνηση του είναι βασικά μεταφορική και περιστροφική. Αυτό έχει ως αποτέλεσμα ο αριθμός των συνιστωσών κίνησης που προκύπτει από το σύνολο εκπαίδευσης να είναι σχετικά μικρός : τρία ιδιοδιανύσματα περιγράφουν το χώρο σχήματος.



7.3.1



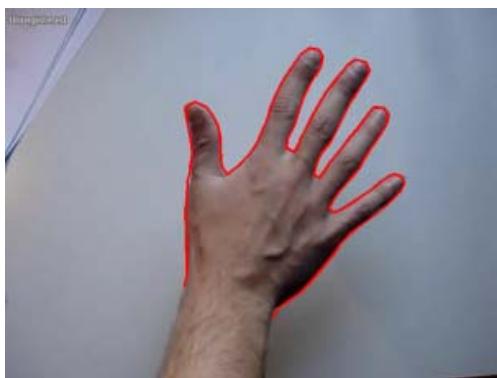
7.3.2



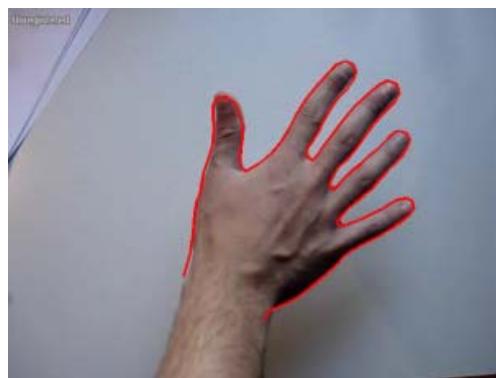
7.3.3



7.3.4

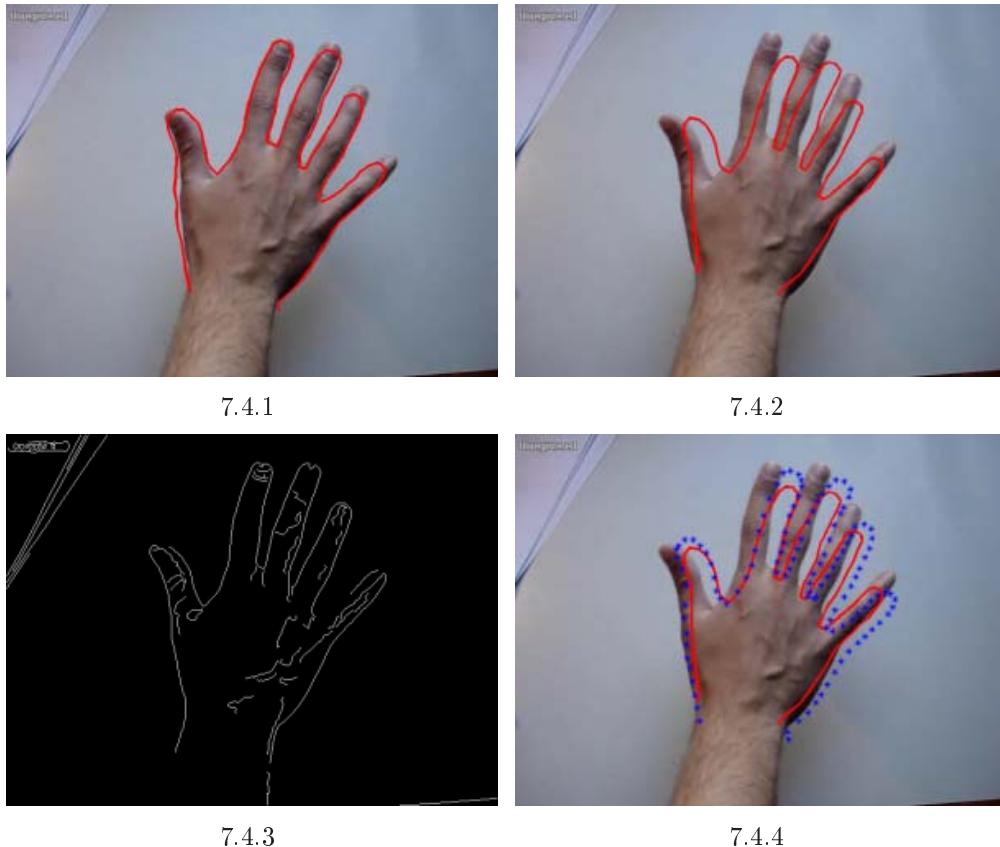


7.3.5



7.3.6

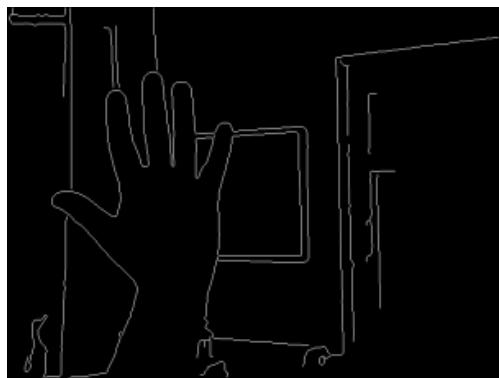
Εικόνα 7.3: Στις εικόνες αυτές απεικονίζονται τα αποτελέσματα του πρώτου πειράματος της παραχολούθησης του χεριού



Εικόνα 7.4: Στις εικόνες αυτές απεικονίζεται η αδυναμία του συστήματος να παρακολουθήσει τη πορεία του χεριού λόγω της γρήγορης ταχύτητας της μεταφορικής κίνησης. Στην εικόνα (7.4.1) εμφανίζεται η τελική θέση του φιδιού στο ακριβώς προηγούμενο χαρέ από το χαρέ εμφάνισης της αδυναμίας. Στην (7.4.3) απεικονίζεται η τελική θέση της καμπύλης στην εικόνα αδυναμίας και στην (7.4.3) παραθέτουμε την αντίστοιχη εικόνα ακμών. Στην (7.4.4) παρουσιάζεται με μπλε αστεράκια η αρχικοποίηση και με κόκκινο η τελική θέση. Παρατηρούμε λοιπόν ότι το πρόβλημα είναι περίπλοκο λόγω της μεγάλης μετατόπισης: η αρχική θέση της καμπύλης δεν περικλύει σχεδόν κανένα δάκτυλο, κατά συνέπεια αδυνατεί στην συνέχεια να τα προσεγγίσει. Αξιοσημείωτο είναι ότι το σχήμα της καμπύλης παραμένει στον επιτρεπτό χώρο σχήματος και στα ακόλουθα χαρέ.

7.2.2 Πείραμα 2^o

Η ακολουθία των εικόνων του δευτέρου πειράματος παρουσιάζει πολύπλοκο φόντο, ενώ ο φωτισμός είναι τέτοιος ώστε το περίγραμμα του χεριού να διαγράφεται με ευχρίνεια. Η παρακολούθηση του χεριού γίνεται σχετικά δύσκολη στις περιοχές οπού οι ακμές του χεριού περιπλέκονται με τα σύνορα άλλων αντικειμένων, παρόλα αυτά όπως φαίνεται και στις παρακάτω εικόνες η ενεργή καμπύλη τείνει να περιγράψει την μορφολογία του χεριού.



7.5.1



7.5.2



7.5.3



7.5.4

Εικόνα 7.5: Στις εικόνες αυτές απεικονίζονται τα αποτελέσματα της παρακολούθησης του χεριού σε μη ουδέτερο φόντο



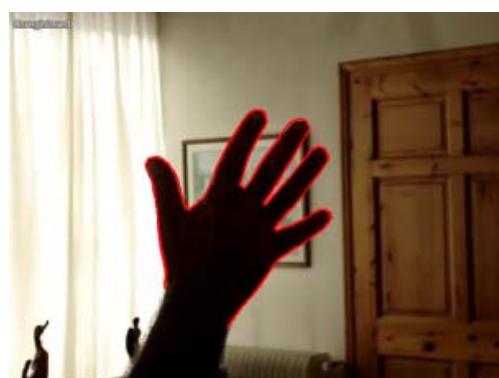
7.6.1



7.6.2



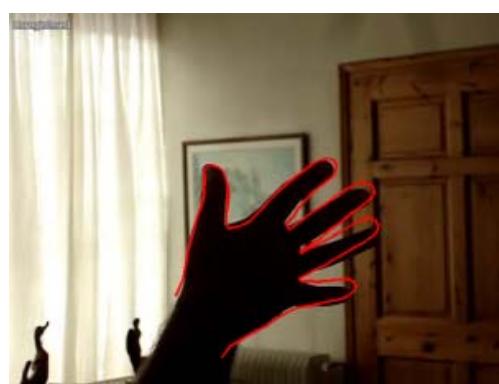
7.6.3



7.6.4



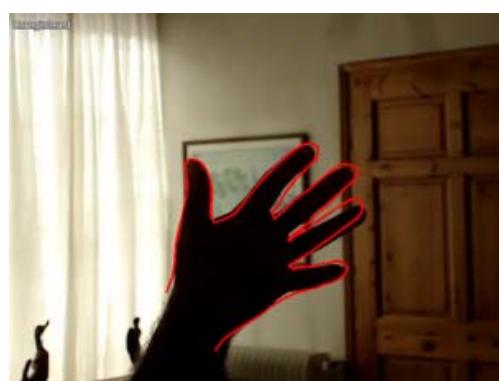
7.6.5



7.6.6



7.6.7



7.6.8

Εικόνα 7.6: Παρατίθενται στην αριστερή στήλη η εικόνα ακμών και στην δεξιά το αποτέλεσμα της μεθόδου

7.2.3 Πείραμα 3^o

Το τρίτο πείραμα εξετάζει την περίπτωση της παρακολούθησης του χεριού, το οποίο παραμορφώνεται σχηματίζοντας μια δεδομένη χειρονομία. Επειδή κατά την κίνηση αυτή η διακύμανση των σημείων είναι μεγάλη, η ανάλυση πρωτευουσών συνιστώσων οδηγεί στην εύρεση 7 κυρίων αξόνων μεταβολής. Το φόντο και η φωτεινότητα είναι κατάλληλες ώστε η εύρεση των ακμών να είναι ακριβής. Τα αποτελέσματα της μεθόδου παρουσιάζονται στην συνέχεια:



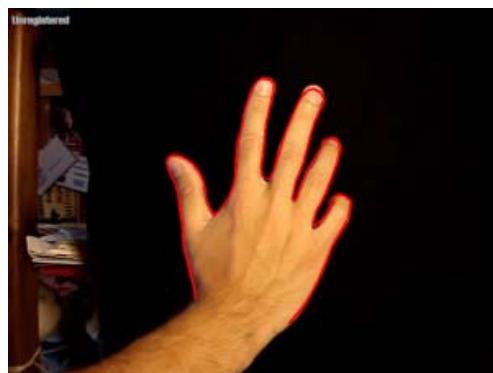
7.7.1



7.7.2

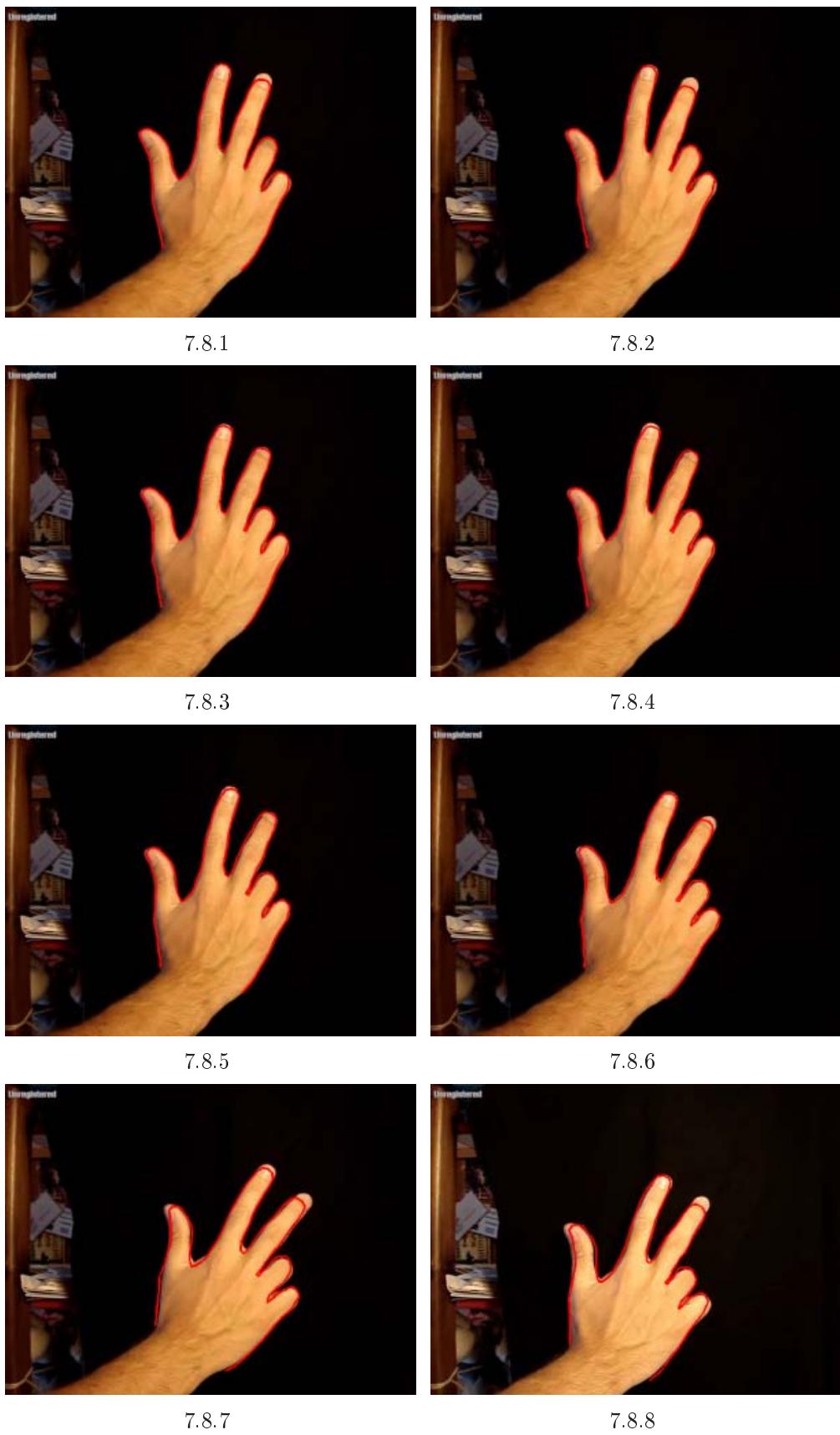


7.7.3



7.7.4

Εικόνα 7.7: Στις εικόνες αυτές απεικονίζεται το αποτέλεσμα της παρακολούθησης μιας χειρονομίας



Εικόνα 7.8: Ακολουθία εικόνων του 3^{ου} πειράματος

Στο σημείο αυτό πρέπει να επισημάνουμε ότι οι τιμές των παραμέτρων της ενεργής καμπύλης και στα τρία πειράματα είναι ίδιες: $a = 0.05, b = 0, \gamma = 1$ και ο συντελεστής βαρύτητας των εξωτερικών δυνάμεων είναι ίσος με 2.2. Αυτό που αλλάζει αλλά όχι σημαντικά είναι το χρονικό διάστημα που μεσολαβεί για την προβολή της καμπύλης στον επιτρεπτό χώρο σχήματος. Στο πρώτο πείραμα η προβολή γίνεται μετά από 10 εξέλιξεις του μοντέλου του φιδιού. Στο δεύτερο και στο τρίτο ανά 5 εξέλιξεις. Όσο πιο πολλές επαναλήψεις χωρίς προβολή γίνονται τόσο πιο ελεύθερο είναι το φίδι να λάβει διαφορετική μορφολογία από αυτή που παρουσιάζεται στο σύνολο εκπαίδευσης.

Κεφάλαιο 8

Συμπεράσματα και Κατευθύνσεις για Μελλοντική Έρευνα

8.1 Συμπεράσματα

Η ιδέα χρησιμοποίησης του μοντέλου των ενεργών καμπύλων σε συνδυασμό με το στατιστικό μοντέλο σχήματος αποδεικνύεται ικανή για την επιτυχή παρακολούθηση χειρονομιών. Φυσικά, η εφαρμογή παρουσιάζει πλεονεκτήματα και μειονεκτήματα, τα οποία παρουσιάστηκαν αναλυτικά στα προηγούμενα κεφάλαια. Συνοπτικά, παρατίθενται τα βασικότερα από αυτά.

8.1.1 Πλεονεκτήματα

Η μεθοδολογία προσδιορισμού των συνόρων αντικειμένων με την βοήθεια ενεργών καμπύλων ισχυροποιείται με την εισαγωγή στατιστικού μοντέλου. Με το περιορισμό της μορφολογίας της καμπύλης στον επιτρεπτό χώρο σχήματος, όπως αυτός έχει προκύψει από το σύνολο εκπαίδευσης, η ενεργή καμπύλη εξειδικεύεται στις παραμορφώσεις του σχήματος του χεριού. Λόγω της εξειδίκευσης αυτής η καμπύλη μπορεί :

- Να επιτύχει το προσδιορισμό των συνόρων του χεριού ακόμα και όταν η ακμές δεν είναι με σαφήνεια ορισμένες.
- Να μπορεί να προσεγγίσει τα σύνορα των δακτύλων ακόμα και όταν η αρχικοποίηση της καμπύλης τέμνει τα σύνορα αυτά. Αρκεί μέρος του μέσου άξονα του συνόρου να ανήκει εντός της αρχικοποίησης της καμπύλης.

- Να περιγράφει με ακρίβεια παραμορφώσεις για τις οποίες έχει εκπαιδευτεί ακόμα και όταν το φόντο της εικόνας είναι πολύπλοκο.

Οδηγούμαστε λοιπόν στο συμπέρασμα ότι το στατιστικό μοντέλο είναι απαραίτητο για την επίτευξη της προσέγγισης και παρακολούθησης του περιγράμματος ενός σχήματος τόσο πολύπλοκου όπως το χέρι. Επιπλέον, ένα θετικό στοιχείο της εφαρμογής είναι ότι κατά την πειραματική διαδικασία αποδείχθηκε ότι τα αποτελέσματα είναι ικανοποιητικά χωρίς να είναι απαραίτητη η μεταβολή των παραμέτρων του φιδιού. Όπως αναφέρθηκε και στην αντίστοιχη ενότητα, η παράμετρος της τάσης, a , διατηρείται ίση με 0.05 και η παράμετρος της ακαμψίας, b ίση με 0, ενώ ο συντελεστής τριβής γ είναι ίσος με 1 και ο συντελεστής βαρύτητας των εξωτερικών δυνάμεων ίσος με 2.

8.1.2 Μειονεκτήματα και Περιορισμοί

Η εφαρμογή που αναπτύχθηκε κληρονομεί μέρος των περιορισμών του μοντέλου ενεργών καμπύλων. Η εικόνα χρειάζεται ένα βαθμό επεξεργασίας ώστε η τοπολογία ακμών που παράγεται να είναι κατάλληλη ώστε να χρησιμοποιηθεί ως δυναμική επιφάνεια πάνω στην οποία θα κινηθεί η καμπύλη. Για το λόγο αυτό πρέπει στα διάφορα πειράματα να μεταβάλλεται το κατώφλι της μεθόδου εύρεσης ακμών. Γενικότερα, όμως, η εφαρμογή δεν μπορεί να επιτύχει ακριβή παρακολούθηση του χεριού όταν το φόντο της ακολουθίας των εικόνων είναι ιδιαίτερα περίπλοκο. Για παράδειγμα, εάν κλείσουμε κάποιο δάκτυλο και στην θέση του στο φόντο υπάρχει κάποιο άλλο αντικείμενο που δημιουργεί κάποια ακμή στο σημείο εκείνο, το πεδίο δυνάμεων θα αδυνατεί να κατευθύνει την καμπύλη προς το επιθυμητό σημείο.

Ένας άλλος περιορισμός, και πιθανόν ο σημαντικότερος, είναι η ανάγκη αυστηρής διαδικασίας αρχικοποίησης. Η ενεργή καμπύλη δεν μπορεί να τοποθετηθεί στα άκρα της εικόνας και στην συνέχεια να προσδιορίσει τα σύνορα των αντικειμένων στην εικόνα, όπως τα γεωδαιτικά ενεργά περιγράμματα. Η ενεργή καμπύλη πρέπει, είτε να είναι σχετικά κοντά στο περίγραμμα του χεριού και να περιέχει μέρος των δακτύλων, είτε να είναι ολόκληρη ή στο εσωτερικό του χεριού ή στον εξωτερικό χώρο χωρίς όμως να υπάρχουν έντονες ακμές σε αυτών εκτός του περιγράμματος του χεριού. Ο περιορισμός αυτός οδηγεί αρχικά σε πρόβλημα έναρξης της μεθόδου παρακολούθησης, επειδή στο πρώτο καρέ η καμπύλη πρέπει να τοποθετηθεί χειρωνακτικά, και στην συνέχεια οδηγεί σε αδυναμία παρακολούθησης γρήγορων μετατοπίσεων του χεριού.

Ακόμα, το στατιστικό μοντέλο εισάγει, εκτός από τα διάφορα πλεονεκτήματα, το μειονέκτημα ότι η ενεργή καμπύλη πλέον δεν μπορεί να προσεγγίσει κάποια τυχαία παραμόρφωση του χεριού που δεν περιγράφεται στο σύνολο εκπαίδευσης.

8.2 Συνδρομή Εργασίας

Η πρωτότυπη συνδρομή της παρούσας διπλωματικής εργασίας βρίσκεται στα ακόλουθα σημεία:

- Μέθοδος απαλοιφής των ακμών μικρού μήκους σε σχέση με το μήκος της μεγαλύτερης ακμής που εμφανίζεται στην εικόνα
- Μεθοδολογία προβολής της ενεργής καμπύλης στο επιτρεπτό χώρο σχήματος

Πέρα από αυτά τα στοιχεία στην εργασία αυτή αναλύθηκε θεωρητικά και υλοποιήθηκε αλγορίθμικά το μοντέλο των ενεργών καμπύλων και το στατιστικό μοντέλο σχήματος που αποτελεί την βάση και άλλων σημαντικών εφαρμογών όπως τα Active Shape Models και τα Active Appearance Models.

8.3 Κατευθύνσεις για Μελλοντική Ερευνά

Η έρευνα γύρω από τα θέματα που εξετάστηκαν αλλά και τους γενικότερου προβλήματος παρακολούθησης χειρονομιών φυσικά δεν ολοκληρώθηκε μέσα στα στενά πλαίσια της διπλωματικής εργασίας. Οι περιορισμοί της εφαρμογής που αναπτύχθηκε αλλά και προεκτάσεις αυτής αποτελούν πηγή έμπνευση των μελλοντικών κατευθύνσεων έρευνας:

- Εισαγωγή μεθοδολογίας εύρεσης του πιθανών θέσεων του χεριού για την αρχικοποίηση της καμπύλης στο πρώτο καρέ. Η μεθοδολογία αυτή μπορεί να στηρίζεται στην εύρεση του χρώματος του χεριού.
- Εισαγωγή πιθανοτικού μοντέλου για την πρόβλεψη της επόμενης θέσης της καμπύλης με σκοπό την επίλυση των προβλημάτων που παρουσιάζονται σε γρήγορες μετατοπίσεις. Η πρόβλεψη αυτή θα χρησιμοποιείται ως αρχικοποίηση του φιδιού για το επόμενο καρέ. Ήδη έχει γίνει έρευνα γύρω από extended και unscented kalman φίλτρα στα οποία όμως είναι αναγκαία η υπόθεση ενός μοντέλου κίνησης, π.χ. σταθερή ταχύτητας, με

αποτέλεσμα ούτε αυτά να εγγυώνται την επίλυση του προβλήματος. Μια νέα κατεύθυνση θα μπορούσε να είναι το Particle filtering ή Bayesian EM μέθοδος.

- Εύρεση μεθόδου προβολής της καμπύλης στον επιτρεπτό χώρο σχήματος χωρίς να εξαλείφεται η δυνατότητα παρεμβολής ή απαλοιφής σημείων στην καμπύλη.
- Έρευνα για τον τρόπο συνδυασμού του μοντέλου με Κρυφά Μαρκοβιανά μοντέλα με σκοπό την αναγνώριση της χειρονομίας. Μια ιδέα θα μπορούσε να βασιστεί στο ότι το διάνυσμα παρατήρησης του μοντέλου θα ταυτίζεται με το συντέλεση παραμέτρων b που προκύπτει κατά την προβολή της καμπύλης στο χώρο σχήματος.

Αυτές είναι πιθανές κατευθύνσεις έρευνας για να επιλυθούν μειονεκτήματα της μεθόδου αλλά και για να προεκταθούν οι δυνατότητες της.

Παράρτημα A

Δημιουργία Ανοιχτής Ενεργής Καμπύλης

Στο παράρτημα αυτό αναλύεται ο ακριβής τρόπος δημιουργίας ανοιχτής ενεργής καμπύλης, καταργώντας τις αλληλεπιδράσεις μεταξύ δυο διακριτών σημείων του φιδιού.

A.1 Εισαγωγή Ασυνέχειας Θέσης

Η ασυνέχεια μεταξύ δυο διακριτών σημείων (snaxels) $\bar{s} = i$ και $\bar{s} = i + 1$ επιτυγχάνεται με τη κατάργηση όλων των συνδέσμων που σχετίζονται με τα δυο αυτά σημεία. Για αυτό τον λόγο οι περιορισμοί τάσης και ακαμψίας που χαρακτηρίζουν τα δυο συγκεκριμένα σημεία της καμπύλης πρέπει να μεταβληθούν κατά την κατασκευή των εξισώσεων των κόμβων. Καθώς κάθε εξίσωση κόμβου συνδέει τα τέσσερα γειτονικά snaxel, επηρεάζονται οι εξισώσεις των κόμβων $\bar{s} = i - 1$, $\bar{s} = i$, $\bar{s} = i + 1$ και $\bar{s} = i + 2$. Στην υλοποίηση της εφαρμογής μας το σπάσιμο της καμπύλης συμβαίνει μεταξύ του πρώτου $\bar{s} = 0$ και τελευταίου $\bar{s} = M_{\bar{s}} - 1$ διακριτού σημείου. Στο $\bar{s} = 1$ κόμβο πρέπει να “σπάσουμε” έναν από τους συνδέσμους. Στους κόμβους $\bar{s} = 0$ και τελευταίου $\bar{s} = M_{\bar{s}} - 1$, λόγω συμμετρίας, πρέπει να σπάσουμε τρεις συνδέσμους, ενώ στο κόμβο $\bar{s} = M_{\bar{s}} - 2$ πρέπει να διασπαστεί ένα σύνδεσμος.

Από τις τροποποιημένες εξισώσεις των κόμβων προκύπτουν τα νέα στοιχεία

του πίνακα ακαμψίας \mathbf{K} . Στο κόμβο $\bar{s} = 1$ αλλάζουν τρία στοιχεία:

$$\begin{aligned} c_1 &= b(M_{\bar{s}} - 1) + 4b(0) + a(0) + a(1), \\ \beta_1 &= -2b(0) - a(1), \\ \alpha_1 &= 0. \end{aligned} \tag{A.1}$$

Στο κόμβο $\bar{s} = 0$ και τα τέσσερα στοιχεία του πίνακα μεταβάλλονται, το ένα έχει ήδη αναφερθεί (β_1), ως εξής:

$$\begin{aligned} c_0 &= b(1) + a(1), \\ \beta_0 &= -2b(1) - a(1), \\ \alpha_0 &= 0. \end{aligned} \tag{A.2}$$

Όμοια, στο κόμβο $\bar{s} = M_{\bar{s}} - 1$, μεταβάλλονται τέσσερα στοιχεία:

$$\begin{aligned} c_{M_{\bar{s}}-1} &= b(M_{\bar{s}} - 2) + a(M_{\bar{s}} - 2), \\ \beta_{M_{\bar{s}}-1} &= -2b(M_{\bar{s}} - 2) - a(M_{\bar{s}} - 2), \\ \alpha_{M_{\bar{s}}-1} &= 0. \end{aligned} \tag{A.3}$$

Τέλος, για το κόμβο $M_{\bar{s}} - 2$ έχουμε ήδη προσδιορίσει τα $\beta(M_{\bar{s}} - 1)$ και $\alpha(0)$ και το διαγώνιο στοιχείο είναι:

$$c_{M_{\bar{s}}-2} = 4b(M_{\bar{s}} - 2) + b(M_{\bar{s}} - 3) + a(M_{\bar{s}} - 2) + a(M_{\bar{s}} - 3) \tag{A.4}$$

Παράρτημα B

Ευθυγράμμιση Δύο Σχημάτων

Δοσμένων δυο παρόμοιων σχημάτων, x_1 και x_2 επιθυμούμε την εύρεση των κατάλληλων παραμέτρων περιστροφής θ , αλιμάκωσης s και μετατόπισης (t_x, t_y) με σκοπό να οδηγήσουμε το x_2 με τέτοιο τρόπο πάνω στο $M(x)$ ώστε να ελαχιστοποιήσουμε το παρακάτω άθροισμα:

$$E_j = (x_j - M_j[x_j])^T \mathbf{W} (x_j - M_j[x_j]) \quad (\text{B.1})$$

όπου

$$M_j \begin{pmatrix} x_{jk} \\ y_{ik} \end{pmatrix} = \begin{pmatrix} (s_j \cos \theta)x_{jk} - (s_j \sin \theta)y_{jk} + t_{jx} \\ (s_j \sin \theta)x_{jk} + (s_j \cos \theta)y_{jk} + t_{jy} \end{pmatrix} \quad (\text{B.2})$$

και \mathbf{W} είναι ένας διαγώνιος πίνακας με βάρη για κάθε σημείο-ορόσημο.

Αν θέσουμε

$$\alpha_x = s \cos \theta \quad \alpha_y = s \sin \theta \quad (\text{B.3})$$

τότε η προσέγγιση ελαχίστων τετραγώνων (παραγωγίζοντας ως προς τις μεταβλητές $\alpha_x, \alpha_y, t_x, t_y$) οδηγεί σε ένα σύνολο από τέσσερις γραμμικές εξισώσεις:

$$\begin{pmatrix} X_2 & -Y_2 & W & 0 \\ Y_2 & X_2 & 0 & W \\ Z & 0 & X_2 & Y_2 \\ 0 & Z & -Y_2 & X_2 \end{pmatrix} \begin{pmatrix} \alpha_x \\ \alpha_y \\ t_x \\ t_y \end{pmatrix} = \begin{pmatrix} X_1 \\ Y_1 \\ C_1 \\ C_2 \end{pmatrix} \quad (\text{B.4})$$

όπου

$$X_i = \sum_{k=0}^{n-1} w_k x_{ik} \quad Y_i = \sum_{k=0}^{n-1} w_k y_{ik} \quad (\text{B.5})$$

$$Z = \sum_{k=0}^{n-1} w_k (x_{2k}^2 + y_{2k}^2) \quad W = \sum_{k=0}^{n-1} w_k \quad (\text{B.6})$$

$$C_1 = \sum_{k=0}^{n-1} w_k (x_{1k} x_{2k} + y_{1k} y_{2k}) \quad (\text{B.7})$$

$$C_2 = \sum_{k=0}^{n-1} w_k (y_{1k} x_{2k} - x_{1k} y_{2k}) \quad (\text{B.8})$$

Βιβλιογραφία

- [1] David S. Cohen Alan L. Yuille, Peter W. Hallinan. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 8(2):99–111, Aug 1992.
- [2] Vassilis Athitsos and Stan Sclaroff. 3D hand pose estimation by finding appearance-based matches in a large database of training views. Technical Report 21, Computer Science Department, Boston University, 2001.
- [3] I. Herlin B. Leroy and L. Cohen. Multi-resolution algorithms for active contour models. *Proc. 12th Int'l Conf. Analysis and Optimization of Systems*, 1996.
- [4] A. Blake and A. Zisserman. Visual reconstruction. *Cambridge, MA: MIT Press*, 1987.
- [5] Andrew Blake and M. Isard. *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.
- [6] F.L. Bookstein. Landmark methods for forms without landmarks: Localizing group differences in outline shape. *Medical Image Analysis*, pages 1:225–243, 1997.
- [7] Lars Bretzner, Ivan Laptev, and Tony Lindeberg. Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering.
- [8] J. Canny. A computational approach to edge detection. *IEEE Transaction Pattern Analysis Machine Intelligence*, 8(6):679–698, 1986.

- [9] L. D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing. Image Understanding*, 53(2):211–218, 1991.
- [10] L.D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *PAMI*, 15(11):1131–1147, November 1993.
- [11] T. Cootes and C. Taylor. Active shape models- - smart snakes. 1992.
- [12] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Training models of shape from sets of examples. pages 266–275, 1992.
- [13] Y. Cui and J. Weng. Hand segmentation using learningbased prediction and verification for hand sign recognition, 1996.
- [14] I. L. Dryden and K. V. Mardia. Statistical shape analysis.
- [15] David A. Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
- [16] G. Hamarneh and T. Gustavsson. Combining snakes and active shape models for segmenting the human left ventricle in echocardiographic images, 2000.
- [17] D.C. Hogg. *Model-based vision: A program to see a walking person*. In Image and Vision Computing, 1983.
- [18] Terzopoulos D Kass M, Witkin A. Snakes: active contour models. pages 259–68. In: First international conference on computer vision, 1987.
- [19] Ivan Laptev and Tony Lindeberg. Tracking of multi-state hand models using particle filtering and a hierarchy of multi-scale image features. *Lecture Notes in Computer Science*, 2106:63+, 2001.
- [20] F. Leymarie. Tracking and describing deformable objects using active contour models. Master’s thesis, Montreal, Canada, 1990.
- [21] R. Lockton and A. Fitzgibbon. Real-time gesture recognition using deterministic boosting, 2002.

- [22] John MacCormick and Andrew Blake. A probabilistic exclusion principle for tracking multiple objects. *International Journal of Computer Vision*, 39(1):57–71, 2000.
- [23] John MacCormick and Michael Isard. Partitioned sampling, articulated objects, and interface-quality hand tracking. pages 3–19, 2000.
- [24] I. Matthews and S. Baker. Active appearance models revisited, 2003.
- [25] N. Ray, S. Acton, and K. Ley. Tracking leukocytes in vivo with shape and size constrained active contours.
- [26] Nilanjan Ray. Active contours for cell tracking.
- [27] James M. Rehg and Takeo Kanade. Visual tracking of high DOF articulated structures: an application to human hand tracking. In *ECCV* (2), pages 35–46, 1994.
- [28] James M. Rehg and Takeo Kanade. Visual tracking of high DOF articulated structures: an application to human hand tracking. In *ECCV* (2), pages 35–46, 1994.
- [29] R. Rosales, V. Athitsos, L. Sigal, and S. Sclaroff. 3d hand pose estimation using specialized mappings. Technical report, Los Alamitos, CA, July 9–12 2001.
- [30] R. Stenger. *Model-Based Hand Tracking Using A Hierarchical Bayesian Filter*. PhD thesis.
- [31] B. Stenger, A. Thayanathan, P. Torr, and R. Cipolla. Hand pose estimation using hierarchical detection. 2004.
- [32] Bjorn Stenger. 3d model-based hand tracking.
- [33] D. Terzopoulos. On matching deformable models to images. In *Optical Society of America, Topical Meeting on Machine Vision p 160-163 (SEE N89-19145 11-74)*, pages 160–163, 1980.
- [34] Demetri Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(4):413–242, 1986.

- [35] Demetri Terzopoulos and Richard Szeliski. Tracking with kalman snakes. pages 3–20, 1993.
- [36] DőArcy W. Thompson. On growth and form.
- [37] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.
- [38] C. Xu and J. Prince. Gradient vector flow: A new external force for snakes. In *Proceedings of Computer Vision and Pattern Recognition (CVPR '97)*, pages 66–71, San Juan, Puerto Rico, June 1997. IEEE.